



## Working Paper

02/2010

Peter Gruber  
Claudio Tebaldi  
Fabio Trojani

**Three make a smile – dynamic  
volatility, skewness and term  
structure components in option  
valuation**



# **Three make a smile – dynamic volatility, skewness and term structure components in option valuation**

*by Peter Gruber  
Claudio Tebaldi  
Fabio Trojani*

*n. 02/10  
Milan, March 2010*

*Copyright  
Carefin, Università Bocconi*

*Further research materials can be downloaded at  
[www.carefin.unibocconi.eu](http://www.carefin.unibocconi.eu)*

# INDEX

<i>Abstract</i>	I
<b>1. INTRODUCTION</b>	1
<b>2. MODEL</b>	4
<b>2.1 Return dynamics, stochastic volatility, and jumps</b>	4
<b>2.2 State decomposition</b>	6
<b>2.3 Stochastic leverage</b>	7
<b>2.4 Stochastic total factor mean reversion</b>	8
<b>2.5 Long-term volatility</b>	9
<b>2.6 The Laplace transform of the asset returns</b>	9
<b>3. DATA</b>	10
<b>3.1 Characteristics and source</b>	10
<b>3.2 Explorative data analysis</b>	11
<b>4. ESTIMATION PROCEDURE</b>	13
<b>4.1 Identification</b>	13
<b>4.1.1 Parameter identification</b>	13
<b>4.1.2 State identification</b>	14
<b>4.2 Time-series estimation</b>	14
<b>5. RESULTS</b>	16
<b>5.1 Implied state</b>	16
<b>5.1.1 State interpretation</b>	18
<b>5.2 Scope and limits of the model</b>	21
<b>5.2.1 Nested models</b>	22
<b>5.3 More results</b>	23
<b>5.3.1 Model comparison</b>	24
<b>5.4 Dynamic properties of the model-implied volatility surface</b>	24
<b>5.5 Pricing performance</b>	25
<b>5.6 Risk-neutral parameters M, Q, R</b>	29
<b>6. CONCLUSION</b>	29

<b>A</b>	<b>NESTED MODELS</b>	30
<b>A.1</b>	<b>Diffusive models</b>	30
<b>A.2</b>	<b>Jump models</b>	30
<b>B</b>	<b>ALTERNATIVE MATRIX REPRESENTATIONS</b>	31
<b>B.1</b>	<b>Spectral decomposition</b>	31
<b>B.2</b>	<b><math>v - \zeta - \alpha</math> decomposition</b>	32
<b>B.3</b>	<b>General ellipse in eigenvalue coordinate representation</b>	33

## *Abstract*

*We propose a new modeling framework for the valuation of European options, in which dynamic short and long run volatility components drive the smile dynamics. The model state dynamics is driven by a matrix jump diffusion, provides efficient pricing formulas for plain vanilla options by means of standard transform methods, and it nests as special cases a number of affine option pricing models in the literature. In contrast to other approaches, short and long run volatility components interact dynamically with a further component linked to stochastic skewness, which we show is important in order to capture accurately the joint behavior of the implied volatility skew and the volatility term structure. We estimate our model and a number of competing benchmarks without interactions using S&P 500 index options. We find that models with dynamic interactions provide better pricing performance and a more accurate description of the joint dynamics of the implied volatility skew and term structure, both in-sample and out-of-sample. These findings support the use of option pricing models with (i) at least three dynamic volatility factors and (ii) dynamic interactions between volatility and stochastic skewness components.*

### *Keywords*

*Option Pricing, Stochastic Volatility, Short and Long Term Volatility Risk, Stochastic Leverage, Wishart Diffusion*

## 1 INTRODUCTION

<sup>1</sup> In this paper, we propose a new modeling framework for the valuation of European options, in which dynamic short and long run volatility components are allowed to drive the smile dynamics. In contrast to other modeling approaches, we model short and long run volatility components that interact dynamically with a further component linked to stochastic skewness. We find that this last component is useful to capture more accurately the joint feed-backs of implied volatility skew and term structure, which are not completely spanned by shocks to short and long-term volatility components alone.

The model state dynamics is driven by a matrix jump diffusion, provides efficient pricing formulas for plain vanilla options by means of standard transform methods, and it nests as special cases a number of two- and three-factor affine option pricing models in the literature. Using S&P 500 index options data, we empirically study the pricing performance of our model and a number of competing benchmarks without interactions. We find that our setting with dynamic interactions indeed provides better pricing performance and a more accurate description of the implied volatility skew and term structure joint dynamics, both in-sample and out-of-sample.

A large literature has shown that equity index volatility varies over time and negatively co-moves with index returns. It is also well understood that such asymmetric volatility behavior is directly linked to the negative skewness of index returns. Within modern stochastic volatility models, these properties naturally motivate well-known biases with respect to the Black-Scholes models, such as the steep implied volatility skew and the emergence of an implied volatility term structure. However, while many empirical studies using continuous-time and discrete time stochastic volatility models tend to explain in a qualitative way the pricing biases of Black-Scholes model, they often come short in producing a quantitatively accurate explanation of these deviations.

For instance, it is well known that Heston (1993) stochastic volatility model can generate an option implied volatility smirk and a negative skewness through a negative correlation between innovations in returns and volatility, the so-called leverage effect. Similarly, the mean reverting features of Heston's volatility process can also generate a simple form of an implied volatility term structure. However, the single factor features of the volatility in this model implied a strong link between implied volatility skew and term structure, which is not supported by the data. Empirical evidence shows that index implied volatility level, slope and term structure partly co-move, which explains why Heston's model can provide on average a pricing correction in the right direction relative to Black-Scholes option prices. In our sample of thirteen years of S&P 500 option data, the correlations are +0.74 for level/smirk, -0.77 for level/term structure and -0.52 for smirk/term structure. Fig. 1 depicts combinations of level, smirk and term structure obtained on monthly data. A principal component analysis shows that the first three factors explain 96.7, 2.0 and 0.9%

---

<sup>1</sup>Peter Gruber is at the University of Lugano. Claudio Tebaldi is at Bocconi University, Department of Finance is a Research Fellow of CAREFIN. Fabio Trojani is a Research Fellow of the Swiss Finance Institute. The authors gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK and grants 101312-103781/1 and 100012-105745/1) and of the CAREFIN center of Bocconi University. The usual disclaimer applies.

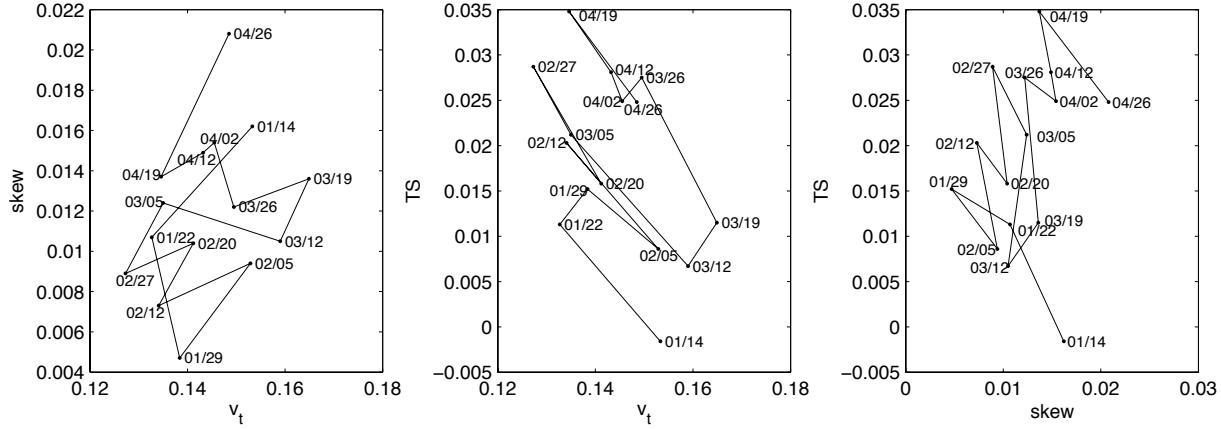


Figure 1: Combinations of level, skew and term structure for the first four months in 2004.

respectively of the variance in implied volatility. The eigenvectors of these components can again be interpreted as level, smirk and term structure factors<sup>2</sup>, see Fig. 2. These three factors are mirrored in the main option strategies employed by participants: straddles to speculate on changes in the level, risk reversals to speculate on changes in the smirk and calendar spreads to speculate on changes in the term structure. These strategies exist independently and cannot be used to hedge for one another.

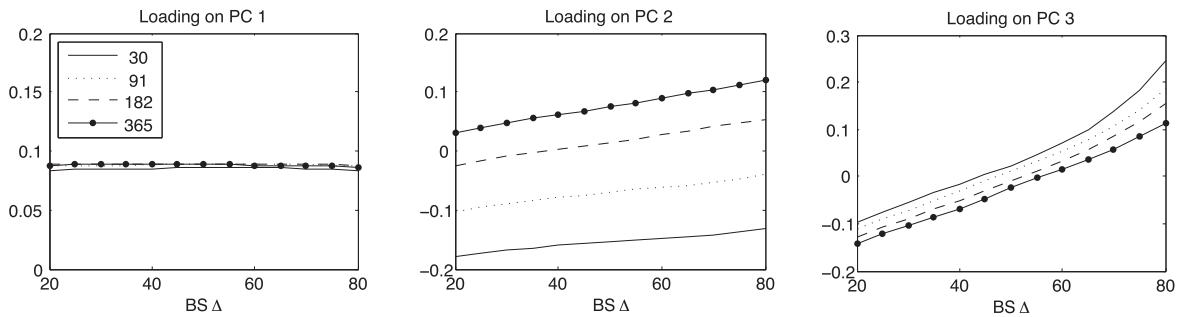
All this empirical evidence suggests that the implied volatility surface is more complex than captured by a one or two factor model like Heston (1993) or Bates (2000). In order to allow for realistic, stochastically varying smirk and term structure patterns, an option pricing model needs to incorporate a stochastic correlation between returns and volatility and a stochastic factor mean reversion. Furthermore, these two properties need to be modeled independently to allow for a certain degree of independence between skewness and term structure effects. While stochastic correlation is already obtained in two factor models, a third factor is needed to allow for an independent component in the smirk and term structure.

This paper studies the determinants of a realistic model for the volatility surface and its dynamics in the class of affine parametric models. We assess the theoretical and empirical value of adding a third factor and/or jumps in returns as well as the role of matrix-valued volatility factors. To do so, we use build on the framework of matrix affine jump diffusion processes recently proposed by Leippold and Trojani (2008). This framework offers full tractability and closed form solutions up to a Fourier inversion. It encompasses a list of interesting pure diffusive and jump-diffusive models such as Heston (1993), Bates (1996) and (2000), Christoffersen, Heston and Jacobs (2007) and DaFonseca, Grasselli and Tebaldi (2008), thus supplying a common tool for comparing the properties and the performance of seven different models.

As our main analysis tool we introduce a new volatility decomposition that makes it possible to compare the structural properties and boundaries of matrix models to models

<sup>2</sup>Similar results are obtained by Skiadopoulos, Hodges and Clewlow (2000) who identify two to four significant components, depending on the criterion applied. Fengler, Härdle and Villa (2001) interpret the first three components of the DAX implied volatility surface as level, skew (slope) and smile (curvature).

### Panel A: projection on strike



### Panel B: projection on time to maturity

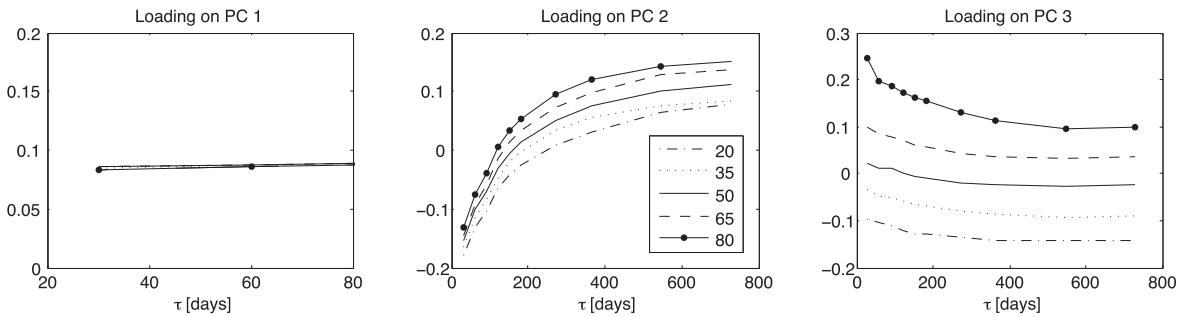


Figure 2: First four eigenvectors of the PCA of the implied volatility surface, shown for maturities of 30, 91, 182 and 365 days (panel A) and for Black-Scholes deltas of  $-80, -65, -50, -35$  and  $-20$  (panel B). PC1 can be interpreted as level, PC2 as maturity slope, PC3 as moneyness slope.

with multiple scalar factors. More specifically, we decompose the state of our benchmark  $2 \times 2$  matrix model into three components: a Heston-like component, a component that measures the distance from the Heston model and a rotation component that measures the “direction” of this deviation. We furthermore link these three state components to the observable volatility factors level, smirk and term structure and are thus able to identify the economic significance of each state component as well as to distinguish the roles of diffusive processes and jumps in creating skewness.

We show both theoretically and empirically that, in order to model independent skew and term structure components, matrix model with a rotation component is a better choice than the naive addition of a third Heston-like factor. While the three factor matrix model generally outperforms all other models, we find that this effect is largest when the direction of deviation points towards increasing both skewness and the term structure.

We also develop an identification strategy by rigorously separating time-independent parameters from the time-varying state. We solve the problem of jointly estimating parameters and state using a computationally intensive nested optimization. The flexibility of the matrix state space allows us to estimate it on five years of S&P 500 option data with a stable set of parameters, with a mean absolute pricing error of \$0.77 while almost perfectly reproducing the unconditional correlations between level, smirk and term structure.

The remainder of this paper is organized as follows: Section 2 presents our framework and discusses key properties of some model specifications. Section 3 describes the data used and provides some empirical analysis to pinpoint its regularities. Section 4 discusses our identification strategy. Section 5 presents the results of the time series estimation and discussed the results.

## 2 MODEL

### 2.1 Return dynamics, stochastic volatility, and jumps

The standard approach to model return skewness in the continuous-time option pricing literature assumes a correlation between shocks to returns and shocks to volatility while a volatility term structure is modelled using multiple factors with different mean reversion speeds. As it includes both features as well as jumps in returns, the model of Bates (2000) is a good starting point for our framework. The returns process is assumed as

$$\frac{dS}{S} = rdt + \sqrt{v_1}dz_1 + \sqrt{v_2}dz_2 + dL_t \quad (2.1)$$

with the two stochastic volatility processes

$$dv_i = (\alpha_i - \beta_i v_i) dt + \sigma_i \sqrt{v_i} dz_{vi} \quad (2.2)$$

All Brownian motions  $dz_i$  and  $dz_{vi}$  are independent except for  $\langle dz_i, dz_{vi} \rangle = \rho_i$ . The jump process  $dL$  has a poisson distributed arrival rate with intensity  $\lambda = \lambda_0 + \lambda_1 v_1 + \lambda_2 v_2$  and a lognormally distributed jump size  $k$  with  $\ln(1 + k) \sim N(\ln(1 + \bar{k}) - \frac{\delta^2}{2}, \delta^2)$ .

This model can produce stochastic leverage and a stochastic term structure, however, there is a strong link between the two, as will be shown later. This is a severe limitation as there is no economic reason for a rigid relation between short-term out of the money puts and long-term options. Two extensions can be considered in order to break the link. First, the addition of a third volatility factor. This is done by simply adding the term  $\sqrt{v_3}dz_3$  to (2.1) with an according third volatility process in (2.2).

Alternatively, the third factor can be introduced as a stochastic correlation between the two volatility factors. Gourieroux (2006) has introduced a matrix-based process that can serve as elegant way to produce correlated factors. To see the link to the Bates model, we rewrite the two volatility processes in matrix form without changing the returns process:

$$\begin{pmatrix} dv_1 & 0 \\ 0 & dv_2 \end{pmatrix} = \left[ \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} - \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \sqrt{v_1} & 0 \\ 0 & \sqrt{v_2} \end{pmatrix} \begin{pmatrix} dz_{v1} & 0 \\ 0 & dz_{v2} \end{pmatrix} \quad (2.3)$$

From here it is only a matter of introducing a few new symbols to generalize this diagonal model to a full matrix model.

$$X_t = \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix}, \quad \Omega\Omega' = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}, \quad M = -\frac{1}{2} \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix}, \quad Q = \frac{1}{2} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \quad (2.4)$$

Now we can express the two volatility factors in the Bates model in matrix form:

$$dX_t = (\Omega\Omega' + MX'_t + X_t M') dt + \sqrt{X_t} dWQ + Q'(dW') \sqrt{X_t} + dJ_t \quad (2.5)$$

where the apex ' denotes the transpose of a matrix,  $dW$  is a matrix Brownian motion and  $dJ_t$  possible jumps in volatility, which are not discussed here. A change to the returns process (2.1) is not needed, if we assume the jump intensity to be  $\lambda = \lambda_0 + Tr[\Lambda X_t]$  and write  $v_1$  and  $v_2$  for the eigenvalues of  $X_t$ , a notation that will be kept throughout this paper.

The volatility process (2.10) is the matrix jump diffusion process presented in Leippold and Trojani (2008). By construction,  $X_t$  is a positive definite matrix. In the pure diffusion setting, the process is called Wishart process, if  $M$  is negative definite and the condition  $\Omega\Omega' = kQ'Q$ , with  $k > n - 1$  is met (Bru (1991) and Grasselli and Tebaldi (2008)). To ensure nesting of the models of Christoffersen et al (2007) and Bates (2000), we generalize the process using  $\Omega\Omega' = Q'QK$ , with  $K$  an upper triangular matrix.

**Nested models.** The matrix affine jump diffusion notation may be uncommon in the option pricing literature, but apart from allowing for correlated volatility factors it offers the advantage of providing a common framework to study models with multiple factors, factor correlations and jumps. In analogy to (ref), we propose the notation  $SV(J)_{n,r}$  for option pricing models with  $n$  denoting the total number of factors,  $r$  denoting the number of factor correlations and the letter  $J$  denoting jumps in returns. The table below presents a few interesting jump-diffusive models in this classification<sup>3</sup>.

<sup>3</sup>The structure of the matrix-valued state allows only for (a)  $n = i$  and  $r = 0$ , (b)  $n = \frac{1}{2}j(j+1)$  and  $r = \frac{1}{2}j(j-1)$  or (c) any combination of (a) and (b); with  $i$  and  $j$  positive integers.

$n$	$r$	Pure diffusion	Jump-diffusion
1	0	$SV_{1,0}$ Heston (1993)	$SVJ_{1,0}$ Bates (1996)
2	0	$SV_{2,0}$ Christoffersen et al (2007)	$SVJ_{2,0}$ Bates (2000)
3	0	$SV_{3,0}$ this paper	$SVJ_{3,0}$ (not considered)
3	1	$SV_{3,1}$ da Fonseca et al (2008)	$SVJ_{3,1}$ this paper

$n$  = total number of factors,  $r$  number of factor correlations

Appendix A lists in detail how the parameters and the state of the nested models can be expressed in the form of our framework (2.10).

## 2.2 State decomposition

The state  $X_t$  in (2.10) is usually expressed as a standard cartesian matrix. This representation highlights the affine structure of the model and is useful for most algebraic manipulations. It does not, however, lend hand to any economic interpretation and makes comparison to the nested models difficult. The out-of-diagonal factor  $X_{12}$  is especially difficult to handle, while being important for the model. Its admissible range, for example, is  $-\sqrt{X_{11}X_{22}} \leq X_{12} \leq \sqrt{X_{11}X_{22}}$ , which conceals its true nature as a correlation factor.

The problem of rendering  $X_t$  interpretable is linked to identifying a representation that separates volatility structure from volatility level, i.e. that is homogenous in  $v_t$ . We propose a new representation that unveils the structure of the state and renders possible its interpretation in economic terms.

**Proposition 2.1 ( $v - \xi - \alpha$  decomposition)** *The positive-definite state matrix  $X_t$  can be decomposed in a volatility part and a structural part. In the  $2 \times 2$  case, this decomposition is given as*

$$X_t = \begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix} = \frac{v_t}{2} \left[ Id + (2\xi_t - 1) \begin{pmatrix} \cos(2\alpha_t) & \sin(2\alpha_t) \\ \sin(2\alpha_t) & -\cos(2\alpha_t) \end{pmatrix} \right] \quad (2.6)$$

with  $v_t = Tr[X_t] = v_{1t} + v_{2t}$  being the total variance,  $\xi = \frac{v_{1t}}{v_{1t} + v_{2t}}$  being the fraction of total variance explained by the first volatility factor and  $\alpha$  being a polar coordinate representation of the matrix' eigenvectors. Proof: see Appendix B.

This decomposition has a few useful properties. First, it separates volatility level from structure. Second, it nests the one and two factor models listed above. By setting  $\alpha = 0$ , we recover the two-factor  $SV(J)_{2,0}$  model. By additionally letting  $\xi = 1$ , we recover  $SV(J)_{1,0}$ .

Notice that the expressions  $(2\xi - 1)$  and  $\cos(2\alpha)$  are bound between  $-1$  and  $1$ . It will be later shown that this property can be linked to correlation processes and portfolio weights. It is therefore equivalent to denote the state space as  $\mathcal{S}_n^+$  (the space of  $n$ -dimensional positive definite matrices) or  $\mathbb{R}_+^m \times [-1, 1]^{m(m-1)/2}$ .

## 2.3 Stochastic leverage

In order to discuss the leverage effect, we need to rewrite the returns process as

$$\frac{dS}{S} = rdt + Tr[\sqrt{X_t}dB_t] + dL \quad (2.7)$$

where  $Tr[\cdot]$  denotes the trace operator. This expression still nests (2.1), but allows for a matrix-values correlation between shocks in returns and volatility:  $\langle Z_t, B_t \rangle = R$ .

We anticipate here the point estimates of the parameters of the  $SV_{3,1}$  matrix model, as they serve us to illustrate important features of our framework:

$$M = \begin{pmatrix} -0.34 & 0 \\ -9.36 & -5.07 \end{pmatrix} \quad Q = \begin{pmatrix} 0.0340 & 0 \\ 0 & 0.48 \end{pmatrix} \quad R = \begin{pmatrix} -0.98 & 0.16 \\ 0 & -0.69 \end{pmatrix} \quad \beta = 1.00$$

As can be easily shown, the correlation between innovations in returns and volatility is

$$\text{corr}\left(\frac{dS}{S}, dV\right) = \frac{Tr[RQX]}{\sqrt{Tr[X]}\sqrt{Tr[Q'QX]}} \quad (2.8)$$

We can now use decomposition (2.6) to analyze this correlation structure

$$\text{corr}\left(\frac{dS}{S}, dV\right) = \frac{Tr[RQ] + (2\xi - 1)\left[\cos(2\alpha)(RQ_{11} - RQ_{22}) + \sin(2\alpha)(RQ_{12} + RQ_{21})\right]}{\sqrt{2}\sqrt{Tr[QQ] + (2\xi - 1)\left[\cos(2\alpha)(QQ_{11} - QQ_{22}) + \sin(2\alpha)(QQ_{12} + QQ_{21})\right]}} \quad (2.9)$$

We can learn several things from (2.9). Since  $v_t$  does not appear in (2.9), it is readily confirmed that the leverage effect is independent of the volatility level. Second, the correlation structures of the nested models can be easily produced by setting  $\alpha = 0$  for a two factor model or  $\alpha = 0$  and  $\xi = 1$  for a one factor model.

In the  $SV_{3,1}$  model, correlations fluctuate around a central value of  $Tr[RQ]/\sqrt{2Tr[QQ]}$ , which is  $-0.455$  at our parameter estimates. Large deviations from this value obtain if  $|2\xi - 1| \rightarrow 1$ , i.e. for  $v_1 \ll v_2$  or vice-versa. Even though the matrix model is quite flexible, it has to be noted that  $\text{Corr}(dS/S, dV)$  cannot take any value on  $[-1, 1]$ . Using our parameter estimates, we solve numerically for the minimal and maximal possible values and obtain  $-0.86$  and  $+0.36$ . This compares favorably to the  $-0.713$  obtained from the one factor diffusive model and the range of  $[-0.785, -0.602]$  for two factor case.

To facilitate the interpretation of (2.9), we plot a decomposition of the model-implied correlations for the  $SV(J)_{3,1}$  model in Fig. 10. The dashed lines depict border values of the model given any state. The red and blue lines depict the maximal and minimal possible correlations given the two volatility factors, i.e. given  $\xi = \frac{v_{1t}}{v_{1t} + v_{2t}}$ . The actual model-implied correlation depends on the correlation factor  $\alpha$  and is depicted by a black line. Notice that the range between the red and blue line is very wide on most of the days, indicating a wide range of possible leverage effects. Fig. 3 presents an alternative representation of (2.9). Note that a wide range of values for the correlation obtains only for  $\xi \ll 0.5$ .

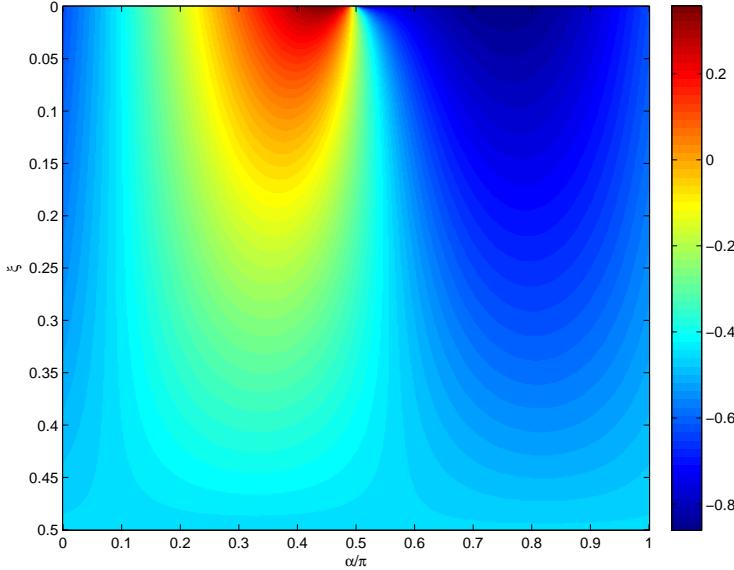


Figure 3: Illustration of the possible values for  $\text{corr}(dS/S, dV)$  in 2.9 as a function of  $\alpha$  and  $\xi = \frac{\lambda_1}{\lambda_1 + \lambda_2}$ .

## 2.4 Stochastic total factor mean reversion

The speed of mean reversion of the volatility factor is important for the volatility term structure. Even though we consider mostly models with multiple volatility factors, we are still able to provide an expression for the spot total factor mean reversion for the total variance  $v_t = v_{1t} + v_{2t} = \text{Tr}[X_t]$ . To do so, we take the trace of the bounded variation part of (2.10) and use the decomposition (B.8) :

$$\begin{aligned} dv_t &= \text{Tr} [\Omega \Omega' + M X'_t + X_t M'] dt + \dots \\ &= [\text{Tr}[\Omega \Omega'] - m_t v_t] dt + \dots \end{aligned} \tag{2.10}$$

$$\text{with} \tag{2.11}$$

$$m_t = - \{ \text{Tr}[M] + (2\xi_t - 1) (\cos(2\alpha_t)(M_{11} - M_{22}) + \sin(2\alpha_t)(M_{12} + M_{21})) \} / 12$$

The expression  $1/m_t$  is the spot mean reversion speed of  $v_t$ <sup>4</sup>. It is stochastic for all but the one factor models. In the other cases, it varies around the central value  $\text{Tr}[M]$ .

More interesting, a specific mean reversion speed can be obtained by an infinite combination of  $\xi$  and  $\alpha$  in the  $SV(J)_{3,1}$  model. In this model the leverage effect and the mean reversion form a system of two questions in  $\xi$  and  $\alpha$ , realizing an independent modelling of the two quantities.

---

<sup>4</sup>We qualify it as a “spot mean reversion”, because  $m_s | X_t, (s > t)$  is a mean-reverting process itself.

## 2.5 Long-term volatility

The infinite-maturity state  $X_\infty$  can be calculated as

$$-\Omega\Omega' = MX_\infty + X_\infty M' \quad (2.13)$$

We solve this linear Ljapunov equation and obtain

$$X_\infty = \begin{pmatrix} 0.0003 & 0.0006 \\ 0.0006 & 0.0238 \end{pmatrix},$$

which implies a long-term average volatility level of  $\text{Tr}[\sqrt{X_\infty}] = 0.1725$  in the  $SV(J)_{3,1}$  model. This compares not perfectly well with the mean one-year ATM implied volatility of 0.2027 in our data, however, we are still match the main feature of a downward sloping average volatility surface, given the short term implied volatility level (0.2071 data-implied and 0.2042 model-implied).

## 2.6 The Laplace transform of the asset returns

Following Duffie, Pan and Singleton (2000), Trojani and Leippold (2007) Da Fonseca et al. (2007) in order to solve the pricing problem for plain vanilla options we just need the Laplace transform of  $Y_t = \ln(S_t)$ . Since the Laplace transform of Wishart processes is exponentially affine (see e.g. Bru 1991), we guess that the conditional moment generating function of the asset returns is the exponential of an affine combination of  $Y_t$  and of the elements of the Wishart matrix  $X_t$ . In other terms, we look for three deterministic functions  $A(t) \in M_n, B(t) \in \mathbb{R}, C(t) \in \mathbb{R}$  that parametrize the Laplace transform:

$$\begin{aligned} \Psi_{\gamma,t}(\tau) &= \mathbb{E}_t [\exp \{\Gamma Y_{t+\tau}\}] \\ &= \exp \{Tr [A(\tau)X_t] + B(\tau)Y_t + C(\tau)\}, \end{aligned} \quad (2.14)$$

where  $\mathbb{E}_t$  denotes the conditional expected value with respect to the risk-neutral measure and  $\gamma \in \mathbb{R}$ . By applying the Feynman-Kac argument, we have

$$\begin{aligned} \frac{\partial \Psi_{\gamma,t}}{\partial \tau} &= \mathcal{L}_{Y,X} \Psi_{\gamma,t} \\ \Psi_{\gamma,t}(0) &= \exp \{\Gamma Y_t\}, \end{aligned} \quad (2.15)$$

where  $\mathcal{L}_{Y,X}$  is the infinitesimal generator for the couple  $(Y_t, X_t)$  and by replacing the candidate (??) we obtain that the coefficients  $A(\tau), B(\tau), C(\tau)$  have to verify the following

set of ordinary differential equations:

$$\begin{aligned}\frac{d}{d\tau}A(\tau) &= A(\tau)(M + \gamma Q^T R^T) + (M^T + \gamma RQ)A(\tau) + 2A(\tau)Q^T Q A(\tau) + \frac{\gamma(\gamma - 1)}{2}\mathbb{I}_n \\ &\quad + \Lambda[\Theta^{X,Y}(\gamma, A(\tau)) - 1]\end{aligned}\tag{2.16}$$

$$\frac{d}{d\tau}B(\tau) = 0,\tag{2.17}$$

$$\frac{d}{d\tau}C(\tau) = Tr[\Omega\Omega^TA(\tau)] + \lambda_0[\Theta^{X,Y}(\gamma, A(\tau)) - 1] + \Gamma r,\tag{2.18}$$

with boundary conditions

$$\begin{aligned}A(0) &= 0 \in M_n, \\ B(0) &= \gamma \in \mathbb{R}, \\ C(0) &= 0.\end{aligned}$$

where  $\Theta^{X,Y}(\gamma, .)$  is the Laplace transform of the jump size. The solution of these equations determines the conditional characteristic function.

**Option Pricing using Cosine-FFT** Fourier inversion of the Laplace transform is computationally quite challenging, even in the pure diffusive setting. We obtain prices by applying the Cosine-FFT method of Fang and Oosterlee (2008).

## 3 DATA

### 3.1 Characteristics and source

We use call options on the S&P 500 index traded at the CBOE (ticker symbol SPX). Thanks to their liquidity, their European style and the absence of institutional regularities they have become a de-facto standard for testing new option pricing models. We obtained a sample of almost 13 years of end-of-day prices (Jan 1996 to Sept 2008) with maturities up to one year from Optionmetrics.

We apply the cleaning procedures outlined in Bakshi, Cao and Chen (1997) and eliminate all options with midquotes below \$0.375, with zero bid price and with a bid price larger than the ask. We further eliminate stale quotes (i.e. bid or ask identical to the previous trading day), prices that violate arbitrage bounds, duplicate entries and prices where the bid-ask spread is smaller than the minimum tick size (i.e. five cents for prices below \$3 and ten cents else). As in Bakshi et al., we also drop options with a time to maturity of less than 10 days. We disregard trading volume, being aware that our data may occasionally be based on quotes or exchange models. Apart from this data cleaning, we perform no cuts in moneyness in order to ensure a data set that is as rich and as challenging as possible.

Sample	monthly	daily	extended
Trading days in sample	59	1256	3205
Total call option prices	21'993	189'285	546'971
Price observations per day	160.5	150.7	170.6
Maturities per day	6 or 7	5 – 7	5 – 7
Unique maturities	24	255	255
Average time to maturity	130 days	145 days	145
Average moneyness ( $S/K$ )	1.06	1.05	1.05
Average strike price	\$107.05	\$104.81	\$ 104

Table 1: Summary statistics of the data.

On average, we retain 171 observations per trading day. For the estimation, we split the data into three sub-samples for computational reasons and to allow for true out-of-sample evaluation, see section 4.2. For a summary statistics of the data, see Tab. 1.

Interest rate are taken as observed and are obtained by linearly interpolating the US yield curve from Optionmetrics. Implicit dividend yields are calculated from a put-call parity regression of at the money ( $0.9 \leq K/S_0 \leq 1.1$ ) options, separately for each trading day and each maturity.

### 3.2 Explorative data analysis

To substantiate our hypothesis of multiple factors driving the implied volatility surface, we perform a principal component analysis. While this has already been done in Skiadopoulos et al. (2000) and Fengler, Härdle and Mammen (2005), our focus is different, namely on higher components (up to PC4), on the maturity dimension and on the dynamic structure. We follow the argument in Christoffersen et al. (2007) that our state variable  $TrX_t$  is a variance rather than a volatility and use option-implied variances (i.e. the square of the usual implied volatility) from a range of synthetic (i.e. linearly interpolated in implied volatility space) calls with time to maturity  $\tau = \{1, 2, 3, 4, 5, 6, 9, 12\}$  months and a Black-Scholes  $\Delta = \{0.2, 0.25 \dots 0.75, 0.8\}$ , for a total of 104 data points per trading day.

Using the mean eigenvalue criterion, we find two components in a pooled analysis. We further divide our sample into a high volatility and low volatility period, split at the median value of the short term, at the money maturity. Interestingly, we find three components in each of the subsets.

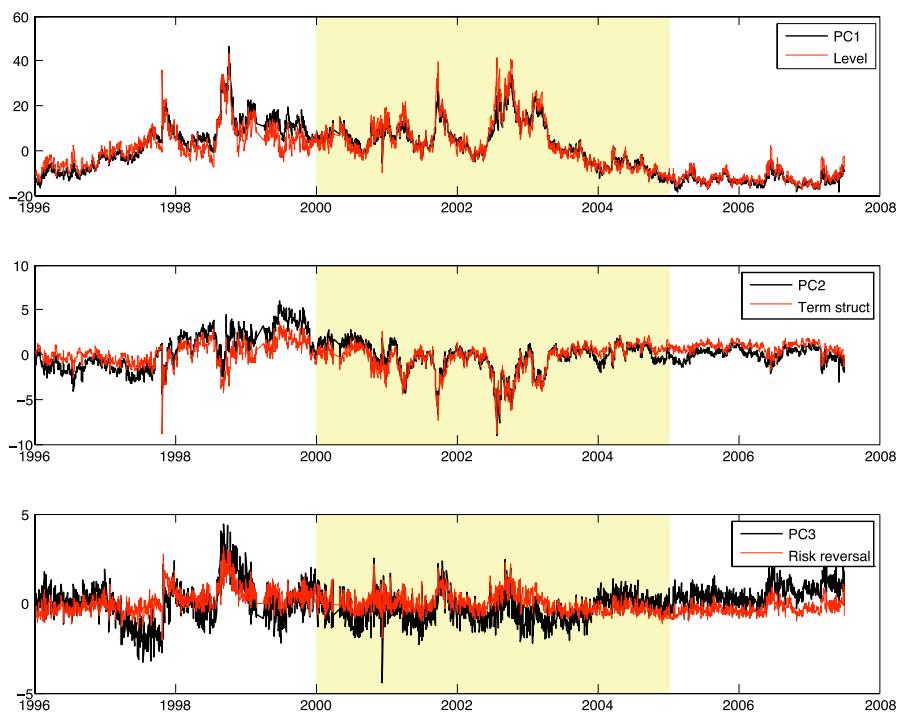


Figure 4: First three principal components (black) and corresponding portfolios of implied volatilities (red).

## 4 ESTIMATION PROCEDURE

### 4.1 Identification

#### 4.1.1 Parameter identification

Every diffusion process is uniquely characterized by its infinitesimal generator. The infinitesimal generator of the joint process of stock returns and volatility is:

$$\mathcal{L}_{Y,X} = \left( r - \frac{1}{2} \text{Tr}X \right) \frac{\partial}{\partial Y} + \frac{1}{2} \text{Tr}X \frac{\partial^2}{\partial Y^2} + \mathcal{L}_X + 2 \text{Tr}[XRQD] \frac{\partial}{\partial Y} \quad (4.1)$$

with the matrix differential operator  $(D)_{ij} = \frac{\partial}{\partial X_{ij}}$  and the infinitesimal generator of the Wishart process:

$$\mathcal{L}_X = \text{Tr} [(kQ'Q + MX + XM') D + 2XDQ'QD] \quad (4.2)$$

Parameter identification requires that the infinitesimal generator be unique for each set of parameters  $k, M, R, Q$  given any state  $X$ . Maximal identification aims at achieving this goal through the minimal set of restrictions on the parameters. Equations (4.1) and (4.2) feature two ambiguities that have to be resolved.

Firstly,  $Q$  and  $R$  appear only in the expressions  $Q'Q$  and  $RQ$ , requiring a choice for their signs and a canonical definition for  $Q$ . We choose  $Q$  to be the unique Choleski decomposition of  $Q'Q$ , i.e. upper triangular and positive definite.

Secondly, (4.1) and (4.2) contain a total of six expressions containing the trace. The trace of a matrix is invariant under transformations with an invertible matrix  $V$ , as  $\text{Tr}[X] = \text{Tr}[XVV^{-1}] = \text{Tr}[V^{-1}XV]$ . For a product of matrices, every element has to be transformed by the same matrix:  $\text{Tr}[XY] = \text{Tr}[V^{-1}XVV^{-1}YV]$ . This reduces the number of admissible transformations to one common transformation, as  $X, M, R, Q$  appear in pairwise products in the trace. The expression  $\text{Tr}[XDQ'QD]$  in (4.2) finally adds the requirement that  $V' = V^{-1}$ , i.e.  $V$  is a rotation matrix.<sup>5</sup> The matrix  $V$  can only be chosen implicitly via a parameter restriction. Among the several possibilities to do so, we choose to set  $M$  lower triangular.<sup>6</sup>

---

<sup>5</sup>One can make a similar argument by analyzing the dynamics of  $X$  in (eq). Again, only the products  $Q'Q$  and  $RQ$  are identified and the trace of  $X$  allows for one (common) rotation. Yet alternatively, the argument can be based on the Laplace transform (eq). Only the products  $Q'Q$  and  $RQ$  appear in the expression for the Laplace transform  $\Psi(\cdot) = \exp\{\text{Tr}[A(\tau)X] + bY + c\}$ , which allows for a common rotation of  $X$  and  $A(\tau)$ . A simple calculation reveals that a rotation of  $A(\tau)$  is equivalent to a common rotation of  $M, R$  and  $Q$ .

<sup>6</sup>When interpreting the time series of the implied state  $X_t$ , one has to keep in mind that it had been rotated by the arbitrary matrix  $V$ . While this does not change the eigenvalues  $\lambda_i$ , the rotation factor  $\alpha_t$  is shifted.

The maximal identified model has 11 parameters:

$$\theta = \{m_{11}, m_{21}, m_{22}, q_{11}, q_{12}, q_{22}, r_{11}, r_{12}, r_{21}, r_{22}, k\}.$$

Some parameters are restricted in sign or size. Identification requires  $q_{11} > 0$  and  $q_{22} > 0$ . To ensure that the process is non-explosive,  $M$  needs to be negative definite (i.e.  $m_{11} < 0$  and  $m_{22} < 0$ ) and  $k > 1$ . Also, the eigenvalues of  $R'R$  have to be lower than 1 in absolute value to ensure the existence of  $\sqrt{1 - R'R}$  in (eq).

To identify the parameters, at least eleven option prices in excess of the three prices needed to infer on the state are required. In practise, these requirements are easily exceeded.

#### 4.1.2 State identification

All option prices  $\widehat{C}(\cdot)_t$  of one trading day are a function of the parameters  $\theta$  and the day's state  $X_t$ . Once the parameters are identified, the pricing function

$$\widehat{C} = \widehat{C}(\cdot; \theta, X)$$

can be inverted for the three distinct components of the state  $X_t$  using at least three option prices. To force the implied state to be symmetric and positive definite, we represent it by its spectral representation  $\{\lambda_1 > 0, \lambda_2 > 0, \alpha_t\}$ , see appendix B.1.

**Ad hoc state restrictions.** There is only one consistent sub-model with a restricted state space: the case with  $X_t$  as well as  $M, R, Q$  diagonal. This model is similar to a two-factor Heston model and reported as “diagonal” in the results table. Two other interesting state restrictions are  $\cos \alpha_t = 1$  and  $\lambda_1 = 0$ . Both of them are inconsistent with the state dynamics, therefore we choose not to discuss them in depth. Our estimation results, however, reveal that the first model behaves similar to a two-factor Heston model, while the latter outperforms all other two factor models. We take this as a hint for the importance of the rotation factor  $\alpha_t$ .

## 4.2 Time-series estimation

Our estimation goal is to obtain a constant parameter set and daily estimates of the state. We refrain from re-estimating our model in regular intervals, e.g. yearly, because any change in parameters would open arbitrage opportunities and violate the underlying assumptions of every latent factor model.

Estimating a latent factor model poses the problem of the joint estimation of parameters and state. Usually, this is done by iterating between parameter estimation keeping the state

fixed and state estimation keeping the parameters fixed, see for example Christoffersen et al. (2007). A flexible three factor model requires a new estimation strategy because the addition of the rotation factor  $\alpha$  makes the estimation of parameters and state much more intertwined and there is no guarantee that the iterative approach finds the global minimum.

We estimate the parameters in two steps. First we perform a simple least squares estimation where we minimize the sum of quadratic pricing errors (implied volatility errors). For every candidate parameter vector  $\theta$  we calculate the time series of optimal state variables  $\{X_t^*(\theta)\}$ . We evaluate the pricing formula for a given  $\theta$  at the optimal state  $\{X_t^*(\theta)\}$  and minimize over  $\theta$ .

$$\text{Step 1: } \hat{\theta}_0 = \arg \min_{\theta} \left( \hat{C}_i(\theta, X_t^*(\theta)) - C_i \right)^2 \quad (4.3)$$

$$\text{with } X_t^*(\theta) = \arg \min_{\{X_t\}} \left( \hat{C}_i(\theta, X_t) - C_i \right)^2 \quad (4.4)$$

We use the result of this estimation as starting values for a maximum likelihood estimation. To correct for heteroskedasticity, we use a weighting matrix that is derived from the residuals of the first step. This procedure is similar to the method presented in Bates (2000), except that we do not correct for autocorrelated errors.

$$\text{Step 2: } \hat{\theta} = \arg \max_{\theta} -\frac{1}{2} \sum_t \ln |\Omega_t| + \mathbf{e}'_t \Omega_t^{-1} \mathbf{e}_t \quad (4.5)$$

with the pricing errors  $\mathbf{e}_t = \hat{C}_i(\theta, X_t^*(\theta)) - C_i$ , the conditional variance-covariance matrix  $\Omega_t$  and the conditionally implied state (4.4).

We choose the five years from 2000 to 2004 as the core sample to estimate our parameters, leaving the years before and after for out-of-sample analysis. We use monthly data within this core sample for parameter estimation to reduce the computational burden. Every month, we select the day on which the shortest time to maturity is ten days, yielding a sample of 59 Wednesdays<sup>7</sup>.

We then invert for the state by performing (4.4) independently for every trading day in the sample, i.e. we do not impose the dynamics (eq) on the implied state. This would not matter if the model were perfectly specified, but in a real world application this introduces a minor inconsistency. The estimate of the time series of the latent state is then

$$\{\hat{X}_t\} = \{X_t^*(\hat{\theta})\} \quad (4.6)$$

The result of our estimation are the 11 values for  $\theta$  and  $3 \times 3206 = 9618$  values for  $\{X_t\}$ . With these, we price 546'971 options in our sample.

**Estimation of reference models.** We use exactly the same data and estimation strategy for the reference models, i.e. the two-step nested optimization and the monthly data for the parameter estimation.

---

<sup>7</sup>There is no observation for September 2001 as US exchanges were closed from Sept 11 to Sept 16, 2001.

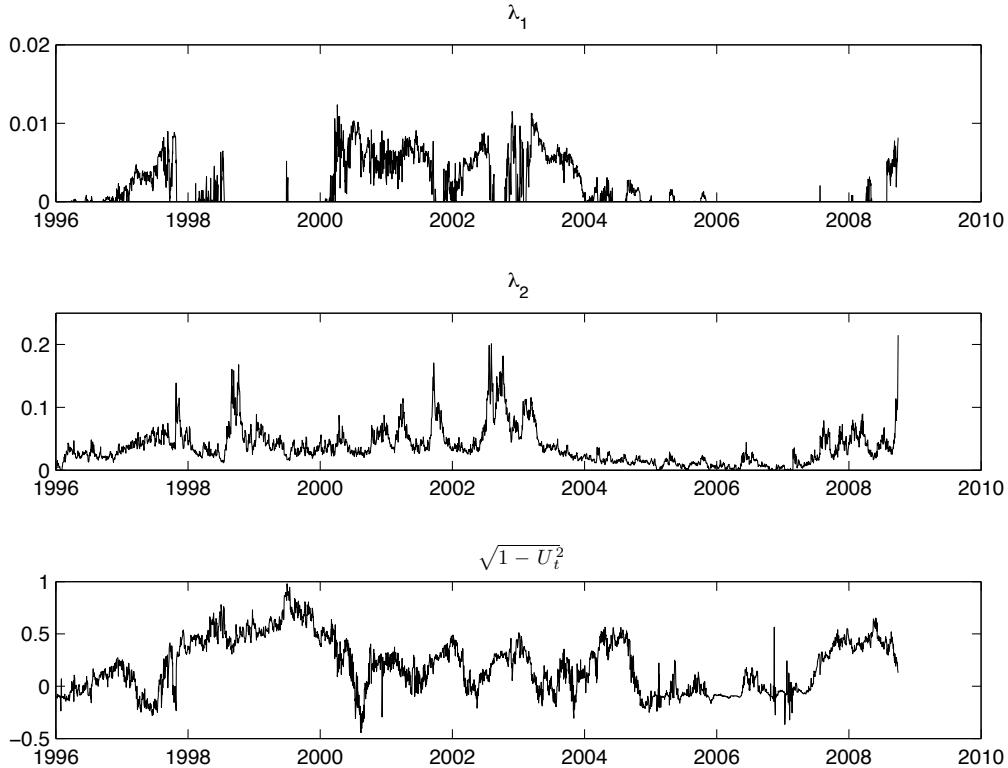


Figure 5: The daily state in eigenvalue representation.

## 5 RESULTS

### 5.1 Implied state

We obtain the implied state  $(v_1, v_2, \alpha)$  by inverting (4.4), separately for each of the 3206 trading days and report it in Fig. 5. Two things immediately strike the eye. First, the factor  $\lambda_2$  is much larger than  $\lambda_1$  and for some periods,  $\lambda_1$  is zero, making our model switch between one and two volatility factors. This is in contrast to existing two-factor models, in which both volatility factors are of the same order of magnitude. Second, the rotation factor  $\cos \alpha$  has a rich structure and changes sign.

As our model is equivalent to a Wishart option pricing model (see appendix), we also report the three unique elements of the state matrix under the cartesian Wishart specification in Fig. 6.

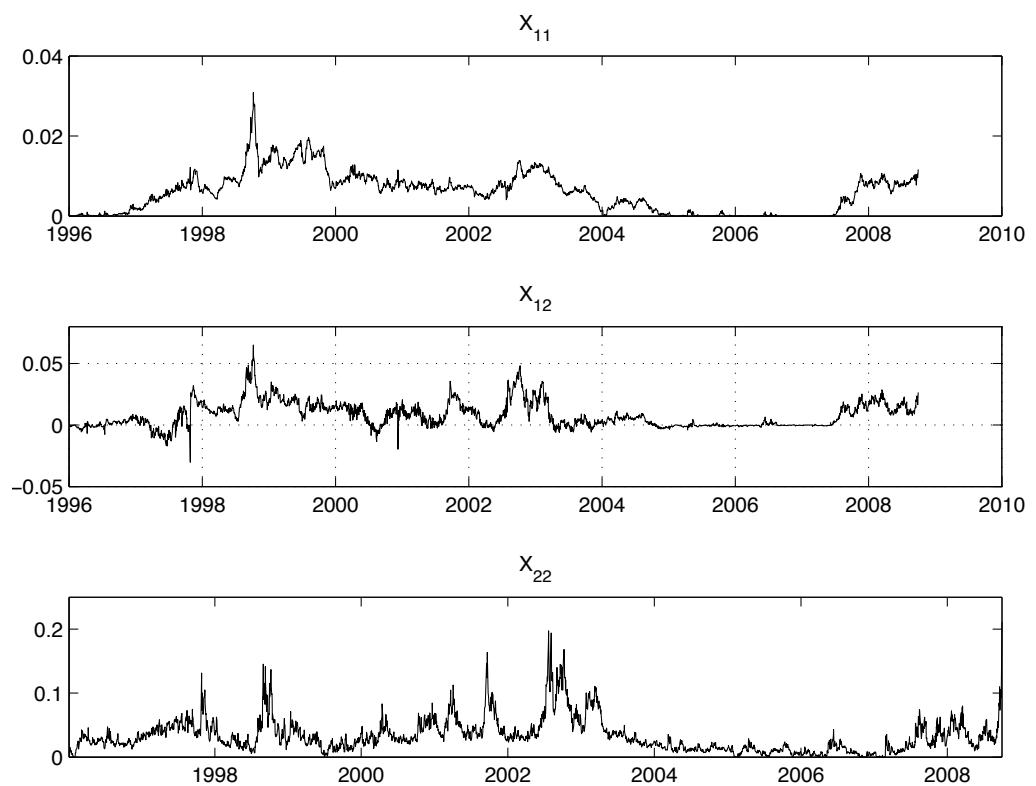


Figure 6: Implied state for the time series estimation in cartesian coordinates.

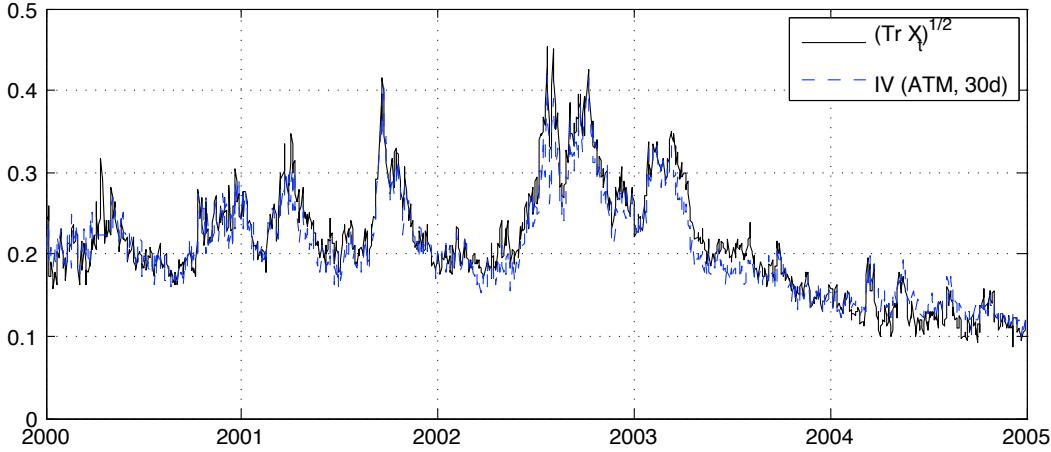


Figure 7: The level factor  $\sqrt{\text{Tr}[X_t]}$  is equal to the spot ATM implied volatility. When comparing to the observed 30-day, ATM implied volatility, the two curves are almost identical.

### 5.1.1 State interpretation

Even though the model is a latent state model, it would be desirable to find a link between the components of the state and the observable factors (PCs). While there is no direct interpretation for the state's components in our model (nor for in the cartesian formulation), we can give clear interpretations for three simple transformations of these elements and relate them to the principal components identified in 3.2. For this task, we use the expansion of Durrleman (2004) and Durrleman and Karoui (2007), which had been initially developed to approximate the implied volatility surface numerically.

**Level.** The expression  $\sqrt{v_t} = \sqrt{v_{1t} + v_{2t}}$  can be interpreted as a level factor. This can be shown by evaluating eq. (1a) in Durrleman and Karoui (2007) or by taking the limit  $\tau \rightarrow 0$  of the Laplace transform (eq), which becomes the Black-Scholes Laplace transform. The identification of a level factor is not specific to our model, in fact, all benchmark models can be interpreted in the way that the sum of the volatility factors is the short-term implied variance. Fig. 7 illustrates this relationship. Minor deviations are explained (a) by the difference between the observed 1-month implied volatility and the theoretical spot implied volatility and (b) by the nature of the least-squares estimate. To sum up, the level factor can be expressed as:

$$\sqrt{v_t} = \sqrt{v_{1t} + v_{2t}} \quad (5.1)$$

As already mentioned, our eigenvalue composition factor  $\xi$  and rotation factor  $\alpha$  are truly independent of the level. This is mirrored by the fact that they do not appear in (5.1).

**Skewness and term structure** The skewness  $\mathcal{S}_t$  and term structure  $\mathcal{M}_t$  need to be analyzed jointly. The Karoui-Durleman expansion of the Wishart option pricing model yields the following expressions:

$$\mathcal{S}_t = \frac{\partial IV}{\partial K} = \frac{Tr[RQX_t]}{2Tr[X_t]^{3/2}} \quad (5.2)$$

$$\mathcal{M}_t = \frac{\partial IV}{\partial \tau} = \frac{Tr[kQ'Q] - 2Tr[MX_t]}{4Tr[X_t]^{1/2}} - Tr[X_t]\mathcal{C}_t - 3Tr[X_t]^{1/2}\mathcal{S}_t^2 \quad (5.3)$$

$$\begin{aligned} \mathcal{C}_t &= \frac{\partial^2 IV}{\partial K^2} = \frac{1}{3} \frac{Tr [Q^T Q X_t]}{Tr [X_t]^{3/2}} - \left( \frac{Tr [RQ X_t]}{2Tr [X_t]^{3/2}} \right) - \left( \frac{Tr [RQ X_t]^2}{2Tr [X_t]^{7/2}} \right) \\ &\quad + \frac{2}{3} \frac{Tr \left[ \left( X_t RQ + (RQ)^T X_t \right) \left( RQ - \frac{3Tr[RQX_t]}{2Tr[X_t]} \right) \right]}{2Tr [X_t]^{5/2}} \end{aligned} \quad (5.4)$$

where  $\sqrt{v_t} = Tr[X_t]^{1/2}$  is the volatility level and  $\mathcal{C}_t$  is a convexity adjustment.

We rewrite these expressions using the eigenvalue representation (B.6), the symbol  $\xi_t = \frac{\lambda_1}{\lambda_1 + \lambda_2}$  and the trace decomposition (B.8) and obtain a parametric representation of  $\mathcal{S}_t$  and  $\mathcal{M}_t$  in elliptical coordinates:

$$\mathcal{S}_t = \frac{1}{\sqrt{v_t}} \left( \mathcal{S}_0 + (2\xi_t - 1) [f_1 \cos(2\alpha_t) + f_2 \sin(2\alpha_t)] \right) \quad (5.5)$$

$$\mathcal{M}_t = \sqrt{v_t} \left( \mathcal{M}_0 + (2\xi_t - 1) [f_3 \cos(2\alpha_t) + f_4 \sin(2\alpha_t)] \right) \quad (5.6)$$

with

$$\begin{aligned} \mathcal{S}_0 &= \frac{1}{4} (RQ_{11} + RQ_{22}) & \mathcal{M}_0 &\approx \frac{1}{4} (M_{11} + M_{22} + Tr[kQ'Q]v_t^{-2}) \\ f_1 &= \frac{1}{4} (RQ_{11} - RQ_{22}) & f_3 &\approx \frac{1}{4} (M_{11} - M_{22}) \\ f_2 &= \frac{1}{4} (RQ_{12} + RQ_{21}) & f_4 &\approx \frac{1}{4} (M_{12} + M_{21}) \end{aligned}$$

Note that we have ignored the second and third expressions in (5.3) when calculating  $\mathcal{M}_0$ ,  $f_3$  and  $f_4$ , as they contribute relatively little to the term structure but complicate the exposition notably. We still use them in our numerical results.

**Heston-Plus Interpretation.** Formulation (5.5-5.6) makes it possible to analyze the influence of the factors  $(\sqrt{v_t}, \xi, \alpha)$  on the skewness  $\mathcal{S}_t$  and term structure  $\mathcal{M}_t$  in comparison to a one-factor Heston model. Just like in the Heston case, both quantities are homogeneous in the level  $\sqrt{v_t}$ . Equations (5.5-5.6) have a similar structure with two major terms: the Heston values  $\mathcal{S}_0$  and  $\mathcal{M}_0$  and a deviation term. Our additional factors  $\xi$  and  $\alpha$  determine the size and direction of the deviation from the Heston model.

**Role of the eigenvalue composition  $\xi$ .** The factor  $\xi = \frac{\lambda_1}{\lambda_1 + \lambda_2}$  can be interpreted as a measure of distance from the Heston model. It is largest for  $\xi = 0$  or  $\xi = 1$ , i.e. when one of the eigenvalues is zero. Our model would exactly match the Heston case for  $\xi = 0.5$  (i.e.  $\lambda_1 = \lambda_2$ ), but this point can only be reached asymptotically, as the eigenvalues never cross in a Wishart process.

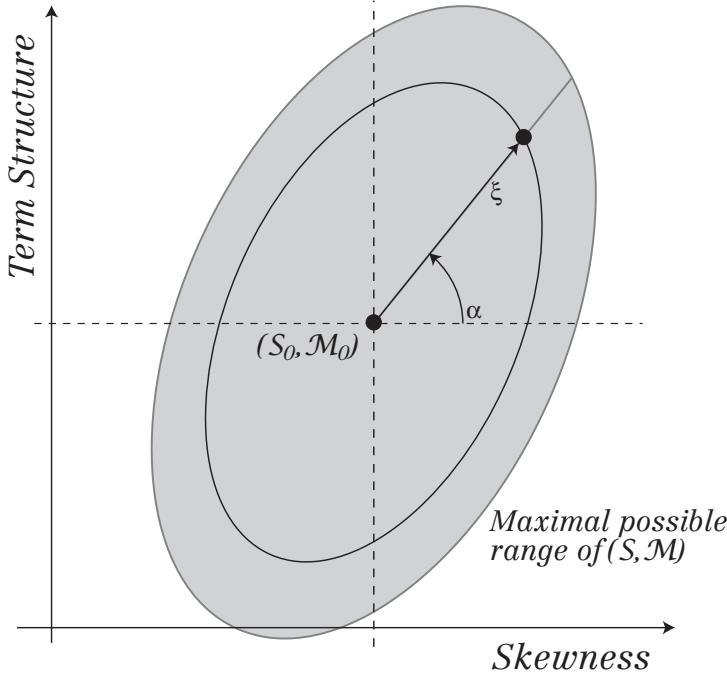


Figure 8: The state components  $\alpha$  and  $\xi$  jointly determine the skewness  $\mathcal{S}$  and the term structure  $\mathcal{M}$ . The radius is fixed by  $\xi$ , the angle by  $\alpha$ .

[Todo] Additional interpretation of  $\xi$  with respect to events. Note that expressions of the type  $TR[MX_t]$  are homogenous in  $v_t$ , thus in times of crisis, when the vola is high, the model has a much wider range. In these times, we do not need the full range of the model and obtain a  $\xi$  diverse from 0. For calm periods, when the vola is lower, we have two effects that push  $\xi$  towards the boarders. First the fact that overall vola is lower and second the fact that in calm times all sorts of sophisticated trades happen. In the crisis, traders try to service, correlations increase and everything follows the Heston model.

**Role of the rotation factor  $\alpha$ .** The factor  $2\alpha$  determines the direction of the deviation from the Heston model. If  $2\alpha = \text{atan}(-f_1/f_2)$ , only the term structure is changed, if  $2\alpha = \text{atan}(-f_4/f_3)$ , only the skewness is changed. For all other values, there is a deviation in both  $\mathcal{S}$  and  $\mathcal{M}$ . See Fig. 8 for an illustration.

The relevant feature of our three-factor model is not the fact that skewness and term structure are stochastic – this can be achieved even in the one factor Heston model – but the presence of the two components  $\xi$  and  $\alpha$  that are independent of the level.

**Numerical values.** Inserting our preferred parameters into (5.5-5.6) gives:

$$\begin{aligned}\mathcal{S}_t &= \frac{1}{\sqrt{v_t}} \left( -0.063 + (2\xi_t - 1)[0.063 \cos(2\alpha_t) - 0.059 \sin(2\alpha_t)] \right) \\ \mathcal{M}_t &= \sqrt{v_t} \left( -1.207 + (2\xi_t - 1)[1.036 \cos(2\alpha_t) + 2.218 \sin(2\alpha_t)] \right) + 0.0511 v_t^{-1}\end{aligned}\quad (5.7)$$

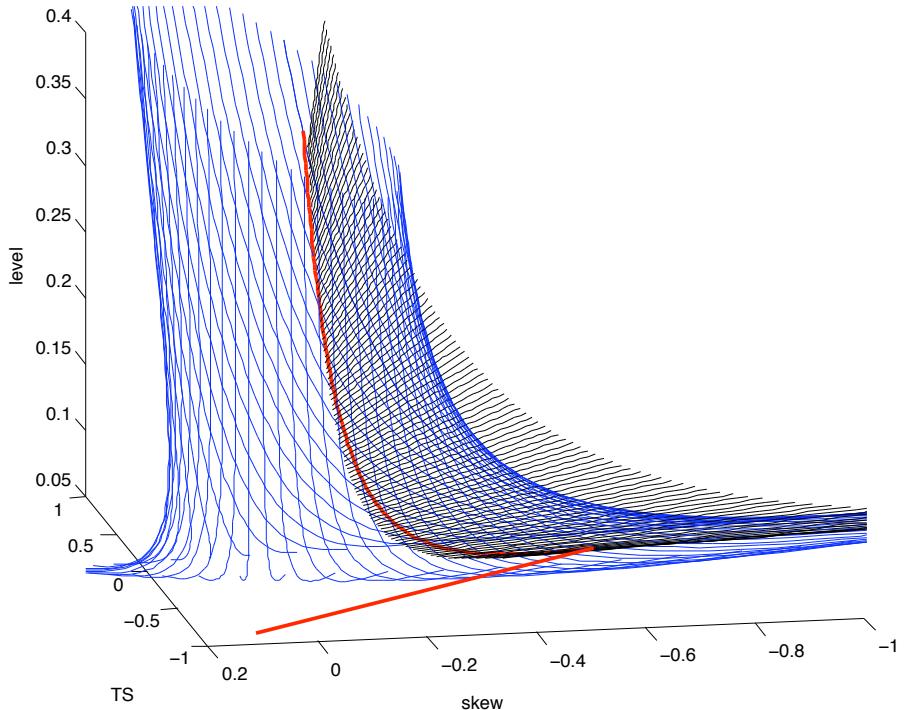


Figure 9: The set of possible combinations of level,  $\mathcal{S}_t$  and  $\mathcal{M}_t$  is a volume in the Wishart model (blue), a surface in the Double Heston model (black) and a line for the Heston model (red).

## 5.2 Scope and limits of the model

How flexible is our model? In order to assess its limits, we study the set of possible combinations of  $(\sqrt{v_t}, \mathcal{S}_t, \mathcal{M}_t)$  for a fixed parameter set  $\theta = \{M, R, Q, k\}$ .

**Proposition 5.1** *The set of achievable combinations of  $(\sqrt{v_t}, \mathcal{S}_t, \mathcal{M}_t)$  is a distorted elliptical prisma.*

**Proof.** We focus on the projection  $\sqrt{v_t} = v_0$ . Setting  $\xi = 0$  we obtain the maximal possible surface of combinations of  $\mathcal{S}_t$  and  $\mathcal{M}_t$ . Appendix B.3 shows that this surface is an ellipse. The size of the ellipse scales continuously with  $\sqrt{v_t}$  for  $\sqrt{v_t} > 0$ , so stacking together all ellipses for  $\Sigma_0, \Sigma_0 + d\Sigma$  and so on produces a distorted elliptical prisma, see Fig. 9.

**Expanding the reach.** At first sight it seems that the model's reach as described in (5.5-5.6) could be expanded almost infinitely by scaling  $f_1 \dots f_4$ . Algebraically it is possible to achieve this while leaving  $\mathcal{S}_0$  and  $\mathcal{M}_0$  unchanged. Such a result would, however, require one of the diagonal elements of  $M$  or  $RQ$  to change sign, which is economically not sensible. Furthermore, the long-term mean (2.13) would change.

**Degenerate solutions.** The solution for a given volatility level degenerates to a line if one axis length of the ellipse  $(\mathcal{S}, \mathcal{M})$  is zero. This happens when the matrix  $\begin{pmatrix} f_1 & f_2 \\ f_3 & f_4 \end{pmatrix}$  has reduced rank, for example when  $RQ = \alpha M$  or when  $M_{11} = M_{22}$  and  $(RQ)_{11} = (RQ)_{22}$ .

### 5.2.1 Nested models

**Two factor Heston model** As discussed [...], the double Heston model is nested in our model by setting the parameter matrices  $(M, R, Q)$  diagonal and expanding the degrees of freedom from the scalar  $k$  to a diagonal matrix  $K$ . Furthermore, the state has only the two components  $\lambda_1, \lambda_2$  which correspond to  $\Sigma^{ATM}$  and  $\xi$ , with  $\alpha = 0$ . For a given level  $\Sigma^{ATM}$ , (5.5-5.6) becomes a line parameterized by  $\xi$ . The set of possible combinations of  $(\sqrt{v_t}, \mathcal{S}_t, \mathcal{M}_t)$  becomes a surface, which is illustrated in Fig. 9.

$$\mathcal{S}_t = (\sqrt{v_t})^{-1} \left( \mathcal{S}_0 + (2\xi - 1)f_1 \right) \quad (5.8)$$

$$\mathcal{M}_t = \sqrt{v_t} \left( \mathcal{M}_0 + (2\xi - 1)f_3 \right) \quad (5.9)$$

inserting our preferred parameters, this gives:

$$\mathcal{S}_t = (\sqrt{v_t})^{-1} \left( -0.128 - 0.066(2\xi - 1) \right) \quad (5.10)$$

$$\mathcal{M}_t = \sqrt{v_t} \left( -1.540 - 1.470(2\xi - 1) \right) + 0.073\sqrt{v_t} \quad (5.11)$$

These values match the values of the Wishart in (5.7) model remarkably well, if we take into the range of values of  $\cos(2\alpha)$ .

**One factor Heston model.** The only remaining factor is  $\Sigma^{ATM}$ , which means that for a given volatility level only the combination  $\mathcal{S}_0, \mathcal{M}_0$  is achievable. The set of all possible becomes a line, shown in red in Fig. 9.

$$\mathcal{S}_t = (\sqrt{v_t})^{-1} \mathcal{S}_0 \quad (5.12)$$

$$\mathcal{M}_t = \sqrt{v_t} \mathcal{M}_0 \quad (5.13)$$

**Alternative three factor models** Once we have accepted that three factors are an important ingredient to modeling the implied volatility surface, the question remains how to invest them. One possibility is the Wishart model with two volatility and one rotation factors. An alternative would be a three factor Heston model. Such a model can be seen as a  $3 \times 3$  diagonal Wishart model parameterized with  $\xi = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3}$  and  $\eta = \frac{\lambda_2}{\lambda_2 + \lambda_3}$ . The surface  $(\mathcal{S}_t, \mathcal{M}_t)$  becomes a triangle, a shape which is much less useful to represent the data than an ellipse.

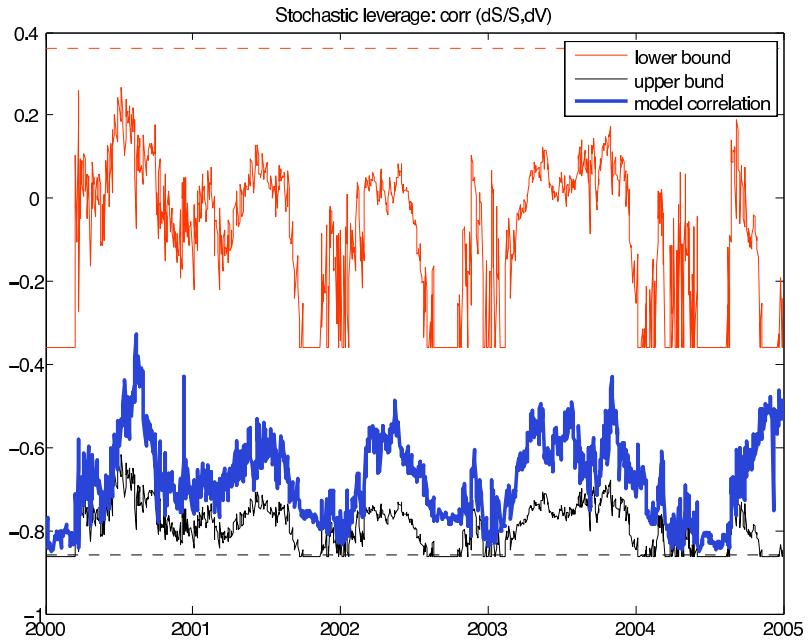


Figure 10: Limits of the correlation structure. Maximum and minimum possible correlation given the parameter set (red lines) and maximum/minimum possible correlation on a day given  $\lambda_1$  and  $\lambda_2$  (black lines). The actual model-implied correlation is shown as thick black line.

### 5.3 More results

**Boundaries of  $\text{Corr}(dS/S, dV)$ .** Fig. 10 aims at decomposing the relative contributions of  $\xi$  and  $\alpha$  to the correlation. For every trading day, we take  $\xi$  as fixed and plot the maximal and minimal possible values (for any  $\alpha$ ) for the correlation in red and black. We then plot the daily model-implied correlation in blue and add the theoretical maximal and minimal limits as dashed lines.

This plot can be interpreted as follows: the dashed horizontal red line would have been produced by the Heston model. The black line would have been produced by the model of Christoffersen et al. (2007). Our model – illustrated by a thick blue line – enjoys additional flexibility from the rotation factor  $\alpha$ .

**Nested models.** The expressions for the one and two factor Heston models can be obtained from (2.9) by the nesting property. Using our parameters, the range of possible correlations is  $[-0.785, -0.602]$  for the two factor Heston model, and a constant  $-0.713$  for the one factor Heston model.

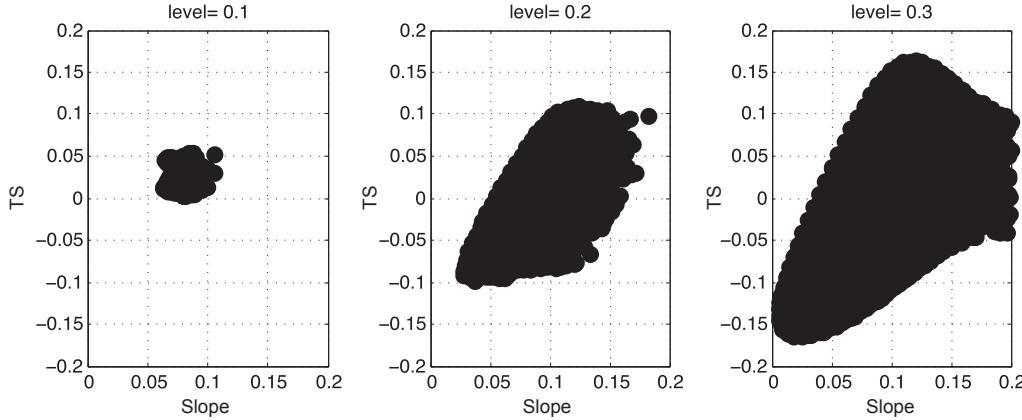


Figure 11: Montecarlo simulation of possible combinations of smirk and term structure for three different volatility levels. Increasing the volatility level expands and moves the achievable region.

### 5.3.1 Model comparison

It is obvious that each additional factor adds to the flexibility of a model. The pricing performance of the three factor model clearly beats all other models. But what is the merit to adding a third factor to an option pricing model? To understand this better, we plot pairwise projections of level, skew and term structure of the model-implied values to the data for each model in Fig. 11. Even three factors do not cover the whole space spanned by the data, but they provide a significant improvement. While the level/slope and slope/term structure relations are close to lines for the one factor models, especially the level/TS relation becomes much richer in the two factor case. Note the triangular shape of the model-implied points: for low values of the level, the range of possible combinations of  $\lambda_1$  and  $\lambda_2$  is very limited, thus there is not much variation. This changes in the three factor model, which covers the level/TS space quite well. Some room for improvement remains in the level/slope space.

## 5.4 Dynamic properties of the model-implied volatility surface

To analyze the dynamic properties of the model-generated volatility surfaces, we perform the same principal component analysis that we have performed on the data in section 3.2 on model-implied volatilities and report the results in Tab. 2. It is no big surprise that the one (two) factor models show a very low fraction of variance explained beyond the first (second) principal component. Our model matches the fractions explained in the data better than

	<i>l</i>	PC 1	PC 2	PC 3	PC 4
Data	2	96.76	2.00	0.88	0.13
Heston	1	99.68	0.24	0.05	0.01
Bates	2	98.95	0.97	0.04	0.01
Christoff	1	99.79	0.16	0.02	0.01
Bates	2	98.27	1.65	0.04	0.03
Full Wishart	2	97.54	2.20	0.14	0.07

Table 2: Number of factors identified by the mean eigenvalue criterion and percentage of variance explained by different principal components for whole sample (1996-2008).

	Data	$SV_{1,0}$	$SVJ_{1,0}$	$SV_{2,0}$	$SVJ_{2,0}$	$SV_{3,1}$
Factors		1	1	2	2	3
Level-Smirk	0.74	-0.75	0.64	0.19	0.28	0.66
Level-TS	-0.77	-0.99	-0.90	-0.99	-0.92	-0.83
Smirk-TS	-0.52	0.75	-0.78	-0.23	-0.57	-0.51

Table 3: Unconditional factor correlations for the extended sample (1996-2008).

any other model, with the third component considerably larger than in any other model, but still smaller than in the data.

In a further step we analyze the structure of correlations between level, smirk, curvature and term structure for the data, our model and the benchmark models, as they constitute an important dynamic property and capture distinct features of the models. In the one-factor Heston model, in which the term structure is entirely determined by the (one) level factor, we observe a correlation of almost  $-1$  between the two. Again, our model matches the data remarkably well for level-smirk, level-TS and smirk-TS. Correlations involving the curvature are slightly worse, but this should be of no worry as they are beyond the scope of a three factor model.

## 5.5 Pricing performance

As one could expect, the three-factor model produces the lowest rms pricing error of all considered models<sup>8</sup>. An overview of the pricing performance in comparison can be found in Tab. 4. We split the analysis in three parts. First the sub-sample of 59 Wednesdays, on which our parameter estimation was based. As expected, performance is best for this sample. Second, we include all days in between the 59 Wednesdays, extending our sample to all trading days in 2000-2004. The daily rms pricing error varies in this sample between 46 cents and \$1.78. Remarkably, there is little difference in performance for this “semi out of sample” analysis. We see this as a confirmation of our estimation strategy. In a third

---

<sup>8</sup>We were unable to find a three factor option pricing model as comparison.

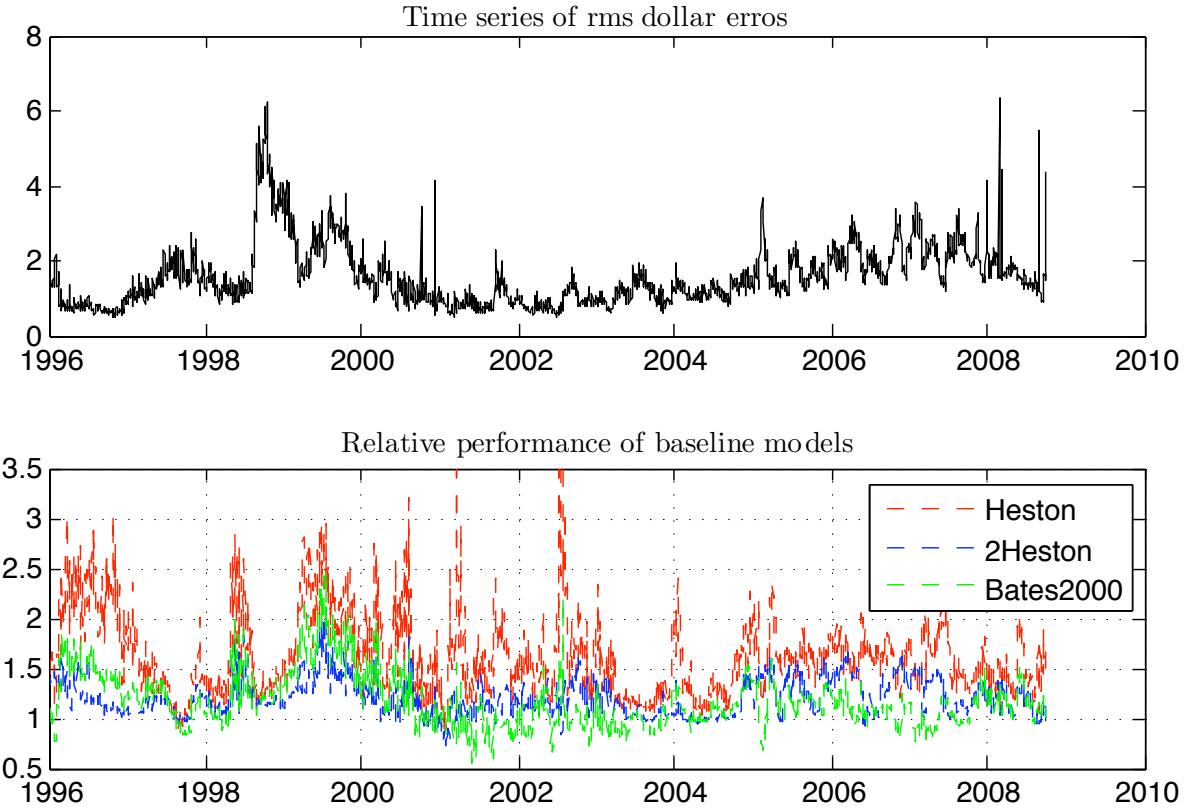


Figure 12: Daily rms pricing error across all maturities and all strikes, in USD.

sample, as an out-of-sample exercise, we used all available data (i.e. from 1996 to 2008). A time series of the daily rms pricing error and the performance relative to baseline models is shown in Fig. 12.

Out of sample performance is worse for all models, especially for late 1998 and 2007. These periods of turbulence have shown an extreme amount of skewness not seen in the sample period. It is therefore no surprise that a model which has been calibrated in a relatively more calm period cannot reproduce these extreme skews and will therefore perform worse. The performance gains must be understood along the analysis in section 5.1.1. Even our model cannot match the above-mentioned crisis situations, because the range of possible skews and term structures is limited (see Fig. 8). Our model shines in times of moderately extreme skews and term structures, such as in mid-1998 and early 1999, when skew and term structure are already too extreme for the baseline models but still within the range of possible values of our model.

**Specification analysis.** Carr, Geman, Madan and Yor (2003) use a regression of the residuals  $\hat{C}_i - C_i$  on the standard explanatory variables in option pricing (moneyness and its square, time to maturity and a cross-term) to assess a model's (mis-)specification. This line of analysis is also picked up by Li and Pearson (2005). If a model could explain all systematic

Factors	$SV_{1,0}$	$SV_{J_{1,0}}$	$SV_{2,0}$	$SV_{J_{2,0}}$	$SV_{3,0}$	$SV_{J_{3,1}}$
	1	1	2	2	3	3
<i>rms</i> dollar error	1.5713	1.4424	1.1800	1.1151	1.1284	1.0476
..semi-OOS	1.656	1.533	1.230	1.1736	1.092	0.9054
..OOS (1996-2008)	2.493	2.382	1.937	1.862	1.570	
IV error	3.02	2.41	2.25	2.72	2.19	
\$ bias	0.021	0.057	0.062	0.136	0.055	
within bid-ask	0.439	0.489	0.575	0.602	0.618	
best day dd/mm/yy	30/01/01	08/03/01	14/02/01	08/03/01	08/03/01	
rms \$ error	0.69	0.58	0.51	0.38	0.47	

Table 4: Performance comparison. The first three rows report the root mean square dollar pricing error for (1) the sample of 59 Wednesdays, (2) all trading days in 2000-2004, (3) all trading days in 1996-2008. The fourth row reports the root mean square implied volatility error for the sample of 59 Wednesdays. The fifth row reports the mean dollar pricing error (bias) for the same sample. The sixth row reports the fraction of prices that come within the bid-ask spread. The last row reports the best performing day and the root mean square dollar pricing error.

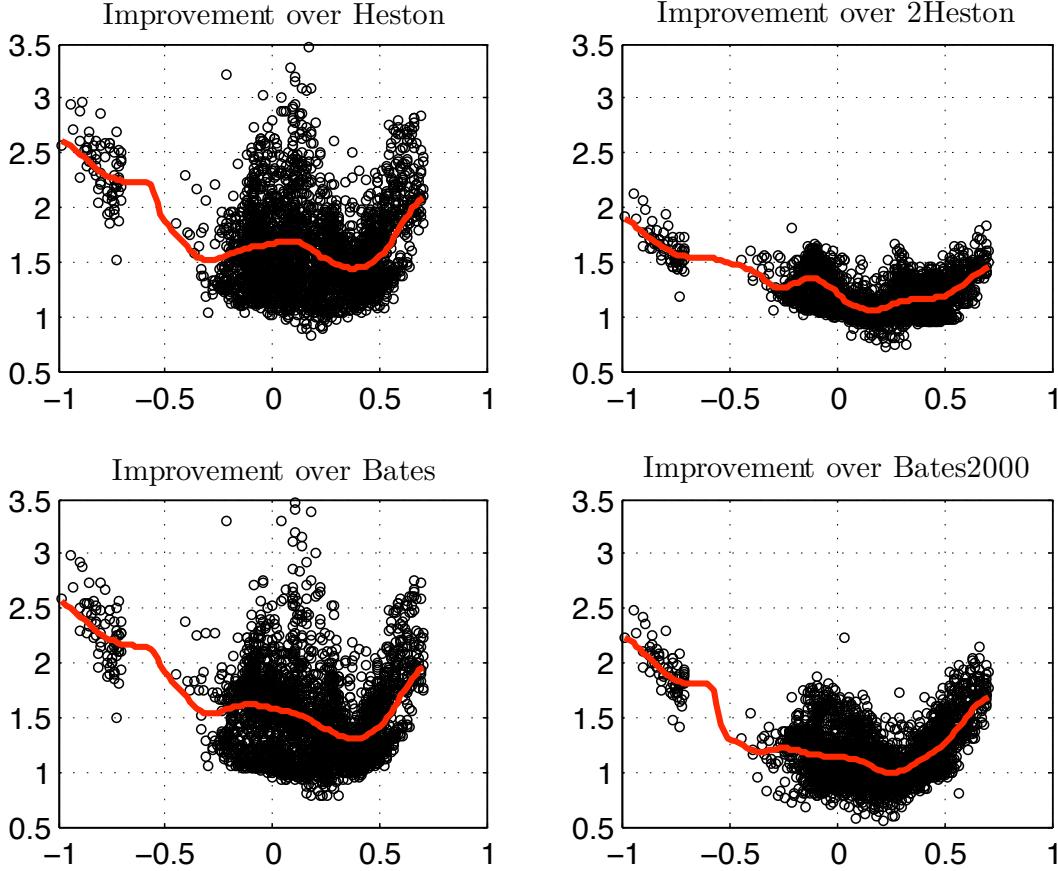


Figure 13: Improvement of model performance as a function of the rotation factor  $\sin \alpha$ . For  $\sin^2 \alpha \approx 0$ , the  $SV_{3,1}$  model behaves like a two-factor model and no or little improvement over a two-factor model can be observed. The largest improvements are observed for extreme values of  $\sin \alpha$ . Each dot represents the factor by which the *rms* dollar pricing error for this day is smaller in the  $SV_{3,1}$  compared to the reference model. The red line is a nonparametric regression to guide the eye.

components of the implied volatility surface, only idiosyncratic errors uncorrelated to the above-mentioned explanatory variables would remain. Thus a high  $R^2$  can be interpreted as a sign of misspecification. Our model produces an  $R^2$  of 4.1%, indicating that our model captures almost all explainable features of the data. This result compares favorably to [Carr et al. \(2003\)](#), who obtain  $R^2$ -values of up to 50%, albeit in a somehow different setting.

**Performance gain through rotation factor.** In Fig. 13, we stratify the gain in pricing performance by our rotation factor  $\cos \alpha$ . Remember that for  $\cos \alpha \approx 1$ , our model behaves similar to a two-factor model, while the rotation factor is most potent for absolute values of  $\cos \alpha$  close to zero. The improvement over the benchmark two-factor models ranges from 10-20% for low absolute values of  $\cos \alpha$  to more than a factor of 2 for extreme values of  $\cos \alpha$ , yet another indication for the importance of this rotation factor.

$$M = \begin{pmatrix} -0.3426 (0.0218) & 0 \\ -9.3628 (0.0603) & -5.0719 (0.0180) \end{pmatrix} \quad Q = \begin{pmatrix} 0.0340 (0.0049) & 0 \\ 0 & 0.4789 (0.0079) \end{pmatrix}$$

$$R = \begin{pmatrix} -0.9757 (0.1362) & 0.1586 (0.0335) \\ 0 & -0.6905 (0.0186) \end{pmatrix} \quad \beta = 1.0001 (0.0427)$$

$$RQ = \begin{pmatrix} -0.0332 & 0.0760 \\ 0 & -0.3307 \end{pmatrix}$$

Table 5: Risk-neutral parameters of the  $SV_{3,1}$  diffusion process, standard errors in brackets.

## 5.6 Risk-neutral parameters $M, Q, R$

We report the parameter estimates and standard errors in Tab. 5.

**Mean reversion parameter  $M$**  The inverse values of the diagonal elements of  $M$  can be directly interpreted as mean reversion speeds. They are 2.9 years and 2.7 months.

**Vol-of-vol parameter  $Q$**  The vol-of-vol matrix  $Q$  has one element ( $Q_{22}$ ) which is an order of magnitude larger than the other elements. This large difference is a key ingredient to producing a realistic (i.e. time-varying and mostly negative) correlation structure and will be more thoroughly discussed in the next section.

**Correlation parameter  $R$**  Interpretation of  $RQ$  in context of stochastic skewness decomposition.

## 6 CONCLUSIONS

We have implemented a new framework for option pricing with matrix affine diffusions. While this is still work in progress, there is already strong evidence that a model with multiple correlated factors can greatly improve standard affine option pricing, both from a conceptual and from an empirical point of view. Jumps bear the potential for further improvements, especially for very short maturities.

The tree-factor specification, with one factor not loading on the volatility level, increases the model's flexibility and versatility. We can price five years of options on the S&P500 with a single set of parameters and acceptable pricing errors. We obtain a long time-series of the model-implied state and can derive three clearly interpretable factors from it.

## A NESTED MODELS

The following five models are nested in our model. For the convenience of the reader, we list below how the parameters as described in the original papers can be converted into our model.

### A.1 Diffusive models

For all nested diffusive models, the jump parameters are zero:  $J_t = 0; L_t = 0$ .

**One volatility factor** The model in [Heston \(1993\)](#) in the usual notation obtains by setting

$$n = 1, \quad M = \frac{-b}{2}, \quad R = \sqrt{\rho}, \quad Q = \frac{\sigma}{2}, \quad K = \frac{a}{\sigma^2}$$

The state is a positive scalar  $X_t = v_t$ .

**Two volatility factors** The model of Christoffersen et al. (2007) obtains by setting

$$n = 2, \quad M = \frac{1}{2} \begin{pmatrix} -b_1 & 0 \\ 0 & -b_2 \end{pmatrix}, \quad R = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix}, \quad Q = \frac{1}{2} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \quad K = \begin{pmatrix} \frac{a_1}{\sigma_1^2} & 0 \\ 0 & \frac{a_2}{\sigma_2^2} \end{pmatrix}$$

using their notation. The state is a positive diagonal matrix  $X_t = \begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}$ . In our decomposition  $v_t = V_1 + V_2$ ;  $\xi = \frac{V_1}{V_1+V_2}$ ;  $\alpha = \pi$ .

The model of DaFonseca et al. (2008) obtains by setting  $n = 2$  and  $K = \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix}$ .

### A.2 Jump models

**One volatility factor** The model of Bates (1996) obtains by setting

$$n = 1, \quad M = \frac{-\beta^*}{2}, \quad R = \rho, \quad Q = \frac{\sigma_v}{2}, \quad K = \frac{\alpha}{\sigma_v^2}, \quad \Lambda_0 = \lambda, \quad \Lambda_1 = 0$$

The state is a positive scalar  $X_t = V_t$ .

**Two volatility factors** The model of Bates (2000) obtains by setting

$$n = 2, \quad M = \frac{1}{2} \begin{pmatrix} -\beta_1 & 0 \\ 0 & -\beta_2 \end{pmatrix}, \quad R = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix}, \quad Q = \frac{1}{2} \begin{pmatrix} \sigma_{v,1} & 0 \\ 0 & \sigma_{v,2} \end{pmatrix}, \quad K = \begin{pmatrix} \frac{\alpha_1}{\sigma_{v,1}^2} & 0 \\ 0 & \frac{\alpha_2}{\sigma_{v,2}^2} \end{pmatrix}$$

and the jump parameters

$$\lambda_0 = \lambda_0, \quad \Lambda_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

The state is a positive diagonal matrix  $X_t = \begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}$ . In our decomposition  $v_t = V_1 + V_2$ ;  $\xi = \frac{V_1}{V_1 + V_2}$ ;  $\alpha = 0$ .

## B ALTERNATIVE MATRIX REPRESENTATIONS

In this section, we present two equivalent representations of the Wishart state matrix. The polar eigenvalue representation makes it possible to disentangle the role of long term, short term and rotation factors. The  $v - \xi - \alpha$  representation separates the volatility level from the structure of the state matrix. Therefore it makes it possible to draw comparisons to the one- and two-factor Heston models.

### B.1 Spectral decomposition

Any  $n$ -dimensional matrix of full rank can be expressed in terms of  $n$  eigenvalues and  $n(n - 1)$  angles, as any square matrix  $A$  with linearly independent eigenvectors can be written as

$$A = PDP^{-1} \tag{B.1}$$

with  $D$  being a diagonal matrix of the  $n$  eigenvalues and  $P$  being a matrix whose column vectors are the normalized eigenvectors. Any  $n$ -dimensional vector can be expressed in polar coordinates using its norm and  $n - 1$  angles. Eigenvectors are normalized, so the  $n - 1$  angles are sufficient.

**The  $2 \times 2$  case.** A symmetric  $2 \times 2$  matrix in polar coordinates is therefore represented as

$$X_t = \begin{pmatrix} \cos^2(\alpha_t)\lambda_t^1 + \sin^2(\alpha_t)\lambda_t^2 & \cos(\alpha_t)\sin(\alpha_t)(\lambda_t^1 - \lambda_t^2) \\ \cos(\alpha_t)\sin(\alpha_t)(\lambda_t^1 - \lambda_t^2) & \sin^2(\alpha_t)\lambda_t^1 + \cos^2(\alpha_t)\lambda_t^2 \end{pmatrix} \tag{B.2}$$

This is equivalent to the following coordinate transformation

$$\{X_{11}, X_{12}, X_{22}\} \longrightarrow \{v_1, v_2, \alpha\}$$

## B.2 $v - \xi - \alpha$ decomposition

Next step: take  $PDP^{-1}$  decompsition and do

$$A = PDP^{-1} = Tr[A] \left( P \frac{D}{Tr[A]} P^{-1} \right) \quad (\text{B.3})$$

To allow for a better interpretation within the framework of our PCs/observed factors, we want make the role of the volatility level  $v_t = v_{1,t} + v_{2,t}$  more visible. To so so, we introduce a new symbol for the fraction of implied variance that is taken by the first eigenvalue:

$$\xi_t = \frac{v_{1t}}{v_{1t} + v_{2t}} = \frac{v_{1t}}{v_t} \quad (\text{B.4})$$

with  $v_{1t} = \xi_t v_t$  and  $v_{2t} = (1 - \xi_t) v_t$ . The goal is to represent our state matrix with three new variables:

$$\{X_{11}, X_{12}, X_{22}\} \longrightarrow \{v_t, \xi, \alpha\}$$

Equation (B.2) becomes:

$$X_t = v_t \begin{pmatrix} \cos^2(\alpha_t)\xi + \sin^2(\alpha_t)(1 - \xi_t) & \cos(\alpha_t)\sin(\alpha_t)(2\xi_t - 1) \\ \cos(\alpha_t)\sin(\alpha_t)(2\xi_t - 1) & \sin^2(\alpha_t)\xi_t + \cos^2(\alpha_t)(1 - \xi_t) \end{pmatrix} \quad (\text{B.5})$$

Using the standard identities  $\cos^2 \alpha = \frac{1}{2}(\cos 2\alpha + 1)$ ;  $\sin^2 \alpha = \frac{1}{2}(1 - \cos 2\alpha)$  and  $\sin \alpha \cos \alpha = \frac{1}{2} \sin 2\alpha$ , we are finally able to separate the volatility level  $v_t$  from the structure of  $X_t$ :

$$\begin{aligned} X_t &= \frac{v_t}{2} \begin{pmatrix} 1 + \cos(2\alpha_t)(2\xi_t - 1) & \sin(2\alpha_t)(2\xi_t - 1) \\ \sin(2\alpha_t)(2\xi_t - 1) & 1 - \cos(2\alpha_t)(2\xi_t - 1) \end{pmatrix} \\ &= \frac{v_t}{2} \left[ Id + (2\xi_t - 1) \begin{pmatrix} \cos(2\alpha_t) & \sin(2\alpha_t) \\ \sin(2\alpha_t) & -\cos(2\alpha_t) \end{pmatrix} \right] \end{aligned} \quad (\text{B.6})$$

Note the above expression contains two trace-invariant transformations, namely a rotation by  $\alpha$  and a re-distribution of the eigenvalues via  $\xi$ . Further note that  $(2\xi - 1)$  is bound between  $-1$  and  $1$ .

Notice that  $\sqrt{X_t}$  can be equally decomposed using the same  $\alpha$  and  $\tilde{\xi} = \frac{\sqrt{v_{1t}}}{\sqrt{v_{1t}} + \sqrt{v_{2t}}}$  as  $X_t$ :

$$\sqrt{X_t} = \frac{\sqrt{v_{1t}} + \sqrt{v_{2t}}}{2} \left[ Id + (2\tilde{\xi}_t - 1) \begin{pmatrix} \cos(2\alpha_t) & \sin(2\alpha_t) \\ \sin(2\alpha_t) & -\cos(2\alpha_t) \end{pmatrix} \right] \quad (\text{B.7})$$

**Trace decomposition.** Equation (B.6) can be used to decompose expressions of the type  $\text{Tr}[MX_t]$  with  $M$  a (general) parameter matrix and  $X_t$  the symmetric, positive definite state matrix:

$$\begin{aligned}\text{Tr}[MX_t] &= \frac{v_t}{2} \text{Tr} \left\{ M \cdot \left[ Id + (2\xi_t - 1) \begin{pmatrix} \cos(2\alpha_t) & \sin(2\alpha_t) \\ \sin(2\alpha_t) & -\cos(2\alpha_t) \end{pmatrix} \right] \right\} \\ &= \frac{v_t}{2} \left\{ \text{Tr}[M] + (2\xi_t - 1) \left( \cos(2\alpha_t)(M_{11} - M_{22}) + \sin(2\alpha_t)(M_{12} + M_{21}) \right) \right\}\end{aligned}\quad (\text{B.8})$$

We note that this expression is again homogenous in  $v_t$ .

### B.3 General ellipse in eigenvalue coordinate representation

We start with the equation of the ellipse in parameter form:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 + a \cos(\beta) \\ y_0 + b \sin(\beta) \end{pmatrix} \quad (\text{B.9})$$

Next we note that  $\beta$  is not unique, i.e. can be shifted by a constant angle  $\delta$ , i.e. the following two parameterizations produce the same shape

$$\begin{pmatrix} x_0 + a \cos(\beta) \\ y_0 + b \sin(\beta) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x_0 + a \cos(\beta + \delta) \\ y_0 + b \sin(\beta + \delta) \end{pmatrix} \quad (\text{B.10})$$

Applying the appropriate identities, we obtain

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 + a(\cos(\beta)\cos(\delta) - \sin(\beta)\sin(\delta)) \\ y_0 + b(\sin(\beta)\cos(\delta) + \cos(\beta)\sin(\delta)) \end{pmatrix} \quad (\text{B.11})$$

Next we apply a rotation by an angle  $\phi$ :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} x_0 + a(\cos(\beta)\cos(\delta) - \sin(\beta)\sin(\delta)) \\ y_0 + b(\sin(\beta)\cos(\delta) + \cos(\beta)\sin(\delta)) \end{pmatrix} = \dots \quad (\text{B.12})$$

$$\begin{pmatrix} x_0 \cos(\phi) - y_0 \sin(\phi) + \cos(\beta)[a \cos(\phi)\cos(\delta) + b \sin(\phi)\sin(\delta)] + \sin(\beta)[-a \cos(\phi)\sin(\delta) + b \sin(\phi)\cos(\delta)] \\ -x_0 \sin(\phi) + y_0 \cos(\phi) + \cos(\beta)[-a \sin(\phi)\cos(\delta) + b \cos(\phi)\sin(\delta)] + \sin(\beta)[a \sin(\phi)\sin(\delta) + b \cos(\phi)\cos(\delta)] \end{pmatrix} = \dots \quad (\text{B.13})$$

$$\begin{pmatrix} \tilde{x}_0 + f_1 \cos(\beta) + f_2 \sin(\beta) \\ \tilde{y}_0 + f_3 \cos(\beta) + f_4 \sin(\beta) \end{pmatrix} \quad (\text{B.14})$$

We can now use the following identities to determine the size of the ellipse:

$$\begin{aligned}
 (a^2 + b^2) &= f_1^2 + f_2^2 + f_3^2 + f_4^2 \\
 \cos(2\delta)(a^2 - b^2) &= f_1^2 - f_2^2 + f_3^2 - f_4^2 \\
 \cos(2\phi)(a^2 - b^2) &= f_1^2 + f_2^2 - f_3^2 - f_4^2 \\
 \frac{1}{2} \tan 2\delta &= \frac{f_1 f_2 + f_3 f_4}{-f_1^2 + f_2^2 - f_3^2 + f_4^2} \\
 \frac{1}{2} \tan 2\phi &= \frac{f_1 f_3 + f_2 f_4}{f_1^2 + f_2^2 + f_3^2 + f_4^2}
 \end{aligned} \tag{B.15}$$

## References

- Bakshi, Gurdip, Charles Cao, and Zhiwu Chen (1997) ‘Empirical performance of alternative option pricing models.’ *Journal of Finance* 52 (5), 2003–2049
- Bates, David (1996) ‘Jumps and stochastic volatility: Exchange rate processes implicit in deutschmark options.’ *Review of Financial Studies* 9, 69?108
- (2000) ‘Post-’87 crash fears in the S&P 500 futures option market.’ *Journal of Econometrics* 94(1), 181–238
- Bru, M.-F. (1991) ‘Wishart processes.’ *Journal of Theoretical Probability* 4, 725–751
- Carr, P., H. Geman, D. B. Madan, and M. Yor (2003) ‘Stochastic volatility for levy processes.’ *Mathematical Finance* 13, 345–382
- Christoffersen, Peter, Steve Heston, and Kris Jacobs (2007) ‘The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well.’ *Working paper, January 23, 2007*
- DaFonseca, J., M. Grasselli, and C. Tebaldi (2008) ‘A multifactor volatility heston model.’ *Quantitative Finance* 8 (6), 591–604
- Durrleman, Valdo (2004) ‘From implied to spot volatilities.’ *Working paper*
- Durrleman, Valdo, and Nicole El Karoui (2007) ‘Couplingsmiles.’ *Working Paper, March 20, 2007*
- Fang, Fang, and Kees Oosterlee (2008) ‘A nocel pricing method for european options based on fourier-cosine expansions.’ *MPRA working paper, 10 March 2008*
- Fengler, Matthias R., Wolfgang K. Härdle, and Enno Mammen (2005) ‘A dynamic semi-parametric factor model for implied volatility string dynamics.’ *SFB 649 Discussion Paper 2005-020*
- Fengler, M.R., W. Härdle, and C. Villa (2001) ‘The dynamics of implied volatilities: A common principle components approach.’ *Discussion Paper 38, SFB 373, Humboldt University*
- Gourieroux, Christian (2006) ‘Continuous time wishart process for stochastic risk.’ *Econometric Reviews* 25, 177–217
- Grasselli, M., and C. Tebaldi (2008) ‘Solvable affine term structure models.’ *Mathematical Finance* 18, 135–153
- Heston, S. (1993) ‘A closed-form solution for options with stochastic volatility with applications to bond and currency options.’ *Review of Financial Studies* 6, 327–343

Leippold, Markus, and Fabio Trojani (2008) ‘Asset pricing with matrix affine jump diffusions.’ *Working paper, April 22, 2008*

Li, Minqiang, and Neil D. Pearson (2005) ‘Price deviations of s&p 500 index options from the black-scholes formula follow a simple pattern.’ *Working paper*

Skiadopoulos, George, Stewart Hodges, and Les Clewlow (2000) ‘The dynamics of the s&p 500 implied volatility surface.’ *Review of Derivatives Research* 3, 263–282

## WORKING PAPERS IN THE CAREFIN SERIES

- 1/08** La stima del capitale economico a fronte del portafoglio crediti: un'introduzione alle nuove metodologie
- 2/08** Imperfect Predictability and Mutual Fund Dynamics: How Managers Use Predictors in Changing the Systematic Risk
- 3/08** Il sistema dualistico per la governance di banche e assicurazioni
- 4/08** Precautionary investments and vertical externalities: the role of private insurers in intergovernmental relations
- 5/08** I derivati climatici per il settore vitivinicolo
- 6/08** Market discipline in the banking industry. Evidence from spread dispersion
- 7/08** A Survey on Risk Management and Usage of Derivatives by Non-Financial Italian Firms
- 8/08** Hedge Funds: Ability Persistence and Style Bias
- 9/08** Cross-Industry Diversification: Integration, Bubble and Predictability
- 10/08** The Impact of Government Ownership on Banks' Ratings: Evidence from the European Banking Industry
- 11/08** A Framework for Assessing the Systemic Risk of Major Financial Institutions
- 12/08** Multidimensional Distance to Collapse Point and Sovereign Default Prediction
- 13/08** Search Costs and Mutual Fund Fee Dispersion
- 14/08** The Choice of Target's Advisor in Mergers and Acquisitions: the Role of Banking Relationship
- 15/08** Stochastic Mortality: the Impact on Target Capital
- 1/09** Regular(ized) Hedge Fund Clones
- 2/09** Loads and Investment Decisions
- 3/09** The role of Basel II in the subprime financial crisis: guilty or not guilty?
- 4/09** Why Larger Lenders obtain Higher Returns: Evidence from Sovereign Syndicated Loans
- 5/09** Crisi finanziaria, controlli interni e ruolo delle Autorità
- 6/09** Why do foreign banks withdraw from other nations
- 7/09** Do derivatives enhance or deter mutual fund risk-return profiles? Evidence from Italy
- 8/09** External Debt to the Private Sector and the Price of Bank Loans
- 9/09** Pricing insurance contracts following the cost of capital approach: some conceptual issues
- 10/09** Listed Private Equity Funds: IPO Pricing, J-Curve and Learning Effects
- 11/09** Rating changes: the European evidence
- 12/09** Beyond macroeconomic risk: the role of contagion in the italian corporate default correlation
- 13/09** Revisiting corporate growth options in the presence of state-dependent cashflow risk
- 14/09** Borrowing in Buyouts
- 15/09** Explaining Returns in Private Equity Investments
- 16/09** Banks' Loan Loss Provisioning: Procyclical Behaviour and Potential Solutions
- 17/09** Informed intermediation of longevity exposures
- 18/09** Regulations and soundness of insurance firms: international evidence
- 19/09** Project Finance Collateralised Debt Obligations: An Empirical Analysis on Spreads Determinants
- 20/09** Crashes and Bank Opaqueness
- 21/09** Investors' Distraction and Strategic Re-pricing Decisions
- 22/09** Italian labourers participation in private pension plans: an analysis of closed funds targeting specific industries
- 23/09** Diritto delle assicurazioni e diritto europeo: la prospettiva italiana
- 24/09** The Common Frame of Reference (CFR) of European Insurance Contract Law
- 01/10** Optimal time of annuitization in the decumulation phase of a defined contribution pension scheme
- 02/10** Three make a smile – dynamic volatility, skewness and term structure components in option valuation

The working papers by Carefin, Bocconi's Centre for Applied Research in Finance  
are made possible by the following Institutions

ABI	INTESA SANPAOLO
ALLIANZ	INTESA VITA
ARCA VITA	PIONEER INVESTMENTS MANAGEMENT SGR
ARCA SGR	SARA ASSICURAZIONI
ASSICURAZIONI GENERALI	STATE STREET GLOBAL ADVISORS LIMITED
AVIVA ITALIA HOLDING.	
BANCA CARIGE	TOWERS WATSON
BANCA POPOLARE DI MILANO	UNICREDIT
BANCASSURANCE POPOLARI	UNICREDIT BANCASSURANCE MANAGEMENT & ADMINISTRATION
BNL VITA	VENETO BANCA
CARIGE AM SGR	WILLIS RE SOUTHERN EUROPE
DELOITTE CONSULTING	

*Copyright  
Carefin, Università Bocconi*

**CAREFIN**  
Centre for Applied Research in Finance  
Università Bocconi  
via Roentgen 1  
I-20136 Milano  
tel. +39 025836.5908/07/06  
fax +39 025836.5921  
[carefin@unibocconi.it](mailto:carefin@unibocconi.it)  
[www.carefin.unibocconi.it](http://www.carefin.unibocconi.it)