

Leverage Effect, Volatility Feedback, and Self-Exciting Market Disruptions

Disentangling the Multi-dimensional Variations in S&P 500 Index Options

Peter Carr^{a,b*} Liuren Wu^{c†}

^a*Bloomberg L.P., 731 Lexington Avenue, New York, NY 10022, USA*

^b*Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA*

^c*Baruch College, Zicklin School of Business, One Bernard Baruch Way, New York, NY 10010, USA*

This version: March 12, 2009; First draft: September 4, 2007

Abstract

The equity index and index volatility interact through several distinct channels. First, holding business risk fixed, an increase in the level of financial leverage raises the level of the equity volatility. Second, regardless of the level of financial leverage, a positive shock to business risk increases the cost of capital and reduces the valuation of future cash flows, generating an instantaneous negative correlation between asset returns and asset volatility. Finally, the market experiences both small continuous movements and large market disruptions. The large and negative market disruptions often generate self-exciting behaviors. The occurrence of one disruption induces more disruptions to follow, thus raising market volatility. We propose an equity index dynamics that capture all three channels of interactions through the separate modeling of the asset return dynamics and the financial leverage variation. We analyze how the different sources of variations impact the index options behaviors differently across a wide range of strikes, maturities, and calendar days.

JEL Classification: F34, G12, G13.

Keywords: Option pricing; implied volatility; leverage effect; volatility feedback; self-exciting; market disruptions; jumps; constant elasticity of variance; time-changed Lévy processes; Fast Fourier Transform; Gauss-Hermite quadrature rule; unscented Kalman filter.

We thank Bruno Dupire, John Hull, Dilip Madan, Tom McCurdy, Matthew Richardson, Allen White, participants at Bloomberg, University of Toronto, and the 2008 Princeton conference on implied volatility models at Huntington Beach, California for comments. We also thank Sergey Nadtochiy for research assistance and Richard Holowczak for computing support. We welcome comments, including references to related papers we have inadvertently overlooked. All remaining errors are ours.

*Tel.: +1-212-617-5056; fax: +1-917-369-5629. *E-mail address:* pcarr4@bloomberg.com.

†Tel.: +1-646-312-3509; fax: +1-646-312-3451. *E-mail address:* liuren_wu@baruch.cuny.edu.

1. Introduction

Equity index returns and volatilities show negative co-movements. The academic literature mostly models the co-movements through a negative instantaneous correlation between innovations in equity returns and return variance, e.g., Heston (1993), Bates (1996, 2000), Bakshi, Cao, and Chen (1997), Heston and Nandi (2000), Pan (2002), Carr and Wu (2004), Eraker (2004), and Huang and Wu (2004). All these models maintain scale-free dynamics: Changing the scales or units of the equity index does not alter the dynamic specification for returns and volatilities.

On the other hand, models with direct index level dependence are popular in the industry. The local volatility model of Dupire (1994) specifies the return volatility as a function of the stock index level and time, and determines this local volatility function by the ratio of calendar spreads to butterfly spreads. A parsimonious functional choice for the local volatility is the constant elasticity of variance (CEV) specification, where the local volatility is a time-homogeneous power function of the index level, e.g., Beckers (1980), Cox (1996), Emanuel and MacBeth (1982), and Schroder (1989). By analyzing the time-series behavior of the Black and Scholes (1973) implied volatilities on the S&P 500 index options, Derman (1999) concludes that the data show different regimes, under which the implied volatility and the equity index show different dependence structures.

In this paper, we argue that the equity index and its volatility interact through several distinct channels. First, the equity return volatility does exhibit level dependence, but the dependence is not on the level of the stock price or market capitalization, but rather on the level of financial leverage. Holding fixed the business risk, increasing the level of financial leverage leads to a direct increase in the level of the equity return volatility. Furthermore, the financial leverage variation can come from the stock price variation when the debt level is fixed (Black (1976b)), but it can also come from active managerial decisions on the company's capital structure (Titman and Tompaidis (2005) and Adrian and Shin (2008)).

A second source of interaction can come from the volatility feedback effect on asset returns. Positive shocks to business risk increase the cost of capital and reduce the business valuation. Such a volatility feedback effect generates a negative correlation between shocks in asset returns and asset return volatilities. The effect exists regardless of the financial leverage level or its variation.

Yet another source of interaction can come from market turmoils. The financial market experiences both continuous movements and discontinuous disruptions. Large negative market disruptions can show self-exciting behaviors. The occurrence of one negative disruption often induces more disruptions to follow, thus raising market volatility.

We propose a model for the equity index dynamics that capture all three channels of interactions. We do so by separately modeling the asset return dynamics and the financial leverage variation. Specifically, we decompose the forward equity index level as a product of two nonnegative martingales that measure the forward asset value and the equity to asset ratio, respectively. We model the equity to asset ratio as following a constant elasticity of variance process. As the ratio declines, the level of financial leverage increases, and the index return volatility increases, thus generating the leverage effect. Furthermore, we decompose shocks in the asset return into two components, a continuous component and a discontinuous component. We apply separate time changes to the two components so that stochastic volatilities can come from both components separately. We specify the activity rate underlying the time change of the continuous component as a continuous Markov process, with its Brownian innovation negatively correlated with the Brownian innovation in the return component. This negative instantaneous correlation captures the volatility feedback effect: A positive shock to the business risk reduces the present valuation of future cash flows and is hence associated with a negative shock on business return. We specify the activity rate underlying the time change of the discontinuous component as a purely discontinuous Markov process, and we let the negative jumps in the asset return drive the positive shocks in the activity rate to generate the self-exciting behavior: A large downside jump on the asset return leads to a spike in the arrival rate of discontinuous movements in the future.

Under this specification, the index return volatility can be written as a direct function of the index level, thus generating the index-level dependence. Nevertheless, the index level in the volatility function is scaled by the asset value so that the volatility function becomes a unitless quantity and the return dynamics maintain the scale-free property, thus circumventing the major concerns on the consistency and stability of local volatility-type specifications. In addition to the index level dependence, the index return variance also varies with three additional state variables, including the asset level and the two activity rates underlying the continuous and discontinuous business shocks, respectively.

Through change of variables, we can represent the index return dynamics alternatively as driven by three

independent stochastic variance processes, with no direct index level dependence. Historically, researchers try to distinguish models with “sticky delta,” under which return volatility does not depend directly on the price level, from the local volatility-type models, under which the return variance has a direct dependence on the price level. Our specification generates a mixture behavior of both types and can be represented either in the sticky delta format or with direct price level dependence while maintaining the dynamics scale free. Thus, our specification bridges the gap between the two distinct streams of literature.

Given the specified dynamics, we propose a tractable procedure for pricing European options on the equity index. Conditional on the terminal value of the equity to asset ratio, options on the equity index can be written as options on the asset value. Since we model the asset return as time-changed Lévy processes, we can derive the generalized Fourier transform of the asset return following Carr and Wu (2004) and compute the conditional option value on the asset through fast Fourier inversion. Then, we numerically integrate the conditional option values on assets over the possible values of the equity to asset ratio. With the constant elasticity of variance specification, the probability distribution of the equity to asset ratio is known analytically as a transformation of the density for a standard Bessel process. The numerical integration can be performed efficiently using Gauss-Hermite quadrature rules.

We estimate the model using about over 12 years of over-the-counter option implied volatility quotes on the S&P 500 index from January 8 1997 to March 5, 2008. At each date, we have 40 quotes on a grid of five relative strikes from 80% to 120% of the spot index level and eight fixed time to maturities from one month to five years. Data analysis reveals three major sources of variation on the implied volatility surface that dictates the variation of the average implied volatility levels, the implied volatility term structure, and the implied volatility skew along the strike dimension. Our three-factor stochastic volatility structure can accommodate all three sources of variations. The model estimation results show that our model can price the equity index options well across the wide range of strikes and maturities. Through an analysis of the model parameter estimates and the state variables, we highlight the different roles played by the three different risk sources in the variation of the index option implied volatility surface.

The rest of the paper is organized as follows. The next section discusses the equity index dynamics specification, its different representations, and option pricing under the specified dynamics. Section 3 analyzes the data set and performs principal component analysis to determine the major sources of variations in the

implied volatility surface. Section 4 discusses the estimation strategy. Section 5 summarizes the estimation results. Section 6 concludes.

2. Separate modeling of asset return dynamics and financial leverage variation

We fix a filtered complete probability space $\{\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0}\}$ satisfying the usual technical conditions. We assume no-arbitrage in the economy. Then, under certain technical conditions, there exists at least one risk-neutral probability measure \mathbb{Q} , absolutely continuous with respect to \mathbb{P} , such that the gains process associated with any admissible trading strategy deflated by the riskfree rate is a martingale (Cochrane (2004), Duffie (1992), Harrison and Kreps (1979)).

Let F_t denote the time- t forward level of the equity index over some fixed time horizon and A_t the corresponding time- t forward asset value underlying the equity index, with $X_t \equiv F_t/A_t$ measuring the forward equity to asset ratio. We assume deterministic interest rates and dividend yields. By specifying the dynamics on the forward instead of the spot, we avoid the notational clustering associated with interest rates and dividends. Furthermore, instead of directly specifying the equity index dynamics, we propose to separately model the asset value dynamics and the financial leverage variation through the following multiplicative decomposition,

$$F_t = X_t A_t. \quad (1)$$

2.1. Equity-to-asset ratio dynamics and the leverage effect

To capture the financial leverage effect on equity valuation and equity volatility, we propose to model the equity-to-asset ratio by a simple constant elasticity of variance process. Assuming orthogonality between dX_t and dA_t , the equity-to-asset ratio is a martingale under the risk-neutral measure \mathbb{Q} , with its variation controlled by the following stochastic differential equation,

$$dX_t/X_t = \delta X_t^{-p} dW_t, \quad p > 0, \quad (2)$$

where W_t denotes a standard Brownian motion, δ is a positive constant, and the power coefficient p determines how the equity index return volatility varies on the level of financial leverage. When $p = 0$, the level of

financial leverage does not have an impact on the equity return volatility. When the power coefficient is positive $p > 0$, holding the asset value fixed, a decline in the equity-to-asset ratio X_t lowers the equity level, raises the financial leverage, and increases the equity return volatility, thus generating the well-known *leverage effect* described by Black (1976a). We maintain $p > 0$ to generate the leverage effect. For $0 < p \leq 1/2$, the origin (zero) is an exit boundary. For $p > 1/2$, the origin is a regular boundary point and is specified as a killing boundary by adjoining a killing boundary condition (Davydov and Linetsky (2001)).

The constant elasticity of variance process in (2) is related to a standard Bessel process of order $\nu = 1/(2p)$ through the change of variable, $z_t = X_t^p/(\delta p)$. From the well-known expression for the transition density of the Bessel process (see Borodin and Salminen (1996) and Revuz and Yor (1999) for details on Bessel processes), we can derive the probability transition density of X_T conditional on X_t , with $T > t$, as,

$$f(X_T|X_t) = \frac{X_T^{2p-\frac{3}{2}} X_t^{\frac{1}{2}}}{\delta^2 p(T-t)} \exp\left(-\frac{X_t^{2p} + X_T^{2p}}{2\delta^2 p^2(T-t)}\right) I_\nu\left(\frac{X_t^p X_T^p}{\delta^2 p^2(T-t)}\right), \quad (3)$$

where $I_\nu(x)$ is the modified Bessel function of the first kind of order ν . Since the value of the modified Bessel function increases quickly once its argument x becomes large, a modified version of the function $J_\nu(x) = I_\nu(x)e^{-x}$ can be calculated with more numerical stability, especially when x is large. In this case, we can rewrite the density function as,

$$f(X_T|X_t) = \frac{X_T^{2p-\frac{3}{2}} X_t^{\frac{1}{2}}}{\delta^2 p(T-t)} \exp\left(-\frac{(X_t^p - X_T^p)^2}{2\delta^2 p^2(T-t)}\right) J_\nu\left(\frac{X_t^p X_T^p}{\delta^2 p^2(T-t)}\right). \quad (4)$$

The constant elasticity of variance process has been used in the literature to model the equity price dynamics. Directly applied to equity prices, the process can generate the observed negative co-movements between equity prices and equity return volatilities, but it also generates scale-dependence in the dynamics: A re-scale of the equity price level necessitates a corresponding re-scaling of the dynamics to maintain stability. This scale-dependence issue does not show up in our specification as we are applying the process to a scale-free quantity, the equity-to-asset ratio.

2.2. Asset return dynamics with volatility feedback and self-exciting disruptions

To specify the asset value dynamics, we decompose the shocks to the asset return into two distinct types: (i) small continuous shocks, which we model as a standard Brownian motion Z_t , and (ii) large discontinuous shocks, which we model as a jump process J_t that includes both downside jumps J_t^- and upside jumps J_t^+ , with $J_t = J_t^+ + J_t^-$. We allow the two types of economic shocks to exhibit different stochastic intensities by applying separate stochastic time changes to the two Lévy components. Formally, we model the return on the asset under the risk-neutral measure \mathbb{Q} as,

$$\ln A_t/A_0 = \left(Z_{\mathcal{T}_t^Z} - \frac{1}{2} \mathcal{T}_t^Z \right) + \left(J_{\mathcal{T}_t^J} - \mathcal{T}_t^J k_J(1) \right), \quad (5)$$

where (Z_t, J_t) denote the two Lévy components that drive the two types of economic shocks on the asset return, and $(\mathcal{T}_t^Z, \mathcal{T}_t^J)$ denote the two stochastic time changes separately applied to the two Lévy components. We further assume that the two stochastic time changes are locally deterministic and defined through their respective activity rates (Carr and Wu (2004)),

$$\mathcal{T}_t^Z \equiv \int_0^t v_s^Z ds, \quad \mathcal{T}_t^J \equiv \int_0^t v_s^J ds. \quad (6)$$

We can think of t as the calendar time and $(\mathcal{T}_t^Z, \mathcal{T}_t^J)$ as the business clocks on the two types of economic shocks. The business clock runs faster when the underlying business activity is heavier and thus the activity rate is higher.

The terms $\frac{1}{2} \mathcal{T}_t^Z$ and $\mathcal{T}_t^J k_J(1)$ in equation (5) represent the convexity adjustments of the two shocks such that the forward asset value is an exponential martingale under the risk-neutral measure. The term $k_J(s)$ denotes the cumulant exponent of the Lévy jump process J_t defined as

$$k_J(s) \equiv \frac{1}{t} \mathbb{E} [e^{sJ_t}], \quad (7)$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator under the risk-neutral measure \mathbb{Q} . The cumulant exponent of a standard Brownian motion Z is $k_Z(s) = \frac{1}{2}s^2$. When the cumulant exponent $k_L(s)$ of a Lévy process L_t is well-defined, the compensated Lévy process $L_t - k_L(1)t$ becomes an exponential martingale as $\mathbb{E} [e^{L_t - k_L(1)t}] = 1$. By virtue of the optional stopping time theorem, replacing the calendar time t with a locally deterministic time

change \mathcal{T}_t as defined in (6) does not alter the martingale behavior of the exponential (Küchler and Sørensen (1997), p. 230).

We specify the Lévy jump components through their Lévy densities under the risk-neutral measure \mathbb{Q} ,

$$\pi_{J^+}(x) = e^{-x/v_{J^+}} x^{-1}, \quad \pi_{J^-}(x) = e^{-|x|/v_{J^-}} |x|^{-1}, \quad (8)$$

which govern the arrival rate of jumps of size x . The specification describes the variance-gamma Lévy jump process studied in Madan and Seneta (1990) and Madan, Carr, and Chang (1998). The jump arrival rate centers around zero and declines monotonically as the absolute jump size declines, with v_{J^+} and v_{J^-} controlling the scale of up and down jumps, respectively. The variance-gamma process is an infinite-activity jump specification that generates an infinite number of jumps within any finite interval, a specification that has generated better empirical performance in pricing equity index options than finite-activity jump specifications (Carr and Wu (2003) and Huang and Wu (2004)). Under the variance-gamma specification, the cumulant exponents of the upside and downside jumps are,

$$k_{J^+}(s) = -\ln(1 - sv_{J^+}), \quad k_{J^-}(s) = -\ln(1 + sv_{J^-}). \quad (9)$$

We specify the risk-neutral dynamics of the two activity rates (v_t^Z, v_t^J) via the following stochastic differential equations,

$$dv_t^Z = \kappa_Z (\theta_Z - v_t^Z) dt + \sigma_Z \sqrt{v_t^Z} dZ_t^v, \quad \mathbb{E}[dZ_t^v dZ_t] = \rho dt, \quad (10)$$

$$dv_t^J = \kappa_J (\theta_J - v_t^J) dt - \sigma_J dJ_{\mathcal{T}_t}^-. \quad (11)$$

Equation (10) models the activity rate v_t^Z underlying the continuous shock Z_t by a continuous mean-reverting square-root process. The two Brownian innovations are allowed to be correlated. A negative correlation $\rho < 0$ generates the desired *volatility feedback* effect: A positive shock to the business risk increases the cost of capital and reduces the asset value.

Equation (11) models the activity rate v_t^J underlying the discontinuous shock J_t , as a mean-reverting pure jump process. In particular, a downside jump in the asset value leads to an upside jump in the activity rate,

which in turns controls the arrival rate of the downside and upside jumps. Thus, the downside jump in the asset value generates a *self-exciting* behavior: The occurring of a downside jump event increases the arrival rate of future jump events. Eraker (2004) and Eraker, Johannes, and Polson (2003) use synchronized finite-activity jumps to model the equity index return and volatility. In our specification, the jumps are not only synchronous in timing, but also in jump sizes. Azizpour and Giesecke (2008) and Ding, Giesecke, and Tomecek (2008) use a self-exciting process to model the contagion effect in corporate defaults: The default of one company can increase the arrival rate of default for other companies.

Summarizing the specifications, we can write the risk-neutral asset value dynamics in terms of the following set stochastic differential equations,

$$\begin{aligned}
dA_t/A_t &= \sqrt{v_t^Z} dZ_t + \int_0^\infty (e^x - 1) (\mu^+(dx, dt) - \pi_{J^+}(x) dx v_t^J dt) + \int_{-\infty}^0 (e^x - 1) (\mu^-(dx, dt) - \pi_{J^-}(x) dx v_t^J dt), \\
dv_t^Z &= \kappa_Z (\theta_Z - v_t^Z) dt + \sigma_Z \sqrt{v_t^Z} dZ_t^v, \quad \mathbb{E}[dZ_t^v dZ_t] = \rho dt, \\
dv_t^J &= \kappa_J (\theta_J - v_t^J) dt - \sigma_J \int_{-\infty}^0 x (\mu^-(dx, dt) - \pi_{J^-}(x) dx v_t^J dt),
\end{aligned} \tag{12}$$

where $\mu^+(dx, dt)$ and $\mu^-(dx, dt)$ denote the counting measures of the upside and downside jumps. For each downside jump of size x on $\ln A_t$, the activity rate v_t^J jumps up by $-\sigma_J x$. The arrival rate of a downside jump in $\ln A_t$ of size x at time t is governed by $\pi_{J^-}(x) v_t^J$. Similarly, the time- t arrival rate of an upside jump of size x is governed by $\pi_{J^+}(x) v_t^J$.

Taken together, our specification incorporates three distinct channels of dynamic interactions between the equity index and its volatility,

1. **The financial leverage effect:** A lower equity-to-asset ratio (and hence a higher financial leverage level) leads to higher volatility levels for the equity return, holding business risk constant.
2. **The asset volatility feedback effect:** A rise in business risk increases the cost of capital and lowers business valuation, thus generating an instantaneous negative correlation between innovations in asset return and return volatility.
3. **The self-exciting behavior for market disruptions:** The occurring of a large negative market disruption increases the arrival rate of future discontinuous movements.

2.3. Alternative representations

Combining the dynamics for the equity-to-asset ratio in equation (2) with the dynamics for the asset value in equation (12), we can write the forward equity index dynamics in terms of the asset value A_t and the two activity rates (v_t^Z, v_t^J) ,

$$\begin{aligned} dF_t/F_t &= \delta \left(\frac{F_t}{A_t} \right)^{-p} dW_t + \sqrt{v_t^Z} dZ_t \\ &+ \int_0^\infty (e^x - 1) (\mu^+(dx, dt) - \pi_{J^+}(x) dx v_t^J dt) + \int_{-\infty}^0 (e^x - 1) (\mu^-(dx, dt) - \pi_{J^-}(x) dx v_t^J dt). \end{aligned} \quad (13)$$

Holding (A_t, v_t^Z, v_t^J) fixed, the index return variance level depends inversely on the forward index level F_t , similar to traditional constant elasticity of variance specifications for equity dynamics. Different from the traditional specification, however, our specification in (13) scales the index level by the level of the asset value to make the index return dynamics scale free, circumventing the common concern against scale-dependent dynamic specifications. In addition to the level dependence, our specification also generates three additional sources of variations for the index return through the variation of the three state variables (A_t, v_t^Z, v_t^J) .

Alternatively, we can regard (X_t, v_t^Z, v_t^J) as the state variable and write the forward index dynamics as,

$$\begin{aligned} dF_t/F_t &= \delta X_t^{-p} dW_t + \sqrt{v_t^Z} dZ_t \\ &+ \int_0^\infty (e^x - 1) (\mu^+(dx, dt) - \pi_{J^+}(x) dx v_t^J dt) + \int_{-\infty}^0 (e^x - 1) (\mu^-(dx, dt) - \pi_{J^-}(x) dx v_t^J dt). \end{aligned} \quad (14)$$

Under this alternative specification, the index return variance no longer shows explicit dependence on the index level. The variations of the index return variance are solely determined by the variations of the three state variables, (X_t, v_t^Z, v_t^J) . If we further perform a change of variable $v_t^X = \delta^2 X_t^{-2p}$, we obtain a three-factor stochastic volatility model of the traditional sense,

$$\begin{aligned} dF_t/F_t &= \sqrt{v_t^X} dW_t + \sqrt{v_t^Z} dZ_t \\ &+ \int_0^\infty (e^x - 1) (\mu^+(dx, dt) - \pi_{J^+}(x) dx v_t^J dt) + \int_{-\infty}^0 (e^x - 1) (\mu^-(dx, dt) - \pi_{J^-}(x) dx v_t^J dt), \end{aligned} \quad (15)$$

where the first stochastic volatility factor v_t^X follows a 3/2-process,

$$dv_t^X = \kappa_X (v_t^X)^2 dt - \sigma_X (v_t^X)^{3/2} dW_t, \quad (16)$$

with $\kappa_X = p(2p + 1)$ and $\sigma_X = 2p$ and its innovation is perfectly but negatively correlated with its corresponding return innovation component. Under this specification, the index return is driven by two Brownian motion components and a jump component. The instantaneous variances on the two Brownian motions v_t^X and v_t^Z and the arrival rate underlying the jump process v_t^J are all stochastic and are driven by three separate dynamic processes (16), (10), and (11). While we specify v_t^Z as following affine diffusion and v_t^J as following affine jump dynamics, the variance rate underlying the constant elasticity of variance process v_t^Z follows a 3/2 process, the behaviors of which have been studied by several authors, e.g., Heston (1997), Lewis (2000), and Carr and Sun (2007). Within the one-factor diffusion context, several empirical studies find that a 3/2 specification on the variance rate dynamics performs better than the square-root specification. Favorable evidence from time-series returns includes Chacko and Viceira (2003), Ishida and Engle (2002), Javaheri (2005), and Jones (2003). Supporting evidence from equity index options include Jones (2003), Medvedev and Scaillet (2003), and Bakshi, Ju, and Ou-Yang (2006).

Equations (15) and (16) reveal that from the equity index forward level F_t and index options, we cannot identify the scale parameter δ for the equity-to-asset ratio X_t . Accordingly, we henceforth normalize $\delta = 1$ and regard X_t as a proportionally scaled version of the actual equity-to-asset ratio. Furthermore, although the two activity rates (v_t^Z, v_t^J) follow mean-reverting processes, the 3/2-process in equation (16) is mean-repelling. This mean-repelling behavior generates long-run dependence between index return and volatility and contributes to the option implied volatility skews at very long option maturities.

Historically, researchers distinguish models with “sticky delta,” under which return volatility does not depend directly on price level, from the local volatility and constant elasticity of variance models, under which the return variance has direct dependence on the price level. By contrast, our specification generates a mixture behavior of both types and can be represented either in the sticky delta format as in (15) or with direct price level dependence as in (13). Thus, our model represents a unification and reconciliation of the two distinct streams of literature.

2.4. Option pricing

Consider the time- t forward value of a European call option on the equity index with strike price K and expiry date T , conditional on the time- t values of the index forward and the three state variables (X_t, v_t^Z, v_t^J) . Since the equity index F_t is driven by two orthogonal sources of variations A_t and X_t , we can perform the valuation through the law of iterated expectations,

$$\begin{aligned}
c(F_t, K, T) &= \mathbb{E}[(F_T - K)^+ | (F_t, X_t, v_t^Z, v_t^J)] \\
&= \mathbb{E}[(X_T A_T - K)^+ | (A_t, X_t, v_t^Z, v_t^J)] \\
&= \mathbb{E}[\mathbb{E}[(X A_T - K)^+ | (X_T = X, A_t, v_t^Z, v_t^J)] | X_t] \\
&= \mathbb{E}[X \cdot C(A_t, K/X, T) | X_t],
\end{aligned} \tag{17}$$

where $C(A_t, K, T)$ is defined as

$$C(A_t, K, T) \equiv \mathbb{E}[(A_T - K)^+], \tag{18}$$

which can be regarded as the forward value of a call option on the asset. In equation (17), we first write F_T as the product of A_T and X_T . Then, given the orthogonality of the two components, we perform an expectation operation conditional on fixed values of $X_T = X$, which results in the forward call value on the asset $C(A_t, K/X, T)$. The last steps involves integrating the product of X and the forward call on the asset over the probability density of $X_T = X$ conditional on X_t , which is known in analytical form as in (4). Conceptually, the procedure is analogous to the procedure advocated in Hull and White (1987), where the option value on a pure-diffusion process with independent stochastic volatility is written as an integration of the Black and Scholes (1973) formula over the conditional density of the aggregate return variance.

2.4.1. Fourier transform on asset returns and FFT valuation of options on asset

To compute the forward call value on asset $C(A_t, K, T)$, we first derive the Fourier transform of the log asset return $\ln A_T/A_t$,

$$\phi(u) \equiv \mathbb{E}_t \left[e^{iu \ln A_T/A_t} \right], \quad u \in \mathcal{D} \subseteq \mathbb{C}, \tag{19}$$

where \mathcal{D} denotes a subset of the complex plane under which the expectation in (19) is well defined. Once we obtain this transform $\phi(u)$, we can compute the option value C via fast Fourier transform (FFT) following

the procedure first proposed by Carr and Madan (1999).

Since the log asset return is modeled as a linear combination of two time-changed Lévy processes in equation (5), we follow Carr and Wu (2004) and write the Fourier transform as the Laplace transforms of the two stochastic time changes under a new measure \mathbb{M} ,

$$\begin{aligned}\phi(u) &= \mathbb{E}_t \left[\exp \left(iu \left(Z_{\mathcal{T}_{t,T}^Z} - \frac{1}{2} \mathcal{T}_{t,T}^Z \right) + iu \left(J_{\mathcal{T}_{t,T}^J} - k_J(1) \mathcal{T}_{t,T}^J \right) \right) \right] \\ &= \mathbb{E}_t^{\mathbb{M}} \left[\exp \left(-\Psi_Z(u) \mathcal{T}_{t,T}^Z - \Psi_J(u) \mathcal{T}_{t,T}^J \right) \right]\end{aligned}\quad (20)$$

where the new measure \mathbb{M} is defined by the following complex-valued exponential martingale,

$$\frac{d\mathbb{M}}{d\mathbb{Q}} \Big|_t = \exp \left(iu \left(Z_{\mathcal{T}_{t,T}^Z} - \frac{1}{2} \mathcal{T}_{t,T}^Z \right) + \Psi_Z(u) \mathcal{T}_{t,T}^Z + iu \left(J_{\mathcal{T}_{t,T}^J} - k_J(1) \mathcal{T}_{t,T}^J \right) + \Psi_J(u) \mathcal{T}_{t,T}^J \right), \quad (21)$$

and $(\Psi_Z(u), \Psi_J(u))$ are the characteristic exponents of the two convexity-adjusted Lévy components $(Z_t - \frac{1}{2}t, J_t - k_J(1)t)$ prior to the time change,

$$\begin{aligned}\Psi_Z(u) &= \frac{1}{2}(iu + u^2), \\ \Psi_J(u) &= \ln(1 - iuv_{J^+})(1 + iuv_{J^-}) - iu \ln(1 - v_{J^+})(1 + v_{J^-}).\end{aligned}\quad (22)$$

Under measure \mathbb{M} , the dynamics of the two activity rates become,

$$\begin{aligned}dv_t^Z &= (\kappa_Z \theta_Z - \kappa_Z^{\mathbb{M}} v_t^Z) dt + \sigma_Z \sqrt{v_t} dZ_t^{\mathbb{M}}, \quad \kappa_Z^{\mathbb{M}} = \kappa_Z - iu \rho \sigma_Z, \\ dv_t^J &= (\kappa_J \theta_J - \kappa_J^{\mathbb{M}} v_t^J) dt - \sigma_J \int_{-\infty}^0 x (\mu^-(dx, dt) - \pi_{J^-}^{\mathbb{M}}(x) dx v_t^J dt), \quad \kappa_J^{\mathbb{M}} = \kappa_J - \sigma_J (v_{J^-}^{\mathbb{M}} - v_{J^-}).\end{aligned}\quad (23)$$

The exponential martingale in (21) generates an exponential tilting in the Lévy density of J_t^- under the new measure \mathbb{M} ,

$$\pi_{J^-}^{\mathbb{M}}(x) = e^{iux} e^{-|x|/v_{J^-}} |x|^{-1} = e^{-|x|/v_{J^-}^{\mathbb{M}}} |x|^{-1}, \quad v_{J^-}^{\mathbb{M}} = \frac{v_{J^-}}{1 + iuv_{J^-}}. \quad (24)$$

The cumulant exponent of the downside jump J_t^- under measure \mathbb{M} becomes,

$$k_{J^-}^{\mathbb{M}}(s) = -\ln(1 + sv_{J^-}^{\mathbb{M}}). \quad (25)$$

With the affine activity rate dynamics in equation (23), the Laplace transform in (20) can be solved in

exponential-affine forms (Filipović (2001) and Duffie, Filipović, and Schachermayer (2003)),

$$\phi(u) = \exp(-a_Z(\tau) - b_Z(\tau)v_t^Z - a_J(\tau) - b_J(\tau)v_t^J), \quad \tau = T - t, \quad (26)$$

where the affine coefficients solve the following ordinary differential equations,

$$\begin{aligned} b_Z'(\tau) &= \psi_Z(u) - \kappa_Z^M b_Z(\tau) - \frac{1}{2}\sigma_Z^2 b_Z(\tau)^2, & a_Z'(\tau) &= b_Z(\tau)\kappa_Z\theta_Z, \\ b_J'(\tau) &= \psi_J(u) - (\kappa_J + \sigma_J v_{J-}) b_J(\tau) - k_{J-}^M (\sigma_J b_J(\tau)), & a_J'(\tau) &= b_J(\tau)\kappa_J\theta_J, \end{aligned} \quad (27)$$

starting at $a_Z(0) = b_Z(0) = a_J(0) = b_J(0) = 0$. The ordinary differential equations governing the coefficients $(a_Z(\tau), b_Z(\tau))$ can be solved analytically,

$$\begin{aligned} b_Z(t) &= \frac{2\psi_Z(u)(1-e^{-\xi\tau})}{2\xi - (\xi - \kappa_Z^M)(1-e^{-\xi\tau})}, & \xi &= \sqrt{(\kappa_Z^M)^2 + 2\sigma_Z^2\psi_Z(u)}, \\ a_Z(t) &= \frac{\kappa_Z\theta_Z}{\sigma_Z^2} \left[2\ln\left(1 - \frac{\xi - \kappa_Z^M}{2\xi}(1-e^{-\xi\tau})\right) + (\xi - \kappa_Z^M)\tau \right]. \end{aligned} \quad (28)$$

The ordinary differential equations governing the coefficients $(a_Z(\tau), b_Z(\tau))$ can be solved numerically using the standard Runge-Kutta 4th order method.

With the Fourier transform $\phi(u)$ on the asset return, we first re-scale the forward call value on the asset $c(k) = C(A_t, K, T)/A_t$ with $k = \ln K/A_t$. That is, we represent the forward call value in percentages of the forward asset value as a function of moneyness defined as the log strike over forward. Then, we derive the Fourier transform on the re-scaled forward call $c(k)$ in terms of the Fourier transform on the asset return,

$$\chi(u) \equiv \int_{-\infty}^{\infty} e^{iuk} c(k) dk = \frac{\phi(u-i)}{(iu)(iu+1)}, \quad (29)$$

which is well-defined when u contains an imaginary component $u = u_r - iv$, with u_r being real and v being a real positive number. With the transform in (29), the call value can be computed via the following inversion,

$$c(k) = \frac{e^{-vk}}{\pi} \int_0^{\infty} e^{-iu_r k} \chi(u_r - iv) du_r. \quad (30)$$

To perform the inversion, we first discretize the integral using the trapezoid rule:

$$c(k) \approx \frac{e^{-vk}}{\pi} \sum_{m=0}^N \delta_m e^{-iu_m k} \chi(u_m - iv) \Delta u, \quad (31)$$

where $\delta_m = \frac{1}{2}$ when $m = 0$ and 1 otherwise. We perform the summation using discrete fast Fourier transform (FFT). FFT is an efficient algorithm for computing discrete Fourier coefficients. The discrete Fourier transform is a mapping of $\mathbf{f} = (f_0, \dots, f_{N-1})^\top$ on the vector of Fourier coefficients $\mathbf{d} = (d_0, \dots, d_{N-1})^\top$, such that

$$d_j = \frac{1}{N} \sum_{m=0}^{N-1} f_m e^{-jm \frac{2\pi}{N} i}, \quad j = 0, 1, \dots, N-1. \quad (32)$$

FFT allows the efficient calculation of \mathbf{d} if N is an even number, say $N = 2^n, n \in \mathbb{N}$. The algorithm reduces the number of multiplications in the required N summations from an order of 2^{2n} to that of $n2^{n-1}$, a very considerable reduction.

To map the inversion in equation (31) to the FFT form in (32), we set the summation grid by $\eta = \Delta u$ and $u_m = \eta m$, and we set the relative strike grid by $k_j = -b + \lambda j$ with $\lambda = 2\pi/(\eta N)$ and $b = \lambda N/2$. Then, the call value becomes

$$c(k_j) \approx \frac{1}{N} \sum_{m=0}^{N-1} f_m e^{jm \frac{2\pi}{N} i}, \quad f_m = \delta_m \frac{N}{\pi} e^{-vk_j + iu_m b} \chi(u_m - iv) \eta, \quad (33)$$

with $j = 0, 1, \dots, N-1$. The inversion has the FFT form and can hence be computed efficiently across the whole spectrum of strikes k_j .

2.4.2. Numerical integration with Gauss-Hermite quadrature

Once we have computed the forward call value on asset across the whole spectrum of strikes using the FFT method, we approximate the integration in the last line of equation (17) by a weighted sum of a finite number (M) of forward asset call values at equity-to-asset ratio values $X_j, j = 1, 2, \dots, M$,

$$c(F_t, K, T) = \int_0^\infty f(X|X_t) X C(A_t, K/X, T) dX \approx \sum_{j=1}^M \mathcal{W}_j X_j C(A_t, K/X_j, T), \quad (34)$$

where we choose the points X_j and their corresponding weights based on the Gauss-Hermite quadrature rule.

The Gauss-Hermite quadrature rule is designed to approximate an integral of the form $\int_{-\infty}^\infty h(x) e^{-x^2} dx$,

where $h(x)$ is an arbitrary smooth function. After some re-scaling, the integral can be regarded as an expectation of $h(x)$ where x is a normally distributed random variable with zero mean and variance of one half. For a given target function $h(x)$, the Gauss-Hermite quadrature rule generates a set of weights w_i and nodes x_i , $i = 1, 2, \dots, M$, that are defined by

$$\int_{-\infty}^{\infty} h(x) e^{-x^2} dx = \sum_{j=1}^M w_j h(x_j) + \frac{M! \sqrt{\pi}}{2^M} \frac{h^{(2M)}(\xi)}{(2M)!} \quad (35)$$

for some $\xi \in (-\infty, \infty)$. The approximation error vanishes if the integrand $f(x)$ is a polynomial of degree equal or less than $2M - 1$. See Davis and Rabinowitz (1984) for details.

To apply the quadrature rules, we need to map the quadrature nodes and weights $\{x_i, w_j\}_{j=1}^M$ to our choice of X_j and the weights \mathcal{W}_j . Given the constant elasticity of variance dynamics, one reasonable choice is,

$$X(x) = X_t e^{\sqrt{2V_X}x - \frac{1}{2}V_X}, \quad V_X = X_t^{-2p}(T-t). \quad (36)$$

The choice is motivated by a log-normal approximation of the density of X by assuming that the instantaneous return variance $\delta^2 X_t^{-2p}$ is fixed. Then, given the Gauss-Hermite quadrature $\{w_j, x_j\}_{j=1}^M$, we choose the X_j points as

$$X_j = X_t e^{\sqrt{2V_X}x_j - \frac{1}{2}V_X}, \quad (37)$$

and the summation weights as

$$\mathcal{W}_j = \frac{f(X_j|X_t) X'(x_j)}{e^{-x_j^2}} w_j = \frac{f(X_j|X_t) X_j \sqrt{2V_X}}{e^{-x_j^2}} w_j, \quad (38)$$

with the transition density $f(X_j|X_t)$ given in (4).

2.5. Market prices of risks and the statistical dynamics

The economy that we have described includes four sources of risk: (i) W_t , which controls the variation of the financial leverage, (ii) Z_t , which controls the continuous variation of the asset return, (iii) J_t , which controls the discontinuous movements of the asset return, and (iv) Z_t^v , which controls the continuous variation of the activity rate v_t^Z . The activity rate v_t^J is assumed to be driven by the same downside jump variation in the

return jump and hence does not add additional variation sources. Furthermore, since Z_t and Z_t^v are correlated, we perform the following decomposition on Z_t ,

$$Z_t = \rho Z_t^v + \sqrt{1 - \rho^2} \tilde{Z}_t, \quad (39)$$

where \tilde{Z}_t captures the return diffusion variation component that is independent of the diffusion variance rate variation.

To derive the statistical dynamics of the equity index, we use the following exponential martingale to define the measure changes from \mathbb{P} to \mathbb{Q} ,

$$\frac{d\mathbb{P}}{d\mathbb{Q}} \Big|_t = \mathcal{E} \left(- \int_0^t \gamma_s^X dW_s \right) \mathcal{E} \left(- \gamma^Z \tilde{Z}_{\mathcal{T}^Z} \right) \mathcal{E} \left(- \gamma^J J_{\mathcal{T}^J} \right) \mathcal{E} \left(- \gamma^v Z_{\mathcal{T}^Z}^v \right), \quad (40)$$

where $\mathcal{E}(\cdot)$ denotes the stochastic exponential operator. The specification in (40) assumes constant market prices $\gamma^Z, \gamma^J, \gamma^v$ for the independent diffusion return risk \tilde{Z} , the jump risk J , and the diffusion variance rate risk Z^v , respectively. The market price for the equity-to-asset ratio risk is assumed to be time varying and take the following linear form,

$$\gamma_t^X = (a_X - \kappa_{XX} X_t - \kappa_{XZ} v_t^Z - \kappa_{XJ} v_t^J), \quad (41)$$

under which the \mathbb{P} -dynamics of the equity-to-asset ratio becomes,

$$dX_t = X_t^{1-p} (a_X - \kappa_{XX} X_t - \kappa_{XZ} v_t^Z - \kappa_{XJ} v_t^J) dt + X_t^{1-p} dW_t^{\mathbb{P}}, \quad (42)$$

where the statistical drift of the X_t process accommodates managerial decisions on financial leverage. The specification allows the managers to adjust the financial leverage as a function of the current leverage level X_t , the current business diffusion risk level v_t^Z , and the current business jump risk level v_t^J . We use $\kappa_L \equiv [\kappa_{XX}, \kappa_{XZ}, \kappa_{XJ}]^\top$ to denote the leverage loading coefficient vector on the three state variables (X_t, v_t^Z, v_t^J) . The constant term a_X allows the manager to set a long-run target on the equity-to-asset ratio.

With the market price of the Brownian risk W_t specified in (41), we can also derive the \mathbb{P} -dynamics for the variance rate $v_t^X = X_t^{-2p}$,

$$dv_t^X = (v_t^X)^2 (\kappa_X - \sigma_X X_t^p (a_X - \kappa_{XX} X_t - \kappa_{XZ} v_t^Z - \kappa_{XJ} v_t^J)) dt - \sigma_X (v_t^X)^{3/2} dW_t^{\mathbb{P}}. \quad (43)$$

The constant market price (γ^Z) of the independent diffusion return risk (\tilde{Z}_t) generates an instantaneous risk premium on the asset return of $\gamma^Z \sqrt{1 - \rho^2} v_t^Z$. The market price (γ^J) of the jump return risk (J) generates an exponential tilting on the Lévy density under the statistical measure \mathbb{P} ,

$$\pi_{J+}^{\mathbb{P}}(x) = e^{\gamma^J x} e^{-x/v_{J+}} x^{-1} = e^{-x/v_{J+}^{\mathbb{P}}} x^{-1}, \quad \pi_{J-}^{\mathbb{P}}(x) = e^{\gamma^J x} e^{-|x|/v_{J-}} |x|^{-1} = e^{-|x|/v_{J-}^{\mathbb{P}}} |x|^{-1}, \quad (44)$$

with $v_{J+}^{\mathbb{P}} = v_{J+}/(1 - \gamma^J v_{J+})$ and $v_{J-}^{\mathbb{P}} = v_{J-}/(1 + \gamma^J v_{J-})$. The exponential tilting generates an instantaneous asset return risk premium given by $k_J^{\mathbb{P}}(1) - k_J(1)$, the cumulant exponent differences under the two probability measures evaluated at one. Furthermore, the market price of jump risk induces a drift adjustment on v_t^J under the statistical measure \mathbb{P} , $\sigma_J (v_{J-}^{\mathbb{P}} - v_{J-}) v_t^J$. Finally, the constant market price (γ^ν) of the diffusion variance risk (Z_t^ν) generates a drift adjustment term $\gamma^\nu \sigma_Z v_t^Z$ for v_t^Z . It also generates an instantaneous asset return risk premium $\rho \gamma^\nu v_t^Z$.

Taken together, we can write the dynamics of the asset value and its underlying activity rates under the statistical measure \mathbb{P} as,

$$\begin{aligned} dA_t/A_t &= (\sqrt{1 - \rho^2} \gamma^Z + \rho \gamma^\nu) v_t^Z + (k_J^{\mathbb{P}}(1) - k_J(1)) v_t^J \\ &\quad + \sqrt{v_t^Z} dZ_t^{\mathbb{P}} + \int_0^\infty (e^x - 1) (\mu^+(dx, dt) - \pi_{J+}^{\mathbb{P}}(x) dx v_t^J dt) + \int_{-\infty}^0 (e^x - 1) (\mu^-(dx, dt) - \pi_{J-}^{\mathbb{P}}(x) dx v_t^J dt), \\ dv_t^Z &= (\kappa_Z \theta_Z - \kappa_Z^{\mathbb{P}} v_t^Z) dt + \sigma_Z \sqrt{v_t^Z} dZ_t^{\mathbb{P}}, \\ dv_t^J &= (\kappa_J \theta_J - \kappa_J^{\mathbb{P}} v_t^J) dt - \sigma_J \int_{-\infty}^0 x (\mu^-(dx, dt) - \pi_{J-}^{\mathbb{P}}(x) dx v_t^J dt), \end{aligned} \quad (45)$$

with $\kappa_Z^{\mathbb{P}} = \kappa_Z - \gamma^\nu \sigma_Z$ and $\kappa_J^{\mathbb{P}} = \kappa_J - \sigma_J (v_{J-}^{\mathbb{P}} - v_{J-})$.

2.6. A reduced-form benchmark for stock return dynamics

The current state of the art in option pricing is to specify the underlying security return dynamics as driven by two return components, each driven by a separate stochastic volatility process. Several studies show that two stochastic volatility factors perform much better than one stochastic volatility factor does in pricing equity index options (Bates (2000), Huang and Wu (2004), and Christoffersen, Heston, and Jacobs (2007)) and currency options (Carr and Wu (2007)). The two return components with separate stochastic volatilities

can readily accommodate stochastic skewness in the underlying return risk-neutral distribution. Furthermore, when the two volatility factors show different risk-neutral persistence, they can also help capture variations along the term structure dimension.

To compare our model with the state of the art in the option pricing literature, we create a reduced-form benchmark model for the equity index return dynamics with two stochastic volatility factors by setting $X_t = 1$ and hence $S_t = A_t$. Thus, we can regard the benchmark as a restricted version of our model that ignores the leverage effect. By comparison, Bates (2000) specifies two jump-diffusion return components, each driven by a separate stochastic volatility process. The jump components are useful in generating the short-term implied volatility smiles while the two stochastic volatility processes generate stochastic skewness in the return distribution. Christoffersen, Heston, and Jacobs (2007) consider a special case of the Bates model by removing the jump components. Huang and Wu (2004) enhances the identification of the Bates model by limiting one return component to a pure diffusion and the other component to be a pure jump process. The specification allows the stochastic skew behavior while also having the capability of generating short-term smiles through the jump component. All these specifications incorporate a negative instantaneous correlation between the index return and return variance. The negative correlation is often labeled as the leverage effect, without further distinction between the actual leverage effect and the volatility feedback effect.

Our benchmark is similar to the Huang-Wu specification, but with an additional layer of interactions between the index return and return volatility through the self-exciting behavior of downside return jumps. Thus, our reduced-form benchmark represents a refined representation of the state of the art in reduced-form equity index option pricing. We use the term “reduced-form” to highlight the fact that the benchmark, as in most extant models in the option pricing literature, does not separately model asset value dynamics and finance leverage, and hence does not differentiate the leverage effect and the volatility feedback effect. By comparing the empirical performance of our full model with the reduced-form benchmark, we gauge the benefits of allowing distinct channels of interactions between equity returns and volatilities.

Under the benchmark specification for the equity index, we can price equity index options $c(F_t, K, T)$ using the same Fourier transform method as we have used in pricing options on asset, $C(A_t, K, T)$, discussed in Section 2.4.1.

3. Data analysis

We obtain over-the-counter implied volatility quotes from a broker dealer on S&P 500 index options. The sample is from January 8, 1997 to March 5, 2008. The data are available daily, but we sample the data weekly every Wednesday to avoid weekday effects. At each date, the quotes are available at eight fixed time to maturities at one, three, six, 12, 24, 36, 48, and 60 months. At each maturities, there are five quotes at 80, 90, 100, 110, and 120 percent of the spot index level. All together, we have 40 implied volatility series on a grid of five relative strikes and eight fixed time to maturities over 583 weeks.

Compared to the exchange market (the Chicago Board of Options Exchange), the over-the-counter market trades options over a much wider spectrum of maturities. At one particular maturity, any one of the many different mechanisms can be used to generate an implied volatility skew along the strike dimension; yet these different mechanisms have different implications on how the implied volatility skew evolve with the option time to maturity. By using the over-the-counter data that cover maturities from one month to five years, we can achieve a better distinction of the different mechanisms that our model incorporates.

3.1. Summary statistics

Table 1 reports the summary statistics of the implied volatility quotes. Panel A reports the sample averages of the implied volatilities. At each fixed maturity, the average implied volatilities decline with increasing strike prices, generating the well-documented implied volatility skew pattern for stock indexes. At each fixed relative strike (\mathcal{K}), the implied volatilities increase with the time to maturity for at-the-money options (100% strike) and at high relative strikes (110% and 120% strikes), but the term structure becomes downward sloping for options at 80% and 90% relative strikes.

[Table 1 about here.]

Figure 1 plots the mean implied volatility surface in the left panel as a function of the relative strikes and time to maturities. Since an 80% relative strike is far more out-of-the-money at the one-month maturity than at the five-year maturity, the implied volatility skew looks much steeper at short than at long maturities. To adjust for the maturity and implied volatility level difference, we define a standardized moneyness measure, $d \equiv \ln(\mathcal{K}/100)/(IV\sqrt{\tau}/100)$, where IV denotes the implied volatility (in percentage points) at the relative

strike \mathcal{K} (in percentages of spot) and the time to maturity τ (in years). The standardized moneyness measure approximates the number of standard deviations that the strike is away from the spot. With the standardized moneyness measure d , we construct a comparable implied volatility skew measure at each date and maturity as,

$$SK_{t,T} = \frac{IV_{t,T}(80\%) - IV_{t,T}(120\%)}{|d_{t,T}(80\%) - d_{t,T}(120\%)|}. \quad (46)$$

which measures the implied volatility difference in percentage points at 80% and 120% strikes, divided by the absolute difference in their respective standard moneyness. Since the 80% strike implied volatility is universally higher than the corresponding 120% strike implied volatility across all dates and maturities, our skew measure generates positive estimates. The higher the estimates, the steeper the negative slope of the implied volatility plot against the standardized moneyness measure. The right panel of Figure 1 plots the sample averages of the skew measure at different maturities in the solid line and their 10th and 90th percentiles in the two dashed lines. The skew measure is on average lower at short maturities due to the smile pattern (curvature). As the option maturity increases, the smile becomes a pure skew and the slope increases in absolute magnitude.

[Fig. 1 about here.]

Table 1 reports the standard deviation estimates of the implied volatility series in Panel B. We also visualize the standard deviation variation along the strike and maturity dimension in the left panel of Figure 2. Consistent with the negative skew in the implied volatility levels, the standard deviation estimates also decline with increasing strikes. Along the maturity dimension, the standard deviation estimates show a steeply downward sloping term structure, suggesting the presence of a highly mean-reverting (under the risk-neutral measure) stochastic volatility factor. Nevertheless, even at five-year maturity, the implied volatilities still have a large amount of variation left, suggesting the presence of a nearly non-stationary component in the implied volatility variation.

[Fig. 2 about here.]

Panel C of Table 1 reports the weekly autocorrelation estimates of the implied volatility series. The estimates range from 0.936 to 0.989. Figure 2 visualizes the autocorrelation variation along the relative

strike and time to maturity dimension. The plot reveals a prominent upward sloping term structure for the autocorrelation estimates. The upward-sloping term structure suggests the presence of multiple stochastic volatility factors, with a relatively transient factor contributing mainly at the short maturities and generating the lower autocorrelation estimates, and a more persistent factor affecting implied volatilities of all maturities.

3.2. *Principal component analysis*

To understand the driving forces of the implied volatility surface variation, we perform a principal component analysis on the 40 implied volatility series. We first estimate the covariance matrix of the weekly changes on the 40 series, and then compute the eigenvalues and eigenvectors of the covariance matrix. The normalized eigenvalues can be interpreted as the percentage variation explained by each principal component, and the mimicking portfolio for each component can be formed by using the eigenvector corresponding to the eigenvalue as the portfolio weight. The bar charts in the first panel of Figure 3 represent the normalized eigenvalues for the first ten principal components. The first principal component explains 85.09% of the variation, the second component explains 8.24% of the variation, and the third component explains 3.30% of the variation. Thus, the first three principal components explain over 96.62% of the variation in the implied volatility surface. Therefore, a three-factor stochastic volatility structure can explain the majority of the variations in the implied volatility surface.

[Fig. 3 about here.]

The three lines in the second panel of Figure 3 plot the eigenvectors corresponding to the first three eigenvalues, which capture the loading of the first three principal components on the 40 implied volatility series. In the plot, the 40 implied volatility series are ranked first in five strikes from 80% to 120% at each maturity, and then in the eight maturities from one month to five years. The solid line graphs the loadings of the first principal component (P1), which are positive across all 40 series. Thus, this first principal component captures the overall variation of the implied volatilities. The loading decline with increasing maturity reflects the declining variation of the implied volatility series at longer maturities.

The dashed line in the second panel of Figure 3 plots the loadings of the second principal component (P2), which are relatively flat across the five strikes at each maturity, but decline monotonically as the time to

maturity increases. The loadings are positive at short maturities (one and two months), but become negative when the maturities are six months and longer, thus creating a term structure slope factor. This distinct term structure pattern becomes clearer when we plot the factor loading of the second principal component in a three-dimensional surface in the third panel in Figure 3.

Finally, the dash-dotted line in the second panel of Figure 3 plots the loadings of the third principal component (P3), which show strong variations along the strike dimension. The pattern becomes all the more evident when we plot the loadings in a three-dimensional surface in the last panel in Figure 3. At each maturity, the loadings are positive at low strikes but negative at high strikes. The strike dimension variation is particularly strong at short maturities, but the variation diminishes as the maturity increases. The smaller variation along the strike dimension at longer maturities is partly due to the scaling because the same relative strike range implies a smaller range of standard deviations at longer maturities.

Compared to the over-the-counter implied volatility quotes, the exchange-listed index options trade over a wider strike range but a narrower maturity range, with the maximum time to maturity being less than three years. When we perform principal component analysis on standardized data on the exchange-listed options, we obtain a similar three-factor structure, but with the strike-dimension variation contributing more to the total variance than the term structure variation does. Taken together, the data ask for a three-factor volatility structure that capture the level, the term structure, and the strike-dimension variation, respectively. Our model specification can in principal accommodate the three dimensions of variation in the implied volatility surface through the variations of the financial leverage ratio (X_t) and the two activity rates on the diffusion and jump components of the asset return (v_t^Z, v_t^J), respectively.

3.3. Dynamic interactions between index returns and option implied volatilities

To understand the interactions between the stock index and the option implied volatility level and the slopes along the maturity and moneyness dimension, we use the average of the eight at-the-money implied volatility series to proxy the volatility level, the difference between five-year and one-month at-the-money implied volatilities to proxy the term structure variation, and the average of the eight implied volatility skew series as defined in equation (46) to proxy the moneyness dimension variation. When we measure the correlation between the weekly index returns and weekly changes in the three implied volatility dimensions, we

obtain highly negative estimates between the index return and the volatility level and skew, at -0.8114 and -0.707 , respectively. On the other hand, the index return correlation estimate with the term structure series is highly positive at 0.7643 . These correlation estimates suggest that when the stock index drops and hence the index return is negative, (i) the implied volatility level tends to go up; (ii) the negative implied volatility skew along the moneyness dimension becomes steeper (more negatively skewed), and (iii) the implied volatility term structure becomes flatter or more downward sloping. Thus, the stock index movements interact strongly with all three dimensions of the option implied volatility surface.

The three panels on the left side of Figure 4 plot the weekly changes in the three implied volatility series against weekly returns on the stock index, with the circles denoting data points and the solid lines denoting local linear regression fits with a Gaussian kernel and a default bandwidth choice. The plots confirm the correlation estimates in showing strongly negative co-movements between the index return and the implied volatility level and skew, and strongly positive co-movements between the index return and the implied volatility term structure. Further inspection of the plots reveals a certain degree of asymmetry for the implied volatility response when the index return experiences large positive versus negative movements. Overall, the responses of all three implied volatility series are stronger to large negative index movements than to large positive movements. This asymmetry is the strongest for the implied volatility skew. For example, the correlation between the index return and the implied volatility skew is -0.5425 when conditioning on positive index returns, but becomes -0.6370 when conditioning on negative index returns. Furthermore, the correlation estimate is -0.6612 when conditioning on the index return being lower than -5% , but the estimate becomes virtually zero at 0.03 when conditioning on the index return being greater than 5% . Similar asymmetries exist on the implied volatility levels, but not on the term structure. This asymmetric behavior on the implied volatility level and skew is consistent with our self-exciting downside return jump specification. A downside jump in the asset return leads to a positive jump in the arrival rate of jumps and thus a positive jump in index return variance. Furthermore, when the negative jump sizes are on average larger than the positive jump sizes ($v_{J^-} > v_{J^+}$), the jump component in the asset return generates negative skewness in the index return and hence contributes to the negatively sloped implied volatility skew. When the negative return jump leads to a spike in the jump arrival rate, it also leads to a spike in the return negative skewness and the option implied volatility skew.

[Fig. 4 about here.]

The three panels on the right side of Figure 4 plot the cross-correlations at different leads and lags between the weekly index returns and weekly changes in the three implied volatility series. For all three pairs, the instantaneous correlations are the strongest. In addition, lagged index returns show some positive predictions on the at-the-money implied volatility level and the implied volatility skew, but negative predictions on the implied volatility term structure. On the other hand, lagged weekly changes in the at-the-money implied volatility and the implied volatility skew seem to predict the index return positively over a two-month horizon.

Figure 5 plots the time series of the equity index in the first panel. The equity index experienced sustained growth between 1996 and 2000, followed by a three-year retreat until the rebound starting in 2003. The rebound reached its peak in mid 2007. The second panel of Figure 5 plots the at-the-money implied volatility at one-month (solid line) and five-year (dashed line) maturities. The variation of the one-month implied volatility is much stronger than the five-year volatility. The implied volatility levels are relatively low around 15% at the beginning of our sample period in 1997 and reached the lowest point in 2007. In between, the volatility series have experienced numerous spikes to as high as 40% and have a sustained period of high long-term implied volatility from 1999 to 2003. The implied volatility started to go up again since 2007.

[Fig. 5 about here.]

The third panel of Figure 5 plots the implied volatility term structure, defined as the difference in five-year and one-month at-the-money implied volatilities. The term structures are mostly positive and hence upward sloping, except during the occasional spikes in the short-term implied volatility, such as during the 1997-1998 Asian crises, the 1999 Nasdaq bubble, and the high volatility period during the 2000-2003 recession. The term structure started to become downward sloping again since the beginning of 2008. The last panel in Figure 5 plots the skew measure defined in (46) at one-month (solid line) and five-year (dashed line) maturities. The estimates are universally positive, suggesting that the option implied volatility are always negatively sloped against strike during our sample period. The skew estimates tend to be high at high volatility days.

Our model specification can in principle explain the three principal sources of variation on the implied volatility surface and their dynamic interactions with the index return. First, the two activity rates (v_t^Z, v_t^J) and the financial leverage X_t (or its transform $v_t^X = X_t^{-2p}$) all contribute to the volatility level variation for the

index return. Second, shocks to the two activity rates can have different impacts across the implied volatility term structure when the risk-neutral mean-reversion speeds are different. In addition, shocks on the mean-repelling 3/2-process v_t^X have long-lasting impacts across the whole spectrum of maturities. The different responses generate variations in the term structure of the implied volatility that dominate the second principal component. Third, the negative jump component J^- in the asset value dynamics generates negative skewness in the equity index return at very short maturities, although its contribution dies away as maturity increases. The relative variation between the two activity rates (v_t^Z, v_t^J) alters relative composition of continuous and discontinuous components in the asset return and thus generates variations on the return skewness, particularly at short to moderate horizons. By contrast, the return skewness generated from the financial leverage effect affects both short and long maturities. Finally, all three sources of variations (v_t^X, v_t^Z, v_t^J) have negative co-movements with the index return and thus contribute to the observed negative correlation between index returns and implied volatility level changes. In addition, the self-exciting behavior embedded in the specification of v_t^J dynamics can induce the observed asymmetry in the relation between the implied volatility skew and the index return.

4. Estimation strategy

The variation of the index options is controlled by four state variables, represented by (F_t, X_t, v_t^Z, v_t^J) . Among the four state variables, the forward index level is observable, whereas the other three (X_t, v_t^Z, v_t^J) are not directly observable. To estimate the model parameters on the observed option prices, we cast the model into a state-space form and infer the three unobserved state variables $V_t \equiv [X_t, v_t^Z, v_t^J]^\top$ at each date t from the observed option prices using a filtering technique. We estimate the model parameters by maximizing the sum of weekly likelihood values on options.

In the state-space form, we specify the state propagation equation on V_t as an euler approximation of their statistical dynamics:

$$V_t = f(V_{t-1}; \Theta) + \sqrt{Q_{t-1}} \varepsilon_t, \quad (47)$$

where ε_t denotes the standardized forecasting error vector, and the forecasting function $f(V_{t-1}; \Theta)$ and the

forecasting error variance are given by,

$$f(V_{t-1}; \Theta) = \begin{bmatrix} X_t + X_t^{1-p}(a - \kappa_L^\top V_{t-1})\Delta t \\ \kappa_Z \theta_Z \Delta t + (1 - \kappa_Z^{\mathbb{P}} \Delta t) v_{t-1}^Z \\ \kappa_J \theta_J \Delta t + (1 - \kappa_J^{\mathbb{P}} \Delta t) v_{t-1}^J \end{bmatrix}, \quad Q_{t-1} = \left\langle \begin{array}{c} X_{t-1}^{2-2p} \\ \sigma_Z^2 v_{t-1}^Z \\ \sigma_J^2 (v_{t-1}^{\mathbb{P}})^2 v_{t-1}^J \end{array} \right\rangle \Delta t \quad (48)$$

with $\Delta t = 7/365$ denoting the weekly frequency of the data, $\langle \cdot \rangle$ denoting a diagonal matrix, and Θ denoting the parameter set.

We specify the measurement equations on the implied volatility quotes, assuming additive, normally-distributed measurement errors:

$$y_t = h(V_t; \Theta) + \sqrt{R} e_t, \quad (49)$$

where y_t denotes the time- t forward value of the out-of-the-money options computed from the implied volatility quotes, scaled by the Black-Scholes vega of the option, $h(V_t; \Theta)$ denotes the corresponding model value as a function of the state vector V_t and the parameter set Θ . We assume that the pricing errors on the scaled option prices are i.i.d. normal with zero mean and constant variance. Hence, we can write the pricing error covariance matrix as, $R = \sigma_e^2 I_{40}$, with σ_e being a scalar and I_{40} denoting an identity matrix of dimension 40. The dimension of the measurement equation corresponds to the 40 implied volatility series across the five relative strikes at each maturity and eight time to maturities.

When the state propagation and measurement equations are Gaussian linear, the Kalman filter provides efficient forecasts and updates on the mean and covariance of the state vector and observations. Our state-propagation equations and measurement equations do not satisfy the Gaussian and linear conditions. We use an extended version of the Kalman filter, the unscented Kalman filter, to handle the deviations. Specifically, let $\bar{V}_t, \bar{y}_t, \bar{\Sigma}_{xy,t}$ denote the time- $(t-1)$ ex ante forecasts of time- t values of the state vector, the measurement series, and the covariance between series x and y ; let $\hat{V}_t, \hat{y}_t, \hat{\Sigma}_{xy,t}$ denote the corresponding ex post update, or filtering, on the state vector, the measurement, and the covariances. The unscented Kalman filter uses a set of deterministically chosen (sigma) points to approximate the state distribution. At each time t , if we use k to denote the number of states (three in our model) and use $\eta > 0$ denote a control parameter, we first generate a set of $2k+1$ sigma vectors χ_{t-1} from the time $(t-1)$ updated mean \hat{V}_{t-1} and covariance $\hat{\Sigma}_{VV,t-1}$ of the state

vector according to the following equations,

$$\begin{aligned}\chi_{t-1,0} &= \widehat{V}_{t-1}, \\ \chi_{t-1,i} &= \widehat{V}_{t-1} \pm \sqrt{(k+\eta)(\widehat{\Sigma}_{VV,t-1})_j}, \quad j=1,\dots,k; \quad i=1,\dots,2k,\end{aligned}\tag{50}$$

with the corresponding weights w_i given by,

$$w_0 = \eta/(k+\eta), \quad w_i = 1/[2(k+\eta)], \quad i=1,\dots,2k.\tag{51}$$

These sigma vectors form a discrete distribution with w_i being the corresponding probabilities. We propagate these sigma points through the propagation equation (47) to compute the forecasted mean and covariance of the state vector at time t ,

$$\begin{aligned}\bar{\chi}_{t,i} &= f(\chi_{t-1,i}; \Theta), \quad \bar{V}_t = \sum_{i=0}^{2k} w_i \bar{\chi}_{t,i}, \\ \bar{\Sigma}_{VV,t} &= \sum_{i=0}^{2k} w_i (\bar{\chi}_{t,i} - \bar{V}_t)(\bar{\chi}_{t,i} - \bar{V}_t)^\top + Q_{t-1}.\end{aligned}\tag{52}$$

We then re-generate the sigma points $\tilde{\chi}_t$ based on the forecasted mean \bar{V}_t and covariance $\bar{\Sigma}_{VV,t}$ in very much the same procedure as in equations (50) and (51). With the re-generated sigma points, we compute the forecasted mean and covariances of the measurements,

$$\begin{aligned}\bar{\xi}_{t,i} &= h(\tilde{\chi}_{t,i}; \Theta), \quad \bar{y}_t = \sum_{i=0}^{2k} w_i \bar{\xi}_{t,i}, \\ \bar{\Sigma}_{yy,t} &= \sum_{i=0}^{2k} w_i (\bar{\xi}_{t,i} - \bar{y}_t) (\bar{\xi}_{t,i} - \bar{y}_t)^\top + R, \\ \bar{\Sigma}_{Vy,t} &= \sum_{i=0}^{2k} w_i (\tilde{\chi}_{t,i} - \bar{V}_t) (\bar{\xi}_{t,i} - \bar{y}_t)^\top.\end{aligned}\tag{53}$$

With these moment conditions, we perform the filtering step the same as in the Kalman filter assuming normal distributions,

$$\widehat{V}_t = \bar{V}_t + \mathcal{K}_t (y_t - \bar{y}_t), \quad \widehat{\Sigma}_{VV,t} = \bar{\Sigma}_{VV,t} - \mathcal{K}_t \bar{\Sigma}_{yy,t} \mathcal{K}_t^\top,\tag{54}$$

where the Kalman gain is $\mathcal{K}_t = \bar{\Sigma}_{Vy,t} (\bar{\Sigma}_{yy,t})^{-1}$. We refer the reader to Wan and van der Merwe (2001) for general treatments of the unscented Kalman filter.

Given the forecasted option prices \bar{y}_t and their conditional covariance matrix $\bar{\Sigma}_{yy,t}$ obtained from the unscented Kalman filtering, we compute the quasi-log likelihood value for each week's observation on the option

prices assuming normally distributed forecasting errors,

$$l_t(\Theta) = -\frac{1}{2} \log |\bar{\Sigma}_{yy,t}| - \frac{1}{2} \left((y_t - \bar{y}_t)^\top (\bar{\Sigma}_{yy,t})^{-1} (y_t - \bar{y}_t) \right). \quad (55)$$

We choose model parameters to maximize the sum of the weekly log likelihood values on the options,

$$\Theta \equiv \arg \max_{\Theta} \mathcal{L}(\Theta, \{y_t\}_{t=1}^N), \quad \text{with} \quad \mathcal{L}(\Theta, \{y_t\}_{t=1}^N) = \sum_{t=1}^N l_t(\Theta), \quad (56)$$

where $N = 583$ denotes the number of weeks in our sample.

The model has ten parameters $(p, \kappa_Z, \theta_Z, \sigma_Z, \rho, \kappa_J, \theta_J, \sigma_J, v_{J^+}, v_{J^-})$ and three state variables (X_t, v_t^Z, v_t^J) to price 40 options each week for 583 weeks (all together 23,320 options). In addition, we have six parameters $(a, \kappa_{XX}, \kappa_{XZ}, \kappa_{XJ}, \gamma^v, \gamma^J)$ that control the market prices of risks and hence the statistical dynamics and one auxiliary parameter σ_e^2 for the measurement error variance. Without using the index return data, we leave the market price of the independent diffusion return risk (γ^Z) un-identified. Thus, the procedure estimates 17 parameters and filters out the time series of the three state variables.

The reduced-form benchmark model can be estimated using the same procedure. The model has nine parameters $(\kappa_Z, \theta_Z, \sigma_Z, \rho, \kappa_J, \theta_J, \sigma_J, v_{J^+}, v_{J^-})$ and two state variables (v_t^Z, v_t^J) to price the options, two parameters (γ^v, γ^J) to define the market prices of risks from the two activity rates, and one auxiliary parameter σ_e^2 for the measurement error variance, a total of 12 parameters.

5. Estimation results

First, we compare the performance of the full model and the reduced-form benchmark in pricing the index options across different strikes and maturities. Then, from the estimates of the structural parameters, we discuss the dynamics of the risk factors and how they contribute to the index option implied volatility skew and term structure.

5.1. Performance analysis

Table 2 reports the summary statistics of the pricing errors from the two estimated models, the full model on the left side and the reduced-form benchmark on the right side. The pricing errors are defined as the difference between the implied volatility quotes and the corresponding model values, in percentage points. Panel A in the table reports the sample averages of the pricing errors. A positive mean pricing error suggests that the model underprices the option on average. The mean pricing errors from the full model are mostly small except at the one-month maturity and do not show obvious patterns. The mean pricing errors from the reduced-form benchmark are larger overall and show an obvious term structure pattern: The mean pricing errors are mostly positive at short and long maturities but are negative at intermediate maturities. The pattern suggests that the model generates more curvature on the term structure than observed from the data.

[Table 2 about here.]

Panel B of Table 2 reports the root mean squared pricing error (rmse). The estimates from the full model are mostly smaller than those from the reduced-form benchmark. The average root mean squared error from the 40 implied volatility series is 0.83 for the full model and 1.187 for the reduced-form benchmark. Thus, the full model generates a 30% reduction over the benchmark in the average root mean squared pricing error.

Panel C reports the explained variation, defined as one minus the variance of the pricing errors over the variance of the original implied volatility series. The full model can explain all but two implied volatility series by over 90% and explain all but seven series by 95%. By contrast, the reduced-form benchmark leaves five series explained by less than 90% and leaves 15 series explained by less than 95%.

Finally, the last row of the table reports the maximized log likelihood values from the two models. We can think of the reduced-form benchmark as a constrained version of the full model. Then, the constrained version has 5 fewer parameters and one fewer state variable, and we can regard twice the difference between the log likelihood values as having a Chi-squared distribution with 588 degrees of freedom. The Chi-squared test static is 16,773. The constrained version is strongly rejected.

5.2. Parameter estimates

Table 3 reports the parameter estimates and their standard errors for both the full model and the reduced-form benchmark. Given the large amount of options data (23,320 observations) used for the model estimation, all parameters are estimated with strong statistical significance. We focus our discussion on the equity index dynamics implied by the estimated full model.

[Table 3 about here.]

The power coefficient p that governs the constant elasticity of variance behavior is estimated at 2.8427, suggesting a strong dependence of the index return variance on the leverage level X_t . When using a pure constant elasticity of variance model to fit the implied volatility skew in the S&P 500 index options, Jackwerth and Rubinstein (1996) find that the fitted values for the power coefficient can be as high as $p = 5$. Our estimate is much smaller as we have several other components in the model that also contribute to the interactions between the index return and return variance. The estimates also tend to be much smaller when the dynamics are fitted to the time-series of the index returns (Beckers (1980) and Christie (1982)) instead of the index option prices.

The activity rate v_t^Z underlying the diffusion return component Z is driven by a mean-reverting square-root process (10), with κ_Z measuring the risk-neutral mean-reverting speed and θ_Z the risk-neutral long-run mean. The estimate of $\kappa_Z = 3.0114$ indicates that this variance rate process is highly mean-reverting under the risk-neutral measure. As a result, shocks on the activity rate v_t^Z dissipates quickly as the option maturity increases. The mean estimate of $\theta_Z = 0.0244$ implies a return volatility contribution of $\sqrt{\theta_Z} = 15.62\%$ under the risk-neutral measure. The activity rate v_t^Z shows high instantaneous volatility of volatility at $\sigma_Z = 0.5988$ and highly negative instantaneous correlation with the corresponding return innovation component $\rho = -0.8354$. The high volatility of volatility contributes to the curvature of the implied volatility smile, whereas the negative instantaneous correlation adds to the negative implied volatility skew.

In contrast to the strong risk-neutral mean-reverting behavior of the diffusion variance rate v_t^Z , the arrival rate of the jump return component v_t^J has a very small risk-neutral mean-reverting speed, $\kappa_J = 0.0009$. As a result, the impact of the jump component persists across all option maturities. The estimate for the risk-neutral mean of the arrival rate is extremely large, magnifying the contribution of the component at long

option maturities. On the other hand, the large estimate on $\sigma_J = 5.6355$ indicates that a downside jump in the asset return evokes a much larger upside jump in the return variance.

The coefficients (v_{J+}, v_{J-}) measure the average sizes of the upside and downside return jumps under the risk-neutral measure. The average downside jump size is estimated to be $v_{J-} = 0.1926$ whereas the average upside jump is close to zero. The large difference in the average jump sizes generate another layer of skewness in the instantaneous stock return distribution under the risk-neutral measure.

We use four parameters to capture how the equity-to-asset ratio responds to the levels of financial leverage and business risk. The parameter κ_{XX} allows mean reversion in the capital structure decision. The small estimate suggests that the leverage ratio is a very persistent process. The parameter κ_{XZ} captures how the capital structure decision responds to the current level of diffusion business risk. Interestingly, the estimate is strongly positive, suggesting that X_t declines and hence financial leverage increases with increasing diffusion business risk. By contrast, the capital structure response to the jump business risk is captured by κ_{XJ} , which is estimated to be negative. Thus, the aggregate economy reduces financial leverage when the expected jump risk increases but increases financial leverage when the diffusion risk increases.

Finally, the estimation generates a significantly negative market price (γ^V) on the diffusion variance risk and a significantly positive market price (γ^J) on the jump return risk. Several studies, e.g., Bakshi and Kapadia (2003a,b) and Carr and Wu (2009), have documented negative variance risk premiums on stock indexes. Our model decomposes the index return into three risk sources (W_t, Z_t, J_t) , all contributing to the stochastic return variance. Based on the parameter estimates and the extracted state variables (X_t, v_t^Z, v_t^J) , we can calculate the average contribution to return variance and variance risk premium from each risk source. The financial leverage factor X_t generates an instantaneous index return variance of $v_t^X = X_t^{-2p}$, which has a sample average of 0.0119, or 10.91% in volatility term. The instantaneous risk-neutral drift of v_t^X is $\mu(v_t^X)^{\mathbb{Q}} = p(2p+1)(v_t^X)^2$, and the corresponding statistical drift is $\mu(v_t^X)^{\mathbb{P}} = (v_t^X)^2 (\kappa_X - \sigma_X(X_t^p)(a_X - \kappa_{XX}X_t - \kappa_{XZ}v_t^Z - \kappa_{XJ}v_t^J))$. The sample averages of the two drifts are 0.0037 and 0.0066, generating a slightly positive average instantaneous variance risk premium of 0.0028.

The diffusive component of the asset return contributes to the instantaneous index return variance by v_t^Z , which generates a sample average of 0.0231, or 15.19% in volatility term. We specify v_t^Z as following a square-root diffusion dynamics, with the instantaneous risk premium on v_t^Z given by $\gamma^V \sigma_Z v_t^Z$. With a negative

estimate on the market price γ' , the instantaneous risk premium averages at -0.2412 .

Finally, the jump component of the asset return contributes to the instantaneous index return variance by $(v_{J+}^2 + v_{J-}^2)v_t^J$ under the risk-neutral measure and $((v_{J+}^{\mathbb{P}})^2 + (v_{J-}^{\mathbb{P}})^2)v_t^J$ under the statistical measure, with the difference induced by the market price of jump risk γ^J . The sample averages of the two variance series are 0.0137 (11.72% in volatility term) and 0.0116 (10.79% in volatility term), respectively, thus generating a negative variance risk premium of -0.0021 . Furthermore, the market price of the jump risk also induces a difference between the statistical and risk-neutral drifts of the jump arrival rate v_t^J at $\sigma_J(v_{J-}^{\mathbb{P}} - v_{J-})v_t^J$, which averages at -0.0318 .

Figure 6 plots the time series of the extracted state variables. The financial leverage and its contribution to the index return variance (v_t^X) reached historical highs before the burst of the Nasdaq bubble, followed by a deleveraging process. The diffusion return variance (v_t^Z) reached its highest point during the 2003 recession, but had been low during the years after. The jump risk (v_t^J) spiked to its highest level during the hedge fund crisis and the ensuing Asian crises in 1999.

[Fig. 6 about here.]

5.3. Multiple sources of index return variance variation

According to the index return dynamics representation in equation (15), the variation of the index return is governed by three random sources, (W_t, Z_t, J_t) . Both the variance rates for the two Brownian motions (W_t, Z_t) and the arrival rate for the jump (J_t) are stochastic and are governed by separate dynamic processes, (v_t^X, v_t^Z, v_t^J) , respectively. The risk-neutral dynamics of the three state variables determine how shocks to the three sources of variance risks dissipate across the return variance term structure. A transient shock mainly affects the short-term return variance and hence short-term options, whereas a persistent shock affects return variance and option prices at both short and long maturities. The $v_t^X = X_t^{-2p}$ process is mean-repelling under the risk-neutral measure. Thus, shocks on X_t have long-lasting impacts on the return variance term structure. Furthermore, the risk-neutral mean-reverting speed for v_t^J is estimated very small at $\kappa_J = 0.0009$, suggesting that shocks on v_t^J are also long-lasting. On the other hand, the risk-neutral mean-reverting speed for v_t^Z is quite large at $\kappa_Z = 3.0114$. As a result, shocks on v_t^Z tend to dissipate quickly as the return variance horizon

increases.

In Figure 7, we plot the responses of the implied volatilities against shocks on each of the three risk factors across different strikes and maturities. Panels in the first, second, and the third row plot responses to the equity-to-asset ratio X_t , the activity rate underlying the diffusion return component v_t^Z , and the activity rate underlying the jump return component v_t^J , respectively. In each row, panels on the left plot the responses of one-month option implied volatilities across different strikes, panels in the middle plot the responses of the five-year option implied volatilities across different strikes, and panels on the right plot responses of at-the-money (100% strike) option implied volatilities across different time to maturities. In each panel, the solid lines represent the model-generated implied volatilities when evaluated at the sample averages of the state variables; the dashed lines represent the model-generated values when evaluated at the 90th-percentile of the state variable in question while holding the other two state variables at their sample averages; the dash-dotted lines represent responses to a shift to the 10th percentile for the state variable in question while holding the other two state variables at their sample averages.

[Fig. 7 about here.]

When we shift the equity-to-asset ratio X_t from its sample average to its 90th-percentile value while holding everything else constant, the financial leverage is reduced and the contribution of the factor to the index return variance (X_t^{-2p}) is reduced accordingly. Thus, as shown in the first row of Figure 7, the dashed lines are below the solid line across all strikes and maturities. On the other hand, when the equity-to-asset ratio is reduced to its 10th-percentile value, the financial leverage increases, and the option implied volatilities (dash-dotted lines) increase and hence move above the solid line. At short maturities, shocks to X_t affect the implied volatilities more at-the-money than out-of-the-money. The implied volatility response becomes more uniform across strikes at long maturities. When we plot the at-the-money implied volatility term structure in the right panel, we observe that the responses of the at-the-money implied volatilities to the X_t risk factor are relatively uniform across maturities. A shock in X_t induces a near parallel shift in the at-the-money implied volatility term structure.

When the activity rate v_t^Z experiences a positive shock, the return variance contribution from the diffusion component (Z_t) increases. In response, the option implied volatilities move up. Panels in the second row of Figure 7 show that the dashed lines are all above the solid line. A negative shock to v_t^Z lowers the option

implied volatility (dash-dotted lines). At short maturities, the responses are stronger at high strikes than at low strikes. At long maturities, the implied volatility responses become much smaller but more uniform across different strikes. The term structure plot on the right side highlights the transient nature of the v_t^Z risk factor: Shocks to v_t^Z induce large responses at short maturities, but the response declines quickly as the option maturity increases.

When the activity rate v_t^J experiences a positive shock, the return variance increases from the contribution of the jump component. Furthermore, as v_t^J governs the intensity of the negatively skewed return jump component, a positive shock in v_t^J also increases the negative skewness of the instantaneous return distribution. In response, a positive shock to v_t^J not only raises the implied volatility level, but also increases the negative skew, especially at short maturities, by raising the low-strike implied volatility more than the high-strike implied volatility. Furthermore, given the low mean-reversion estimate, shocks to v_t^J have a persistent impact on the at-the-money implied volatility term structure.

5.4. Different channels of implied volatility skew

It is well-known that the equity index option implied volatilities show a persistent negative skew pattern along the moneyness dimension. Several mechanisms have been proposed in the literature to account for a negative skew at a particular maturity, including downside jumps in equity returns, negative correlations between equity returns and return volatilities, and local-volatility type price level dependence. Our model accommodates all these mechanisms and more. Under our specified dynamics for the equity index, the index option implied volatility skew can come from four distinct sources: (i) the leverage effect through the strictly positive power coefficient p , (ii) the instantaneous negative correlation $\rho < 0$ between diffusion return and volatility risks, (iii) the return jump component that generates much larger downside jumps than upside jumps ($v_{J-} \gg v_{J+}$), and (iv) the self-exciting downside jumps in the asset return that lead to more frequent arrival of jumps. Based on the parameter estimates, we assess the contribution of each source to the implied volatility skews at different maturities.

Figure 8 plots the term structure of the model-generated implied value skew, calculated according to equation (46) based on model-generated implied volatility estimates at different strikes and maturities. The model values are generated at the sample averages of the three state variables. The solid lines in all panels

plot the same term structure generated from the model parameter estimates reported in Table 3. The dashed lines in each panel are generated by increasing the absolute magnitude of one particular parameter and the dashed-dotted lines are generated by decreasing the absolute magnitude of the same parameter, while holding all other parameters to their original estimates.

The first panel shows the effect of varying the power coefficient p that governs the constant elasticity of variance behavior. Table 3 reports an estimate of $p = 2.8427$. When we increase the power coefficient to 3.3 for the dashed line in the first panel, the implied volatility skew becomes steeper except at very long maturities. On the other hand, lowering the coefficient to 2.3 lower the skew at short maturities, but raises the skew at long maturities, as shown by the dash-dotted line. The fact that the effect switch directions from short to long maturities indicates that the power coefficient interacts with other components of the index return dynamics to generate the implied volatility skew.

[Fig. 8 about here.]

The second panel in Figure 8 plots the effect of the instantaneous correlation ρ . The correlation is estimated to be strongly negative at -0.8354 . When we alter the correlation estimate, its effect on the implied volatility skew is relatively uniform across all maturities. Increasing the absolute magnitude of the negative correlation increases the implied volatility skew at both short and long maturities.

To capture the effect of the return jumps, we alter the average size of the downside jump (v_{J^-}) and plots the effect of this perturbation on the skew term structure in the third panel of Figure 8. Reducing the average size of downside jumps reduces the negative skewness of the jump return component and hence lowers the skew estimate at the short maturities. Furthermore, due to the persistent risk-neutral dynamics of v_t^J , this jump size effect persists to long option maturities.

In the last panel, we capture the self-exciting feature through the σ_J parameter. Setting σ_J to zero would completely remove the self-exciting feature as negative return jumps would not be able to increase the arrival rate of future jumps. When we vary the parameter σ_J , the skew estimates virtually do not change at short maturities, but the skew variations at long maturities are quite significant. Increasing the intensity of the self-exciting feature by increasing σ_J generates stronger negative skews in the option implied volatility plots along the moneyness dimension.

6. Concluding remarks

The movements of the stock index and index volatility can interact through several distinct channels. First, holding the aggregate debt level and business risk fixed, a decline in the index level increases the financial leverage of the aggregate economy and raises the equity volatility through the increased financial leverage. Second, holding the leverage ratio fixed, an increase in the business risk increases the discounting and hence reduces the valuation of future cash flows, thus generating an instantaneous negative correlation between asset return and risk. Finally, the stock market experiences both continuous movements and discontinuous large disruptions. The large disruptions, especially negative ones, often exhibit a self-exciting feature in that the occurrence of one disruption induces more disruptions to follow, thus again raising the index return volatility. In this paper, we capture all three channels of interactions by separately modeling the asset return dynamics and the financial leverage variations. We analyze the implications of the dynamic specifications on pricing the equity index options across different strikes and maturities.

We propose a numerically tractable procedure to price equity index options under the specified dynamics, and we estimate the dynamics using about a decade worth of over-the-counter equity index options data that span five different strikes at each fixed time to maturity and eight different fixed time to maturities from one month to five year. The estimation results show that the model performs well in pricing the equity index options. From the parameter estimates, we analyze the dynamics of each risk sources and how they affect the implied volatility behaviors across strikes and maturities.

Our analysis highlights the virtue of separating the effect of financial leverage from the dynamics of asset returns that underly the fundamental business risk. Although our analysis is on the equity index, the separate treatment is all the more appropriate for individual companies. In a recent working paper, Choi and Richardson (2008) study the separate variations of firm volatilities and financial leverages across different companies. A line for future research is to link the cross-sectional differences in firm volatilities and financial leverages to the different behaviors of individual stock options across strike and maturities.

References

- Adrian, T., Shin, H. S., 2008. Liquidity and leverage. *Journal of Financial Intermediation* forthcoming.
- Azizpour, S., Giesecke, K., 2008. Self-exciting corporate defaults: Contagion vs. frailty. Working paper. Stanford University.
- Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *Journal of Finance* 52, 2003–2049.
- Bakshi, G., Ju, N., Ou-Yang, H., 2006. Estimation of continuous-time models with an application to equity volatility. *Journal of Financial Economics* 82, 227–249.
- Bakshi, G., Kapadia, N., 2003a. Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies* 16, 527–566.
- Bakshi, G., Kapadia, N., 2003b. Volatility risk premium embedded in individual equity options: Some new insights. *Journal of Derivatives* 11, 45–54.
- Bates, D., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies* 9, 69–107.
- Bates, D., 2000. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics* 94, 181–238.
- Beckers, S., 1980. The constant elasticity of variance model and its implications for option pricing. *Journal of Finance* 35, 661–673.
- Black, F., 1976a. The pricing of commodity contracts. *Journal of Financial Economics* 3, 167–179.
- Black, F., 1976b. Studies of stock price volatility changes. In: *Proceedings of the 1976 American Statistical Association, Business and Economical Statistics Section American Statistical Association, Alexandria, VA.*
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.
- Borodin, A. N., Salminen, P., 1996. *Handbook of Brownian Motion*. Birkhauser, Boston, MA.
- Carr, P., Madan, D., 1999. Option valuation using the fast Fourier transform. *Journal of Computational Finance* 2, 61–73.

- Carr, P., Sun, J., 2007. A new approach for option pricing under stochastic volatility. *Review of Derivatives Research* 10, 87–150.
- Carr, P., Wu, L., 2003. Finite moment log stable process and option pricing. *Journal of Finance* 58, 753–777.
- Carr, P., Wu, L., 2004. Time-changed Lévy processes and option pricing. *Journal of Financial Economics* 71, 113–141.
- Carr, P., Wu, L., 2007. Stochastic skew in currency options. *Journal of Financial Economics* 86, 213–247.
- Carr, P., Wu, L., 2009. Variance risk premiums. *Review of Financial Studies* 22, ??
- Chacko, G., Viceira, L., 2003. Spectral GMM estimation of continuous-time processes. *Journal of Econometrics* 116, 259–292.
- Choi, J., Richardson, M., 2008. The volatility of the firms assets. Working paper. New York University.
- Christie, A. A., 1982. The stochastic behavior of common stock variances: Value, leverage, and interest rate effects. *Journal of Financial Economics* 10, 407–432.
- Christoffersen, P. F., Heston, S. L., Jacobs, K., 2007. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. Working paper. McGill University.
- Cochrane, J., 2004. *Asset Pricing*. Princeton University Press, Princeton, NJ.
- Cox, J. C., 1996. The constant elasticity of variance option pricing model. *Journal of Portfolio Management* 23, 15–17.
- Davis, P. J., Rabinowitz, P., 1984. *Methods of Numerical Integration*. Academic Press, New York.
- Davydov, D., Linetsky, V., 2001. Pricing and hedging path-dependent options under the cev process. *Management Science* 47, 949–965.
- Derman, E., 1999. Regimes of volatility. *Risk* April, 55–59.
- Ding, X., Giesecke, K., Tomecek, P. I., 2008. Time-changed birth processes and multi-name credit derivatives. *Operations Research* forthcoming.
- Duffie, D., 1992. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey.
- Duffie, D., Filipović, D., Schachermayer, W., 2003. Affine processes and applications in finance. *Annals of Applied Probability* 13, 984–1053.

- Dupire, B., 1994. Pricing with a smile. *Risk* 7, 18–20.
- Emanuel, D. C., MacBeth, J. D., 1982. Further results on the constant elasticity of variance call option pricing model. *Journal of Financial and Quantitative Analysis* 17, 533–554.
- Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance* 59, 1367–1404.
- Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in equity index volatility and returns. *Journal of Finance* 58, 1269–1300.
- Filipović, D., 2001. A general characterization of one factor affine term structure models. *Finance and Stochastics* 5, 389–412.
- Harrison, J. M., Kreps, D. M., 1979. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20, 381–408.
- Heston, S., 1993. Closed-form solution for options with stochastic volatility, with application to bond and currency options. *Review of Financial Studies* 6, 327–343.
- Heston, S., 1997. A simple new formula for options with stochastic volatility. Working paper. University of Maryland.
- Heston, S., Nandi, S., 2000. A closed-form GARCH option valuation model. *Review of Financial Studies* 13, 585–625.
- Huang, J., Wu, L., 2004. Specification analysis of option pricing models based on time-changed Lévy processes. *Journal of Finance* 59, 1405–1440.
- Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42, 281–300.
- Ishida, I., Engle, R. F., 2002. Modeling variance of variance: The square root, the affine, and the CEV GARCH models. Working paper. New York University.
- Jackwerth, J. C., Rubinstein, M., 1996. Recovering stochastic processes from option prices. Working paper. University of California–Berkeley.
- Javaheri, A., 2005. *Inside Volatility Arbitrage: The Secrets of Skewness*. John Wiley & Sons, London.

- Jones, C. S., 2003. The dynamics of stochastic volatility: Evidence from underlying and options markets. *Journal of Econometrics* 116, 181–224.
- Küchler, U., Sørensen, M., 1997. *Exponential Families of Stochastic Processes*. Springer, New York.
- Lewis, A. L., 2000. *Option Valuation under Stochastic Volatility*. Finance Press, Newport Beach, California, USA.
- Madan, D., Seneta, E., 1990. The variance gamma (V.G.) model for share market returns. *Journal of Business* 63, 511–524.
- Madan, D. B., Carr, P. P., Chang, E. C., 1998. The variance gamma process and option pricing. *European Finance Review* 2, 79–105.
- Medvedev, Scaillet, O., 2003. A simple calibration procedure of stochastic volatility models with jumps by short term asymptotics. Working paper. HEC, University of Geneva.
- Pan, J., 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.
- Revuz, D., Yor, M., 1999. *Continuous Martingales and Brownian Motion*. Springer, Berlin, Germany.
- Schroder, M., 1989. Computing the constant elasticity of variance option pricing formula. *Journal of Finance* 44, 211–219.
- Titman, S., Tompaidis, S., 2005. Determinants of credit spreads in commercial mortgages. *Real Estate Economics* 33, 711–738.
- Wan, E. A., van der Merwe, R., 2001. The unscented Kalman filter. In: Haykin, S. (Eds.), *Kalman Filtering and Neural Networks*. Wiley & Sons Publishing, New York.

Table 1
Summary statistics of implied volatility quotes

\mathcal{K}/m	1	3	6	12	24	36	48	60
A. Mean								
80	33.84	28.95	26.69	25.18	24.53	24.46	24.56	24.77
90	25.91	23.97	23.11	22.63	22.70	22.97	23.30	23.66
100	18.81	19.39	19.75	20.20	20.93	21.54	22.08	22.60
110	15.11	16.02	16.96	18.01	19.29	20.20	20.94	21.59
120	14.28	14.59	15.25	16.36	17.86	18.97	19.87	20.64
B. Standard deviation								
80	6.65	5.90	5.25	4.80	4.63	4.41	4.21	4.08
90	6.22	5.61	5.06	4.66	4.48	4.32	4.17	4.06
100	6.06	5.41	4.92	4.57	4.37	4.25	4.14	4.04
110	5.09	4.92	4.70	4.47	4.29	4.20	4.12	4.02
120	4.52	4.24	4.19	4.18	4.12	4.11	4.08	4.01
C. Weekly autocorrelation								
80	0.937	0.965	0.974	0.981	0.986	0.987	0.988	0.988
90	0.937	0.963	0.973	0.980	0.986	0.987	0.988	0.988
100	0.936	0.962	0.972	0.980	0.985	0.987	0.988	0.988
110	0.944	0.961	0.972	0.980	0.985	0.987	0.988	0.988
120	0.963	0.969	0.972	0.980	0.984	0.987	0.988	0.989

Entries report the summary statistics of 40 implied volatility series (in percentage points) across five relative strikes (\mathcal{K} , in percentages of the spot index level) and eight time to maturities (m , in months). Panel A reports the sample average, panel B reports the standard deviation, and panel C reports the weekly autocorrelation estimates. The data are sampled weekly (on Wednesdays) from January 8, 1997 to March 5, 2008, 583 observations for each series.

Table 2
Summary statistics of pricing errors

\mathcal{K}/m	Full model								Reduced-form benchmark							
	1	3	6	12	24	36	48	60	1	3	6	12	24	36	48	60
A. Mean pricing error																
80	-0.595	0.508	0.200	0.056	-0.061	-0.208	-0.285	-0.260	0.867	1.141	0.253	-0.358	-0.414	-0.201	0.112	0.499
90	0.389	0.250	0.189	0.141	0.068	-0.003	-0.019	0.046	0.152	0.225	-0.062	-0.329	-0.341	-0.136	0.172	0.554
100	-0.237	0.030	0.118	0.086	0.043	0.055	0.108	0.220	0.328	0.149	0.007	-0.233	-0.312	-0.123	0.182	0.563
110	-0.654	-0.112	0.015	-0.021	-0.048	0.032	0.149	0.307	0.948	0.369	0.132	-0.151	-0.314	-0.143	0.160	0.541
120	-1.297	0.215	0.227	0.062	-0.099	-0.019	0.132	0.326	-0.442	-0.157	-0.252	-0.299	-0.358	-0.196	0.103	0.483
B. Root mean squared error																
80	2.216	1.103	1.050	0.836	0.607	0.550	0.770	1.064	3.904	2.155	1.515	1.301	1.131	0.902	0.908	1.181
90	1.225	0.727	0.701	0.641	0.445	0.279	0.445	0.758	2.252	1.112	0.852	0.810	0.708	0.591	0.711	1.056
100	1.111	0.409	0.474	0.555	0.418	0.286	0.377	0.657	1.290	0.601	0.641	0.678	0.528	0.425	0.592	0.966
110	1.499	0.720	0.720	0.717	0.551	0.436	0.465	0.674	2.486	0.982	1.067	1.025	0.711	0.488	0.554	0.901
120	4.014	1.081	1.057	0.984	0.714	0.561	0.572	0.735	3.836	1.488	1.299	1.327	1.014	0.663	0.583	0.856
C. Explained variation																
80	0.897	0.972	0.961	0.970	0.983	0.987	0.971	0.936	0.672	0.904	0.919	0.932	0.948	0.960	0.954	0.931
90	0.965	0.985	0.982	0.982	0.990	0.996	0.989	0.965	0.869	0.962	0.972	0.975	0.981	0.982	0.973	0.951
100	0.968	0.994	0.991	0.986	0.991	0.996	0.992	0.977	0.957	0.988	0.983	0.981	0.990	0.991	0.981	0.962
110	0.930	0.979	0.977	0.974	0.984	0.989	0.989	0.978	0.795	0.966	0.949	0.949	0.978	0.988	0.983	0.968
120	0.293	0.938	0.939	0.945	0.971	0.981	0.981	0.973	0.289	0.878	0.907	0.904	0.947	0.976	0.980	0.969
Average	0.810	0.974	0.970	0.971	0.984	0.990	0.984	0.966	0.717	0.940	0.946	0.948	0.969	0.979	0.974	0.956
\mathcal{L}	46,651								38,265							

Entries report the summary statistics of the pricing errors from both our full model (left side) and the reduced-form benchmark (right side). The pricing errors are defined as the difference between the implied volatility quotes and the corresponding model values, both in percentage points. Panel A reports the sample averages of the pricing errors. Panel B reports the root mean squared pricing errors. Panel C reports the explained variation, defined as one minus the ratio of the pricing error variance to the variance of the original implied volatility series. The last panel reports the maximized log likelihood values for the two models.

Table 3
Maximum likelihood estimates of model parameters

Θ	Full model		Reduced-form benchmark	
	Estimates	Std. Error	Estimates	Std. Error
p	2.8427	0.0074		
κ_Z	3.0114	0.0127	1.0933	0.0076
θ_Z	0.0244	0.0000	0.0479	0.0001
σ_Z	0.5988	0.0015	0.4923	0.0025
ρ	-0.8354	0.0016	-0.7242	0.0027
κ_J	0.0009	0.0000	0.0000	0.0000
θ_J	113.8562	0.0715	0.0394	0.0006
σ_J	5.6355	0.0272	19.6423	0.3165
ν_{J^+}	0.0000	0.0000	0.0197	0.0004
ν_{J^-}	0.1926	0.0002	0.0779	0.0004
a_X	0.0003	0.0000		
κ_{XX}	0.0001	0.0000		
κ_{XZ}	17.5360	0.3087		
κ_{XJ}	-0.0774	0.0000		
γ^v	-17.4507	0.3048	-18.7061	0.5781
γ^f	0.4468	0.0023	-0.0000	0.0000
σ_e^2	0.0052	0.0000	0.0120	0.0000

Entries report the maximum likelihood estimates of the model parameters and their standard errors for both the full model and the reduced-form benchmark.

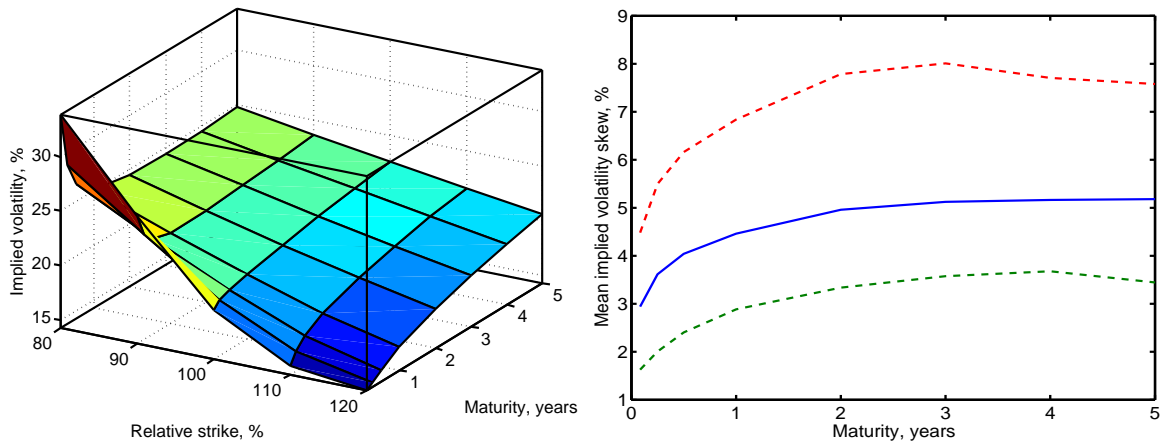


Fig. 1. The mean implied volatility surface and implied volatility skew. The left panel plots the sample averages of the implied volatilities as a function of relative strikes and time to maturities. The right panel plots the sample averages (solid line) and the 10th and 90th percentiles (dashed lines) of the implied volatility skew estimates across different maturities. The skew is defined as the implied volatility differences at 80% and 120% strikes divided by the absolute difference in the corresponding standardized moneyness measure $d \equiv \ln(\mathcal{K}/100)/(IV\sqrt{\tau}/100)$, where IV is the implied volatility quote (in percentage points) at the percentage relative strike \mathcal{K} and time to maturity τ .

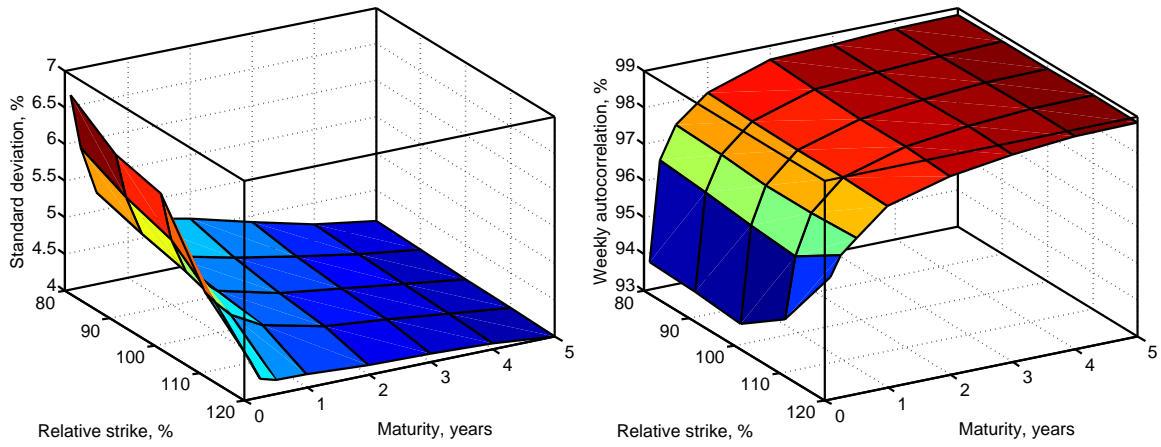


Fig. 2. The weekly standard deviation (left panel) and autocorrelation (right panel) of the implied volatility series are plotted against the relative strikes and time to maturities.

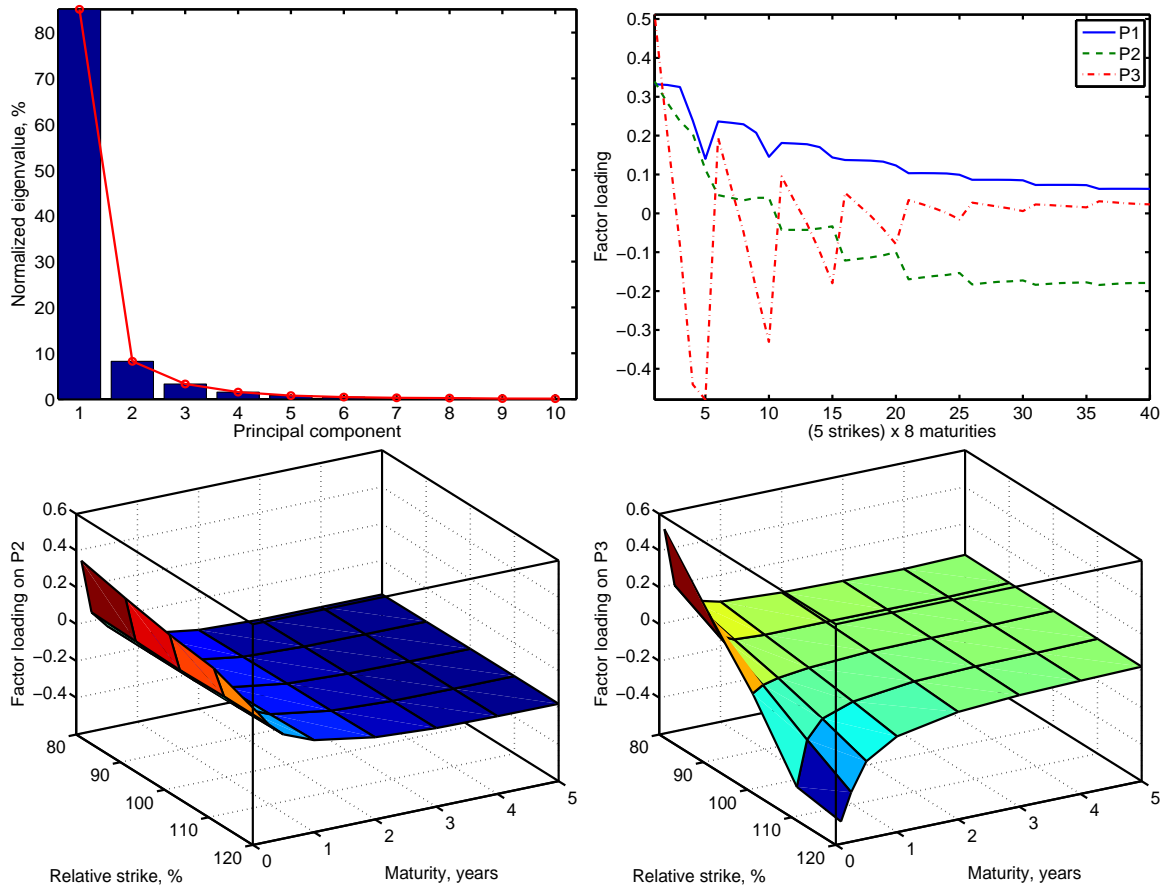


Fig. 3. Principal component analysis on the implied volatility surface. The bar chart in the left panel plots the first ten normalized eigenvalues of covariance matrix of weekly changes on the 40 implied volatility series. They can be interpreted as the percentage variation explained by each principal component. The right panel plot the eigenvectors of the first three eigenvalues (P1, P2, P3, respectively), which can be interpreted as the loading coefficients of the three principal components on the 40 implied volatility series. The 40 implied volatility series are ranked first in five strikes from 80% to 120% at each maturity, and then in the eight maturities from one month to five years.

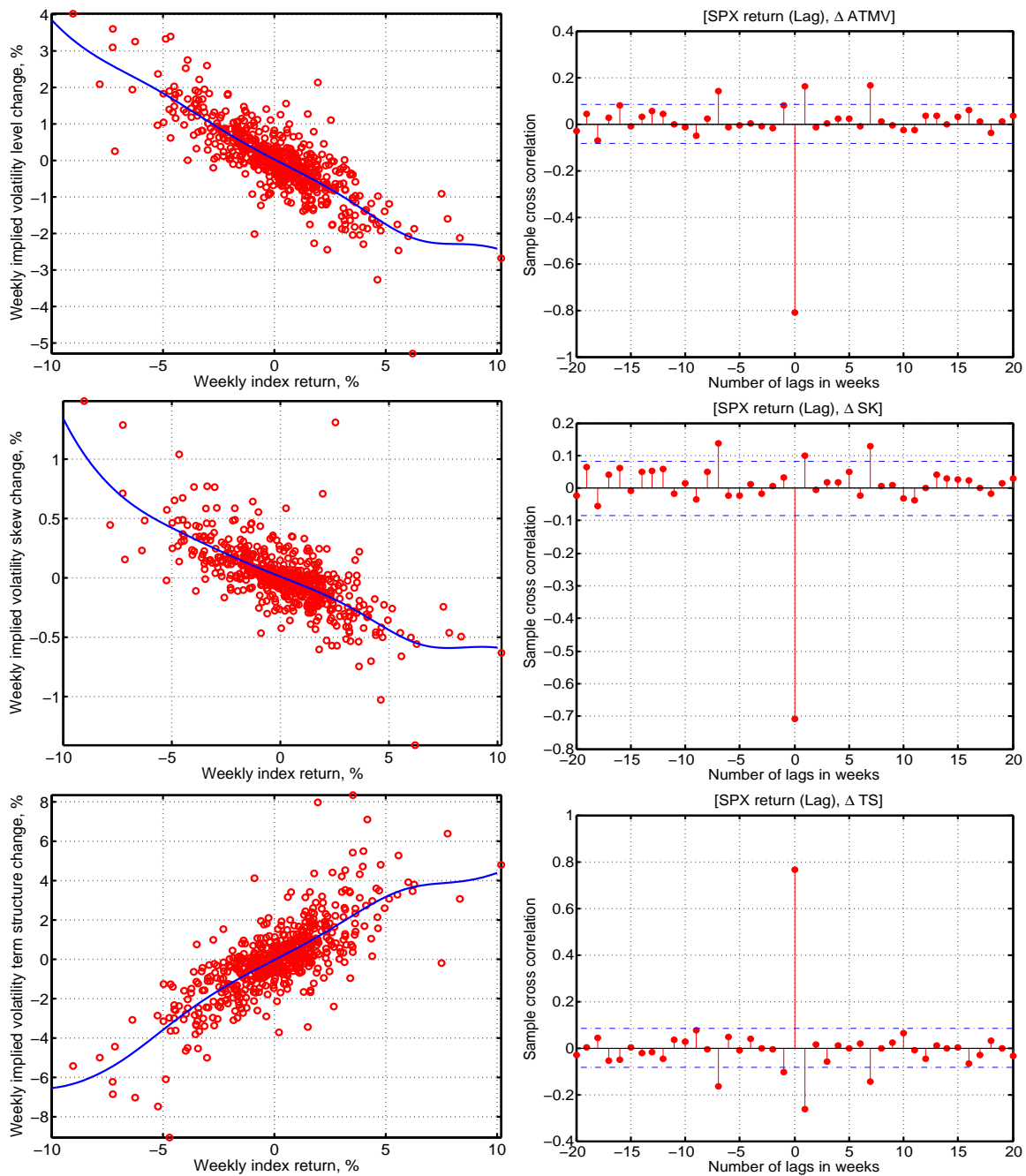


Fig. 4. Co-movements between the index returns and the implied volatility surface. The three panels on the left side plot weekly changes in the average at-the-money implied volatility level (top panel), the average implied volatility skew (middle panel), and the at-the-money implied volatility term structure (bottom panel) against weekly index returns. Circles denote data points and solid lines represent local linear regression fits with a Gaussian kernel and a default bandwidth choice. The three panels on the right side plot the cross-correlogram at different leads and lags between the weekly index return and weekly changes in the three implied volatility series.

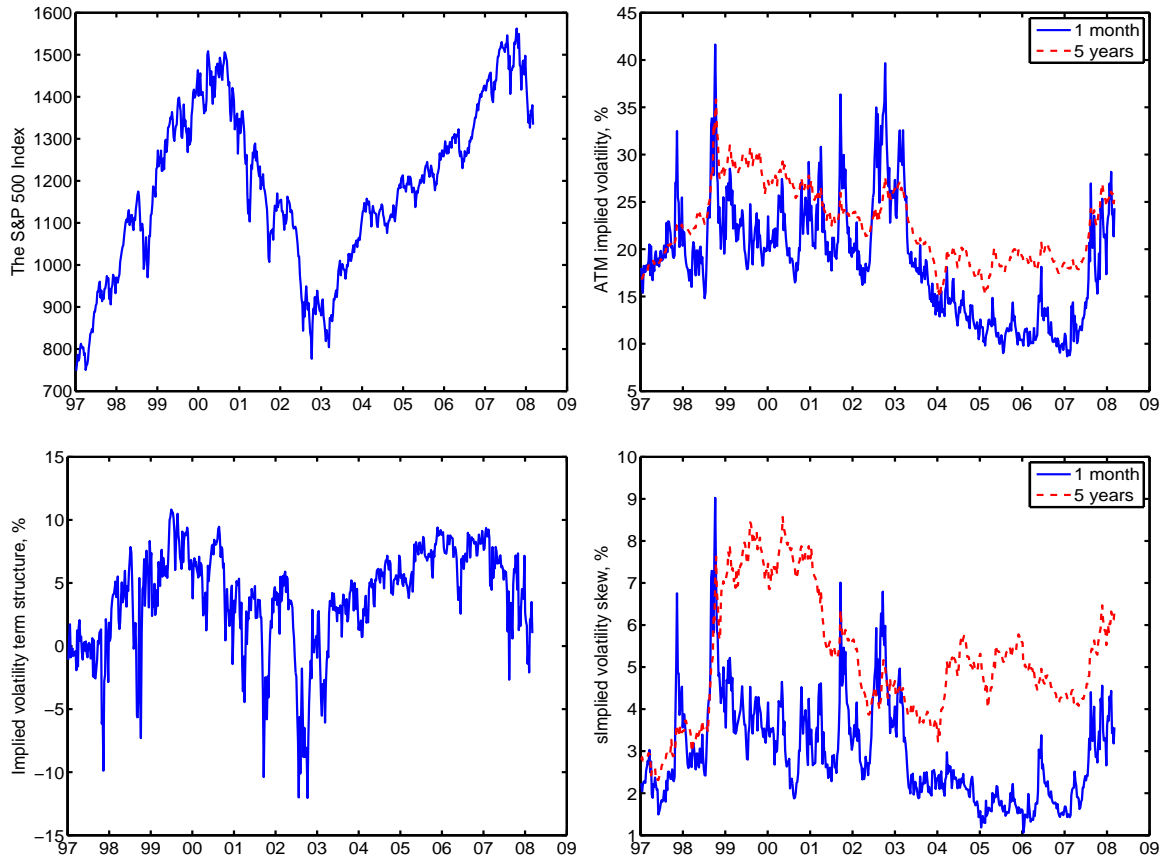


Fig. 5. The time series of the S&P 500 index and the option implied volatility combinations. The first panel plots the time series of the S&P 500 index. The second panel plots the at-the-money implied volatility at one month (solid line) and five-year (dashed line) maturities. The third panel plots the at-the-money implied volatility term structure, defined as the difference between five-year and one-month at-the-money implied volatilities. The last panel plots the implied volatility skew (SK_t) series at one-month and five-year maturities.

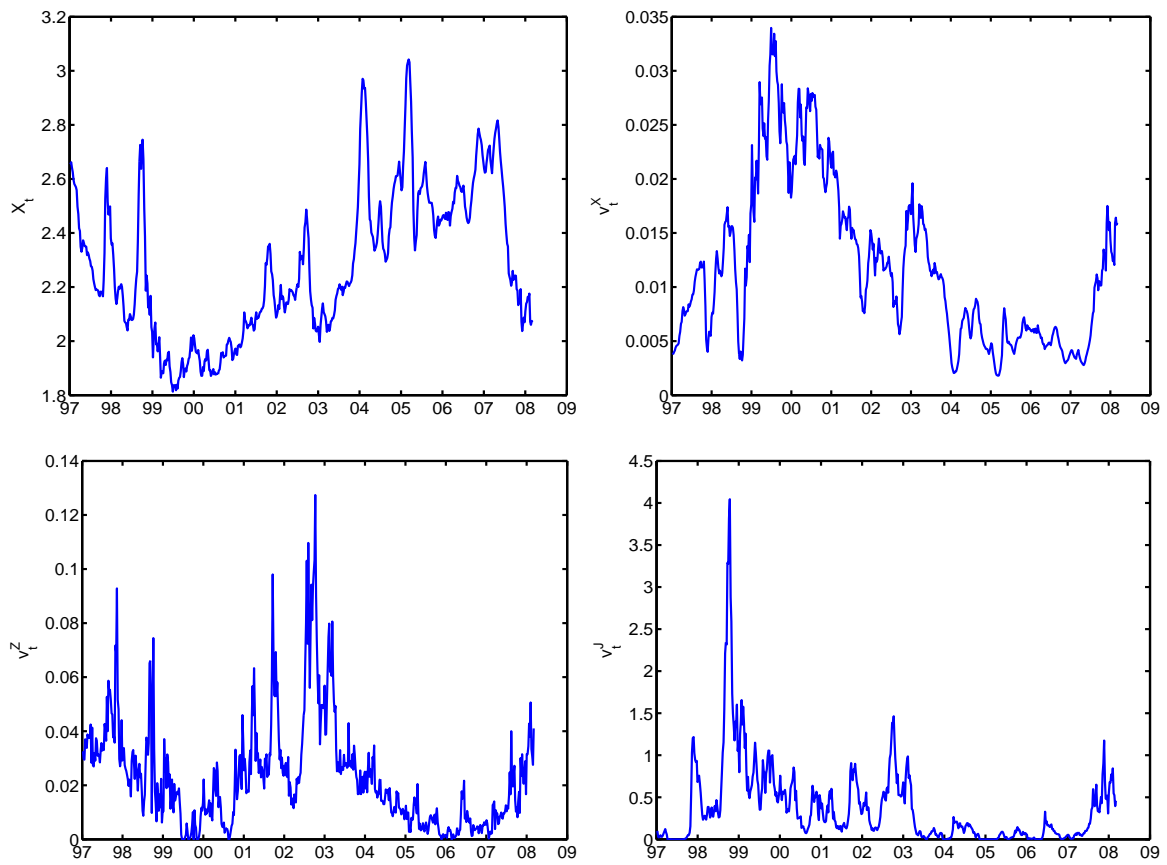


Fig. 6. The time series of the state variables. The state variables time series are extracted using the unscented Kalman filter under the estimated model parameters.

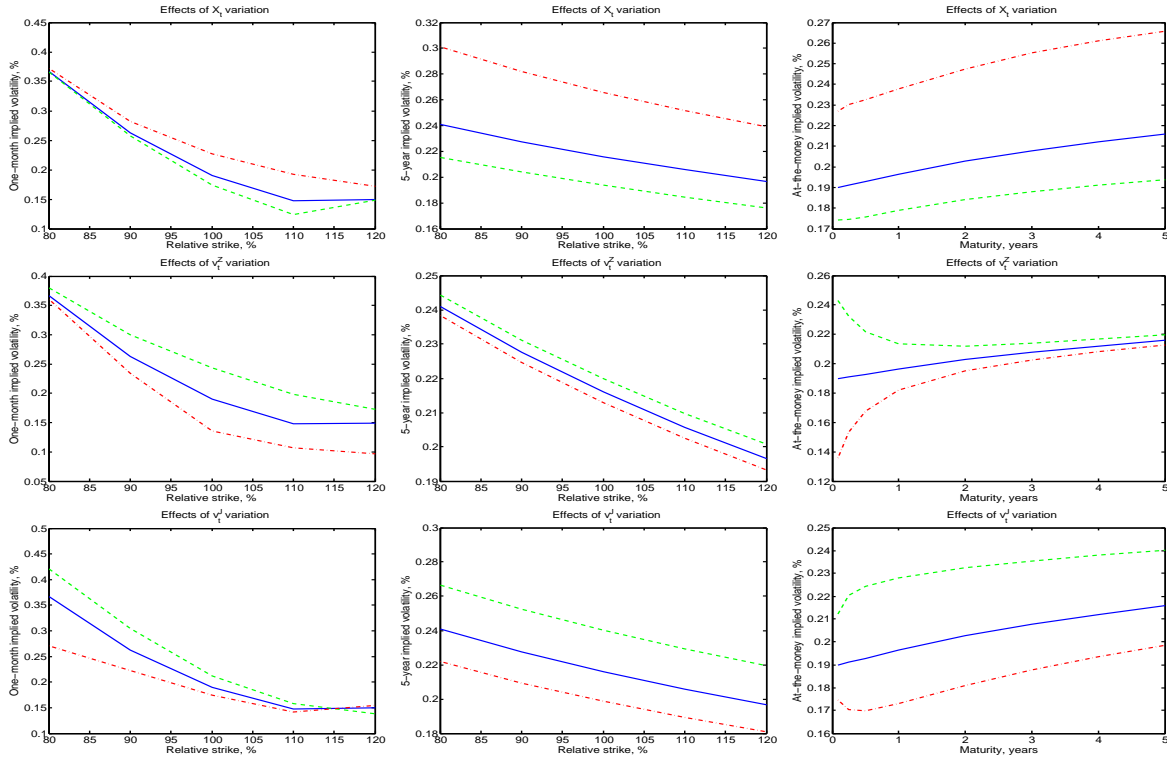


Fig. 7. Shocks and implied volatility responses. The solid lines in each panel represent the implied volatility generated from the estimated model when evaluated at the sample average of the state variables. The dashed lines are obtained by setting one state variable to its 90th percentile while holding the other two to their average values. The dashed-dotted lines are obtained by setting one state variable to its 10th percentile while holding the other two to their average. The three rows represent implied volatility responses to the three state variables. The first two columns plot the responses of one-month and five-year implied volatilities, respectively, across different strikes. The last column plots the responses of at-the-money implied volatilities at different maturities.

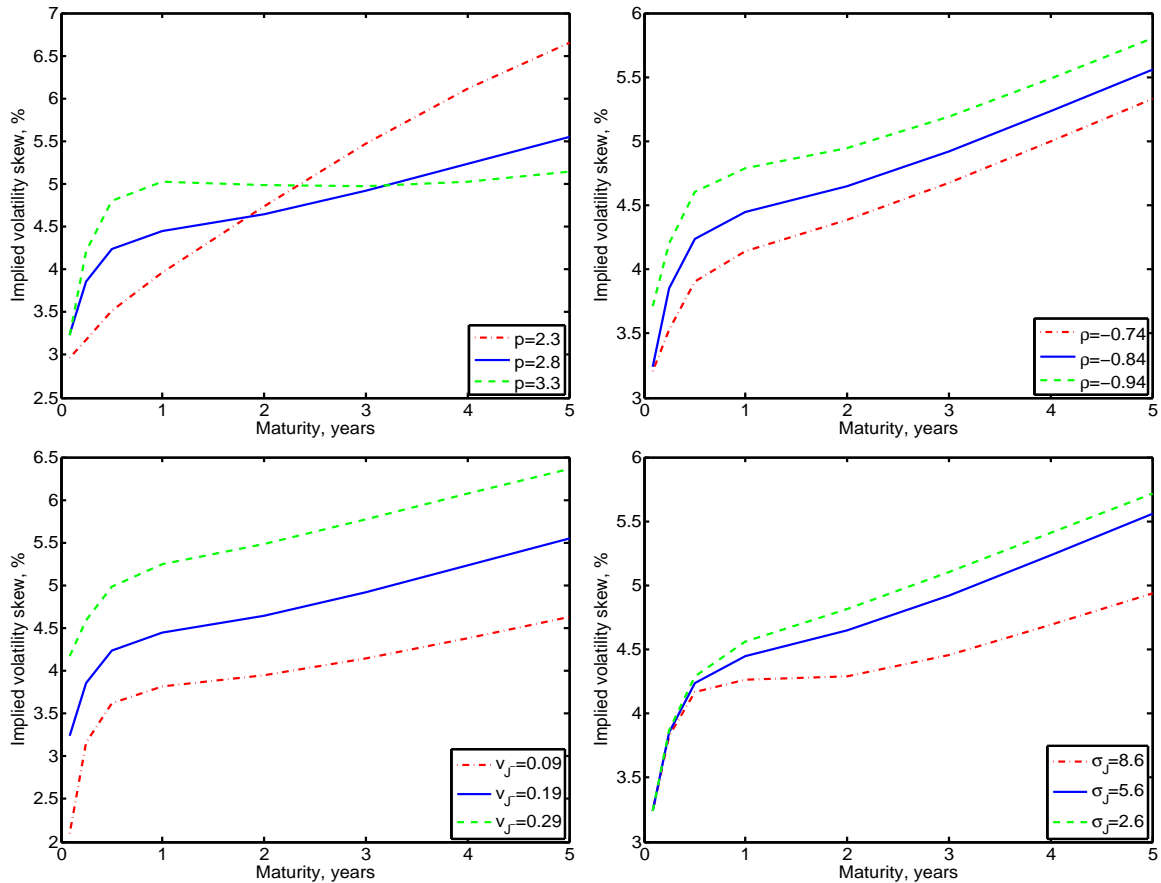


Fig. 8. Multiple sources of implied volatility skews. Each panel plots the effect of varying one parameters on the term structure of the option implied volatility skew while holding other parameters to their maximum likelihood estimates. The skew is defined according to equation (46) based on model-generated implied volatility values.