Three make a dynamic smile – unspanned skewness
and interacting volatility components
in option valuation

PETER H. GRUBER, CLAUDIO TEBALDI, and FABIO TROJANI*

First version: June 2009. This version: November 4, 2010

ABSTRACT

We study a new class of three-factor affine option pricing models with interdependent volatility
dynamics and a stochastic skewness component unrelated to volatility shocks. These
properties are useful in order (i) to model a term structure of implied volatility skews more
consistent with the data and (ii) to capture comovements of short and long term skews largely
unrelated to the volatility dynamics. We estimate our models using about fourteen years of
S&P 500 index option data and find that on average they improve the out-of-sample pricing
accuracy of benchmark two- and three-factor affine models by 20%. Using an appropriate
decomposition of volatility and skewness, highlighting the main directions of improvements
produced by our setting, we show that the enhanced fit results from an improved modeling
of the term structure of implied-volatility skews. The largest pricing improvements tend
to concentrate during periods of financial crises or market distress, suggesting volatility-
unrelated skewness as a potentially useful reduced-form risk factor for reproducing some of
the crisis-related dynamics of index option smiles.

JEL classification: G10, G12, G13

Keywords: Option Pricing, Stochastic Volatility, Stochastic Leverage, Short and Long Run
Volatility Risk, Matrix Affine Jump Diffusions.

*Peter H. Gruber is at the University of Lugano. Claudio Tebaldi is from Bocconi University, IGIER and
a CAREFIN fellow. Fabio Trojani is at the University of Lugano and a Research Fellow of the Swiss
Finance Institute. All authors gratefully acknowledge the financial support of the CAREFIN foundation.
Peter Gruber and Fabio Trojani gratefully acknowledge the financial support of the Swiss National Science
Foundation (NCCR FINRISK and grants 101312-103781/1 and 100012-105745/1). We thank the participants
of the 2008 Finrisk Research Day in Gerzensee, the X. Workshop on Quantitative Finance in Milano, the
FAM research seminar at Vienna University of Technology, the finance seminar of Bocconi University, the
6th World Congress of the Bachelier Finance Society in Toronto, the 2010 European Summer Symposium in
Financial Markets in Gerzensee and the 2010 EFA meeting in Frankfurt for valuable comments. The usual
disclaimer applies.
I. Introduction

We study a new class of three-factor affine option pricing models, featuring interdependent volatility risks and a stochastic skewness component conditionally unrelated to the volatility. These properties allow us to improve on the modeling of the implied volatility smile of standard affine models along two main dimensions. First, they enhance the specification of short and long term skew dynamics largely unrelated to volatility shocks. Second, they produce a broader range of term structures of implied volatility skews, which are potentially more consistent with the data.

We specify our models using three distinct stochastic components for the joint dynamics of return volatility and skewness: Two components capture short and long run volatility risks, while the third component captures stochastic skewness effects not related to volatility shocks. In contrast to standard affine models, we introduce interdependent risks, specified by a multivariate dynamics, in which the persistence and local variance of the volatility components depends on the degree of return skewness, and vice versa. Methodologically, we borrow from Leippold and Trojani (2008) and specify our model using a matrix affine jump diffusion (AJD). In this way, we achieve two objectives at the same time. First, we preserve a good degree of model tractability, with efficient pricing formulae for plain vanilla options, computed by means of standard transform methods. Second, we can nest as special cases a number of two- and three-factor affine models in the literature, such as Bates (2000) jump diffusion model or two-factor Heston (1993)-type models. These benchmark models have (i) independent volatility components and (ii) a skewness dynamics that is a function only on the volatility dynamics, which excludes volatility-unrelated stochastic skewness effects. Therefore, they are natural models against which we can benchmark the incremental pricing accuracy of our option valuation framework.

We estimate our three-factor models together with the benchmark models, using S&P 500 index options data from January 1996 to September 2009, and obtain a number of novel findings for the option pricing literature. First, interdependent volatility and volatility-unrelated skewness dynamics are well consistent with S&P 500 index option smiles: A test of the null hypothesis that these features are not present in S&P 500 index option data is rejected with a high degree of statistical significance. Second, our models improve on benchmark affine two-factor option valuation models, by reducing pricing errors on average by 20% out-of-sample. The reliability in pricing performance is also improved, with on average lower standard deviations of pricing errors by 27% out-of-sample. Therefore, these fit improvements are unlikely a consequence of overfitting effects related to the higher-dimensional (three-factor) state space underlying our models. Third, standard affine three-factor models with independent volatility components tend to have poor pricing performance out-of-sample,
indicating a likely misspecification of their state dynamics. For instance, we find that the improvement in out-of-sample pricing performance of our three-factor models relative to a three-factor Heston-type model is on average about 21%. Fourth, while the improvements in pricing performance relative to benchmark models are quite consistent over time, the largest improvements tend to arise during periods of financial crises or market distress, like, e.g., during the Russia debt crisis in 1998 and the more recent Subprime crisis: While the average improvement in out-of-sample model fit is about 20%, daily pricing improvements in such crisis periods are often larger that 30% and can be, in some cases, even above 50%. This finding suggests volatility-unrelated skewness as a potentially useful reduced-form risk factor, which is better able to reproduce some of the crisis-related dynamics of index option smiles. Fifth, while the improved fit of our models is consistent along both the moneyness and maturity dimensions, we find that it is largely due to the improved modeling of the whole term structure of implied-volatility skews: For jump diffusion models, we observe monotonically decreasing out-of-sample pricing improvements, from out-of-the-money put options of moneyness less than 0.8 (where average improvements are 35%) to out-of-the-money call options of moneyness between 1.3 and 1.4. (where average improvements are about 17%).

To understand the motivation of modeling (i) a stochastic skewness component unrelated to volatility shocks and (ii) interdependent volatility dynamics, we plot in Figure 1, top panels, the one-month maturity skew of the implied volatility surface of S&P 500 index options, against the at-the-money volatility term structure (grey points). In order to better isolate smile components that are largely unrelated to variations of the level of the volatility, each scatter plot is stratified with respect to different levels of the at-the-money implied volatility, ranging from 0.16 (first left Panel) to 0.28 (last right Panel).

[Insert Figure 1 about here.]

First, we see that in each panel the variability of the one month skew and term structure proxies is quite substantial. This feature suggests that a fraction of these variations is not explained by the level of the volatility alone, indicating the potential presence of an implied volatility dynamics driven by several sources of time-varying risk. Second, stochastic volatility models with stochastic skewness dynamics that are function of the volatility dynamics tend to fail in generating (i) the large variability of the skew of index options and (ii) the joint relation between slope of the smile and term structure across different volatility states. As a first simple illustration of this point, we estimate a two-factor Heston-type model and plot in Figure 1 top panels, the model-implied values for the one-month implied volatility

---

1See Section III for details on the definition and the computation of the implied volatility skewness and term structure proxies used in Figure 1.
skew and term structure (black points). We find that while the model tends to generate a degree of term structure variation similar to the one in the data, it tends to imply (i) a rather limited degree of variation along the skewness dimension and (ii) a comovement of skewness and term structure that is not well supported by the data. This empirical evidence suggests the presence of interesting skew dynamics that are only weakly linked to either the level or the term structure of the implied volatility smile. Therefore, models in which return skewness is completely spanned by shocks to the volatility dynamics might by overly restrictive for adequately specifying these dynamic aspects of the smile. Third, the tight link between skew and term structure of two-factor option valuation models with independent volatility components might also imply an overly simplified term structure of volatility skews. This feature is illustrated in Figure 2 top panels, where we plot the twelve month skew of the smile against the one month skew in the data (gray points). Each scatter plot is again stratified with respect to different levels of the at-the-money implied volatility.

In the data, the degree of variability of the skew at one and twelve months maturities is similar, especially for the low volatility state (left Panel in Figure 2). The model-implied twelve month and one month skews of a two-factor Heston-type model (black points) feature, as expected, a lower variability than the data. At the same time, it appears that the model has an even larger difficulty in generating (i) a sufficient variability of twelve month skews and (ii) a degree of comovement between short and long term skews similar to the one data. Overall, this simple preliminary empirical evidence motivates our interest in option valuation models featuring (i) an interdependent volatility dynamics and (ii) a skewness component not related to volatility shocks.

Our work borrows from a large literature documenting the time variation of the equity volatility and its negative co-movement with returns. In the option pricing literature, stochastic volatility processes linked to a negative skewness of returns are key ingredients of most valuation models aiming at generating well-known pricing biases of Back-Scholes model. The current state of the art in this literature specifies the underlying return dynamics as driven by several components that follow independent volatility processes, each negatively related to return shocks. A number of recent studies shows that these models tend to perform better than single factor stochastic volatility models in pricing equity index options.

---

2 The literature in this domain is too large to be reviewed exhaustively here. Early papers introducing single-factor volatility models with a correlation between returns and volatility include Heston (1993), Hull and White (1988) and Melino and Turnbull (1990). Leverage effects and jump driven skewness for modeling volatility have been studied in Bakshi, Cao and Chen (1997), Bates (1996), Backus, Foresi, Li and Wu (1997), Nandi (1998), Chernov and Ghysels (2000), Pan (2002), Jones (2003), Eraker, Johannes and Polson (2003),
In contrast to single-factor models, multiple component models can generate a degree of stochastic skewness that can help to capture part of the time variation of the smile along the moneyness dimension. Moreover, when the distinct volatility components feature different persistence properties, these models also tend to better capture the behavior of the implied volatility surface along the maturity dimension. Bates (2000) specifies two jump-diffusion components driven by independent volatility processes and studies empirically the relative performance of pure diffusion models and models augmented by Poisson-normal jumps. Using S&P 500 futures option data from 1988 to 1993, he documents the negative skew of the smile after the 1987 crash and concludes that models with jumps better reconcile return and option data. Christoffersen et al. (2009) focus on the ability of a pure diffusion version of the Bates (2000) model to explain the option implied volatility dynamics. They document that these models improve the pricing performance relative to single factor volatility settings, both in-sample and out-of-sample, because they imply a higher degree of flexibility in modeling conditional skewness and kurtosis of returns in dependence of the overall level of the volatility. Using time-changed Levy processes, Huang and Wu (2004) study two-component jump diffusion models with different types of jump specifications. They document that models with high frequency jumps and volatility variations deriving from both the instantaneous variance of the diffusion component and the arrival rate of the jump component better capture the behavior of S&P 500 index options. They also find that the diffusion induced volatility exhibits a larger instantaneous variation, but the jump induced volatility features a much higher persistence. Finally, Carr and Wu (2008) propose a three-component model based on three different sources of variation in volatility: Time varying financial leverage, time-varying business risk and self-exciting market behavior. The first component follows a CEV-type dynamics in order to model a dependence of the volatility on the level of financial leverage. The second component specifies volatility feedback effects modeled by a Heston (1993)-type volatility model. The third component models self-exciting market behavior using a high frequency pure-jump Levy-process. The model is estimated using about a decade of over-the-counter equity index options data and is shown to perform well in pricing equity index options.

All these specifications model the equity index return as a sum of independent components driven by separate and independent volatility processes. While all other models are based on a two factor volatility dynamics, the specification in Carr and Wu (2008) models

three distinct sources of independent volatility variation. Their empirical results support three factor models as convenient settings to describe the overall shapes of the implied volatility surface of equity index options. We borrow from this insight and specify a three factor state dynamics for returns volatility and skewness, but we use a completely different modeling approach with distinct implications, starting from the family of matrix AJD introduced in Leippold and Trojani (2008). First, we specify a two-component model for the volatility and introduce a third component linked to stochastic skewness variations that are not spanned by shocks to the two volatility factors. In this sense, our model comprises a component for volatility-unrelated (unspanned) stochastic skewness. This model feature allows us in the first place to obtain a wider range of model-implied degrees of risk neutral skewness, thus improving along the moneyness dimension in the description of the implied volatility smile. Second, we specify dynamics for the volatility and skewness components that admit feedback effects in the persistence and local variance of volatility and skewness shocks: The risk neutral skewness and term structure of volatility in our model interact dynamically and imply a more flexible specification for the dynamics of implied volatility skew term structures. Third, we introduce a new representation of the state space of matrix AJD, which is more convenient to interpret the structural pricing implications of our framework. Using this methodology, we identify broad directions of improvements, relative to benchmark affine models, for a more accurate modeling of the implied volatility skew term structure.

The article proceeds as follows. Section 2 introduces our modeling approach and discusses key properties of our model specifications. Starting from a matrix AJD, it derives a class of three-component option valuation models, with interacting two-factor volatility dynamics and a volatility-unrelated stochastic skewness component. It also shows that a variety of multi-factor affine option pricing models in the literature are special cases of our setting. Section 3 introduces our model estimation procedure, presents estimation results, as well as the in-sample and out-of-sample model fit analysis. Using the estimated parameters, Section 4 analyzes in more detail the main structure and theoretical features of our setting, relative to a number of benchmark affine models. Section 5 concludes. All proofs are in the Appendix.

II. Model

In this section, we introduce the modeling methodology that allows us to specify a class of option valuation models with (i) volatility components that interact dynamically and (ii) skewness components that are unspanned by volatility components. To provide a simple intuition for our approach and to link it to benchmark models in the literature, we start from Bates (2000) two factor volatility model and extend it by considering state dynamics
within the class of matrix AJD proposed in Leippold and Trojani (2008).

A. A Two-Component Benchmark Volatility Model with Poisson-Normal Jumps

Bates (2000) proposes a two-component jump diffusion model with tractable pricing formulas for European options, in which returns are driven by two independent volatility factors and Poisson-Normal jumps that have a stochastic intensity. The model features two different important channels for generating a stochastic return skewness: The standard feed-back effect between returns and volatility and a time-varying probability of return jumps.

We denote by \( S_t \) the value of an equity index at time \( t \), by \( r \) and \( q \) the (constant) interest rate and dividend yield, and by \( v_{1t}, v_{2t} \) the two volatility components. Under the risk neutral probability measure, the return dynamics is:

\[
\frac{dS_t}{S_t} = (r - q - \lambda_t)dt + \sqrt{v_{1t}}dz_{1t} + \sqrt{v_{2t}}dz_{2t} + kdN_t
\]

where \( z_1, z_2 \) are independent standard Brownian motions and the volatility components have dynamics:

\[
dv_{it} = (\alpha_i - \beta_i v_{it})dt + \sigma_i \sqrt{v_{it}}dw_{it}; \quad i = 1, 2
\]

where \( w_1 \) and \( w_2 \) are independent standard Brownian motions, having correlation \( \rho_1 \) and \( \rho_2 \) with \( z_1 \) and \( z_2 \), respectively. Poisson-Normal jumps \( kdN_t \) feature a stochastic jump probability

\[
\lambda_t := P_t(dN_t = 1)/dt = \lambda_0 + \lambda_1 v_{1t} + \lambda_2 v_{2t}
\]

and a jump size \( k \) distributed as \( \ln(1 + k) \sim N \left( \ln (1 + \bar{k}) - \frac{\delta^2}{2}, \delta^2 \right) \). The well-known volatility feedback effect is captured by the (stochastic) correlation between returns and diffusive volatility \( v_{1t} + v_{2t} \):

\[
\text{Corr}_t \left( \frac{dS_t}{S_t}, d(v_{1t} + v_{2t}) \right) = \frac{\rho_1 v_{1t} + \rho_2 v_{2t}}{\sqrt{(v_{1t} + v_{2t} + \lambda_t E(k^2))(\sigma_1^2 v_{1t} + \sigma_2^2 v_{2t})}}.
\]

In addition, the time varying jump probability \( \lambda_t \) generates a direct channel for (stochastic) jump-driven return skewness. Finally, distinct mean reversion speeds or volatilities of volatility associated with volatility components \( v_{1t} \) and \( v_{2t} \) generate a stochastic term structure of volatility, as the composition of the volatility varies over time.

For our analysis, two features of Bates (2000) model are key. First, the time varying jump intensity \( \lambda_t \) and the volatility-feedback \( \rho_1 v_{1t} + \rho_2 v_{2t} \) are functions only of the volatility components \( v_{1t} \) and \( v_{2t} \). Therefore, shocks to risk neutral skewness are perfectly spanned by shocks
to volatility. As a consequence, it is unlikely that the model can generate a large degree of skewness variations not linked to volatility shocks. Second, volatility components $v_{1t}, v_{2t}$ are independent processes with no interaction. Intuitively, this feature implies that the channels by which the return skewness and volatility term structure can interact are separated between short and long term volatility effects. Figure 1 provides simple preliminary evidence that these model properties might overly restrict the model-implied degree of joint smile variability, along the maturity and moneyness dimensions, relative to the one present in the data. This motivates us to study a broader class of option valuation models, in which return skewness can feature unspanned stochastic components interacting dynamically with all volatility factors.

B. Unspanned Skewness and Dynamic Interactions in a Model with Poisson-Normal Jumps

We model an unspanned skewness component in option valuation by a third state variable $v_{12t}$ and specify its interactions with volatility components $v_{1t}, v_{2t}$ using the dynamics of a $2 \times 2$ symmetric and positive definite matrix diffusion $X_t$, where:

$$X_t := \begin{pmatrix} v_{1t} & v_{12t} \\ v_{12t} & v_{2t} \end{pmatrix}.$$  

(5)

Overall, we study a family of stochastic volatility models driven by, at most, three different sources of stochastic risk ($v_{1t}, v_{2t}, v_{12t}$).

B.1. State Dynamics

Positivity of $X_t$ ensures that $v_{1t}, v_{2t}$ are well defined volatility processes. Therefore, it is a natural choice to consider symmetric matrix processes taking values in the positive definite cone.

**Assumption 1** Symmetric positive definite matrix process $X_t$ follows the affine dynamics

$$dX_t = [\Omega \Omega' + MX_t + X_t M']dt + \sqrt{X_t} dB_t Q + Q' dB'_t \sqrt{X_t},$$  

(6)

where $\Omega, M, Q$ are $2 \times 2$ parameter matrices and $B_t$ is a $2 \times 2$ standard Brownian motion. $\sqrt{X_t}$ denotes the unique symmetric square root of $X_t$. 

7
$X_t$ is the affine Wishart diffusion process first introduced by Bru (1991). In general, when matrices $M$ or $Q$ are not diagonal, $(v_{1t}, v_{2t}, v_{12t})$ feature dynamic interactions, because their drift and volatility functions depend on all state variables in equation (6). In the special case where matrices $\Omega, Q$ and $M$ are diagonal, $(v_{1t}, v_{2t})$ is an autonomous Markov process with components distributed as independent Heston (1993)-type volatility models. Therefore, under these constraints the state dynamics in Bates (2000) option valuation model arises as a particular case of our setting.

### B.2. Return Dynamics and Nested Models

Given the matrix state dynamics (6), we specify returns by the following matrix AJD process.

**Assumption 2** Under the risk neutral probability measure, the dynamics of $S_t$ is given by:

$$\frac{dS_t}{S_t} = (r - q - \lambda_t \bar{k})dt + tr(\sqrt{X_t}dZ_t) + kdN_t$$

where $X_t$ follows the dynamics (6),

$$Z_t = B_tR + W_t\sqrt{I_2 - RR'},$$

with $tr(\cdot)$ denoting the trace operator, $W$ another $2 \times 2$ standard Brownian motion, independent of $B$, and $R$ a $2 \times 2$ matrix such that $I_2 - RR'$ is positive semi-definite. Return jumps follow a Poisson-Normal process $kdN_t$ with jump intensity $\lambda_t = \lambda_0 + tr(\Lambda X_t)$, for $\lambda_0 \geq 0$ and a positive definite $2 \times 2$ matrix $\Lambda$, and an iid jump size $k$ distributed as $\ln(1 + k) \sim N(\ln(1 + \bar{k}) - \frac{\delta^2}{2}, \delta^2)$.

The return dynamics (7) features a stochastic return skewness not spanned by (diffusive) volatility shocks. To see this, first note that:

$$Var_t\left(\frac{dS_t}{S_t}\right) = tr(X_t) + \lambda_tE(k^2) = v_{1t} + v_{2t} + \lambda_tE(k^2).$$

Thus, the diffusive volatility $tr(X_t) = v_{1t} + v_{2t}$ does not depend on out-of-diagonal component $v_{12t}$. However, $v_{12t}$ can impact on jump-driven skewness, because the stochastic intensity $\lambda_t$ is a function of $v_{12t}$ when $\Lambda$ is not diagonal. $v_{12t}$ is also related to the volatility feed-back

\footnote{Positive semi-definiteness (definiteness) of $X_t$ follows if $\Omega\Omega' \succ Q^\prime Q (\Omega\Omega' \succ 3Q^\prime Q)$, ensuring that the volatility components cannot cross (reach) the zero boundary.}
effect, because it influences the covariance of returns and (diffusive) volatility:

\[ \text{Cov}_t \left( \frac{dS_t}{S_t}, d(v_{1t} + v_{2t}) \right) = 2tr(R'QX_t) \]  

(10)

Therefore, when matrices \( RQ' \) and \( \Lambda \) are not diagonal, \( v_{12t} \) drives both the jump-driven and diffusive stochastic skewness, acting as a skewness factor not spanned by volatility shocks. When matrices \( R \) and \( Q \) are diagonal, equation (10) collapses to Bates (2000) specification of volatility feed-backs. Similarly, when matrix \( \Lambda \) is diagonal, \( \lambda_t \) coincides with Bates (2000) specification of a stochastic intensity. Thus, if \( M, Q, R \) and \( \Lambda \) are all diagonal, Assumption 2 yields Bates (2000) option valuation model. If in addition \( \lambda_0 = 0 \) and \( \Lambda = 0 \), a two factor Heston (1993)-type volatility model is obtained, which has been recently studied empirically in Christoffersen et al. (2009). By construction, these diagonal models feature independent volatility components and a return skewness spanned by volatility shocks. Overall, it follows that Assumption 2 nests a number of affine option valuation models in the literature, thus providing a consistent framework for (i) comparing the performance of nested models in capturing the behavior of the index option implied volatility surface and (ii) studying the incremental pricing accuracy of models featuring dynamic volatility interactions or unspanned return skewness. Table 1 provides an overview of models nested by Assumption 2 which we analyze more extensively in the empirical part of this paper.

We denote by \( SV_{r,q} \) pure diffusion and by \( SVJ_{r,q} \) jump diffusion models, according to their number \( r \) and \( q \) of state variables and unspanned skewness components, respectively.

### B.3. Option Valuation

The model in Assumption 2 belongs to the class of matrix AJD introduced in Leippold and Trojani (2008). Therefore, it yields closed-form transform expressions for returns, which are useful in order to efficiently compute the prices of plain vanilla options by transform methods, as proposed by Carr and Madan (1999) and Duffie et al. (2000), among others. Closed form expressions for the risk neutral Laplace transform of returns are available when \( \Omega \Omega' = \beta Q'Q \) for some \( \beta > 1 \). In this case, Assumption 2 implies an exponentially affine conditional Laplace transform for \( Y_T := \log(S_T) \), given by (see Leippold and Trojani (2008)):

\[
\Psi(\tau; \gamma) := E_t \left[ \exp (\gamma Y_T) \right] = \exp \left( \gamma Y_t + tr \left[ A(\tau)X_t \right] + B(\tau) \right),
\]

(11)

\footnote{Precisely, in order to nest the diagonal Bates (2000) and two-factor Heston-type models, we allow \( \beta \) to be a diagonal matrix \( K \) when both \( Q \) and \( M \) are diagonal.}
where \( \tau = T - t \), \( A(\tau) = C_{22}(\tau)^{-1}C_{21}(\tau) \) and the \( 2 \times 2 \) matrices \( C_{ij}(\tau) \) are the \( ij \)-th blocks of the matrix exponential:

\[
\begin{pmatrix}
C_{11}(\tau) & C_{12}(\tau) \\
C_{21}(\tau) & C_{22}(\tau)
\end{pmatrix} = \exp \left[ \tau \begin{pmatrix}
M + \gamma Q'R & -2Q'Q \\
0 & -\gamma(M' + \gamma R'Q)
\end{pmatrix} \right].
\]

The explicit expressions for the \( 2 \times 2 \) matrix \( C_0(\gamma) \) is:

\[
C_0(\gamma) = \frac{\gamma(\gamma - 1)}{2} I_2 + \Lambda \left[ (1 + \bar{k})^\gamma \exp \left( \gamma(\gamma - 1)\frac{\delta^2}{2} \right) - 1 - \gamma\bar{k} \right]
\]

and real-valued function \( B(\tau) \) is given by:

\[
B(\tau) = \begin{cases}
\frac{\beta}{2} \text{tr}\left[ \log C_{22}(\tau) + \tau(M' + \gamma R'Q) \right] \\
\end{cases} \tau - r - q + \lambda_0 \left[ (1 + \bar{k})^\gamma \exp \left( \gamma(\gamma - 1)\frac{\delta^2}{2} \right) - 1 - \gamma\bar{k} \right]
\]

where \( \log(\cdot) \) is the matrix logarithm. In contrast to diagonal Bates (2000)-type models, computation of the return transform in the full model cannot be reduced to calculations that involve only scalar exponential and logarithmic functions, because coefficients \( C_{ij}(\tau) \) and \( B(\tau) \) depend on a matrix exponential and a matrix logarithm, respectively. This feature makes the computation of Laplace transform typically at least two orders of magnitude more costly than in diagonal models. We obtain an efficient computation of the pricing transform for the full model using the Cosine-FastFourierTransform (CosFFT) method proposed by Fang and Oosterlee (2008).

Let

\[
\Psi(\tau, \gamma) = \exp \left( \text{tr} [A(\tau)X_t] + B(\tau) \right).
\]

The price \( C_t(x, T) \) \((P_t(x, T))\) of a plain vanilla call (put) option with relative moneyness \( x = \ln(S_t/K) \) and maturity \( T \) is computed by means of a trigonometric series with pay-off dependent coefficients \( (V_k)_{k \in \mathbb{N}} \) detailed in the Appendix. For instance, for calls the CosFFT expression reads:

\[
C(x, T) = e^{-rt} \sum_{k=0}^{N-1} \left. \frac{1}{\sqrt{N}} \phi \left( \frac{k\pi}{b - a} \right) \exp \left( -\frac{i k\pi a}{b - a} \right) \right\} V_k
\]

5 The superior convergence properties of this method allows us to compute accurate transform inversions using on average only \( N = 250 \) Laplace transform evaluations, which is a considerable improvement in comparison to the in average \( N = 2^{12} \) evaluations necessary for standard FFT methods. To control potential discontinuities of the complex matrix logarithm in \( B(\tau) \), we also use a modification of the rotation count algorithm in Kahl and Jäckel (2006) for our matrix setting.
where $a$ and $b$ are suitable integration boundaries, $i = \sqrt{-1}$, $Re(\cdot)$ denotes the real part of a complex number and symbol $\sum^{\prime}$ denotes a sum in which the first element counts half.

### III. Empirical Analysis

We estimate the models listed in Table 1, using about fourteen years of S&P 500 index option data, and study the added value of models with unspanned skewness components or dynamic volatility interactions in explaining the dynamic and cross sectional behaviour of S&P 500 index option implied volatility surfaces.

#### A. Data source and characteristics

We collect from OptionMetrics daily data of end-of-day prices of S&P 500 index options, traded at the Chicago Board Options Exchange, for the sample period January 1996 to September 2009 and maturities up to one year. We then apply a number of standard filtering procedures outlined in Bakshi et al. (1997). First, we eliminate options with midquote premia below 0.375 dollars and options with zero bid price or with bid price larger than the ask price. Second, we eliminate options with stale quotes (i.e., prices identical to the prices of the previous trading day), options with prices that violate arbitrage bounds, options with duplicate entries and options where the bid-ask spread is smaller than the minimum tick size (i.e., five cents for options having prices below 3 dollars and ten cents for all other options). Third, we drop options with a time to maturity less than 10 days, in order to avoid pricing effects largely driven by short term liquidity features. Note that we do not apply additional filters that cut options with extreme moneyness, in order to obtain a data set as rich and challenging as possible, with respect to the empirical features of the term structure of implied volatility skews. On average, we obtain about 185 option prices per trading day, having an average time to maturity of 133.5 days and an average moneyness $S/K = 1.07$. The interest rate $r$ is computed by linearly interpolating the US treasuries yield curve supplied by OptionMetrics. The dividend yield $q$ is computed by minimizing each day the put-call parity error of nearly at the money options ($0.9 \leq K/S \leq 1.1$):

$$q = \arg \min_q (C - P - Se^{-\tau q} + Ke^{-\tau r})^2,$$

where $K$ is the option strike price, $C$ and $P$ the prices of call and put options, $S$ the underlying spot price and $\tau$ the time to maturity of the option.

We estimate all models in Table 1 by Non Linear Least Squares (NLLS) and obtain estimates of each model’s risk neutral dynamics and latent state variables. Parameter esti-

\[\text{We obtain end-of-day midquotes as simple averages of end-of-day bid and ask call prices and force the put-call parity to hold when calculating the implied dividend yields.}\]
mation is based on monthly observations of S&P 500 index option prices in the sample period from January 2000 to December 2004. Each month, we select the Wednesday of the week before the expiry date. Thus, the shortest time to maturity in our estimation sample is fixed at ten days. Focusing on the monthly sample for estimation purposes has several reasons. First, it reduces the computational costs implied by NLLS estimation of models in Table 1, which can have up to three latent states and sixteen parameters. Second, it leaves a large fraction of our data available for out-of-sample model evaluation. This aspect is important in our context, in order to obtain a fair comparison of models in Table 1, which can feature different dimensions of both state dynamics and parameter space. Out-of-sample pricing performance evaluation also controls for possible overfitting effects that naturally tend to favor higher dimensional models, based on pure in-sample pricing performance. Out-of-sample evaluation is performed by (i) fixing parameter estimates at the parameter values estimated using monthly data from January 2000 to December 2004 and (ii) estimating by NLLS only the missing latent state for each daily observation from January 1996 to September 2009. Note that for our out-of-sample analysis we keep model parameter values fixed at one single set of parameters, implying that out-of-sample pricing performance is completely driven by the ability of the specified state space to account for all variations in out-of-sample option prices, both cross-sectionally and over time. This approach is different from a number of out-of-sample option pricing exercises in the literature, in which models are often repeatedly estimated using different subsamples. The drawback of the latter approach for evaluating the pricing performance of models in Table 1 is that a potential deterioration of out-of-sample pricing performance, due to an inadequate state space specification, can be partly masked by the time-variation in the repeatedly estimated model parameters. However, such a time-variation is already indication of a potential misspecification of the specified state dynamics.

Summary statistics of our data sets are reported in Table 2. Panel A presents basic aggregate statistics, while Panel B describes the structure of our data set along the maturity and moneyness dimensions.

[Insert Table 2 about here.]

Overall, our sample consists of 638,365 contracts with average time to maturity of about 133.5 days and moneyness ranges between $S/K = 0.7$ and $S/K = 1.5$. Most frequently traded option maturities are between 20 and 80 days, followed by options of maturity higher than 180 days and options of maturity between 80 and 180 days, respectively. Options with

7There is no option price available for September 2001, as US exchanges were closed from September 11 to September 16, 2001.
maturity less than 20 days are those less frequently traded. Near at the money options of moneyness $1 < S/K < 1.1$ ($0.9 < S/K < 1.1$) are most frequently traded for maturities less than 80 days (higher than 80 days).

B. Estimation Method

The main challenge when estimating the stochastic volatility models in Table 1 is to estimate the model structural parameters together with the time series of latent states $\{X_t\}_{t=1,...,T}$. Several approaches are available in the literature and have been applied to a variety of option pricing models with independent volatility components. A popular approach in single-factor models is to treat the spot volatility as an additional parameter that has to be re-estimated with a recursive procedure; see Bakshi et al. (1997), among others. Other approaches filter the volatility states using time series information on underlying returns, thus ensuring consistency of physical and risk-neutral probabilities. This is achieved for a number of single-factor stochastic volatility models in Jones (2003) and Eraker (2004), using Monte Carlo Markov Chain methods, in Chernov and Ghysels (2000), who apply a version of the Efficient Method of Moments, in Pan (2002), who introduces filtered state Generalized Method of Moments estimation, and in Christoffersen, Jacobs and Mimouni (2010) and Johannes, Polson and Stroud (2010), who make use of particle filtering techniques, in order to better account for model non-linearities. Multi-factor volatility models with independent components have been studied in Carr and Wu (2007), who apply the standard Kalman filter to estimate a two-factor model with high-frequency jumps, and in Carr and Wu (2008), who estimate with the unscented Kalman filter a three-factor volatility model with self-exciting volatility.

In order to estimate our three-factor volatility models with interdependent volatility components in Table 1, we use a modification of the NLLS approach taken in Bates (2000), Huang and Wu (2004) and Christoffersen et al. (2009), among others. We use NLLS estimation to infer model parameter values and state realizations, which allow us to test whether the inferred risk neutral distributions of models in Table 1, potentially including dynamic volatility interactions or unspanned skewness components, are consistent with the observed cross-sectional and time-series behavior of S&P 500 index option prices. We maximize a Gaussian pseudo likelihood function for observed pricing errors, while allowing for conditional error heteroskedasticity driven by group specific and idiosyncratic shocks.

Let $\theta = \{M, R, Q, \beta, \lambda_0, \Lambda, \bar{k}, \delta\}$ be the parameter of interest and $e_{i,t} = (\hat{C}_{it}(\theta, X^*_{it}(\theta)) - C_{it})/F_t$ be the observed, model-implied, pricing error, where $C_{it}$ and $\hat{C}_{it}$ are observed and model-implied option prices of option $i$ at time $t$, respectively, and $F_t$ is the S&P 500 index.
futures price at time \( t \). The conditional implied state \( X_t^*(\theta) \) is defined by:

\[
X_t^*(\theta) = \arg\min_{\{X_t\}} \sum_{i=1}^{N_t} \left( \left( \hat{C}_i(\theta, X_t) - C_{it} \right) / F_t \right)^2 .
\]  

(17)

For any given \( t = 1, \ldots, T \), vector \( e_t = (e_{1,t}, \ldots, e_{N_t,t})' \) denotes the vector of pricing errors at time \( t \) and \( N_t \times N_t \) matrix \( \Omega_t \) is the covariance matrix of these errors. Our point estimate for parameter \( \theta \) is given by the following pseudo Maximum Likelihood estimator[8]

\[
\hat{\theta} = \arg\max_{\theta} \mathcal{L}_T(\theta) := \arg\max_{\theta} -\frac{1}{2} \sum_{t=1}^{T} \left( \ln |\Omega_t| + e_t' \Omega_t^{-1} e_t \right) .
\]  

(18)

We solve the problem of jointly estimating the model implied states \( \{X_t^*(\theta)\}_{t=1,...,T} \) and parameter \( \theta \) using a full nested optimization: In the first step of the optimization, we compute for any candidate parameter vector \( \theta \) optimal state \( \{X_t^*(\theta)\}_{t=1,...,T} \); In the second step, we maximize the pseudo likelihood criterion \( \mathcal{L}_T(\theta) \) over \( \theta \). We find that, especially for the three-factor models with dynamic factor interactions, this is an important issue in order to control for the presence of local minima in pseudo likelihood function \( \mathcal{L}(\cdot) \). We have also investigated a less computationally demanding two-step optimization approach, used in Huang and Wu (2004) and Christoffersen et al. (2009), among others, which iterates between parameter estimation for a given state and state estimation for a fixed parameter. However, this method does not produce good convergence properties and stable results in the estimation of our option pricing models with state variables driven by matrix AJD.

C. Parameter Estimates and In-Sample Results

In this section, we first present parameter estimates and in-sample pricing results for two- and three-factor models in the context of Assumption 2. In a second step, we study the distinct role of model parameters and latent states in generating implied-volatility skew and term structures effects largely unrelated to the level of the volatility.

C.1. Estimated Risk Neutral Dynamics

Estimated risk neutral parameters for two- and three-factor models in Table 1 are presented in Table 4[9]. NLLS estimation procedure (18) is applied to the in-sample data, consisting of

---

8 Under the given conditions, pseudo Maximum Likelihood and NLLS estimators coincide.

9 For brevity, we omit results for single-factor models. They are available on request. For a simple comparison of results across the nested models, we present parameter estimates using the notation of Assumption
monthly observations of S&P 500 index option prices for all Wednesdays of the weeks before
the expiry date of options with shortest maturity, in the time span from January 2000 to
December 2004.

[Insert Table 4 about here.]

Panel A of Table 4 presents estimates for the diffusive parameters of diagonal models
$SV_{2,0}$ and $SVJ_{2,0}$, together with point estimates for the diagonal elements of matrices $M$, $Q$
and $R$ of models $SV_{3,1}$ and $SVJ_{3,1}$. Estimates for out-of-diagonal elements of these matrices
are presented in Panel B, while panel C summarizes estimation results for the parameters in
the jump components of models $SVJ_{2,0}$ and $SVJ_{3,1}$.

Both diagonal Models $SV_{2,0}$ and $SVJ_{2,0}$ feature two volatility factors with very different
mean reversion and volatilities of volatility, and slightly different leverage effect parameters:
In both models, the less persistent volatility component $v_{2t}$ ($M_{22} < M_{11}$) also features a
higher volatility of volatility ($Q_{22} > Q_{11}$) and a stronger leverage effect, since $R_{11}Q_{11} <
R_{22}Q_{22}$, similar to findings in previous studies; see Bates (2000) and Christoffersen et al.
(2009), among others. Despite the different data sets, the estimated jump component for
model $SVJ_{2,0}$ in our study is broadly consistent with the results in Bates (2000): We obtain
an average jump size $k = -0.043$, a jump size volatility $\delta = 0.0954$ and very different
sensitivities of $\lambda_t$ to the two volatility components: $\Lambda_{11} = 84.56$ and $\Lambda_{22} = 0.94$.10
The non significant point estimate for $\Lambda_{22}$ indicates that, in the context of diagonal model $SVJ_{2,0}$,
linear specifications of time varying intensities based on a single volatility component are
not rejected by the data.

The point estimates of out-of-diagonal elements of matrices $M$, $Q$ and $\Lambda$ for models $SV_{3,1}$
and $SVJ_{3,1}$ in Panels B and C, respectively, provide a number of interesting observations.
First, we can reject the null hypothesis of a diagonal mean reversion matrix $M$, since point
estimate $M_{21}$ is highly statistically significant in both the $SV_{3,1}$ and $SVJ_{3,1}$ models. This
finding is a first indication that volatility components with dynamic interactions are well
supported by the data. Second, all point estimates for the components of matrix $\Lambda$ in model
$SVJ_{3,1}$ are similar in absolute value and highly statistically significant: While volatility
components $v_{11}$, $v_{21}$, load positively on $\lambda_t$, the unspanned skewness component $v_{12}$ loads
negatively. If follows that, within non diagonal models, affine dynamic intensity specifications
driven by multiple (interacting) factors are more consistent with the empirical evidence.
Finally, we test the null hypothesis that the leverage effect in models $SV_{3,1}$ and $SVJ_{3,1}$ is

---

10 For comparison, Bates (2000) estimates are $k = -0.057$, $\delta = 0.102$, $\Lambda_{11} = 81.56$ and $\Lambda_{22} = 0.28$. 

The Appendix provides the link to Bates (2000) notation in the context of diagonal model $SVJ_{2,0}$. 
not driven by the unspanned component \( v_{12t} \):

\[
\frac{1}{2dt} \text{Cov}_t(dS_t/S_t, d(v_{1t} + v_{2t})) = (R'Q)_{11}v_{1t} + (R'Q)_{22}v_{2t} .
\]  

(19)

According to the leverage effect expression (10), this is equivalent to a test of null hypothesis \( H_0 \) against alternative hypothesis \( H_A \):

\[
H_0 : (R'Q)_{21} = 0 ; \quad H_A : (R'Q)_{21} \neq 0 ,
\]  

(20)

where \( (R'Q)_{21} = R_{12}Q_{11} + R_{22}Q_{12} \). Using a standard Wald test, we reject \( H_0 \) in favor of \( H_A \) with a \( p \)-value below 0.005. In summary, the estimation results in Table 4 support a \( SVJ_{3,1} \) model specification with non diagonal matrices \( M, R \) and \( \Lambda \), i.e., a matrix AJD option valuation model with dynamically interacting volatility factors and unspanned skewness features.

### C.2. In-Sample Pricing Results

Table 5, Panel A, presents in-sample pricing results for the different models in Table 1. Consistently with our NNLS estimation criterion (18), we rank pricing performance across models according to the root mean square dollar pricing error (\( \text{RMSE} \)). Rankings according the implied volatility root mean square error (\( \text{IVRMSE} \)) give similar findings. The \( \text{RMSE} \) for day \( t = 1, \ldots, T \), denoted by \( \epsilon_t \), is defined as:

\[
\epsilon_t = \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} (\hat{C}_{t,i} - C_{t,i})^2} ; \quad t = 1 \ldots T ,
\]  

(21)

where \( \hat{C}_{t,i} \) and \( C_{t,i} \) are the model-implied and the observed option prices, respectively, of option \( i = 1, \ldots, N_t \) on day \( t \). The sample \( \text{RMSE} \) is simply the time series average of daily root mean square errors: \( \text{RMSE} = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \). As a measure of overall model reliability, we also compute the standard deviation of daily root mean square errors: \( \sigma_{\text{RMSE}} := \sqrt{\frac{1}{T} \sum_{t=1}^{T} (\epsilon_t - \text{RMSE})^2} \).

Overall, Panel A of Table 5 shows that \( SV_{3,1} \) and \( SVJ_{3,1} \) models with dynamic interactions and unspanned skewness substantially reduce the \( \text{RMSE} \) of diagonal models: Model \( SV_{3,1} \) lowers the \( \text{RMSE} \) of model \( SV_{2,0} \) by approximately 12%, while model \( SVJ_{3,1} \) lowers the \( \text{RMSE} \) of model \( SVJ_{2,0} \) by approximately 19%; the reduction in \( \text{RMSE} \) of model \( SVJ_{3,1} \) over model \( SV_{2,0} \) is about 22%.

[Insert Table 5 about here.]
It is important to note that the RMSE reduction of models $SV_{3,1}$ and $SVJ_{3,1}$ is not simply due to their higher dimensional state space relative to the $SV_{2,0}$ and $SVJ_{2,0}$ models (with three latent components instead of two), but rather to the particular features of their state dynamics. Indeed, fitting a three-factor Heston-type $SV_{3,0}$ model with independent components to our in-sample data set yields a RMSE that is about 23% higher than the RMSE of the $SVJ_{3,1}$ model. Models with dynamic interactions and unspanned skewness components also reduce the in-sample variability of the pricing errors: The RMSE standard deviation of model $SV_{3,1}$ is 23% (18%) lower than the one in model $SV_{2,0}$ ($SV_{3,0}$), while the RMSE standard deviation of model $SVJ_{3,1}$ is 28% lower than the one in model $SVJ_{2,0}$. This finding shows that the matrix state space of $SV_{3,1}$ and $SVJ_{3,1}$ models produces more reliable results also in terms of a less volatile in-sample pricing performance.

C.3. Implied Volatility Skew and Term Structure Variability

In the data, there is a substantial joint variability of implied volatility slope and term structure, which is largely unrelated to the overall level of the volatility; see again Figures 1 and 2 for a simple illustration. Intuitively, we expect models with dynamic volatility interactions and unspanned skewness components to help in qualitatively reproducing such a variability, because the dynamics of the slope of the smile in these models is less tied to short and long term volatility shocks. Whether these models can also quantitatively reproduce the main features of these volatility-unrelated variations, is an interesting question that can be analyzed using estimated models $SV_{3,1}$ and $SVJ_{3,1}$ of Table 4.

In order to study these aspects in more detail, we proceed as follows. First, we fix a benchmark level of volatility for three different volatility ranges, corresponding to a low volatility ($\sqrt{\nu_t} = 0.10$), an average volatility ($\sqrt{\nu_t} = 0.20$) and a high volatility ($\sqrt{\nu_t} = 0.30$). Second, for each benchmark volatility level we consider a limited range of maximal volatility variations, in a range of ±5% around the benchmark. Third, given the fixed highest and lowest volatility levels in the admissible range, we compute the model implied one-month maturity skew ($S$) and at-the-money volatility term structure ($M$), as a function of all admissible matrix AJD states $X_t$ consistent with volatilities in the range $\sqrt{\nu_t}(1 \pm 0.05)$[$^1$]. In this way, we obtain a model-implied version of Figure 1 which can be used to quantify, in the low dimensional ($M, S$)-coordinate system, the degree of volatility-unrelated smile variability that can be produced by a given model. Figure 3 illustrates these features for the $SV_{3,1}$ model.

$^1$Appendix B 4 provides details on the computation of our $S$ and $M$ proxies.
For each given volatility level $\sqrt{V_t}$, the region of admissible $(M, S)$-points produced by the $SV_{3,1}$ model is a surface bounded by a curve similar to an ellipse. Low volatility levels (left panel in Figure 3) are associated, on average, with a more (less) limited range of possible model-implied $M$ ($S$) shapes. As we move toward a higher level of volatility, from the left to the right panel in Figure 3, we tend to obtain a flatter skew and a more negatively-sloped term structure. Since the $SV_{2,0}$ model is nested by the $SV_{3,1}$ model, we can also plot in each graph the admissible range of model-implied $(M, S)$-points generated by this model: These are represented by the approximately piecewise linear lines spanning the main axis of the elliptical curves plotted in Figure 3. Thus, the $SV_{2,0}$ model implies a monotonic relation between slope and term structure in $(M, S)$-coordinates: The range of admissible $(M, S)$ points consistent with a $SV_{3,1}$ model in each volatility regime is broader. For instance, while for a volatility level $\sqrt{V_t} = 0.10$, model $SV_{2,0}$ can achieve a positive term structure $M = 0.2$ only for a short term slope of about $S = -0.5$, model $SV_{3,1}$ can generate the same term structure for all short term slopes in the range between about $S = -1.0$ and $S = -0.05$. It is interesting to compare the structure of the admissible regions implied by models $SV_{2,0}$ and $SV_{3,1}$ with the actual variability observed in the data. To this end, we include in all panels of Figure 3 the $(M, S)$-combinations observed in our sample (black points in Figure 3). Overall, the degree of $(M, S)$-data variability within each volatility regime is quite substantial. For instance, for volatilities of about $\sqrt{V_t} = 0.20$, we can observe term structures of both $M = 0.2$ and $M = -0.2$ for an implied volatility slope of about $S = -0.7$. Such joint $(M, S)$—structures are difficult to explain accurately within $SV_{2,0}$—type models using a single set of estimated parameters, because of the (deterministic) approximate piecewise linear relation between $M$ and $S$ that these models imply.

C.4. Comparative Statics with Respect to Model Parameters

In the context of Assumption 2, which is the relation between model parameters and the appearance of volatility-unrelated skewness and term structure effects? To better understand this link, we can study comparative statics of model-implied proxies $M$ and $S$, with respect to model parameters $M$, $Q$, $R$, $\Lambda$ and $\beta$. For brevity, we focus on comparative statics with respect to the most statistically significant parameters $M_{21}$, $M_{22}$, $Q_{22}$, $R_{12}$, $R_{22}$ and $\beta$, estimated in Table 4 for model $SV_{3,1}$. Similar insights are obtained for model $SV_{J3,1}$. For a fixed volatility level $\sqrt{V_t} = 0.17$, we present in Figure 4 the admissible model-implied combinations of implied volatility term structure and slope ($M$ and $S$), both for the short and the long term segments of the smile. Comparative statics for other volatility levels are similar.
In each Panel of Figure 4, we vary the relevant model parameter by \( \pm 20\% \) and plot in the \((M, S)\)–coordinate system (in red and blue, respectively) the resulting surface of admissible points, bounded by an elliptical curve. We find that the comparative statics of the different parameters are broadly consistent with the main model intuition. Comparative statics with respect to mean reversion parameters \( M_{21}, M_{22} \) mainly influence the range of admissible variability along the term structure \((M)\) dimension. This is intuitive, as these parameters are directly lined to the risk neutral persistence of the volatility factors. Parameter \( Q_{22} \), which is linked both to the leverage effect and the volatility of volatility, tends to affect the admissible range of both implied volatility slopes and term structures \((M \text{ and } S)\). Finally, volatility-feedback parameters \( R_{12}, R_{22} \), virtually only influence the implied volatility slope \((S)\) dimension, while long term volatility parameter \( \beta \) essentially affects only the admissible range of implied volatility term structures \((M)\).

### D. Out of Sample Results and Estimated State Dynamics

We fix the vector of estimated structural model parameters \( \hat{\theta} \) in Table 4 and compute for every daily implied volatility surface the latent state \( X_t^*(\hat{\theta}) \) in equation (17). In this way, we assess the out-of-sample pricing performance of the models in Table 1, by relying exclusively on the ability of each model’s state space specification to reproduce the cross-sectional and time-series patterns of the implied volatility smile of index options. This out-of-sample implementation is similar to Huang and Wu (2004) and Christoffersen et al. (2009).

#### D.1. Out-of-Sample Pricing Performance

Panel B of Table 5 summarizes aggregate out-of-sample pricing results. Overall, \( SV_{3,1} \) and \( SVJ_{3,1} \) models clearly outperform diagonal models also in terms of out-of-sample pricing performance: The \( RMSE \) of model \( SV_{3,1} \) \((SVJ_{3,1})\) is about 19\% (22\%) lower than the \( RMSE \) of model \( SV_{2,0} \) \((SVJ_{2,0})\). Similarly, the \( RMSE \) of model \( SVJ_{3,1} \) is about 21\% lower than the \( RMSE \) of model \( SV_{3,0} \). These out-of-sample \( RMSE \) reductions are on average larger than the in-sample \( RMSE \) reductions in Panel A of Table 5. The \( SV_{3,1} \) and \( SVJ_{3,1} \) models also imply a higher pricing reliability, relative to the \( SV_{2,0} \) and \( SVJ_{2,0} \) benchmarks, with out-of-sample \( RMSE \) standard deviations that are 27\% and 28\% lower, respectively.

The higher performance of \( SV_{3,1} \) and \( SVJ_{3,1} \) models relative to benchmark models is quite consistent over time. Figure 5 Panel A (Panel B) plots the time series of daily \( RMSE \) for model \( SV_{2,0} \) \((SVJ_{2,0})\), together with the percentage reduction in daily \( RMSE \) of \( SV_{3,1} \) and \( SV_{3,0} \) \((SVJ_{3,1})\) models.
The middle plot of Panel A shows that model $SV_{3,1}$ almost always outperforms model $SV_{2,0}$, with a few rare exceptions at the beginning of 2001 and the end of 2008. Relative performance improvements can be large: While their average is about 20%, they are often larger than 30% and, in some cases, even above 40%. Large improvements of pricing performance can arise during some periods of financial crises or market distress, including the Russian debt crisis, the collapse of LTCM, the bursting of the dot-com bubble and the recent Subprime Crisis. In contrast to the results for the $SV_{3,1}$ model, pricing improvements implied by the $SV_{3,0}$ model in the bottom plot of Panel A are often negative, rarely above 30% and substantially more volatile, which is a potential indication of model overfitting. The bottom plot of Panel B shows that model $SVJ_{3,1}$ virtually always outperforms model $SVJ_{2,0}$: There is no distinct period in which the $SVJ_{3,1}$ model systematically behaves worse than the $SVJ_{2,0}$ model. Outperformance is often large and the reduction in daily $RMSE$ can in some cases be even above 50%.

The better pricing performance of $SV_{3,1}$ and $SVJ_{3,1}$ models relative to benchmark models is consistent also across moneyness regions and times to maturity. Panels A1 and B1 of Table 6 present out-of-sample $RMSE$ of $SV_{2,0}$ and $SVJ_{2,0}$ models, respectively, stratified across different degrees of moneyness and different maturities. Panels A2 and B2 summarize the improvements relative to the benchmark $SV_{2,0}$ and $SVJ_{2,0}$ models.

Model $SV_{3,1}$ outperforms model $SV_{2,0}$ for all options with maturity above 20 days. For options having time to maturity of 20 days or less, we find a large outperformance with respect to out-of-the-money put options and a more moderate underperformance for out-of-the-money call options, indicating that for short-maturity options the model faces a pricing trade-off between calls and puts. Model $SVJ_{3,1}$ outperforms model $SVJ_{2,0}$ for all listed options classes. In particular, the stratification of pricing error improvements across moneyness in Panel B2 shows that model $SVJ_{3,1}$ captures much better the skew patterns across maturities, with monotonically decreasing pricing improvements from option of moneyness $S/K < 0.80$ (implying average improvements of 35%) to options of moneyness $1.30 < S/K < 1.4$ (implying average improvements of 17%). In summary, these findings indicate that the better pricing performance of $SV_{3,1}$ and $SVJ_{3,1}$ models is unlikely due to overfitting, but it is rather the consequence of a state space specification that is better able to reproduce the structural dynamics of S&P 500 index option smiles.
D.2. Features of Latent State Dynamics

A key feature of models $SV_{3,1}$ and $SVJ_{3,1}$, relative to benchmarks $SV_{2,0}$ and $SVJ_{2,0}$, is the form of their (matrix) state dynamics, which allows us to model dynamic volatility interactions and unspanned skewness features. In Figure 6, we take a closer look at the latent states $X_{11t} = v_{1t}$, $X_{22t} = v_{2t}$ and $X_{12t} = v_{12t}$, estimated for models $SV_{3,1}$ and $SVJ_{3,1}$ (plotted in black and red, respectively) by the NLLS estimation procedure described in Section 6.

[Insert Figure 6 about here.]

Overall, we estimate a volatility component $v_{2t}$ that is on average larger and less persistent than component $v_{1t}$. For instance, in the $SV_{3,1}$ model, 85% of the average instantaneous variance $V_t = v_{1t} + v_{2t}$ is generated by $v_{2t}$. At the same time, $v_{2t}$ has an unconditional half life of about 16 days, which is about one tenth the half life of the unconditionally more persistent component $v_{1t}$. In the $SVJ_{3,1}$ model, component $v_{1t}$ ($v_{2t}$) is on average responsible for about 39% (24%) of the total variance $V_t = v_{1t} + v_{2t} + \lambda_t E(k^2)$. As a consequence, jump-driven volatility tends to generate on average a quite substantial fraction (37%) of total variance. While the average total variance in the $SV_{3,1}$ model is dominated by a large not very persistent factor, in $SVJ_{3,1}$ model total risk is more evenly distributed across different potential sources of uncertainty. These distinct variance components are associated with a broader variety of persistence features: While $\lambda_t$ has a short unconditional half life of about 18 days, which is comparable to the half life of factor $v_{1t}$ in $SV_{3,1}$ model, $v_{1t}$ and $v_{2t}$ have average half lives of about 24 and 78 days, respectively. Overall, these findings highlight a more pronounced multi-frequency volatility structure in model $SVJ_{3,1}$, which is potentially useful in order to generate more flexible term structures of implied volatility skews from very short to longer maturities.

Estimated state $v_{12t}$ of $SV_{3,1}$ and $SVJ_{3,1}$ models in the middle panel of Figure 6 highlight additional useful features. First, $v_{12t}$ can be both positive or negative: It tends to be positive during phases of high volatility, but it can turn slightly negative in other periods. Second, while estimated states $v_{12t}$ in both models exhibit a roughly similar tendency over time, they can even co-move negatively over short time spans. Within models nested by Assumption 2, $v_{12t}$ can play qualitatively different roles, because it potentially jointly determines (i) the local persistence of volatility factors $v_{1t}$, $v_{2t}$, (ii) the feedback effects of returns and volatility and (iii) the time varying features of jump intensity $\lambda_t$ in the $SVJ_{3,1}$ model. We find that, at the models’ estimated parameters, the largest quantitative implications of out-of-diagonal component $v_{12t}$ arise for the dynamics of risk neutral skewness. Within the $SV_{3,1}$ model, state $v_{12t}$ produces a substantial additional degree of variability of volatility feed-back effects,
which is useful to better capture time varying skewness patterns within these models.

To understand this point more concretely, the top panel of Figure 7 plots the volatility feedback coefficient \( \text{corr}_t(dS_t/S_t, d(v_{1t} + v_{2t})) \) implied by model SV\(_{3,1}\), both for the case where the latent state is fictitiously restricted to be diagonal \((v_{12t} = 0, \text{ blue line})\) and in the unconstrained case \((v_{12t} \neq 0, \text{ black line})\): The large additional variability of volatility feedback effects generated by component \(v_{12t}\) quantifies the impact of unspanned skewness factors in generating (volatility-unrelated) stochastic skewness in model SV\(_{3,1}\). The middle panel of Figure 7 presents similar plots for the volatility feedback coefficient \( \text{corr}_t(dS_t/S_t, d(v_{1t} + v_{2t})) \) within model SV\(_{J,3,1}\). Also in this case, the additional variability of volatility feedback effects produced by component \(v_{12t}\) is apparent. Compared to the \(SV_{3,1}\) setting, the volatility feedback effect in model \(SV_{J,3,1}\) has a similar time series behavior, even if it is slightly less time varying. This last feature is due to the additional role of the stochastic jump intensity \(\lambda_t\) in this model, which is the major driver of the dynamics of short term risk neutral skewness: While the dynamics of the jump intensity is mainly related to time variation in short term risk neutral skewness, time variation of longer term risk neutral skewness is more directly linked to the dynamics of volatility feedbacks. Since risk neutral long term skewness tends to be less time-varying than short term skewness, volatility feedback effects appear as slightly less volatile in \(SV_{J,3,1}\) model. The bottom panel of Figure 7 quantifies the impact of component \(v_{12t}\) on \(\lambda_t\) and clarifies this intuition further, in relation to the dynamics of volatility feedbacks in the middle panel of Figure 7: Ceteris paribus, component \(v_{12t}\) tends to increase (decrease) short and long term skewness together, by means of a higher (lower) intensity and a stronger (weaker) volatility feedback. Thus, as \(v_{12t}\) varies over time, the combined effect of time-varying intensities and stochastic volatility feedbacks produces a dynamic term structure of risk neutral skewness, which is directly reflected by the time-varying features of the term structure of option implied volatility skews.

IV. Model Analysis

In this section, we develop a more formal analysis of the \(SV_{3,1}\) and \(SV_{J,3,1}\) models, in order to better isolate the distinct effects of state variables in matrix state \(X_t\) on the shape of the implied-volatility smile. To achieve this goal, we first introduce a state reparametrization in terms of (i) a volatility level component \(V_t\), (ii) a volatility composition component \(\xi_t\) and (iii) a further component \(\alpha_t\). Using this decomposition, we achieve several purposes. First, we can clearly understand the volatility level-unrelated tradeoff between option implied volatility slope and term structure (\(S\) and \(M\)) implied by our models. Second, we can isolate
and better interpret the contribution of each component \( V_t, \xi_t, \alpha_t \) to the different pieces of the implied volatility smile. Third, we show that with the new parametrization \( SV_{3,1} \) and \( SVJ_{3,1} \) models can be reinterpreted as two-factor volatility models with stochastic coefficients, in which the stochastic coefficients of volatility factors depend exclusively on the third state variable \( \alpha_t \). Therefore, \( \alpha_t \) is a very useful state variable in order to isolate (i) new dynamic volatility interactions and (ii) additional unspanned skewness effects produced by our setting, relative to, e.g., \( SV_{2,0} \) and \( SVJ_{2,0} \) models.

A. A Useful State Reparameterization

We reparameterize state \( X_t \) in a more convenient coordinate system, starting from the standard spectral decomposition of a symmetric positive definite matrix:

\[
X_t = O_t V_t O_t' ,
\]

(22)

where \( V_t = \begin{pmatrix} V_{1t} & 0 \\ 0 & V_{2t} \end{pmatrix} \) is a diagonal matrix of ordered positive eigenvalues \( V_{1t} \geq V_{2t} \) and \( O_t = [O_{1t}, O_{2t}] \) is an orthogonal rotation matrix of eigenvectors \( O_{1t}, O_{2t} \) having unit norm, which can be uniquely expressed by means of a single parameter \( \alpha_t \in (-\pi/2, \pi/2) \) using polar coordinates:

\[
O_t = \begin{pmatrix} \cos(\alpha_t) & -\sin(\alpha_t) \\ \sin(\alpha_t) & \cos(\alpha_t) \end{pmatrix} .
\]

(23)

Based on decomposition (22), we introduce the three following state variables: \( V_t = V_{1t} + V_{2t}, \xi_t = (V_{1t} - V_{2t})/(V_{1t} + V_{2t}) \) and \( \alpha_t \). Note that, since \( V_t = tr(X_t) \), this state variable coincides with the level of the (diffusive) spot volatility. Similarly, \( \xi_t \in [0, 1] \) can be interpreted as the composition of the (diffusive) spot volatility. Therefore, state variable \( \alpha_t \) captures volatility effects not linked to volatility level or volatility composition features. The reparametrization of matrix \( X_t \) in terms of these new state variables reads as follows.

**Lemma 1 (V-\( \xi \)-\( \alpha \) decomposition)** Symmetric 2 \( \times \) 2 state matrix \( X_t \) can be decomposed into a volatility part, \( V_t \), a structural part, \( \xi_t \), and a third state variable, \( U(\alpha_t) \), as follows:

\[
X_t = \frac{V_t}{2} \left[ I_{2 \times 2} + \xi_t \cdot U(\alpha_t) \right] ,
\]

(24)

where \( I_{2 \times 2} \) is the 2 \( \times \) 2 identity matrix and

\[
U(\alpha_t) = \begin{pmatrix} \cos(2\alpha_t) & \sin(2\alpha_t) \\ \sin(2\alpha_t) & -\cos(2\alpha_t) \end{pmatrix} .
\]

(25)
Representation (24) features a number of convenient properties. First, since it is homogenous in $V_t$, it can be used to split state $X_t$ in a part directly linked to the volatility level and a part not depending on it. Second, it further decomposes the volatility unrelated part of $X_t$ into a first component that is a simple function of the volatility structure $\xi_t$ and a second component that depends only on the third state variable $\alpha_t$. Thus, $\alpha_t$ captures dynamic properties of state $X_t$, which are not linked to the level and the structure of the volatility in models $SV3,1$ and $SVJ3,1$. Fourth, since $\xi_t$ and $\alpha_t$ are bounded, it is quite easy to study the overall features of the volatility dynamics, while keeping fixed the conditional level of the volatility. Fifth, representation (24) is particularly convenient to study expressions of the form $\text{tr}(HX_t)$, for some given $2 \times 2$ matrix $H$, because of the rotation invariance properties of the trace (see appendix). Such expressions drive many key quantities in $SV3,1$ and $SVJ3,1$ models, like for instance the stochastic covariance of returns and volatility or the time-varying jump intensity.

Remark: Using state reparametrization (24), it is possible to better isolate within models $SV3,1$ and $SVJ3,1$ the incremental pricing effects of the joint presence of dynamic volatility interactions (i.e., non diagonal matrices $M, Q, R$ or $\Lambda$) and a higher dimensional state space (i.e., non diagonal matrix $X_t$), in comparison, e.g., to benchmark single-factor or two-factor state dynamics with independent components. For instance, if we restrict $\alpha_t$ to zero then:

$$X_t = \begin{pmatrix} V_{1t} & 0 \\ 0 & V_{2t} \end{pmatrix},$$

(26)

i.e., we obtain the conditional state of a two-factor model with dynamic interactions. If, furthermore, parameter matrices $M, R, Q, \Lambda$ are all diagonal, then we obtain models $SV2,0$ and $SVJ2,0$, in the sense that the so restricted $SV3,1$ and $SVJ3,1$ models yield the same option prices as the $SV2,0$ and $SVJ2,0$ models. Similarly, if we also restrict $\xi_t$ to 1, then

$$X_t = \begin{pmatrix} V_{1t} & 0 \\ 0 & 0 \end{pmatrix}.$$  

(27)

$SV1,0$ and $SVJ1,0$ models are then obtained by letting only the upper diagonal elements of matrices $M, R, Q$ and $\Lambda$ to be different from zero.

B. The Volatility-Unrelated Tradeoff between Option Implied Volatility Slope and Term Structure

Figure 1 shows a large volatility-unrelated variability of implied volatility slope and term structure ($\mathcal{S}$ and $\mathcal{M}$) in the data, while Figure 3 quantifies the incremental degree of such
variability generated by $SV_{3,1}$ and $SVJ_{3,1}$ models with respect to diagonal models. Using decomposition [24], we can completely characterize, in terms of states $(V_t, \xi_t, \alpha_t)$, the volatility-unrelated tradeoff between implied volatility slope and term structure in these models.

**B.1. Capturing the Implied Volatility Slope-Term Structure Tradeoff**

Given a fixed volatility level $V_t$, we find that each admissible $(M, S)$ combination lies on an elliptical curve, in which $\xi_t$ parameterizes the distance from the center of the ellipse and $\alpha_t$ parameterizes the location of each point on an ellipse with given distance from the center. This is illustrated in Figure 8 where we plot for two different volatility levels $\sqrt{V_t} = 0.1$ and $\sqrt{V_t} = 0.2$ the set of achievable elliptical curves in $(M, S)$ space, generated for $\alpha \in (-\pi/2, \pi/2]$ as $\xi_t$ takes different values between 0 and 1.

[Insert Figure 8 about here.]

In particular, given a volatility composition $\xi_t = 1$, say, we see that $\alpha_t$ can generate a wide degree of variations in $M$ and $S$, which is by construction unrelated to the volatility level and the volatility composition. For instance, as $\alpha_t$ moves from $-\pi/4$ to $\pi/4$ in the right panel of Figure 8 we obtain points that are moved counter-clockwise from regions of strong negative slope $S$ and strong positive term structure $M$ to regions of slightly negative slope $S$ and strong negative term structure $M$. Recalling that $\alpha_t = 0$ identifies the state space of a diagonal two-factor model, we can also directly see the more restricted set of admissible $(M, S)$ combinations, which are all on an approximately piecewise linear function in $(M, S)$ space, produced by these models as volatility composition $\xi_t$ changes.

**Remark:** For the pure diffusion model $SV_{3,1}$, the shape of all feasible points $(S, M)$ in Figure 8 can be studied analytically, using Lemma 1 and the short term asymptotics in Durrleman and Karoui (2007), to obtain the following formulae for the theoretical short term slope and term structure coefficients $(S, M)$, defined in Appendix B:

\[
2V_t^{3/2}S_t = Cov_t(dS_t/S_t, dV_t) = V_t [tr(RQ') + \xi_t \cdot tr(RQ'U(\alpha_t))] \tag{28}
\]

and

\[
2V_t^{1/2}M_t = \frac{1}{dt}E_t(dV_t) - V_t^{3/2}C_t - 3V_tS_t^2 \\
\approx \frac{1}{dt}E_t(dV_t) = tr(\Omega') + V_t [tr(M) + \xi_t \cdot tr(MU(\alpha_t))] , \tag{29}
\]

where $C_t$ is a convexity adjustment term and, for any $2 \times 2$ matrix $H$, $tr(HU(\alpha_t)) = \ldots$
\[ \cos(2\alpha_t)(H_{11} - H_{22}) + \sin(2\alpha_t)(H_{12} + H_{21}). \] These results imply that, with reasonable approximation, equations (28) and (29) parametrize an ellipse in \((M, S)\)-coordinate space. The center of the ellipse is given by

\[ (S_0, M_0) = \frac{1}{2} V_t^{-1/2} [\text{tr}(RQ'), \text{tr}(\Omega\Omega' + V_t M)] \] (30)

and is completely determined by the level \(V_t\) of the volatility: \((S_0, M_0)\) coincides with the single admissible \((S, M)\)-combination implied by the fix parameter choice of a single-factor Heston-type model \((SV_{1,0})\). Consistently with Figure 8, volatility composition \(\xi_t\) parametrizes the distance of each ellipse from its center \((S_0, M_0)\): Given \(\alpha_t = 0\), volatility composition \(\xi_t\) parametrizes the admissible deviations, in \((S, M)\)-coordinates, implied by a two-factor Heston-type model \((SV_{2,0})\) relative to a \(SV_{1,0}\) model. Finally, \(\alpha_t\) parametrizes the possible directions of deviation from \((S_0, M_0)\), i.e., the admissible deviations achievable by \(SV_{3,1}\) models relative to \(SV_{2,0}\) models.

\(SV_{3,1}\) and \(SVJ_{3,1}\) models are three-factor models with interacting volatility dynamics and unspanned skewness features, which nest two-factor \(SV_{2,0}\) and \(SVJ_{2,0}\) models, respectively. Three-factor \(SV_{3,0}\) or \(SVJ_{3,0}\) models are not formally nested by our setting based on a \(2 \times 2\) matrix state dynamics, even if they can be nested within models with a higher dimensional \((3 \times 3)\) matrix state space.\(^{13}\) How does the model-implied slope and term structure tradeoff of our three-factor models compare to the one of three-factor models with independent volatility components and skweness dynamics spanned by volatility dynamics? To investigate this question in more detail, we plot in Figure 9 the admissible slope \((S)\) and term structure \((M)\) points implied by a three-factor Heston-type model \((SV_{3,0})\) estimated using our data set, for different fixed levels of volatility \(\sqrt{V_t} = 0.1, 0.2, 0.3\).

Despite the equal dimension of their state dynamics, the volatility-unrelated slope and term structure tradeoff implied by models \(SV_{3,0}\) and \(SV_{3,1}\) in Figures 9 and 3 are structurally very different. Admissible \(S\) and \(M\) combinations in Figure 9 are bounded by a surface similar to a distorted triangle, in which for any given slope level \(S\) the admissible range of possible term structures \(M\) is somehow restricted. For instance, for an average volatility \(\sqrt{V_t} = 0.2\) and a slope \(S = -0.5\), the range of possible term structures is between about \(M = 0\) and \(M = 0.2\). In contrast, the admissible range of possible term structures for model \(SV_{3,1}\) in Figure 3 is between \(M = 0.4\) and \(-0.4\), which is more similar to the range

\(^{12}\)Detailed derivations are available on request.

\(^{13}\)Such higher dimensional models are not studied in this paper.
of term structures between about $\mathcal{M} = 0.3$ and $\mathcal{M} = -0.4$ in the data. Overall, these findings confirm the key role of volatility interactions and unspanned skewness features for an improved modeling of volatility-unrelated slope and term structure tradeoffs, also with respect to benchmark three-factor affine option valuation models.

### B.2. Volatility-Unrelated Term Structure of Implied Volatility Skews

In models $SV_{3,1}$ and $SVJ_{3,1}$, dynamic volatility interactions and unspanned skewness features produce a whole term structure of implied volatility skew effects. Using our $(V_t, \xi_t, \alpha_t)$ parametrization, it is possible to isolate more clearly these effects, in order to study in more detail the mechanics of the implied volatility skew term structure in these models. Figure 10 illustrates the study of these model-implied mechanics for a fixed volatility level $\sqrt{V_t} = 0.17$ and a fixed volatility composition $\xi_t = 1$.14

![Insert Figure 10 about here.]

In Panel A, it is shown that variations exclusively of the unspanned skewness component $\alpha_t$ can produce a broad variety of effects for the term structure on implied volatility skews in model $SV_{3,1}$. For instance, while for $\alpha_t = \pi/2$ the model generates a steep increasing term structure, combined with a steep skew at both short and long maturities, for $\alpha_t = 0$ it yields flat or decreasing term structures, combined with a pronounced (flat) skew at short (long) maturities. The structure of the implied volatility term structures is even richer within $SVJ_{3,1}$ models, because of the more pronounced short horizon effects produced by the additional jump component. The decomposition of the total implied volatility skew term structure in a component due to the diffusion part of the model (middle row of Panel B) and a residual generated by the jump term (bottom row of Panel B) produces additional insights. As expected, the diffusion part of the model dominates the implied volatility skew term structures for maturities above roughly three weeks, while the jump component has dominating effects for maturities roughly up to three weeks. Interestingly, the model-implied jump-driven segment of the smile can feature both strong short-term skews ($\alpha_t = 0.5$) or a short term smile ($\alpha_t = 0$), independently of both the level and the structure of the volatility. Overall, these results confirm that $SV_{3,1}$ and $SVJ_{3,1}$ models can produce a rich variety of patterns in the term structure of implied volatility skews, which are largely independent of the level and the composition of the volatility.

Is there also a way to isolate volatility-unrelated effects in the term structure of implied

---

14The choice $\xi_t = 1$ is for illustration purposes and produces the most extreme implied volatility smile effects.
volatility skews using a more model-free approach? To address this natural question, we first stratify our sample in quintiles of estimated latent state $\alpha_t$ for $SV_{3,1}$ and $SVJ_{3,1}$ models, respectively. In a second step, we perform within each quintile a standard Principal Component Analysis of the S&P 500 index option implied volatility surface, using a standardized grid of maturities (1,2,3,4,6,9,12 months) and moneynesses (Black-Scholes deltas of 0.2,0.3,0.4,0.5,0.6,0.7 and 0.8). We find on average two principal components in each quintile, which are plotted in Figure [11] Panel A and B for $\alpha$’s estimates in $SV_{3,1}$ and $SVJ_{3,1}$ models, respectively.

Interestingly, within each quintile the first estimated principal component reflects a very stable volatility level effect, which is remarkably unrelated to other moneyness and term structure effects. The second component in each quintile is typically related to both the moneyness and maturity dimension, indicating that it acts as a factor moving the entire term structure of implied volatility skews. Moreover, this second component has quite different properties across quintiles, indicating that the term structure of implied volatility skews reacts differently, i.e., dynamically, in dependence of the estimated proxy $\alpha_t$ for unspanned skewness. In summary, these results support the existence in the data of dynamic patterns in the term structure of implied volatility skews, which are largely independent of the level of the volatility.

**B.3. Interacting Volatility Dynamics and Unspanned Skewness in Action**

The top panels of Figures [1] and [2] highlight the difficulty of two-factor Heston-type model $SV_{2,0}$ to fit a degree of volatility-unrelated variations and comovements in the term structure of option-implied volatility skews, similar to the one in the data. To which extent do $SV_{3,1}$ and $SVJ_{3,1}$ models improve the fit of $SV_{2,0}$ and $SVJ_{2,0}$ along these particular dimensions? The bottom panels of Figures [1] (Figure [2]) present the fitted combinations of short term slope and term structure (short term and long term slope) implied by model $SV_{3,1}$ for different levels of volatility $\sqrt{V_t} = 0.16, 0.2, 0.24, 0.28$. The results indicate that $SV_{3,1}$ improves on $SV_{2,0}$ model mainly in two directions. First, Figure [1] shows that it generates a fitted co-movement of short term skews and term structures more consistent with the data. Second, Figure [2] shows that model $SV_{3,1}$ also implies an additional degree of variability in fitted long term skews, which produces a comovement of model-implied short and long term skews more similar to the one in the data. Figures [12] and [13] produce a corresponding comparison for models $SVJ_{3,1}$ and $SVJ_{2,0}$.

Figures 12 shows that model $SVJ_{3,1}$ can produce a slightly larger variability of fitted short term skews and implied volatility term structures than model $SV_{2,0}$, especially for low and average volatility levels $\sqrt{V_t} = 0.16, 0.2, 0.24$. Figure 13 indicates that overall the fit of model $SVJ_{3,1}$ improves on model $SV_{2,0}$, by producing a useful additional degree of variability in model-implied long term skews, which modulates the dynamics of the term structure of implied volatility skews more consistently with the empirical evidence. Overall, these findings confirm that volatility interactions and unspanned skewness are useful model properties, for specifying volatility-unrelated dynamic effects in model-implied term structures of volatility skews, which can produce an improved fit of the dynamic patterns of option data.

B.4. Stochastic Feedback Effects and Volatility-Unrelated Risk Neutral Skewness

A main driver of stochastic risk neutral skewness in our models is the stochastic volatility feedback between returns and (diffusive) volatility, given by:

$$corr_t(dS_t/S_t, d(v_{1t} + v_{2t})) = \frac{tr(R'QX_t)}{\sqrt{\lambda_0 + tr((I_2 + \Lambda E(k^2))X_t)tr(Q'QX_t)}}$$  \hspace{1cm} (31)

Therefore, the degree of volatility-unrelated variations in risk neutral skewness is naturally linked to the degree of volatility-unrelated variation of volatility feedback effects produced within our model setting. We illustrate these aspects in Figure 14 for model $SV_{3,1}$, in which correlation (31) does not depend on the level of return volatility:

$$corr_t(dS_t/S_t, d(v_{1t} + v_{2t})) = \frac{tr(R'QX_t)}{\sqrt{tr(X_t)tr(Q'QX_t)}} = \frac{tr(R'Q(I_2 + \xi_t \mathcal{U}(\alpha_t)))}{\sqrt{tr(I_2 + \xi_t \mathcal{U}(\alpha_t))tr(Q'Q(I_2 + \xi_t \mathcal{U}(\alpha_t)))}}$$  \hspace{1cm} (32)

where the last equality follows from Lemma 1. We plot correlation (31), for different volatility compositions $\xi_t$, as a function of unspanned skewness parameter $\alpha_t$.

The case $\alpha_t = 0$ corresponds to the range of possible model-implied volatility compositions and volatility feedback effects in model $SV_{2,0}$. The main message we obtain for this case is that the volatility feedback effect is monotonic in the volatility composition: As $\xi_t$ goes from 1 to 0 (i.e., $\mathcal{V}_{1t} - \mathcal{V}_{2t} \rightarrow 0$; $\mathcal{V}_{1t} \geq \mathcal{V}_{2t}$) the correlation between volatility and returns goes from an upper bound of about 0.35 to a lower bound of about -0.45. This feature generates a tight slope and term structure tradeoff in this model, which is reflected by the graphs in the top panels of Figure 1. As the model tries to fit a more negative slope
in the data with a stronger volatility feedback effect, it also forces the implied state to a more equal volatility composition between states $V_{1t}$ and $V_{2t}$. In doing so, it puts higher weights on the more strongly mean reverting factor $V_{2t}$, implying coeteris paribus also a more negative implied volatility term structure. The introduction of an unspanned skewness dimension in model $SV_{3,1}$ weakens this tight link, by allowing correlation (31) to depend on the unspanned skewness component $\alpha_t$. The range of admissible volatility composition and feedback effects in the model is substantially enlarged. For instance, while in $SV_{2,0}$ model (when $\alpha_t = 0$) a volatility feedback below 0.4 is accessible basically only with a volatility composition $\xi_t \in [0, 0.25]$, in model $SV_{3,1}$ this is achievable by any volatility composition when $\alpha_t$ is above approximately 0.1. Overall, these features produce the broader tradeoff between skewness and implied volatility term structure of fitted $SV_{3,1}$ and $SV_{J,3,1}$ models relative to their $SV_{2,0}$ and $SV_{J,2,0}$ benchmarks.

C. Distance Between Models and Unspanned Skewness Features

As shown in the previous sections, models $SV_{3,1}$ and $SV_{J,3,1}$ produce a broader variety of volatility-unrelated skewness features than $SV_{2,0}$ and $SV_{J,2,0}$ models. Therefore, a measure of discrepancy between the volatility surfaces generated by these models could prove useful to summarize the incremental pricing performance of $SV_{3,1}$ and $SV_{J,3,1}$ models, deriving from their ability to produce unspanned skewness effects. Figure 8 provides useful insights along this dimension. In $(\mathcal{M}, \mathcal{S})$-coordinates, the feasible set of admissible points achievable by $SV_{2,0}$ or $SV_{J,2,0}$ models is obtained, for fixed $\sqrt{V_t}$ and when $\xi_t$ moves between -1 and 1, as the main axis within the ellipses of admissible combinations of $SV_{3,1}$ or $SV_{J,3,1}$ models, parametrized by $\alpha_t \in [-\pi/2, \pi/2]$. Therefore, using the low dimensional $(\mathcal{M}, \mathcal{S})$-coordinate system, we can proxy the distance between volatility surfaces of $SV_{3,1}$ ($SV_{J,3,1}$) and $SV_{2,0}$ ($SV_{J,2,0}$) models, as the projection error $\eta_t$ of each admissible point $(\mathcal{M}, \mathcal{S})$ on the main axis of the ellipse produced by $SV_{3,1}$ ($SV_{J,3,1}$) model. Since $\xi_t$ and $2\alpha_t$ are good proxies for the radius and the angle associated with any particular point on the ellipses of $SV_{3,1}$ ($SV_{J,3,1}$) models, a simple proxy for the (signed) size of the projection error $\eta_t$ is:

$$d_t = \xi_t \sin(2\alpha_t)$$

Figure 15 plots $d_t$ over time, using the estimated latent states of $SV_{3,1}$ ($SV_{J,3,1}$) models.

Interestingly, the estimated distance processes $d_t$ are quite persistent, suggesting that unspanned skewness features embedded in S&P 500 index option-implied volatility surfaces
are likely to last for a while when they appear in the data. Estimated process \(d_t\) for \(SV_{3,1}\) and \(SVJ_{3,1}\) models feature a large comovement and a large variability. Over the whole sample, they both reach a minimum during the 1998 Russian crisis and the recent Subprime Crisis.

Let \(I_t\) be the daily relative \(RMSE\) improvement of \(SV_{3,1}\) and \(SVJ_{3,1}\) models, relative to their \(SV_{2,0}\) and \(SVJ_{2,0}\) benchmarks:

\[
I_t = \frac{\epsilon_{t^{\text{ref}}} - \epsilon_t}{\epsilon_{t^{\text{ref}}}},
\]

where \(\epsilon_t\) and \(\epsilon_{t^{\text{ref}}}\) are the daily \(RMSE\) of the \(SV_{3,1}\) (\(SVJ_{3,1}\)) model and their benchmarks, respectively. We expect the \(SV_{3,1}\) and \(SVJ_{3,1}\) models to improve on the pricing performance of the \(SV_{2,0}\) and \(SVJ_{2,0}\) models precisely when \(|d_t|\) is large, because in this case volatility-unrelated skewness features are likely more pronounced in S&P 500 index option smiles. We verify in more detail this intuition using the scatter plots of Figure 16, where we analyze the link between daily \(RMSE\) percentage improvements, \(|d_t|\) and the level of volatility \(V_t\).

In panel A, the explanatory variable on the \(x\)–axis is the daily pricing error of the reference models. Overall, we see that daily \(RMSE\) percentage improvements are increasing in the reference model’s daily \(RMSE\). This evidence indicates that \(SV_{3,1}\) and \(SVJ_{3,1}\) models tend to improve on benchmark models precisely on days where it is most needed.

In panel B, the explanatory variable on the \(x\)–axis is the 30-days option implied volatility. We find no clear relation between volatility level and percentage improvement of reference model’s daily \(RMSE\), which is a remarkable result when considering that in our data a single volatility level factor, extracted by standard Principal Component Analysis, explains about 97% of the (unconditional) variations of S&P 500 index option implied volatility surfaces. Overall, this evidence indicates that, as suggested by our previous analysis, pricing improvements produced by \(SV_{3,1}\) and \(SVJ_{3,1}\) models with respect to \(SV_{2,0}\) and \(SVJ_{2,0}\) models are largely volatility-unrelated.

In panel C, the explanatory variable on the \(x\)–axis is our proxy of (signed) model distance \(d_t\) defined in (33). We find an apparent U-shaped pattern in the percentage improvement of reference model’s daily \(RMSE\), with a minimum approximately around \(d_t = 0\), especially

\[\text{On days where two-factor models already perform well, with pricing errors within bid-ask spreads, the average improvement is only 8% (15%) for the } SV_{3,1} \text{ (} SVJ_{3,1}\text{) model. On days where reference models perform poorly, with daily } RMSE \text{ above 2.25 dollars, the average improvement jumps to about 25% for both } SV_{3,1} \text{ and } SVJ_{3,1}\text{ models. Remarkably, } SV_{2,0} \text{ and } SVJ_{2,0}\text{ models tend to perform poorly on the same dates, on average for about one quarter of our sample.}
\]
for the improvements of SVJ$_{3,1}$ model. Percentage improvement of reference model’s daily RMSE when $d_t$ deviates from zero are potentially large. For instance, in the SVJ$_{3,1}$ model the average price improvement estimated by non parametric regression for $d_t = 1$ ($d_t = -1$) is about 30% (20%). In summary, we find that RMSE pricing improvements provided by SVJ$_{3,1}$ and SVJ$_{3,1}$ models are large, unrelated to the level of the volatility, and strongly related to our proxy $d_t$ of distance between models, which captures unspanned skewness features through its dependence of state variable $\alpha_t$.

D. Stochastic Coefficients Model

The state reparametrization (22)-(23) allows us to better identify volatility-unrelated skewness effects, parametrized by state variable $\alpha_t$. The risk neutral dynamics of volatility factors $V_{1t}$, $V_{2t}$ can provide additional insight into the role of these state variables. Using Itô’s Lemma we obtain, after lengthy calculations\[16\]

$$dV_{1t} = \left( \beta(\tilde{Q}_t\tilde{Q}_t)^{11} + 2(M_t)^{11}V_{1t} + \frac{V_{1t}(\tilde{Q}_t\tilde{Q}_t)^{22} + V_{2t}(\tilde{Q}_t\tilde{Q}_t)^{11}}{V_{1t} - V_{2t}} \right) dt + 2\sqrt{V_{1t}(\tilde{Q}_t\tilde{Q}_t)^{11}}d\nu_{1t} \quad (35)$$

$$dV_{2t} = \left( \beta(\tilde{Q}_t\tilde{Q}_t)^{22} + 2(M_t)^{22}V_{2t} - \frac{V_{1t}(\tilde{Q}_t\tilde{Q}_t)^{22} + V_{2t}(\tilde{Q}_t\tilde{Q}_t)^{11}}{V_{1t} - V_{2t}} \right) dt + 2\sqrt{V_{2t}(\tilde{Q}_t\tilde{Q}_t)^{22}}d\nu_{1t} \quad (36)$$

where $(\nu_1, \nu_2)'$ is a standard bivariate motion and $2 \times 2$ random matrices $\tilde{M}_t, \tilde{Q}_t$ are defined by $\tilde{M}_t = O_t'MO_t$ and $Q_t = O_t'QO_t$. This shows that $(V_{1t}, V_{2t})$ dynamics is driven by two conditionally independent stochastic volatility processes, in which the (stochastic) volatility of volatility and drift parameters depend only on random matrix $O_t$. Conditional on $O_t$, processes $(V_{1t}, V_{2t})$ behave as two independent Bessel processes, in which the linear drift has been perturbed by the nonstandard term:

$$\pm \frac{V_{1t}(\tilde{Q}_t\tilde{Q}_t)^{22} + V_{2t}(\tilde{Q}_t\tilde{Q}_t)^{11}}{V_{1t} - V_{2t}}. \quad (37)$$

This term ensures that the ranking of eigenvalues $V_{1t}, V_{2t}$ is always preserved, but is is typically small most of the time. We can therefore interpret the risk neutral dynamics of the volatility components $V_{1t}$, $V_{2t}$ in our model as a two-factor random coefficient stochastic volatility model, in which the random coefficients are driven exclusively by the stochastic variable $O_t = O(\alpha_t)$, i.e., the unspanned skewness component. It follows that state variable $\alpha_t$ can impact in two ways on the model-implied volatility surface. First, via the stochastic mean reversion and volatility of volatility coefficients in dynamics (35)-(36), it produces a

---

\[16\]See also Benabid, Bensusan and El Karoui (2009). The proof is available on request.
variety of effects on the term structure of the volatility. Second, it impacts on the time varying jump intensity and the volatility feedbacks, as:

$$\lambda_t = \lambda_0 + tr(\Lambda X_t) = \lambda_0 + (\tilde{\Lambda}_t)^{11} \nu_{1t} + (\tilde{\Lambda}_t)^{22} \nu_{2t},$$

(38)

and

$$\frac{1}{2} Cov_t(dS_t/S_t, d(\nu_{1t} + \nu_{2t})) = tr(R'Q X_t) = (\tilde{R}'\tilde{Q}_t)^{11} \nu_{1t} + (\tilde{R}'\tilde{Q}_t)^{22} \nu_{2t},$$

(39)

where $\tilde{\Lambda}_t = O'_t \Lambda O_t$ and $\tilde{R}'\tilde{Q}_t = O'_t R' Q O_t$. Conditional on $O_t$, equations (38) and (39) define the time varying intensity and volatility feedback effect of a two-factor Bates (2000)-type model. Relative to this model, model $SVJ_{3.1}$ produces an additional degree of skewness variability by making the coefficients $(\tilde{\Lambda}_t)^{ii}$ and $(\tilde{R}'\tilde{Q}_t)^{ii}$ in formulae (38) and (39) stochastic. Since this additional variability is completely driven by matrix $O_t = O(\alpha_t)$, these unspanned skewness effects are largely captured by state variable $\alpha_t$.

V. Conclusions

Using a new option valuation framework, featuring interdependent volatility risks and a stochastic skewness component unrelated to the volatility factors, we analyze the pricing of S&P 500 index options. We estimate our models based on S&P 500 index options data from January 1996 to September 2009 and find that they provide superior pricing performance over a number of benchmark two- and three-factor affine volatility models in the literature, with reductions in average root mean square pricing error of about 20% out-of-sample. We find that the improved fit of our models is largely due to their improved modeling of the term structure of implied-volatility skews. In addition to highlighting the usefulness of multi-factor risk specifications for modeling the dynamics of implied volatility smiles, our results emphasize the even more key role of dynamic volatility interactions and volatility-unrelated skewness for option valuation purposes. More generally, they raise the question of how the dynamics of a multi-factor option pricing model ought to be specified. Our findings show that three-factor state dynamics based on the class of matrix AJD in Leippold and Trojani (2008) can provide a convenient framework to specify within a tractable model interacting volatility components and volatility-unrelated skewness effects.
A Nested models

Several well-studied affine option pricing models with independent factors are nested in our framework, if we allow $\beta$ to be a diagonal matrix instead of a scalar. In this case, the independent volatility factors can be written as diagonal elements of $X_t$. Below, we show the equivalence of the processes and how the parameters can be converted from the notation in the original papers into our notation. For the sake of legibility, we suppress the time index on all components of state variables and Brownian motions.

A. Diffusive models

The return dynamics of the $SV_{2,0}$ two-factor Heston model of Christoffersen et al. (2009) is

$$\frac{dS}{S} = (r - q)dt + \sqrt{V_1}dz_1 + \sqrt{V_2}dz_2$$

(40)

where $r$ is the risk-free rate, $q$ the dividend yield and $V_i$ are two independent stochastic volatility factors with the following dynamics:

$$dV_i = (a_i - b_i V_i)dt + \sigma_i \sqrt{V_i}dw_i \quad i = 1, 2$$

(41)

where the correlation between $dz_i$ and $dw_j$ is $\delta_{ij}\rho$.

If we write $X = \begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}$ and $dZ = \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}$, then (40) can be written as

$$\frac{dS}{S} = (r - q)dt + tr[\sqrt{X_t}dZ],$$

which is exactly the diffusive part of our return equation (7).

In order to show the equality of the volatility factors, we first need to establish that the diagonal elements of $X_t$ in (6) are independent CIR processes if the parameter matrices $M, R, Q$ are diagonal. We start by explicitly writing the diagonal elements of $X_t$ in this case:

$$dX_{ii} = (\beta Q_{ii}^2 + 2M_{ii}X_{ii}) dt + \sum_k \sqrt{X_{ki}} dB_{ki}$$

(42)

To eliminate the seeming interdependence of the diagonal elements, we introduce $n$ new independent Brownian motions $dW_i$:

$$dW_i = \frac{1}{\sqrt{X_{ii}}} \sum_k \sqrt{X_{ki}} dB_{ki}$$
This allows us to express (42) as $n$ independent CIR processes:

$$dX_{ii} = (\beta Q_{ii}^2 + 2M_{ii}X_{ii})\,dt + 2Q_{ii}\sqrt{X_{ii}}dW_i$$  \hspace{1cm} (43)

To convert our notation into the notation of (41), simply set

$$a_i = \beta_i Q_{ii}^2, \quad b_i = -2M_{ii}, \quad \sigma_i = 2Q_{ii}, \quad \text{and} \quad \rho_i = R_{ii}.$$

**Remark 3** Our state matrix $X_t$ will generally not remain diagonal, even if all parameter matrices and the initial state $X_0$ are diagonal. This does not void the nesting argument, because $X_{12,t}$ does not enter the pricing equation. There is no economic interpretation for the process $X_{12,t}$, it is a mere artefact of writing a two-dimensional CIR process in matrix form.

**B. Jump parameters**

The jump intensity in Bates (2000) is given as $\lambda_t = \lambda_0 + \lambda_1 V_{1t} + \lambda_2 V_{2t}$, which is already identical to our jump intensity $\lambda_t = \lambda_0 + tr(\Lambda X_t)$, if we write $\lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Our definitions of the jump size distribution is the same as in Bates (2000).

**B Short-maturity smile asymptotics as a framework to study unspanned volatility effects**

We choose a convenient, low-dimensional framework to study the unspanned volatility effects in our model. Let $IV_t(T,K)$ be the Black-Scholes option implied volatility at time $t$ for maturity $T$ and strike price $K$, and consider the following approximation of the implied volatility smile:

$$IV_t(T,K) = V_t^{1/2} + S_t \frac{K - S_t}{S_t} + M_t(T - t) + \frac{1}{2} C_t \left( \frac{K - S_t}{S_t} \right)^2$$  \hspace{1cm} (44)

with

$$S_t = S_t \lim_{T \to t} \frac{\partial IV(T,S_t)}{\partial K}, \quad M_t = \lim_{T \to t} \frac{\partial IV(T,S_t)}{\partial T}, \quad C_t = S_t^2 \lim_{T \to t} \frac{\partial^2 IV(T,S_t)}{\partial^2 K}. \hspace{1cm} (45)$$

\footnote{See, among others, Dumas, Fleming and Whaley (1998), Durrleman (2004) and Durrleman and Karoui (2007).}
Where \( V^{1/2} \) is the short-term, at the money volatility level, \( S_t \) is the skew for short maturities, i.e., the short maturity limit of the derivative of the at-the-money implied volatility with respect to moneyness \( K/S_t \). \( M_t \) is the smile term structure for short maturities, i.e., the short maturity limit of the derivative with respect to maturity \( T \). Finally, \( C_t \) is the smile convexity for short maturities.

A. Construction of Level, Skewness and Term Structure Factors from Data and Model

For our empirical studies, we construct proxies for the skewness and term structure factors using two different methods. Whenever we calculate these quantities from the data or from model fits, we perform regression (44), separately for each day of data (model-implied prices). To obtain the short-term skew \( S_t \) and the short term structure \( M_t \), we consider only options with \( \tau < 73 \) days and \( 0.85 \leq K/S_t \leq 1.15 \). To obtain the long-term skew \( S_t^{\text{long}} \) and the long term structure \( M_t^{\text{long}} \) in Figures 1 and 4, we consider options with \( \tau \geq 122 \) days and \( 0.7 \leq K/S_t \leq 1.3 \).

Whenever we calculate feasible regions as in Figures 3, 4, 8 and 9, we calculate the derivatives (45) numerically. More precisely, we approximate the skew as \( S_t = \frac{\partial IV(T,S_t)}{\partial K} \) at \( \tau = 0.25 \) (6) months for the short (long) term. We approximate the at the money term structure as \( M_t = \frac{IV(\tau_1) - IV(\tau_0)}{\tau_1 - \tau_0} \), where \( \tau_1 = 0.25 \) (12) months for the short (long) term structure and \( \tau_0 = 0 \). In the SV\(_{3,1}\)-model, we evaluate \( IV(\tau_0) = tr(X)^{1/2} \) and use this quantity directly.

C Proofs and additional expressions

A. Proof of Lemma 7

Since \( X_t \) is a symmetric positive definite matrix, we can always write it as

\[
X_t = O_t V_t O_t',
\]

where \( V_t \) is a \( 2 \times 2 \) diagonal matrix of positive eigenvalues \( V_{1t} \) and \( V_{2t} \) and \( O_t = [O_{1t}, O_{2t}] \) is a \( 2 \times 2 \) orthogonal matrix of eigenvectors \( O_{1t}, O_{2t} \) having unit norm. A convenient parametrization of \( O_t \) by means of a single parameter \( \alpha_t \in [-\pi/2, \pi/2] \) is obtained using standard polar coordinates:

\[
O_t = \begin{pmatrix}
\cos(\alpha_t) & -\sin(\alpha_t) \\
\sin(\alpha_t) & \cos(\alpha_t)
\end{pmatrix}.
\]

The sum of \( V_{1t} \) and \( V_{2t} \) naturally parametrizes the spot volatility of returns. Thus
we can define the volatility level factor $V_t$ and a dimensionless factor $\xi_t$ that measures the composition of the return volatility:

\[ V_t := \text{tr}(X_t) = \text{tr}(\mathcal{V}_t) = \mathcal{V}_{1t} + \mathcal{V}_{2t}, \quad \xi_t := \frac{\mathcal{V}_{1t} - \mathcal{V}_{2t}}{\mathcal{V}_{1t} + \mathcal{V}_{2t}} = \frac{V_t}{V_i} \]  

(46)

Using this notation we can derive Lemma 1.

\[
X_t = V_t \left[ \mathcal{O}_t \left( \begin{array}{cc} 1 + \xi_t & 0 \\ 0 & 1 - \xi_t \end{array} \right) \mathcal{O}_t' \right]
\]

\[
= \frac{V_t}{2} \left[ \mathcal{I}d_2 + \xi_t \cdot \mathcal{U}_t \right]
\]

(47)

where $\mathcal{I}d_2$ is the $2 \times 2$ identity matrix and

\[
\mathcal{U}(\alpha_t) = \begin{pmatrix} \cos(2\alpha_t) & \sin(2\alpha_t) \\ \sin(2\alpha_t) & -\cos(2\alpha_t) \end{pmatrix}
\]

is a reflection matrix with trace zero and determinant minus one such that all components are bounded in the interval $[-1, 1]$. Therefore, it can be conveniently used to specify different correlation processes, such as, for instance, those needed to specify stochastic volatility feedback effects.

**Remark 4** In order to make decomposition (24) unique, one has to choose an ordering of the eigenvalues, and thus the sign of $\xi_t$. We choose $\mathcal{V}_{1t} > \mathcal{V}_{2t}$ and therefore $0 \leq \xi_t \leq 1$.

**Remark 5** Lemma 1 can be used to decompose expressions of the form $\text{Tr}[HX_t]$ when $H$ a $2 \times 2$ parameter matrix, as follows:

\[
\text{Tr}[HX_t] = \frac{V_t}{2} \left[ \text{Tr}(H) + \xi_t \cdot \text{Tr}(H \mathcal{U}(\alpha_t)) \right]
\]

\[
= \frac{V_t}{2} \left[ \text{Tr}(H) + \xi_t \cdot \left( \cos(2\alpha_t)(H_{11} - H_{22}) + \sin(2\alpha_t)(H_{12} + H_{21}) \right) \right]
\]

(48)

**B. Parameter identification**

We first discuss the identification of the diffusive parameters. Every stochastic process is uniquely characterized by its infinitesimal generator. The infinitesimal generator of the joint
process for stock returns $Y_t := dS_t/S_t$ and the factor $X_t$ is (see Leippold and Trojani (2008)):

$$\mathcal{L}_{Y,X} = \left( r - q - \frac{1}{2} Tr[X] \right) \frac{\partial}{\partial Y} + \frac{1}{2} Tr[X] \frac{\partial^2}{\partial Y^2} + 2 Tr[X'R'QD] \frac{\partial}{\partial Y} + 
$$

$$+ Tr \left[ (\beta Q'Q + MX + XM')D + 2XDQ'QD \right] \quad (49)$$

where $(D)_{ij} = \frac{\partial}{\partial X_{ij}}$ is the matrix differential operator.

The parameter set of the diffusive process is $\theta = \{\beta, M, R'Q, Q'Q\}$. Parameter identification requires that the infinitesimal generator be unique for each set of parameters given any state $X_t$. Maximal identification aims at achieving this goal through the minimal set of parameter restrictions. Equation (49) contains an ambiguity that has to be resolved. Let

$$Z_t = DX_tD^{-1},$$

then

$$\mathcal{L}(X_t, \theta) = \mathcal{L}(Z_t, \theta_Z)$$

with $\theta_Z = \{\beta, DMD^{-1}, DR'QD^{-1}, DQ'QD^{-1}\}$.

We now want to identify parameter restrictions on $\theta$ that only admit $D$ to be the identity matrix. Without loss of generality, we can assume $|\det(D)| = 1$. Next we observe that the expression $X_t$ is symmetric by construction, thus $Z_t$ needs to be symmetric, as well. Symmetry of $DX_tD^{-1}$ is ensured if $D$ is orthogonal ($D' = D^{-1}$), thus $D$ must be a rotation or mirror matrix.

In a next step, we choose $M$ to be lower triangular. This requires $D$ to be lower triangular, in order to ensure $DMD^{-1}$ to be lower triangular, as well. If $D$ is orthogonal, lower triangular and has a determinant of one, it must be a diagonal matrix $\left( \begin{smallmatrix} \alpha & 0 \\ \beta & \beta \end{smallmatrix} \right)$ with elements $\alpha, \beta = \pm 1$. We now have $DMD^{-1} = \left( \begin{smallmatrix} M_{11} & 0 \\ \beta/\alpha & M_{22} \end{smallmatrix} \right)$. By choosing the sign of $M_{21}$ as being negative, we exclude the case $\alpha \neq \beta$, which concludes the identification of the diffusion parameters.

Remark 6 Our choices for $M$ implicitly identify the state. The choice for $M$ to be lower triangular selects the order of the mean reversion speeds of the eigenvalues of $X_t$. In our setting, this implies that the dominant factor with the fast mean reversion is $X_{22}$. Our choice for $M_{21} > 0$ identifies the sign of $X_{12,t}$.

We now need to relate the composite parameters $Q'Q$ and $R'Q$ to the parameter matrices $Q$ and $R$. We choose $Q$ to be the unique Choleski decomposition of $Q'Q$, i.e. $Q$ upper triangular and positive definite. By simple matrix algebra, we obtain $R = (Q')^{-1}(R'Q)'$. We
add the ad-hoc restriction for \( R \) to be upper triangular, in order to reduce the number of parameters.

The identification of the jump parameter \( \Lambda \) follows a similar argument. The jump intensity (see Assumption 2) is \( \lambda_0 + tr[\Lambda X_t] \), with \( tr[\Lambda X_t] = \Lambda_{11}X_{11,t} + (\Lambda_{12} + \Lambda_{21})X_{12,t} + \Lambda_{22}X_{22,t} \). To identify the out-of diagonal elements of \( \Lambda \), we choose it to be upper triangular, i.e. \( \Lambda_{21} = 0 \).

C. The Cosine-FFT-method

The Cosine-FFT method, introduced by Fang and Oosterlee (2008), is an efficient algorithm to approximate a density \( f(x) \) on a finite support \([a, b]\) via a truncated cosine-series expansion of the characteristic function \( \phi(\omega) \):

\[
 f(x) = \sum_{k=0}^{\infty} 'A_k \cos \left( k\pi \frac{x-a}{b-a} \right) \approx \sum_{k=0}^{N-1} 'A_k \cos \left( k\pi \frac{x-a}{b-a} \right) \quad (50)
\]

with

\[
 A_k = \frac{2}{b-a} \int_{a}^{b} f(x) \cos \left( k\pi \frac{x-a}{b-a} \right) dx \\
 \approx \frac{2}{b-a} \text{Re} \left[ \int_{\mathbb{R}} f(x) \exp \left( ik\pi \frac{x-a}{b-a} \right) dx \right] \\
 \approx \frac{2}{b-a} \text{Re} \left[ \phi \left( \frac{k\pi}{b-a} \right) \exp \left( -i k\pi \frac{a}{b-a} \right) \right] \quad (51)
\]

The price of a contingent claim with payoff \( v(y, T) \) and time to maturity \( \tau \) is then obtained as:

\[
 C(x, T) = e^{-r\tau} \int_{a}^{b} v(y, T) f(y|x) dy \\
 = e^{-r\tau} \int_{a}^{b} v(y, T) \sum_{k=0}^{\infty} 'A_k \cos \left( k\pi \frac{x-a}{b-a} \right) dy \\
 = e^{-r\tau} \sum_{k=0}^{\infty} \frac{1}{2} (b-a) A_k \int_{a}^{b} \frac{2}{b-a} v(y, T) \cos \left( k\pi \frac{x-a}{b-a} \right) \\
 = e^{-r\tau} \sum_{k=0}^{\infty} \frac{1}{2} (b-a) A_k \int_{a}^{b} \frac{2}{b-a} v(y, T) \cos \left( k\pi \frac{x-a}{b-a} \right) \quad (52)
\]

For a plain vanilla call with \( v(y, T) = [K(e^y-1)]^+ \), the integral \( V_k := \int_{\frac{a}{b-a}}^{\frac{b}{b-a}} v(y, T) \cos \left( k\pi \frac{x-a}{b-a} \right) \) evaluates as

\[
 V_k^{\text{call}} = \frac{2}{b-a} K \left( \chi_k(0, b) - \psi_k(0, b) \right)
\]
with
\[
\chi_k(c, d) = \int_c^d e^{y \cos \left( k\pi \frac{y - a}{b - a} \right)} dy \\
= \frac{1}{1 + \left( \frac{k\pi}{b - a} \right)^2} \left[ \cos \left( k\pi \frac{d - a}{b - a} \right) e^d - \cos \left( k\pi \frac{c - a}{b - a} \right) e^c + \frac{k\pi}{b - a} \sin \left( k\pi \frac{d - a}{b - a} \right) e^d - \frac{k\pi}{b - a} \sin \left( k\pi \frac{c - a}{b - a} \right) e^c \right]
\]
\[
\psi_k(c, d) = \int_c^d \cos \left( k\pi \frac{y - a}{b - a} \right) dy \\
= \begin{cases} 
\frac{b - a}{k\pi} \left[ \sin \left( k\pi \frac{d - a}{b - a} \right) - \sin \left( k\pi \frac{c - a}{b - a} \right) \right] & k \neq 0 \\
\frac{b - a}{k\pi} & k = 0 
\end{cases}
\]

The expression for the call price that we use is
\[
C(x, T) = e^{-rT} \sum_{k=0}^{N-1} \text{Re} \left\{ \phi \left( \frac{k\pi}{b - a} \right) \exp \left( -i \frac{k\pi}{b - a} \right) \right\} V_k \tag{53}
\]

We employ the Cosine-FFT method for two reasons. First, its superior convergence properties allow us to reduce the number of required evaluations of the characteristic function from a typical $2^{12} = 4096$ for the FFT method to between 250 and 300. Second, it does not involve an interpolation between strike prices, which makes it possible to calculate the skewness $S_t = S_t \lim_{T \to t} \frac{\partial IV(T, S_t)}{\partial K}$ in (44) with high precision via numerical differentiation.
Panel A: Fit of model $SV_{2,0}$

Panel B: Fit of model $SV_{3,1}$

Figure 1: Short term skew $S$ (short) versus term structure $M$ (short). Grey dots: Data. Black dots: Fitted values of a two factor Heston model ($SV_{2,0}$) and our pure diffusion model ($SV_{3,1}$). In each plot of each panel, we select observations corresponding to a short term at the money implied volatility of ±5% around the observed level, i.e., 19%-21% for the second plot of each panel.
Panel A: Fit of model $SV_{2,0}$

Figure 2: Short term skew $S$ (short) versus long term skew $S$ (long). Grey dots: Data. Black dots: Fitted values of a two factor Heston model ($SV_{2,0}$) and our pure diffusion model ($SV_{3,1}$). In each plot of each panel, we select observations corresponding to a short term at the money implied volatility of ±5% around the observed level, i.e., 19%-21% for the second plot of each panel.
Figure 3: Feasible set of \((M,S)\)-combinations for the \(SV_{3,1}\) model. We plot the feasible set of skewness and term structure combinations for a range of ±5% around three volatility levels: 9.5%-10.5% (left panel), 19%-21% (middle panel), 28.5%-31.5% (right panel). The grey dots in each panel represent \((M,S)\) combinations observed in the data. The lines crossing the ellipses plot the admissible sets of \((M,S)\) points implied by the (nested) \(SV_{2,0}\) model.
Figure 4: Comparative statics of the feasible region with respect to model parameters in $SV_{3,1}$ model. For a volatility level $\sqrt{V_t} = 20\%$, we plot in black the surface of admissible points, bounded by an elliptical shape, for the combinations of short term (long-term) skewness $S$ and term structure $M$, in the left (right) panel. We then vary in each plot one parameter by $-20\% (+20\%)$ and plot in blue (red) the implied surface of admissible points, again bounded by a corresponding curve. In each plot, light grey dots correspond to $(M, S)$-combinations observed in the data.
Panel A: Pure diffusion models

Panel B: Jump diffusion models

Figure 5: Time series of daily RMSE and daily RMSE improvements. Panel A (B) compares pure diffusion (jump diffusion) models. In the top graph of each panel, we plot daily RMSE of benchmark SV2.0 (SVJ2.0) model. In the middle graph, we plot relative RMSE improvements of SV3.1 model over SV2.0 model. In the bottom graph, we plot relative RMSE improvements of SV3.0 (SVJ3.1) model over SV2.0 (SVJ2.0) model. Grey areas in each plot depict NBER recessions in our sample period; important crisis events, indicated as (1) to (5) in each plot, are listed in Table 3.
Figure 6: Estimated time series of implied states $(X_{11t}, X_{12t}, X_{22t})$ in models $SV_{3,1}$ and $SVJ_{3,1}$. Black (red) lines correspond to states estimated for $SV_{3,1}$ pure diffusion ($SVJ_{3,1}$ jump diffusion) model. Grey areas in each plot depict NBER recessions in our sample period; important crisis events, indicated as (1) to (5) in each plot, are listed in Table 3.
Panel A: Volatility feedback effect in the $SV_{3,1}$ model

Panel B: Volatility feedback effect in the $SVJ_{3,1}$ model

Panel C: Jump intensity in the $SVJ_{3,1}$ model

Figure 7: Unspanned stochastic skewness effects in $SV_{3,1}$ and $SVJ_{3,1}$ models. In the different panels, we plot in black the model-implied time series of volatility feedback effects $corr_t(dS_t/S_t, d(v_{1t} + v_{2t}))$ and jump intensity $\lambda_0 + \lambda_{11}X_{11t} + \lambda_{12}X_{12t} + \lambda_{22}X_{22t}$. In blue, we plot the model-implied time series of volatility feedback effects and jump intensities under the additional constraint that $X_{12t} = 0$. Panel A (B) presents volatility feedback effects for $SV_{3,1}$ ($SVJ_{3,1}$) model. Panel C presents time varying jump intensities for model $SVJ_{3,1}$. Grey areas in each plot depict NBER recessions in our sample period; important crisis events, indicated as (1) to (5) in each plot, are listed in Table 3.
Figure 8: Admissible skewness and term structure tradeoff in model $SV_{3,1}$. We plot the admissible model-implied combinations of short term skewness $S$ and term structure $M$, implied for different volatility compositions $\xi := \frac{V_1 - V_2}{V_1 + V_2} = 0.2, 0.4, 0.6, 0.8, 1$, and different unspanned skewness parameters $\alpha = -\pi/4, 0, \pi/4, \pi/2$ in model $SV_{3,1}$. For $\xi = 0$, the set of admissible points collapses to a single point in the center of the different ellipses in the graphs. Due to the periodicity of $\cos(2\alpha)$, model-implied combinations for $\pi/2$ are equal to those for $-\pi/2$. We present plots for volatility levels $\sqrt{V_t} = 10\%$ (left panel) and $\sqrt{V_t} = 20\%$ (right panel).
Figure 9: Admissible set of $({\mathcal{M}},S)$-combinations in the $SV_{3,0}$ model. We plot the feasible set of short term skewness $S$ and term structure $\mathcal{M}$ combinations for a range of ±5% around three benchmark volatility levels: 9.5%-10.5% (left panel), 19%-21% (middle panel) and 28.5%-31.5% (right panel). Grey dots in each plot represent observed $({\mathcal{M}},S)$-combinations in the data.
Panel A: Pure diffusion model $SV_{3,1}$

Panel B: Jump diffusion model $SVJ_{3,1}$

Figure 10: Model-implied volatility surfaces for a volatility level $\sqrt{V} = 0.17$ and a volatility composition $\xi = 1$, in dependence of unspanned skewness parameter $\alpha$. Panel A present plots for pure diffusion $SV_{3,1}$ model. Panel B presents plots for $SVJ_{3,1}$ jump diffusion model. In panel B, the top graphs plot the total volatility surface, the middle graphs the diffusive component of the surface and the bottom graphs the jump component of the surface.
The data into quintiles of model-implied PCA. The following five columns present results of 5 conditional PCAs, in which we stratify the data into quintiles of model-implied state $\alpha$. Panel A (B) applies a stratification with respect to the model-implied $\alpha_t$ of $SV_{3,1}$ ($SVJ_{3,1}$) model.

Figure 11: Factor loadings of the first two principal components of the implied volatility surface as a function of model-implied $\alpha_t$. The left column presents results for an unconditional PCA. The following five columns present results of 5 conditional PCAs, in which we stratify the data into quintiles of model-implied state $\alpha_t$. Panel A (B) applies a stratification with respect to the model-implied $\alpha_t$ of $SV_{3,1}$ ($SVJ_{3,1}$) model.
Figure 12: Short term skew $S$ (short) and term structure $M$ (short) of $SVJ_{2,0}$ and $SVJ_{3,1}$ models. Grey dots in each graph represent $(M, S)$-combinations observed in the data. Black dots reproduce model-implied combinations for the $SVJ_{2,0}$ (panel A) and the $SVJ_{3,1}$ (panel B) model. In each graph, we select observations for an interval of short term at the money implied volatility of $\pm 5\%$ around a benchmark level, i.e., 19%-21% for each second graph from the left.
Panel A: Fit of $SVJ_{2,0}$ model

Panel B: Fit of $SVJ_{3,1}$ model

Figure 13: Short term skew $S$ (short) and long term skew $S$ (long) of $SVJ_{2,0}$ and $SVJ_{3,1}$ models. Grey dots in each graph represent $(M, S)$-combinations observed in the data. Black dots reproduce model-implied combinations for $SVJ_{2,0}$ (panel A) and $SVJ_{3,1}$ (panel B) model. In each graph, we select observations for an interval of short term at the money implied volatility of $\pm5\%$ around a benchmark level, i.e., 19%-21% for each second graph from the left.
Figure 14: Model-implied correlations between returns and volatility in $SV_{3,1}$ model, for different volatility compositions $\xi = 0, 0.25, 0.5, 0.75, 0.9, 1$, as a function of unspanned skewness parameter $\alpha \in (\pi/2, \pi/2]$.

$$d = \xi \cdot \sin(2\alpha)$$

Figure 15: Weekly averages of signed model distance $d = \xi \cdot \sin(2\alpha) = \xi \cdot 2 \sin(\alpha) \cos(\alpha)$, as a proxy for the distance between the volatility surfaces of our 3-factor models and those of the nested 2-factor diagonal models. The black line plots estimated distances for $SV_{3,1}$ pure diffusion model (left scale). The red line plots estimated distances for $SVJ_{3,1}$ jump diffusion model (right scale). Grey areas in each plot depict NBER recessions in our sample period; important crisis events, indicated as (1) to (5) in each plot, are listed in Table 3.
Figure 16: Analysis of RMSE improvements. For each trading day in our sample, we scatter plot the RMSE improvement defined in (34) as a function of different explanatory variables. Left columns compare pure diffusion models ($SV_2,0$ versus $SV_3,1$). Right columns compare jumps diffusion models ($SVJ_2,0$ versus $SVJ_3,1$). The blue line in each graph reproduces a nonparametric regression of RMSE improvements on each explanatory variable. In Panel A, the explanatory variable is the RMSE of the given benchmark model. In Panel B, the explanatory variable is the level of volatility. In Panel C, the explanatory variable in the left (right) plot is the model distance $d = \xi \cdot \sin(2\alpha)$ between $SV_{3,1}$ and $SV_{2,0}$ ($SV_{3,1}$ and $SV_{2,0}$) models.
<table>
<thead>
<tr>
<th>$r$</th>
<th>$q$</th>
<th>$SV_{r,q}$</th>
<th>$SV_{J,r,q}$</th>
<th>Pure diffusion models</th>
<th>Jump-diffusion models</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>$SV_{1,0}$ [Heston (1993)]</td>
<td>$SV_{J1,0}$ [Bates (1996)]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>$SV_{2,0}$ [Christoffersen et al. (2009)]</td>
<td>$SV_{J2,0}$ [Bates (2000)]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>$SV_{3,0}$</td>
<td>$SV_{J3,1}$ [da Fonseca et al. (2008)]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$SV_{3,1}$</td>
<td>$SV_{J3,1}$ [Leippold and Trojani (2008)]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Models considered in our study. $r$ is the total number of state variables and $q$ the number of unspanned stochastic skewness components.

### Panel A: Summary statistics

<table>
<thead>
<tr>
<th>Sample</th>
<th>“monthly”</th>
<th>“full”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling interval</td>
<td>monthly</td>
<td>daily</td>
</tr>
<tr>
<td>Trading days $T$</td>
<td>59</td>
<td>3460</td>
</tr>
<tr>
<td>Total number of observations</td>
<td>21,993</td>
<td>638,365</td>
</tr>
<tr>
<td>Average time to maturity</td>
<td>130 days</td>
<td>133.5 days</td>
</tr>
<tr>
<td>Average moneyness ($S/K$)</td>
<td>1.06</td>
<td>1.07</td>
</tr>
<tr>
<td>Average option price</td>
<td>$107</td>
<td>$115</td>
</tr>
</tbody>
</table>

### Panel B: Number of contracts stratified by moneyness and maturity

<table>
<thead>
<tr>
<th>$S/K$</th>
<th>$\tau \leq 20$</th>
<th>$20 &lt; \tau \leq 80$</th>
<th>$80 &lt; \tau \leq 180$</th>
<th>$\tau &gt; 180$</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt; 0.80$</td>
<td>82</td>
<td>2'113</td>
<td>5'791</td>
<td>16'933</td>
<td>24'919</td>
</tr>
<tr>
<td>$0.80 &lt; S/K &lt; 0.90$</td>
<td>956</td>
<td>16'426</td>
<td>20'016</td>
<td>29'928</td>
<td>67'326</td>
</tr>
<tr>
<td>$0.90 &lt; S/K &lt; 1.00$</td>
<td>14'562</td>
<td>78'895</td>
<td>37'364</td>
<td>37'065</td>
<td>167'886</td>
</tr>
<tr>
<td>$1.00 &lt; S/K &lt; 1.10$</td>
<td>19'199</td>
<td>77'512</td>
<td>33'340</td>
<td>32'301</td>
<td>162'352</td>
</tr>
<tr>
<td>$1.10 &lt; S/K &lt; 1.20$</td>
<td>5270</td>
<td>41'087</td>
<td>21'124</td>
<td>23'406</td>
<td>90'887</td>
</tr>
<tr>
<td>$1.20 &lt; S/K &lt; 1.30$</td>
<td>1139</td>
<td>18'084</td>
<td>14'568</td>
<td>16'525</td>
<td>50'316</td>
</tr>
<tr>
<td>$1.30 &lt; S/K &lt; 1.40$</td>
<td>390</td>
<td>8'258</td>
<td>9'565</td>
<td>10'943</td>
<td>29'156</td>
</tr>
<tr>
<td>$S/K &gt; 1.40$</td>
<td>251</td>
<td>8'590</td>
<td>14'230</td>
<td>22'452</td>
<td>45'523</td>
</tr>
<tr>
<td>all</td>
<td>41'849</td>
<td>250'965</td>
<td>155'998</td>
<td>189'553</td>
<td>638'365</td>
</tr>
</tbody>
</table>

Table 2: Panel A: Summary statistics of the data. The “monthly” column refers to the data set used for parameter estimation and in-sample performance analysis. The “full” column refers to the data set used for out-of sample evaluation. Panel B: Number of contracts stratified by moneyness $S/K$ and maturity in days.
Table 3: Description of major crisis events indicated in our time-series plots with labels (1) to (5).

Panel A: Estimated diagonal diffusion parameters

<table>
<thead>
<tr>
<th></th>
<th>$M_{11}$</th>
<th>$M_{22}$</th>
<th>$Q_{11}$</th>
<th>$Q_{22}$</th>
<th>$R_{11}$</th>
<th>$R_{22}$</th>
<th>$\beta$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SV_{3,1}$</td>
<td>-0.3426</td>
<td>-4.4856</td>
<td>0.0136</td>
<td>0.4116</td>
<td>-0.4878</td>
<td>-0.6279</td>
<td>1.2039</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0120)</td>
<td>(0.0149)</td>
<td>(0.0114)</td>
<td>(0.0041)</td>
<td>(0.5340)</td>
<td>(0.01781)</td>
<td>(0.0339)</td>
<td></td>
</tr>
<tr>
<td>$SV_{2,0}$</td>
<td>-0.0155</td>
<td>-5.9240</td>
<td>0.1583</td>
<td>0.5671</td>
<td>-0.7849</td>
<td>-0.6835</td>
<td>0.0934</td>
<td>0.8997</td>
</tr>
<tr>
<td></td>
<td>(0.0154)</td>
<td>(0.0364)</td>
<td>(0.0040)</td>
<td>(0.0019)</td>
<td>(0.0116)</td>
<td>(0.0124)</td>
<td>(0.0312)</td>
<td>(0.0101)</td>
</tr>
</tbody>
</table>

Panel B: Estimated out-of-diagonal diffusion parameters

<table>
<thead>
<tr>
<th></th>
<th>$M_{21}$</th>
<th>$Q_{12}$</th>
<th>$R_{21}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SV_{3,1}$</td>
<td>8.8713</td>
<td>-0.0137</td>
<td>-0.5900</td>
</tr>
<tr>
<td></td>
<td>(0.0506)</td>
<td>(0.3229)</td>
<td>(0.3747)</td>
</tr>
<tr>
<td>$SV_{J3,1}$</td>
<td>2.1498</td>
<td>0.0366</td>
<td>-0.1039</td>
</tr>
<tr>
<td></td>
<td>(0.0491)</td>
<td>(0.0272)</td>
<td>(0.4037)</td>
</tr>
</tbody>
</table>

Panel C: Estimated parameters of the jump component

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_0$</th>
<th>$\Lambda_{11}$</th>
<th>$\Lambda_{12}$</th>
<th>$\Lambda_{22}$</th>
<th>$\bar{k}$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SV_{J3,1}$</td>
<td>0.0023</td>
<td>51.77</td>
<td>-55.09</td>
<td>55.04</td>
<td>-0.0699</td>
<td>0.0993</td>
</tr>
<tr>
<td></td>
<td>(0.0587)</td>
<td>(1.18)</td>
<td>(3.62)</td>
<td>(1.35)</td>
<td>(0.0028)</td>
<td>(0.0010)</td>
</tr>
<tr>
<td>$SV_{J2,0}$</td>
<td>0.2013</td>
<td>84.56</td>
<td>–</td>
<td>0.94</td>
<td>-0.0436</td>
<td>0.0954</td>
</tr>
<tr>
<td></td>
<td>(0.0455)</td>
<td>(8.18)</td>
<td>(2.36)</td>
<td>(0.0024)</td>
<td>(0.0038)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Point estimates and corresponding standard errors for parameters of different models. Panel A: point estimates and standard errors, in parentheses, for the diagonal components of the diffusion parameter matrices. Panel B: point estimates and standard errors, in parentheses, for the out-of-diagonal components of the diffusion parameter matrices. Panel C: point estimates and standard errors, in parentheses, for the parameters in the jump component.

<table>
<thead>
<tr>
<th>State space dimension</th>
<th>SV\textsubscript{2,0}</th>
<th>SV\textsubscript{3,0}</th>
<th>SV\textsubscript{3,1}</th>
<th>SVJ\textsubscript{2,0}</th>
<th>SVJ\textsubscript{3,1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>1.180</td>
<td>1.127</td>
<td>1.048</td>
<td>1.115</td>
<td>0.913</td>
</tr>
<tr>
<td>(\sigma_{\text{RMSE}})</td>
<td>(0.370)</td>
<td>(0.348)</td>
<td>(0.285)</td>
<td>(0.446)</td>
<td>(0.324)</td>
</tr>
<tr>
<td>% Within bid-ask spread</td>
<td>0.603</td>
<td>0.617</td>
<td>0.640</td>
<td>0.635</td>
<td>0.633</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>State space dimension</th>
<th>SV\textsubscript{2,0}</th>
<th>SV\textsubscript{3,0}</th>
<th>SV\textsubscript{3,1}</th>
<th>SVJ\textsubscript{2,0}</th>
<th>SVJ\textsubscript{3,1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>1.937</td>
<td>1.844</td>
<td>1.570</td>
<td>1.862</td>
<td>1.457</td>
</tr>
<tr>
<td>(\sigma_{\text{RMSE}})</td>
<td>(1.101)</td>
<td>(1.027)</td>
<td>(0.808)</td>
<td>(1.129)</td>
<td>(0.809)</td>
</tr>
<tr>
<td>RMSIVE</td>
<td>2.060</td>
<td>1.974</td>
<td>1.824</td>
<td>1.935</td>
<td>1.772</td>
</tr>
<tr>
<td>(\sigma_{\text{RMSIVE}})</td>
<td>(0.754)</td>
<td>(0.660)</td>
<td>(0.537)</td>
<td>(0.700)</td>
<td>(0.519)</td>
</tr>
<tr>
<td>% Within bid-ask spread</td>
<td>0.437</td>
<td>0.461</td>
<td>0.540</td>
<td>0.452</td>
<td>0.527</td>
</tr>
</tbody>
</table>

Table 5: Performance comparison. RMSE is the root-mean-squared dollar pricing error. \(\sigma_{\text{RMSE}}\) is the sample standard deviation of daily RMSE. RMSIVE is the root-mean-squared implied volatility error. \(\sigma_{\text{RMSIVE}}\) is the sample standard deviation of daily RMSIVE. The row “% Within bid-ask spread” reports the fraction of fitted prices within the bid-ask spread.
Table 6: RMSE and RMSE improvements over benchmark models stratified by maturity and moneyness. For diffusion and jump diffusion models, we present out-of-sample RMSE of benchmark models and percentage out-of-sample RMSE improvements of $SV_{3,1}$ and $SVJ_{3,1}$ models, stratified by moneyness and maturity in days. All performance computations are based on the “full” sample (1996-01/2009-09).

<table>
<thead>
<tr>
<th>Panel A1: RMSE for $SV_{2,0}$ model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau &lt; 20$</td>
</tr>
<tr>
<td>$S/K &lt; 0.80$</td>
</tr>
<tr>
<td>$0.80 &lt; S/K &lt; 0.90$</td>
</tr>
<tr>
<td>$0.90 &lt; S/K &lt; 1.00$</td>
</tr>
<tr>
<td>$1.00 &lt; S/K &lt; 1.10$</td>
</tr>
<tr>
<td>$1.10 &lt; S/K &lt; 1.20$</td>
</tr>
<tr>
<td>$1.20 &lt; S/K &lt; 1.30$</td>
</tr>
<tr>
<td>$1.30 &lt; S/K &lt; 1.40$</td>
</tr>
<tr>
<td>$S/K &gt; 1.40$</td>
</tr>
<tr>
<td>all</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel A2: RMSE improvement of model $SV_{3,1}$ over model $SV_{2,0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau &lt; 20$</td>
</tr>
<tr>
<td>$S/K &lt; 0.80$</td>
</tr>
<tr>
<td>$0.80 &lt; S/K &lt; 0.90$</td>
</tr>
<tr>
<td>$0.90 &lt; S/K &lt; 1.00$</td>
</tr>
<tr>
<td>$1.00 &lt; S/K &lt; 1.10$</td>
</tr>
<tr>
<td>$1.10 &lt; S/K &lt; 1.20$</td>
</tr>
<tr>
<td>$1.20 &lt; S/K &lt; 1.30$</td>
</tr>
<tr>
<td>$1.30 &lt; S/K &lt; 1.40$</td>
</tr>
<tr>
<td>$S/K &gt; 1.40$</td>
</tr>
<tr>
<td>all</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B1: RMSE of model $SV_{2,0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau &lt; 20$</td>
</tr>
<tr>
<td>$S/K &lt; 0.80$</td>
</tr>
<tr>
<td>$0.80 &lt; S/K &lt; 0.90$</td>
</tr>
<tr>
<td>$0.90 &lt; S/K &lt; 1.00$</td>
</tr>
<tr>
<td>$1.00 &lt; S/K &lt; 1.10$</td>
</tr>
<tr>
<td>$1.10 &lt; S/K &lt; 1.20$</td>
</tr>
<tr>
<td>$1.20 &lt; S/K &lt; 1.30$</td>
</tr>
<tr>
<td>$1.30 &lt; S/K &lt; 1.40$</td>
</tr>
<tr>
<td>$S/K &gt; 1.40$</td>
</tr>
<tr>
<td>all</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B2: RMSE improvement of model $SVJ_{3,1}$ over model $SVJ_{2,0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau &lt; 20$</td>
</tr>
<tr>
<td>$S/K &lt; 0.80$</td>
</tr>
<tr>
<td>$0.80 &lt; S/K &lt; 0.90$</td>
</tr>
<tr>
<td>$0.90 &lt; S/K &lt; 1.00$</td>
</tr>
<tr>
<td>$1.00 &lt; S/K &lt; 1.10$</td>
</tr>
<tr>
<td>$1.10 &lt; S/K &lt; 1.20$</td>
</tr>
<tr>
<td>$1.20 &lt; S/K &lt; 1.30$</td>
</tr>
<tr>
<td>$1.30 &lt; S/K &lt; 1.40$</td>
</tr>
<tr>
<td>$S/K &gt; 1.40$</td>
</tr>
<tr>
<td>all</td>
</tr>
</tbody>
</table>
References


