Implied Calibration of Stochastic Volatility Jump Diffusion Models

S. Galluccio[*], Y. Le Cam[‡]

[*]BNPParibas, London, [‡] Université d'Evry Val d'Essonne

Abstract

In the context of arbitrage-free modelling of financial derivatives, we introduce a novel calibration technique for models in the affine-quadratic class for the purpose of contingent claims pricing and risk-management. In particular, we aim at calibrating a stochastic volatility jump diffusion model to the whole market volatility surface at any given time. We numerically implement the algorithm and show that the proposed approach is both stable and accurate.

1 Introduction

A number of studies have demonstrated that the simple Black-Scholes (BS) paradigm of lognormality of asset returns distribution is in contrast with empirical observations. The simplistic assumption of asset dynamics being driven by a Gaussian process must indeed be rejected in favour of more general processes like, for instance, those driving stochastic volatility (SV) models.

One stream of research focuses on the statistical properties of asset returns. In this context, empirical evidence seems to reject SV models since they are not capable of reproducing the observed conditional kurtosis of returns. The presence of jumps is often advocated as a solution to this problem. In fact, evidence of presence of jumps in the asset, in the volatility or in both is reported in Bates (1996), Bakshi *et al.* (1997), Chernov *et al.* (1999), Andersen *et al.*(2002), Pan (2002), Bates (2000), Eraker *et al.* (2003) and Chernov *et al.* (2003), among others.

In another set of studies, departures from the BS model are advocated in relation to the implied volatility smile phenomenon. Deterministic volatility extensions of the BS model were first due to Dupire (1994) and Derman and Kani (1994). These are usually referred to as local volatility (LV) models. Although LV models provide a simple mechanism for smile generation, they are pledged by a number of shortcomings. On the empirical side, they require the knowledge of option prices over a continuum of strikes and maturities, a situation never encountered in practice. On the theoretical side, it is well-known (Rebonato (2000), Andersen and Andreasen (2000)) that models whose volatility evolves deterministically with the underlying state variables generate smiles at future times that are inconsistent with historical observations: while historical smile surfaces display a high degree of time-stationarity, model-implied smiles in LV models tend to flatten very quickly as time goes by. In addition, by imposing a deterministic relationship between volatility and underlying, one is implicitly making a strong assumption on their relative movements, with implications on the risk-management side (Di Graziano and Galluccio (2005)). Empirical literature rejects local volatility models on their impact on hedging, Dumas *et al.* (1997). In a different line of thought, Hull and White (1987), Stein and Stein (1991) and Heston (1993), account for the smile phenomenon through SV models. Finally, modelling the smile through mixed jump-diffusion (JD) processes is proposed in Andersen et al. (2000) (for LV models with jumps) and Duffie et al. (2000) (for SV models with jumps). However, to force a SV or a JD model to be consistent with the whole set of smiles at different maturities (the so-called "implied volatility surface"), model coefficients must be heavily (and unrealistically) time-dependent. In particular, SV models tend to underestimate smile convexity at short maturities while a JD model suffers from the same drawback at large maturities (Section 3).

In summary, empirical studies based on both statistical analysis and market-implied methods

reject SV, JD and LV models in favour of stochastic volatility jump-diffusion models (SVJD) thanks to their superior market explicative power.

A fundamental problem in the applications (e.g. for derivatives pricing and hedging) is the estimation of latent parameters in SVJD models. Estimation based on statistical analysis from historical data series has been given extensive coverage in the past. In particular Bates (2000), Chernov *et al.* (1999), Craine, Lochstoer and Syrtveit (2000) and Deelstra et al. (2002) use simulation-based estimators and/or Efficient Methods of Moments in models with jumps and stochastic volatility. Eraker *et al.* (2003) employ Markov Chain Monte Carlo methods to provide evidence of jumps in both asset and volatility. Finally, extensive empirical studies across several models are conducted in Chernov *et al.* (2003).

As long as option pricing and hedging is concerned, a model must be made consistent with the available market quotations of liquid vanilla options in order to avoid arbitrage opportunities. In this respect statistical estimations must be replaced or, at least, complemented by a reverse engineering process (model calibration) that consists of determining model parameters to reproduce the observed vanilla option prices[1]. Despite the importance of having a fast, robust and accurate model calibration, the literature on the subject is scarce. In this respect, calibration based on short-term asymptotics is studied in Medvedev and Scaillet (2004). Backus *et al.* (1997) and Zhang and Xiang (2005) use an interesting mapping to infer a term structure of market implied cumulants directly from market smiles at different maturities. Unfortunately, this approach does not provide an estimation of the coefficients of a single model that is consistent with the whole volatility surface since different models are needed to match different smiles. Bakshi *et al.* (1997) and Andersen and Andreasen (2000) suggest to calibrate a model by minimizing the sum of the squared errors of all available options across all strikes and maturities. This simple non-linear least squares optimization is usually not convergent and not statistically robust, as shown in Cont and Tankov (2004) and Detlefsen (2005). Cont and Tankov point out that the information contained in the set of available option prices is not sufficient to remove the coefficients degeneracy that is associated to a SVJD process and suggest that calibration can only be achieved provided one adds exogenous information in addition to the available option prices. For, they introduce a calibration algorithm (in the context of exponential-Lévy processes) where the objective function contains a convex functional that is meant to stabilize the (non-convex) optimization problem. The authors are mainly interested in calibrating a single smile at the time, but a straight generalization of Cont-Tankov's approach to more general processes or to cope with the calibration of the whole volatility surface remains, as of

---

[1]When market is complete, model calibration identifies (at least in principle) the unique risk-neutral measure and avoids the problem of determining the market price of risk. This contrasts the case when estimation is conducted from a statistical perspective. When market is incomplete, calibration allows selecting the "market" measure among the infinite set of possible risk-neutral measures.

today, an open issue.

In our paper, inspired by Cont and Tankov (2004) results, we attempt to take the next step in this direction and introduce a novel implied calibration methodology for some SVJD models with time-dependent coefficients in the affine-quadratic class (Piazzesi (2003), Peng and Scaillet (2004)). We recall that the affine-quadratic class contains the affine one as a special case. Our approach, aimed at calibrating the whole volatility surface at any given time, retains (at a qualitative level) some of the interesting features contained in Cont and Tankov's method, namely the regularization of least-squares optimization through addition of a number of constraints to the problem. However, we depart from Cont and Tankov method in both the nature of the problem (we aim at calibrating the whole volatility surface as opposed to a single smile curve) and in the type of dynamics (we don't restrict ourselves to Lévy processes). We apply our method to one of the simplest (yet non-trivial) SVJD model with jumps in the asset and we show that an accurate and financially meaningful calibration to the whole volatility surface is possible and the algorithm is robust. However, our study strongly suggests that the algorithmic complexity is such that generalizing the present approach to more complex models might result impossible to achieve. This is the case, for instance, when jumps in volatility are also present or for more general local volatility forms .[2] These shortcomings expose an intrinsic limitation of SVJD models and clearly indicate that, despite their mathematical and financial appeal, further theoretical developments are needed.

The rest of the paper is organized as follows. In section 2 we introduce model and notations, and we determine closed-form formulae for European options. In Section 3 we analyze the different role played by jumps and stochastic volatility in explaining the market smile. This results in identifying two separate regimes which are instrumental in solving the calibration problem. Section 4 gathers some general ideas about calibration in models mixing jumps and stochastic volatility while Section 5 is devoted to a detailed description of the calibration methodology. Numerical results are presented in Section 6 and Section 7 contains conclusions and prospects for future research.

2 Mathematical setup and option pricing

2.1 The model

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. We shall denote by S_t the price of the Equity asset at time t and r_t the spot rate of interest assumed deterministic. A probability measure $\mathbb{P}^*$, equivalent to the historical probability measure $\mathbb{P}$, is said to be the risk-neutral measure if the relative price $S_t e^{-r_t t}$ follows a local martingale process under $\mathbb{P}^*$.

---

[2] One important exception being hybrid (Equity-IR) modelling with no jumps in volatility since in that case the approach we present here can be applied without major modifications, as shown in Galluccio and LeCam (2005).

We stipulate that the asset dynamics follows a jump-diffusion process under any equivalent risk-neutral measure $\mathbb{P}^*$,[3]

$$\begin{cases} \frac{dS_t}{S_{t^-}} = (r_t - d_t - \mu_t)dt + \eta_t dB_t^1 + dJ_t \\ d\eta_t = \lambda(a_t - \eta_t)dt + \alpha_t dB_t^2, \end{cases} \tag{1}$$

where B^1 and B^2 are two Brownian motions with $d\left\langle B^1, B^2 \right\rangle_t = \rho_{1,2}(t)dt$; λ is a constant and d_t, α_t, a_t are deterministic functions of time. Here, J is a compound Poisson process with stochastic intensity $\xi(t, S_t)$. We shall write $J_t = \sum_{n \geq 1} \left(e^{Y_n} - 1 \right) 1_{\{T_n \leq t\}}$ (T_i are the jump arrival times), and denote by $\mathbb{G}$ the law of the jumps (a *iid* sequence). More precisely, the jump diffusion $Z = (S, \eta)$ is a Feller process with infinitesimal generator $\mathcal{A}$ defined by

$$\begin{aligned} \mathcal{A}\varphi(t,x) &= \partial_t \varphi + \partial_x \varphi(t,x)\zeta(t,x) + \frac{1}{2}tr\left[\partial_{x,x}\varphi(t,x)\Sigma(t,x)\Sigma^\top(t,x) \right] \\ &\quad + \xi(t,x) \int_{\mathbb{R}} \left[\varphi(t, x_1 + u) - \varphi(t,x) \right] d\mathbb{G}(u), \end{aligned} \tag{2}$$

where ζ is the drift vector and Σ the volatility matrix of the diffusion. Let $\mu_t = \xi(t, S_t) \mathbb{E}^*(Y)$ be its associated compensator. An equivalent representation considers state variables driven by a two dimensional vector of independent Wiener processes $\left(W^1, W^2 \right)'$

$$\begin{cases} \frac{dS_t}{S_{t^-}} = (r - d_t - \mu_t)dt + \eta_t dW_t^1 + dJ_t \\ d\eta_t = \lambda(a_t - \eta_t)dt + \alpha_t \left(\rho_1 dW_t^1 + \rho_2 dW_t^2 \right) \end{cases} \tag{3}$$

where $\{\rho_i := \rho_i(t), i = 1, 2\}$ are deterministic functions of time. We will alternate between the first and the second representation depending on the circumstances, if no confusion arises. The relation between the two representations is provided by $\rho_1 = \rho_{1,2}$, $\rho_1^2 + \rho_2^2 = 1$.

As it is well known, the above model is arbitrage-free but is not complete. Model calibration will then be used to select the risk-neutral measure. We also remind that the model defined by Eqs. (1) and (3) does not belong to the affine class, in the sense of Duffie *et al.* (2000), since it corresponds the jumps-augmented version of the Stein-Stein model, (Stein and Stein (1991)). The equivalent of (1) within the affine framework (when $(Y_n)_{n \geq 1}$ are Gaussian r.v.'s) is in fact the Bates (1996) model which is obtained by assuming that η_t^2 follows a one-dimensional CIR-like process.

Our choice is motivated by a number of reasons. First, empirical literature (Jones (2003), Aït-Sahalia and Kimmel (2004)) demonstrates that simple affine models must be rejected in favor of more general processes. Indeed, our model belongs to the so called "linear-quadratic" class (Piazzesi (2003), Peng and Scaillet (2004)), which includes the affine as a special case. Second, our calibration

---

[3] For sake of simplicity, we will thoroughout assume that the divident process d_t is deterministic and that relative dividends are payed continuously in time. Note also that in our formulation no restriction must be imposed on the r.v. Y to ensure that S_t stays positive.

algorithm (Sections 4 and 5) applies without major changes to affine models as well. Finally, SVJD linear-quadratic models can be easily generalized to include quanto and cross-currency features as well as the effect of stochastic interest rates (for the purpose of hybrids modelling) while this is not possible in affine models (Galluccio and Le Cam (2005)). At an intuitive level, the presence of quanto effects or stochastic interest rates induces non-linear terms in the drift of the process. As a consequence, apart from rather unrealistic situations (for instance when some correlations are artificially set to zero so to force these non-linear additional terms to identically vanish, see Galluccio and Le Cam (2005)) the CCY or stochastic interest rates extended SVJD models are not affine.

In the applications, it is useful to recast all equations in a more convenient form by introducing the auxiliary vector diffusion process $X_t := \left(X_t^1 = \ln\left(S_t\right), X_t^2 = \eta_t \right)$ and the jump process $N_t := \sum_{n\geq 1} Y_n 1_{\{T_n \leq t\}}$. In the new setting, the system reads as:

$$
\begin{cases}
dX_t^1 = \left(r - \frac{1}{2}\left(X_t^2\right)^2 - d_t - \mu_t \right) dt + X_t^2 dW_t^1 + dN_t \\
dX_t^2 = \lambda\left(a_t - X_t^2\right) dt + \alpha_t \left(\rho_1 dW_t^1 + \rho_2 dW_t^2 \right).
\end{cases}
\tag{4}
$$

The model can be efficiently handled mathematically. In fact, if we assume that the intensity process takes the form $\xi(t,x) := \xi_t^0 + \xi_t^1 X_t^1 + \xi_t^2 X_t^2 + \xi_t^3 \left(X_t^2\right)^2$, the jump diffusion vector-valued process $dX_t = \varsigma(X_t, t)dt + \Sigma(t, X_t)dW_t + dN_t$ is a semimartingale associated to a triplet of characteristics that are affine-quadratic functions of the state variables, as in Peng and Scaillet (2004).

The presence of jumps in the dynamics is well supported by historical time series analysis, as mentioned in the introduction. Also, Bates (1996) and (2000) suggests that jumps are needed in addition to stochastic volatility to allow matching both long and short-maturity smiles within a single model. Strong evidence in support of this claim (and its implications on calibration) are given in the next Section.

2.2 Option Pricing

To ensure market consistency in pricing and hedging, a model must be "calibrated" to a set of vanilla options. The availability of closed or quasi-closed form formulae for simple European derivatives is of crucial importance to improve speed and avoid numerical convergence problems. This can be easily achieved in our framework. The path we follow is similar to the one adopted in Peng and Scaillet (2004) in the general context of affine-quadratic models.

We introduce two objects. One is the Green function $\psi(u, X_t; t, T)$ from the risk-neutral expectation

$$
\psi(u, X_t; t, T) := \mathbb{E}^* \left[\exp\left(-\int_t^T R(s, X_s)ds \right) e^{u \cdot X_T} \, \middle| \, \mathcal{G}_t \right],
\tag{5}
$$

where $R(t, X_t) = r$ is the spot interest rate which, for simplicity, is assumed constant. The other is the Laplace transform of the law of Y from $\mathcal{L}_f(x) = \int e^{ux} dG(u)$, under the usual conditions of existence and convergence. Next define the functions $\Phi_t^i(x) = \xi_t^i (\mathcal{L}_f(x) - 1)$. Then, the following result holds (where time dependency has been omitted to lighten notation)

Proposition 1 *There exist four functions* $\gamma(t, T), \beta_1(t, T), \beta_2(t, T),$ *and* $\delta(t, T)$ *such that* ψ *can be represented as*

$$\psi(u, X_t; t, T) = e^{\gamma(t,T) + \beta(t,T) \cdot X_t + \delta(t,T)\left(X_t^2\right)^2}$$

for $u = (u_1, u_2)$, $\beta(t, T) := (\beta_1(t, T), \beta_2(t, T))'$, *and* $X_t := \left(X_t^{(1)}, X_t^{(2)}\right)'$. *Moreover the four functions satisfy the following system of ODE's:*

$$
\begin{cases}
\frac{\partial \beta_1}{\partial t} & = -\Phi^1(\beta_1), \\
\frac{\partial \beta_2}{\partial t} & = -\Phi^2(\beta_1) - 2\lambda a \delta - \left(\rho_1 \alpha \beta_1 + 2\alpha^2 \delta\right) \beta_2, \\
\frac{\partial \delta}{\partial t} & = -\Phi^4(\beta_1) + \frac{1}{2}\left(\beta_1 - \beta_1^2\right) + 2\left(\lambda - \alpha\rho_1\beta_1\right)\delta - 2\alpha^2\delta^2 \\
\frac{\partial \gamma}{\partial t} & = -\Phi^0(\beta_1) + (d + \mu - r)\beta_1 - \lambda a \beta_2 - \alpha^2\left(\delta + \frac{1}{2}\beta_2^2\right)
\end{cases}
\tag{6}
$$

with final conditions $\beta(T, T) = u$, $\delta(T, T) = \gamma(T, T) = 0$

Proof. *See Appendix A.* ∎

To allow analytical tractability of the ODE's, avoid model overparametrization and to simplify the calibration without compromising the quality of the result we will assume, from now on, that the stochastic intensity of the compound Poisson process takes a simple form, i.e. $\xi_t^{(0)} \neq 0, \xi_t^{(1)} = \xi_t^{(2)} = \xi_t^{(3)} = 0$. Despite this simplification, the above system contains non-linear second order Riccati equations and cannot be solved in closed form, in general. However, since only a finite set of options at different times to expiry are quoted in the market, we consider a particular specification of the ODE's coefficients. Be $(T_1, \cdots, T_N)$ the set of expiry times associated to the quoted vanilla options. Accordingly, if $\theta(t)$ is a generic time-dependent coefficient in the ODE's system, we will assume that $\theta(t)$ is defined through piecewise constant functions as follows: $\theta(t) = \theta_i$, if $t \in [T_{i-1}, T_i)$, $i = 2, \cdots, N$. With this specification, on every interval $[T_{i-1}, T_i)$, Riccati equations are defined in terms of constant coefficients and then solvable. On every subinterval, terminal conditions are $\beta(T_i) = (u_i^1, u_i^2)', \delta(T_i) = u_i^3, \gamma(T_i) = u_i^4$. We then arrive at the following result, where we have defined $\Psi_i(t, x) = x^{-1}\left[1 - \exp\left(x(T^i - t)\right)\right]$

Proposition 2 *Assume that* $\alpha_t, a_t, k_t, \xi_t$ *are piecewise constant on the intervals* $[T_{i-1}, T_i)$, $i = 2, \cdots, N$ *and that* $\xi_t^{(0)} \neq 0, \xi_t^{(1)} = \xi_t^{(2)} = \xi_t^{(3)} = 0$. *Then the solution of the system of ODE's is*

given, on each $[T_{i-1}, T_i)$, by

$$\beta(t) = \left(u_i^1, \ \ M(t)\left(u_i^2 - K(t)\right)\right)',$$

$$\delta(t) = \frac{1}{\alpha_i^2}\left(-(B_i + \Gamma_i) - \frac{2\Gamma_i C_i}{e^{4\Gamma_i(T_i - t)} - C_i}\right),$$

$$\gamma(t) = u_i^4 - \left(\left(d + \Phi_t^0(1)\right)u_i^1 - \Phi_t^0(u_i^1)\right)(T_i - t) - (B_i + \Gamma_i)(T_i - t),$$

$$-\frac{1}{2}\ln\frac{1 - C_i e^{-4\Gamma_i(T_i - t)}}{1 - C_i} + \int_t^T\left(\lambda a_i + \frac{1}{2}\alpha_i^2\beta_2(s)\right)\beta_2(s)ds,$$

where

$$p_i = -2\lambda a_i, \ \ A_i = u_i^1(1 - u_i^1)/4,$$

$$B_i = \left(\alpha\rho_1 u_i^1 - \lambda\right)/2, \ \ \Gamma_i^2 = B_i^2 + \alpha_i^2 A_i, \ \ C_i = \frac{\alpha^2 u_4^i + B_i + \Gamma_i}{\alpha^2 u_4^i + B_i - \Gamma_i^i}$$

$$M(t) = \frac{1 - C_i}{1 - C_i e^{-4\Gamma_i(T_i - t)}}e^{-\pi_i(T_i - t)}; \ \ \pi_i = 2(B_i + \Gamma_i) - u_1\rho_1\alpha_i;$$

$$K(t) = \frac{1}{1 - C_i}\left[-p_i y_i \Psi_i(t, \pi^i) + (p_i y_i C_i - p_i z_i)\Psi_i(t, \pi^i - 4\Gamma_i)\right]$$

$$z_i = -2\Gamma_i C_i/\alpha_i^2; \ \ y_i = -B_i + \Gamma_i/\alpha_i^2.$$

Proof. *See Appendix B.* ∎

We remark that jumps now appear in the expression of $\gamma(t)$ through the Laplace transform $\mathcal{L}_f(u_i^1)$. With a proper choice of the distribution of the r.v. Y, the transform can be analytically computed. This applies, for instance, when $Y \sim \mathcal{N}(q, v^2)$ is a Gaussian variable so that $\mathcal{L}_f(x) = \exp\left(qx + x^2 v^2/2\right)$. This choice provides a simple and intuitive jumps parametrization and, as we show below, it offers great flexibility in the calibration process.[4]

Our goal is the evaluation of a vanilla call option expiring at T and struck at K written on S, whose arbitrage price at time t is $Call_t(S_t, K, t, T) = \mathbb{E}^*\left\{\exp\left(-\int_t^T R(s, X_s)ds\right)(S_T - K)^+|\mathcal{G}_t\right\}$. To this aim, we use the knowledge of the Green function

$$G(y, \varsigma, \varphi, X_t; t, T) = \mathbb{E}^*\left\{\exp\left(-\int_t^T R(s, X_s)ds\right)e^{\varsigma \cdot X_T}\chi_{(\varphi \cdot X_T \leq y)}|\mathcal{G}_t\right\}, \tag{7}$$

and a number of well-known results on Fourier transforms for option pricing. In fact, as the following Proposition shows, $G(y, \varsigma, \varphi, X_t, t, T)$ can be determined from the knowledge of $\psi(u, X_t; t, T)$ and the pricing problem is then solved.

---

[4]In theory, a simpler solution would be to assume that Y follows a symmetric Laplace distribution with probability density $p(x) = \zeta\exp\left(-\zeta|x|\right)$, defined by a single parameter ζ. This choice, however, does not provide enough flexibility to match the observed smiles.

Proposition 3 *The price of the call option is given by* $Call_t(S_t, K, t, T) = G_1 - KG_2$, *with*

$$G_1 = G(-\ln K, \zeta_1, \nu_1, X_t, t, T), \tag{8}$$
$$\text{with } \zeta_1 = (1, 0, 0), \ \nu_1 = (-1, 0, 0), \ X_t = (\ln S_t, \eta_t, r_t)$$
$$G_2 = G(-\ln K, \zeta_2, \nu_2, X_t, t, T),$$
$$\text{with } \zeta_1 = (0, 0, 0), \ \nu_1 = (-1, 0, 0), \ X_t = (\ln S_t, \eta_t, r_t)$$

and

$$G(y, \varsigma, \varphi, X_t, t, T) = \frac{1}{2}\psi(\varsigma, X_t, t, T) - \frac{1}{\pi}\int_{(R^d)^+} \frac{1}{k} Im\left\{e^{-iky}\psi\left(\varsigma + ik\varphi, X_t, t, T\right)\right\} dk \tag{9}$$

Proof. *Duffie et al. (2000).* ∎

3   Gamma and Vega regimes

In this section we study the relationship between the dynamics Eq.(1) and the associated shape of the volatility surface. The goal is to provide evidence about the different role played by jumps and by stochastic volatility in explaining the observed smile in different portions of the time to expiry dimension. This result is instrumental in understanding the calibration methodology that will be introduced in Section 4.

The analysis of the moments of the asset distribution and their link with the shape of the smile has been already addressed in the literature. In particular, Backus *et al.* (1997) and (in a similar context) Zhang and Xiang (2005) show that if the smile is parametrized through a quadratic polynomial in the "modified moneyness" $m = \frac{\ln(F/K)}{\Sigma_{atm}\sqrt{T-t}} + \frac{\Sigma_{atm}}{2}$, where F is the option's underlying, Σ_{atm} is the at the money (ATM) BS volatility and K is the strike then, approximately, the BS implied volatility at varying m reads as

$$\sigma(m, \tau) \simeq \Sigma_{atm}\sqrt{\tau}\left(1 - \frac{\zeta_1(\tau)}{3!}m - \frac{\zeta_2(\tau)}{4!}\left(1 - m^2\right)\right), \tag{10}$$

where $\zeta_1(t)$ and $\zeta_2(t)$ are the skewness and the kurtosis of the logarithm of the underlying process, and $\tau = T - t$ is the time to expiry. This results from a Edgeworth expansion of the law of the log-asset price and holds for small values of Σ_{atm}. Formula (10) shows the tight link existing between shape of the smile and moments of the underlying asset process. In particular, when skewness and kurtosis are zero, the smile is flat at Σ_{atm} (as in the BS model). In addition, skewness (through the linear term in m) and kurtosis (through the quadratic term in m) act by respectively tilting and bending the smile.

For the sake of simplicity, we will assume in this section that dividends d_t vanish and that all model coefficients are constant since all conclusions are preserved (at a qualitative level) in the general

setup. Even in this simplified scenario the mathematical expression of the characteristic function $\Phi_t(x) := \mathbb{E}_t\left[e^{ix\ln S_T}\right]$ of the log-asset price (and *a fortiori*, that of the associated cumulants) is quite involved. For this reason, we provide explicit formulae in the pure jump case and analyze the general case numerically. When only jumps are present, Eq.(1) reduces to the Merton jump-diffusion model and the characteristic function $\Phi_t^{Mer}(x)$ reads as

$$\Phi_t^{Mer}(\theta) = \exp\left(t\varphi(\theta)\right)$$

with $\varphi(\theta) = i\left(r - \eta^2/2 - \mu\right)\theta - \eta^2\theta^2/2 + \xi\left[\exp\left(i\theta q - \theta^2 v^2/2\right) - 1\right]$. In Appendix C we re-derive this formula and show that the first four cumulants of $Z_t := \ln(S_t)$ are given by

$$\mathbb{E}\left(Z_t\right) \;=\; \Pi_1 = \left(r - \frac{\eta^2}{2}\right)t, \tag{12}$$

$$\mathbb{E}\left[\left(Z_t - \mathbb{E}\left(Z_t\right)\right)^2\right] \;=\; \Pi_2 = Var\left(Z_t\right) = \eta^2 t + \xi t\left(q^2 + v^2\right),$$

$$\mathbb{E}\left[\left(Z_t - \mathbb{E}\left(Z_t\right)\right)^3\right] \;=\; \Pi_3 = \xi t q\left(q^2 + 3v^2\right),$$

$$\mathbb{E}\left[\left(Z_t - \mathbb{E}\left(Z_t\right)\right)^4\right] - 3Var\left(Z_t\right)^2 \;=\; \Pi_4 = \xi t\left(q^4 + 6q^2 v^2 + 3v^4\right).$$

By recalling the expression of the skewness $\zeta_1(t) = \Pi_3\Pi_2^{-3/2}$ and the kurtosis $\zeta_2(t) = \Pi_4\Pi_2^{-2}$, we finally arrive at

$$\zeta_1(t) = \frac{1}{\sqrt{t}}\frac{\xi q\left(q^2 + 3v^2\right)}{\left[\eta^2 + \xi\left(q^2 + v^2\right)\right]^{3/2}}, \qquad \zeta_2(t) = \frac{1}{t}\frac{\xi\left(q^4 + 6q^2 v^2 + 3v^4\right)}{\left[\eta^2 + \xi\left(q^2 + v^2\right)\right]^2}. \tag{13}$$

These equations, in conjunction with Eq.(10), show that the impact of jumps on the volatility smile is restricted at very short times. In particular the jumps-induced smile convexity decays linearly with the time to expiry while smile skewness decays as $t^{-1/2}$. In a similar context to ours, Backus et al. (1997) demonstrate that, on the opposite, the impact of stochastic volatility persists on long-term smiles.

In order to provide quantitative support to this claim and to the different roles played by jumps and stochastic volatility, in Fig.1a we show the term structure of "butterfly spread" prices observed in the market on a generic trading day. A butterfly spread option with expiry T is a combined position in three call options and is usually measured by $H = \Sigma^{BS}(K^{atm}-\Delta) - 2\Sigma^{BS}(K^{atm}) + \Sigma^{BS}(K^{atm}+\Delta)$ where $\Sigma^{BS}(K)$ is the BS implied volatility at K. Butterfly spreads provide the simplest trading strategy to take a position in the smile's convexity. This is due to the fact that (apart from a multiplicative factor) H is the second derivative of the smile (thought of as a function of K) taken at K^{atm}. Thus, the higher H the larger the convexity and viceversa. In Fig.1a we compare the market butterfly with the one generated by our model Eq.(1) in the particular case where no jumps

are present and when all coefficients are time-homogeneous.[5] We see that at intermediate and at long maturities a stochastic volatility model explains the term structure of the smile convexity reasonably well (if mean reversion is carefully chosen) but it fails to do so at short maturities. This fact has a simple and intuitive explanation. Eq.(10) shows that the smile convexity is typically associated to the excess kurtosis of the log-asset returns distribution. Since in our case this is entirely generated by a diffusion process (the volatility η_t), it normally takes time to accumulate enough kurtosis starting from 0 at inception. This implies that a simple diffusive dynamics is not consistent with the market implied smile at short maturities. These observations suggest that jumps and stochastic volatility must be combined together since jumps have a big impact on short term convexity (Eq. (13)) and stochastic volatility on long term one.

These findings can be also interpreted from a different perspective. From a trading point of view, short-term and long-term smiles have a very different origin. Short term convexity is mainly associated to investors risk aversion to unexpected economic and sociopolitical events that might result in sudden jumps in the asset price. On the other side, long-term convexity is usually driven by the law of offer/demand induced by large investors, institutions and hedge/pension funds buying and selling in and out of the money options as a form of leveraged investment. Traders refer to these two regimes as "Gamma" and "Vega" trading, since the option Gamma (resp. Vega) risk is predominant at short (resp. long) maturities and in presence of large (resp. small) asset variations. These considerations indicate that the market smile is implicitly pricing the risk of large fluctuations (indeed, jumps) in the asset dynamics in the short term and of that of unpredictable (indeed, stochastic) asset volatility in the long end. The threshold between the two regimes will be denoted by T^*. This regime "switching" is therefore an intrinsic market characteristic and plays a fundamental role in our calibration approach.

4   SVJD models: the calibration problem

As above mentioned, model calibration consists of solving a multi-dimensional reverse engineering problem. As discussed by many authors it is impossible, in general, to determine a set of parameters such that market prices are exactly reproduced by any model.[6] Throughout this paper, by "model calibration" we then mean a methodology such that

1. The difference between market and model option prices is within the bid/ask spread.

2. The calibrated solution is statistically robust.

---

[5] In this case we recover the Stein and Stein (1991) model.

[6] Even from a pure financial point of view this is impossible to achieve. In fact, market imperfections and inefficiencies do not allow to identify option prices exactly (a bid/ask spread is always present).

Following Cont and Tankov (2004), ideally one would attempt to perform a model calibration by solving the following general non-linear optimization problem (N_S = n. of strikes, N_E = n. of option expires):

$$\{\varsigma_i\}^* = \arg\min_{\{\pi_i\}} \sum_{j=1}^{N_S} \sum_{k=1}^{N_E} w_{jk} \left| \Sigma\left(T_k, K_j^{(k)}; \{\pi_i\}\right) - \Sigma^{BS}\left(T_k, K_j^{(k)}\right) \right|^2 + \psi F(\{\pi_i\}), \qquad (14)$$

where $\{\pi_i\}$ is a set of free model parameters, Σ is the model-implied Black-Scholes volatility, Σ^{BS} is the market-implied Black-Scholes volatility, ψ and w_{jk} are weighting constants, $K_j^{(k)}$ is the j-th. strike for options expiring at T_k and $F(\{\pi_i\})$ is a convex regularization functional. The above minimization problem provides in theory a set of "optimal" free parameters $\{\pi_i\}^*$.

Formally, the methodology we introduce here shares some features in common with that proposal. In our case, however, because of the nature of our problem and because we aim at making the calibrated solution meaningful from a trading perspective, we add a number of additional constraints driven by statistical and risk-management criteria.

First, (condition C1) the influence of the jumps in the dynamics is confined at short times while stochastic volatility mainly acts at medium and long expiries to reflect the transition between "Gamma" and "Vega" regimes, as previously discussed. In this way we enforce smile to be generated by jumps for short maturities and by stochastic volatility for long maturities, allowing for a perfect disentanglement between the two noise sources. This assumption is crucial in the calibration process we address here but it is also beneficial in numerical implementations of PDE's for option pricing, as extensively discussed in Galluccio and LeCam (2005). Second, (condition C2) we favour solutions where the transition between the two regimes is smooth once the model has been calibrated to liquid instruments in order to guarantee a robust risk-management. In particular, jumps will be gradually "switched off" to avoid unreasonable discontinuities in across the two regimes. Third, (condition C2) we enforce solutions where the term structure of the calibrated coefficients is as time-homogeneous as possible. From a statistical point of view, such models are more robust and realistic than those models where all parameters are heavily time-dependent. In addition, when parameters are constant the dynamics of the volatility surface is closer to stationarity and then consistent with empirical observations. This is also beneficial on the risk-management side (Rebonato (2000)).

4.1 Constrained optimization

In our empirical study we consider options up to 5 years time to expiry since the market is rather illiquid at longer maturities.[7] In our empirical tests we consider data from the EuroStoxx 50 equity

---

[7] However, extending our approach beyond 5 years is straightforward.

index, whose ATM volatility matrix is given in Fig 1b. [8] Similar studies conducted on other indeces (S&P 500, FTSE 100, DAX and CAC40) provide similar results to the ones presented here and are available upon request. Two well-known features can be noticed. First, smile convexity is a decreasing function of time to maturity and is extremely high for short maturities. Second, the smile shape is not symmetric around the at-the-money (ATM) strike. In other words, the smile is both convex and "skewed".

In any time interval $[T_{i-1}, T_i)$ between two consecutive option expires our model is specified by a set of 7 independent constant coefficients: the volatility mean reversion level $a(t)$, the volatility of volatility (volvol) $\alpha(t)$, the constant volatility mean reversion rate λ_η, the asset-volatility correlation $\rho_{1,2}(t)$, the stochastic jumps intensity $\xi^{(0)}(t)$, the jumps average $q(t)$ and, finally, the jumps variance $v^2(t)$, with $t \in [T_{i-1}, T_i)$. Obviously, any attempt to perform a global calibration on this 7-dimensional manifold is doomed to failure.[9] Understanding the impact of each single parameter on the shape of the smile is instrumental to the problem's solution. To simplify our discussion in the beginning we will not consider the presence of jumps.

In a pure stochastic volatility framework, the role of coefficients $a(t), \alpha(t)$ and $\rho(t)$ is indeed well established (Hagan *et al.* (2002)). For reader's convenience, we briefly summarize the main points here. At the leading order in the volvol $\alpha(t)$, the at-the-money (ATM) volatility is completely specified by $a(t)$. Thus, $a(t)$ mainly affects the global level of the smile but has little impact on its overall shape. The Equity-IR correlation $\rho(t)$ affects the asymmetry of the smile (or "skew") around the ATM point. Increasing correlation means that assets prices tend to increase when volatility increases and viceversa. Then, out-of-the money call and in-the-money put options become more expensive. The net effect is that implied Black volatilities at high strikes increase while Black volatilities at low strikes decrease, so that the smile takes a positively skewed shape. Similarly, the more negative the correlation, the more negatively skewed the smile. Finally, the volatility of volatility (or "volvol") $\alpha(t)$ rules smile convexity: the higher $\alpha(t)$ the more convex the smile and viceversa. As a secondary effect, the volvol influences the total variance of the process and (just like $a(t)$) it impacts the global level of the smile.[10] Results displayed in Fig.2 (where the influence of

[8] The smooth surface has been obtained by using a BNP Paribas propietary arbitrage-free volatility interpolation algorithm that is capable of matching quoted market prices within their bid-ask spread. Alternative parametrizations have been tried, like the one proposed by Fengler (2005), but results are not significantly affected by this choice. Data correspond to Feb 2nd 2004.

[9] The causes for this are: *i)* the non-linear optimization problem is not strictly convex and, *ii)* some of the model parameters are quasi-degenerate. This implies that: *i)* the objective function has many local minima and, *ii)* it is almost flat in the maximum gradient direction so that both convergence and robustness are at risk (see Cont and Tankov (2004)).

[10] As shown in Fig. 2, by increasing the level of α the ATM volatility increases, as one would intuitively expect. On the opposite, in affine models (Heston) the ATM volatility is inversely proportional to the volvol coefficient. This

the three parameters on the smile are compared) support this view. The picture shows that any among $a(t), \alpha(t), \rho(t)$ plays a different role from the others in explaining possible smile movements. As shown below, this is a fundamental property in the calibration process.

A fundamental role is played by the mean reversion λ which is assumed constant in our model. A mean reverting Ornstein-Uhlenbeck process converges to its ergodic measure after a "characteristic time" $\tau = 1/\lambda$. The quadratic variation of the asset process in Eq.(1) is given by $Q_t = \eta_t^2 S_t^2$. Conditionally to any volatility path, Q_t evolves like the square of a lognormal process, so it increases indefinitely on average as times goes to infinity. Because η is an ergodic process, its variance converges to an asymptotic value after a time $t \simeq \tau$. Therefore, unconditionally to any realization of η_t, the average of Q_t is asymptotically dominated by S_t^2 and the effect of the stochastic volatility becomes negligible at large times. This shows that the presence of the volatility mean reversion λ provides a simple and effective tool to "fine tune" the rate of decrease of the smile convexity at long maturities, as observed in the market (see Fig. 1a, 1b). In models where $\lambda = 0$, like the one proposed in Hagan et $al.$ (2002), it is necessary to artificially impose a decreasing term structure of the volvol to ensure market consistency. We also remind that λ cannot be statistically inferred from historical time series since it is not measure change invariant.

When also jumps are present, the picture becomes much more complex. In fact, although jump parameters $\xi^{(0)}(t)$, $q(t)$ and $v(t)$ play altogether a role similar to $a(t), \alpha(t)$ and $\rho(t)$ in explaining possible smile deformations, their influence on the smile shape cannot be as nicely identified as before since, differently from above, any parameter now plays a "mixed" role. The reason for this can be best understood in the simplified scenario provided by the Merton model. In this case, the cumulants of $Z_t := \ln(S_t)$ are given by Eq.(12). In a BS setting we have $\xi = 0$, so that only mean and variance are different from zero, as expected. In this case the implied smile would be flat and equal to η. When $\xi \neq 0$ second, third and fourth cumulant play altogether a decisive role in moving the implied volatility away from its BS level. Eq. (12) and Eq. (10) show that smile deformations around the BS level can be attributed to either the stochastic intensity ξ, the jumps average q or the jumps standard deviation v. Since different triplets $\{\xi, q, v\}$ can in this case be associated to almost identical smile curves the inverse problem (i.e. determining a unique triplet from a given smile or a set of smiles) is in general ill-defined (see also Cont and Tankov (2004)). When this happens we will refer to the associated parameters as being "degenerate". This identification problem affects a number of studies, including Andersen and Andreasen (2000) and Bakshi et $al.$ (1997). A first step towards the problem solution consists of imposing the above mentioned constraints on the problem. This goes as follows.

1. First (condition C1) we split the calibration problem in two consecutive steps. Initially, the

---

unrealistic behaviour makes affine models less appealing from a trading perspective than affine-quadratic ones.

term structure of diffusion coefficients is kept at a (trial) constant level $\{a_0, \alpha_0, \rho_0\}$ at $t < T^*$, while the smile is calibrated by only adjusting the jump coefficients as shown below. Thus, at $t < T^*$, stochastic volatility does not play any role and the smile is almost entirely generated by jumps. Once jumps calibration has been achieved, we calibrate the remaining smiles by adjusting $\{a(t), \alpha(t), \rho(t)\}$ at $t \geq T^*$ while keeping the jump parameters "frozen" at their previously calibrated levels. In this way jumps and stochastic volatility are not "mixed up" in the optimization procedure and some degeneracies (like those described in Cont and Tankov (2004) are eliminated.

2. Second, (condition C2) we impose that the switch between the two regimes at $t < T^*$ and $t \geq T^*$ is smooth. To meaningfully achieve this, we assume that the stochastic intensity $\xi(t)$ is a continuous (possibly differentiable) strictly decreasing function, i.e. *i)* $\xi(T^*) \in \mathcal{C}^0$, *ii)* $\xi(T^*) = 0$, *iii)* $\xi(t) > \xi(t')$ for $t < t'$. Also, the initial set $\{a_0, \alpha_0, \rho_0\}$ is adjusted to minimize the jump in the value between the two regimes.

3. Third, (condition C3) the volatility mean reversion λ is adjusted to ensure that the calibrated set $\{a(t), \alpha(t), \rho(t)\}$ is as time-homogeneous as possible.

5 Numerical implementation

5.1 Data set

Empirical tests on calibration are performed by using EuroStoxx 50 data on Feb, 2nd 2004. These data correspond to: a) a set of EuroStoxx forward prices for maturities up to 5Y; b) The whole EuroStoxx volatility surface for times to expiry up to 5 years. Both sets correspond to mid-market quotations and have been provided by internal BNP Paribas proprietary systems according to the market information prevailing at that time.

5.2 Jumps calibration

In the paper above mentioned, Cont and Tankov (2004) show that in Merton's model the market-implied calibration of stochastic intensity and diffusion's volatility from a single smile is impossible since the two parameters are degenerate. In addition, the optimization problem is not convex and many different minima exist. Although we never attempt to simultaneously calibrate jumps and

diffusion, Cont and Tankov remark equally applies if one attempts to calibrate a single smile in a Merton model.[11]

Unfortunately, Cont and Tankov (2004) regularization procedure for generic Lévy processes cannot be directly applied to our problem for two reasons. First, we aim at making the model consistent with the whole volatility surface and in doing so we need a term structure of model coefficients. Second, more fundamentally, the process defined by Eq. (1) is not a Lévy process since its increments are independent but not stationary.

Instead of attempting a calibration to each smile individually, we propose an alternative approach aimed at calibrating the whole set of smiles up to (and including) T^* by assuming a suitable parametric form for $\xi(t)$ for given (constant) jumps average q and standard deviation v.

More precisely, we aim at calibrating the three smiles corresponding to options of 1 month, 3 months and 6 months expiry, so that $T^* = \frac{1}{2}$ and introduce a specific form of $\xi(t)$ according to condition C2. We find convenient to define $\xi(t)$ from $\xi(t) = \left(\omega - \frac{\omega}{t}\right)^\delta$ thanks to the fact that (as shown below) it provides excellent calibration results with a minimal number of free parameters, ω and δ. In the applications, $\xi(t)$ must be discretized as follows

$$\xi(t) = \xi(t,\omega,\delta) = \begin{cases} \xi_1(t,\omega,\delta) = \left(\omega - \frac{12\omega}{1}\right)^\delta & \text{for} \quad t \in \left[T_0, \frac{1}{12}\right) \\ \xi_2(t,\omega,\delta) = \left(\omega - \frac{12\omega}{3}\right)^\delta & \text{for} \quad t \in \left[\frac{1}{12}, \frac{1}{4}\right) \\ \xi_3(t,\omega,\delta) = \left(\omega - \frac{12\omega}{6}\right)^\delta & \text{for} \quad t \in \left[\frac{1}{4}, \frac{1}{2}\right) \end{cases} , \quad \xi(t) = 0, \text{ if } t > \frac{1}{2}.$$

With this parametrization, $\xi(t)$ is completely specified once the two constants ω and δ have been assigned.

Initially, suppose that δ is given and that coefficients $\{a_0, \alpha_0, \rho_0\}$ have been fixed to a trial level. Then, the objective function

$$G_{\omega,\delta}(q,v) := \sum_{j=1}^{N_S} \sum_{k=1}^{N_E} w_{jk} \left| \Sigma\left(T_k, K_j^{(k)}; q, v, \omega, \delta\right) - \Sigma^{BS}\left(T_k, K_j^{(k)}\right) \right|^2$$

can be plotted as a function of (q,v) for different values of ω and the results are shown in Fig.3. Two things are worth noticing. First, the function $G_{\omega,\delta}(q,v)$ is strictly convex around a minimum independently of the chosen ω (i.e. independently of the level of the jumps intensity - similar results hold at varying δ). Second, $G_{\omega,\delta}(q,v)$ is not just locally convex around the (single) minimum but, instead, its convex portion extends to a wide region in the (q,v) space where no other local minima exist.

Table 1 provides a more rigorous argument to support this intuitive picture. There, we show the results of a simple Levenberg-Marquardt numerical minimization algorithm on $G_{\omega,\delta}(q,v)$ for different

---

[11] In other words, the triplet $\{\xi, q, v\}$ is degenerate as above pointed out.

choices of the initialization set (q^{init}, v^{init}). In other words we solve the problem $(q^*(\omega), v^*(\omega)) = \arg\min_{q,v} G_{\omega,\delta}(q,v)$ with ω given (in this example $\omega = -0.3$ and $\delta = 1$). We note that the method is very accurate for a wide range of initial conditions, although the low convexity of the surface around the minimum demands a good level of numerical accuracy. The optimization algorithm drifts away from the convex region (and provides no sensible result) only when q^{init} is very badly chosen at inception.

These results have been obtained by fixing the stochastic intensity to a given value. One would be tempted to try estimating the optimal ω jointly with q and v, as a solution of the global least-squares problem $(q^*, v^*, \omega^*) = \arg\min_{q,v,\omega} G_{\omega,\delta}(q,v)$, for a given δ. This way of proceeding was suggested, for instance, in Andersen and Andreasen (2000) and Bakshi *et al.* (1997). Unfortunately, this direct method is not usually viable because in general $\{q, v, \omega\}$ is a degenerate triplet. Table 2 shows that, no matter how ω is selected, it always exists an optimal couple $(q^*(\omega), v^*(\omega))$ such that the objective function $G_\omega(q^*(\omega), v^*(\omega))$ attains the same minimum value. Thus, $G_\omega(q^*(\omega), v^*(\omega))$ is almost flat in ω and the inverse (minimization) problem on ω is ill-posed.

To overcome this potential issue, we introduce a convex penalization term, as in Cont and Tankov (2004). In the present case, however, the relative-entropy (or Kullback-Leibler distance) method used by the authors cannot be applied since ours is not a Lévy process. We then introduce a quadratic penalization term so that the whole inverse problem reads in general as

$$(q^*(\omega), v^*(\omega)) \;=\; \arg\min_{(q,v)} \sum_{j=1}^{N_S} \sum_{k=1}^{N_E} w_{jk} \left| \Sigma\left(T_k, K_j^{(k)}; q, v, \omega, \delta\right) - \Sigma^{BS}\left(T_k, K_j^{(k)}\right) \right|^2, \quad (15)$$

$$(\omega^*, q^*, v^*) \;=\; \arg\min_{\omega} \left[(q^P - q^*(\omega))^2 + (v^P - v^*(\omega))^2 \right]. \quad (16)$$

Simply stated, the problem consists of determining ω^* such that the couple (q^*, v^*) is as close as possible to a "prior" couple $\left(q^P, v^P\right)$ arbitrarily chosen. A good criterion is to estimate jumps average and standard deviation from historical data series and to assign $\left(q^P, v^P\right)$ accordingly. This choice has the advantage that the optimal solution (ω^*, q^*, v^*) is a guarantee that the market-implied model stays "close" (in the probability measure space) to the historically estimated one. [12]

Although the above methodology provides a unique and stable solution, the resulting errors between market and model implied smiles can sometimes be beyond the typical volatility bid/ask spread (i.e., 1% in lognormal units) if δ is badly chosen. In fact, calibration accuracy sensibly depends on the choice of $\xi(t)$. When $\delta = 1$, the simple hyperbolic shape associated to $\xi(t)$ does not

---

[12] As an important remark, we note that the request that couples $\left(q^P, v^P\right)$ and (q^*, v^*) are close is a well defined problem in probabilistic terms. In fact, this is equivalent to enforce that the market-implied jumps probability distribution is as close as possible to the historical (objective) one. Girsanov theorem ensures that the jumps distribution is indeed invariant under changes of probability measure (in our case from the objective to the risk-neutral and viceversa).

provide good results while precision can be improved by modifying δ (i.e., by changing the speed of convergence of $\xi(t)$ to 0 as $t \to T^*$).

The main question is how does the choice of δ impact the above picture and whether any complex interplay between the choice of δ and that of ω exists destroying the above picture. To study this problem in Fig. 4a we report the results of an optimization performed on δ, i.e., for a given set $\{a_0, \rho_0, \alpha_0, \omega\}$ we solve the problem

$$\delta^* = \arg\min_\delta G_{\omega,\delta}(q, v), \tag{17}$$

and we look at how this optimal solution is affected by changes in the trial set $\{a_0, \rho_0, \alpha_0, \omega\}$. Let $G^* = G_{\omega,\delta^*}(q, v)$ be the minimum of the objective function. We consider a typical set of parameters as our base case scenario (Series 1). All other curves in Fig.4a are obtained from the base one by applying a large shock in one single parameter among those in the set $\{a_0, \rho_0, \alpha_0, \omega\}$. Results can be summarized as follows, *i)* δ^* and G^* are not sensibly affected by a shock in ω and a_0, *ii)* a shock in ρ_0 affects calibration accuracy (G^*) but has almost no impact on the optimal δ^*, *iii)* a shock in α_0 affects the optimal δ^* but has almost no impact on the calibration accuracy (G^*). In addition, Fig4b further investigates the dependency of δ^* on α_0. The picture shows that the functional relationship between δ^* and α_0 is linear only for small values of α_0.

These results indicate that the choice of the optimal δ^* is almost entirely dependent on the chosen value of α_0. in other words, once the initial volvol parameter has been set at inception, one can determine an optimal δ^* for any given set of $\{a_0, \rho_0, \omega\}$ by solving problem (17). Even though after calibration to the smile ω will be different, δ^* is guaranteed to stay very close to the new optimal value. The conclusion is that, once the initial volvol α_0 assigned, jumps can be efficiently calibrated to the market.

5.3   Stochastic volatility calibration

In the last section we showed that it is possible to calibrate the jumps once an initial set $\{a_0, \rho_0, \alpha_0\}$ of SV parameters has been assigned. We now address the issue of how optimally select the triplet $\{a_0, \rho_0, \alpha_0\}$. In general, unfortunately, there is no unique answer to this, i.e., there is no obvious way to decide how to fix *a priori* the triplet before any calibration is attempted. Despite finding a unique set is impossible, we introduce a constructive approach that allows calibrating the model efficiently. The good news is that, once this "pre-calibration" study has been carried out, the initial set $\{a_0, \rho_0, \alpha_0\}$ can be taken as granted and one can avoid readjusting it too often (see Section 6 on this point).

The approach we use is based on the intuitive observation that if $\{a_0, \rho_0, \alpha_0\}$ has been badly chosen in the "Gamma" region, the terminal (or cumulative) variance, skewness and kurtosis make

18

impossible to calibrate the remaining smiles for options expiring after T^*. To better understand the link between $\{a_0, \rho_0, \alpha_0\}$ and terminal variance, skewness and kurtosis we have performed a number of tests. As an illustration, we consider three sets of parameters $\left\{a_0, \rho_0, \alpha_0^{(i)}\right\}$, with $i = 1, 2, 3$ corresponding to $\alpha_0^{(1)} = 10\%, \alpha_0^{(2)} = 30\%, \alpha_0^{(3)} = 50\%$. For each given set, the model is then calibrated to all smiles up to 6 months expiry as previously explained. Finally, all smiles with expiry beyond 6 months are generated. This test is aimed at measuring the terminal variance, skewness and kurtosis generated by the initial set $\{a_0, \rho_0, \alpha_0\}$ (and by jumps) and to check whether they are compatible with the market prices. Table 3 gathers the results. There, we show the difference between market and model implied volatility for smiles at 1Y, 2Y, 3Y and 5Y induced by the calibration at shorter maturities. Results show that, as one might expect, if α_0 is assigned a too high value at the beginning, model-implied smiles at 1Y expiry are inconsistent with the market. In our case, if $\alpha_0 = 50\%$ the terminal variance at 1Y is too high: no matter how $\{a(t), \rho(t), \alpha(t)\}$ are selected in the time interval $[6M, 1Y]$ hitting the market smile is impossible. In theory, this could still be achievable by allowing $a(t)$ to take large negative values but this solution is financially meaningless.

We can formally define, for a given a_0, a "critical" value $\widehat{\alpha}_0$ of the volvol coefficient as follows:

$$\widehat{\alpha}_0 = \sup\{\alpha_0 : \text{all smiles are calibrated within the bid/ask spread}; a(t) > 0, \alpha(t) > 0\}.$$

In other words, $\widehat{\alpha}_0$ is the maximum value of the volvol such that, other parameters being given, all smiles can be matched by means of a sequence $\{a(t), \rho(t), \alpha(t)\}$ by keeping both $a(t)$ and $\alpha(t)$ positive. In the next section we show that calibration can be achieved for a wide range of α_0 by performing a full volatility surface calibration (i.e., $\widehat{\alpha}_0$ is normally very high). We finally remark that the above picture is not significantly altered by ρ_0 once its sign has been properly assigned (smiles are usually negatively skewed implying ρ_0 should be always negative). These two properties are extremely important since they indicate that $\{a_0, \rho_0, \alpha_0\}$ can be assigned with great flexibility without compromising the quality of the calibration.

We now assume that $\{a_0, \rho_0, \alpha_0\}$ has been fixed and that an optimal set $\{\omega^*, q^*, v^*, \delta^*\}$ has been determined accordingly. The next step consists of keeping these parameters fixed and calibrate the remaining part of the volatility surface at $t \geq T^*$ by adjusting the stochastic volatility coefficients $a(t), \alpha(t)$ and $\rho(t)$. In other words, starting from the first smile after T^*, we proceed recursively and at each interval in between consecutive smiles we attempt solving the following problem

$$(\alpha^*(t), \rho^*(t), a^*(t)) = \arg \min_{a(t), \alpha(t), \rho(t)} \sum_{j=1}^{N_S} u_{jk} \left| \Sigma\left(T_k, K_j^{(k)}; \alpha(t), \rho(t), a(t)\right) - \Sigma^{BS}\left(T_k, K_j^{(k)}\right)\right|^2 \tag{18}$$

$$\text{for } t \in [T_k, T_{k+1}), \ k = 1, \cdots, L-1, \ \ T_1 = T^*, \tag{19}$$

where $L-1$ is the number of smiles with expiry strictly larger than T^*. As above anticipated, this problem is well posed since $\{\alpha^*(t), \rho^*(t), a^*(t)\}$ are not degenerate.

Finally, λ can be fine tuned so that the calibrated term structure of the volvol $a^*(t)$ is as constant as possible. Finding the optimal λ^* can be easily achieved by solving the following least squares optimization,

$$\lambda^* = \arg\min_\lambda \left[\sum_{j=1}^{L-1} \left(\alpha_j^*(\lambda) - \alpha_{j+1}^*(\lambda) \right)^2 \right], \tag{20}$$

where vector $(\alpha_1^*(\lambda), \alpha_2^*(\lambda), \cdots, \alpha_L^*(\lambda))'$ comprises the piecewise constant term structure of $\alpha^*(t)$, for a given value of λ. In short, λ^* is the volatility mean reversion that corresponds to the least oscillating calibrated term structure $a^*(t)$. As before, the good news is that once optimization problem (20) has been solved it is possible to keep λ^* fixed without significantly altering the result in future calibrations. In this way, we empirically established that optimal values for λ are in the interval $[0.4, 0.7]$, independently on the chosen market.

6   Calibration algorithm and numerical results

6.1   The algorithm

In our example, we calibrate the model on a set of increasing time to expiry options, corresponding to $T_1 = 1$ month, $T_2 = 3$ months , $T_3 = 6$ months, $T_4 = 1$ year, $T_5 = 2$ years, $T_6 = 3$ years, $T_7 = 5$ years. T_0 is the observation date. We define a threshold T^* between the two regimes. In our test we fix $T^* = 0.5$, but similar results can generally be found by fixing T^* anywhere between 3 months and 1 year. We call $\mathcal{T}_<$ the set of option expiries shorter than T^*, that is $\mathcal{T}_< := \{T : T \leq T^*\}$, with $|\mathcal{T}_<| = M_<$. Similarly, $\mathcal{T}_> := \{T : T > T^*\}$, with $|\mathcal{T}_>| = M_>$. The calibration algorithm is based on a recursive procedure that, starting from the shortest expiry T_1, proceeds as follows.

1. Choose a "trial" value for λ. Then run a "pre-calibration" test as described in the previous section to determine, for a given a_0, the critical volvol coefficient $\widehat{\alpha}_0$. Finally, determine an initial set $\{a_0, \rho_0, \alpha_0\}$ by fixing $\alpha_0 < \widehat{\alpha}_0$.

2. Assign a value to ω and determine δ^* such that the calibration errors in $\mathcal{T}_<$ are minimal by solving (17), while $\{a_0, \rho_0, \alpha_0\}$ are kept fixed at their initial values.

3. With δ^* fixed, calibrate the jumps coefficients in $\mathcal{T}_<$ by solving the two optimization problems (15) and (16). In other words, for a given ω solve (15) so that ω^* corresponds to the single ω such that the quadratic distance between $\left(q^H, v^H \right)$ and (q^*, v^*) is minimal. This procedure provides an optimal set $\{\omega^*, q^*, v^*\}$.

20

4. Next, determine the diffusion coefficients by calibrating the smile in the interval $\mathcal{T}_>$. Keep jump parameters frozen at the previously calibrated values, then proceed recursively by sequentially calibrating the remaining smiles starting from the one associated to options with the shortest maturity in $\mathcal{T}_>$. This is done by solving the problem (18) and provides an optimal term structure of SV coefficients $\{\alpha^*(t), \rho^*(t), a^*(t)\}$ for $t \geq T^*$.

5. If the prior mean reversion rate $\lambda^{(0)}$ has been badly chosen, step 4) might provide a too rapidly increasing or decreasing term structure $\{\alpha^*(t), \rho^*(t), a^*(t)\}$, as previously discussed. We then proceed (condition C3) by solving the problem (20): choose a new $\lambda^{(1)}$ and restart from step 1). Then proceed recursively until the optimal λ^* has been found.

As already discussed, it is not necessary to perform all five steps at any new calibration. For instance δ^* and λ^* are very stable with time and, once estimated, can be occasionally readjusted. In addition, if one is not interested in ensuring that couple (q^*, v^*) is close to the historically estimated one, step 3) can be neglected in the algorithm.

Extensive empirical studies performed on S&P and EuroStoxx data in the time period spanning the years 2002 - 2005 (not reported here) suggest that the optimal λ^* must lie in the interval $[0.4, 0.7]$, as above mentioned. Interestingly, this is in contrast with the most recent findings of λ based on historical data series (Eraker *et al.* (2000)) that assign to the mean reversion rate much lower values: $\lambda \in [0.013, 0.025]$. This indirectly indicates that the market price of volatility risk is significant in SVJD models.

6.2   Numerical results

In all our tests we calibrate each smile by selecting three liquid options (i.e., $N_S = 3$) struck at K_i, $i = 1, 2, 3$. These correspond to the at-the-money forward option (K_2), one in-the-money option (K_1), and one out of-the-money option (K_3). All results are however independent from N_S. To ensure selection of liquid (and meaningful) points, for every expiry T_i we fix K_1 (resp. K_3) to a fixed number l of standard deviations from the ATM strike, i.e. $K_1 = K_0 - l\sigma^{ATM}\sqrt{T}$, $K_3 = K_0 + l\sigma^{ATM}\sqrt{T}$. Here, σ^{ATM} is the at-the-money Black implied volatility[13]. Scale parameter l is fixed to 1, although larger values can be considered in case one needs to calibrate wider portions of the smile. Fixing l to a too large value must be avoided since far out of the money or in the money options are illiquid. Spot interest rate is 0.033 and dividends are 0. All weights u_{jk}, w_{jk} are fixed to 1.

Tables 4 and 5 show results of a typical calibration on the EuroStoxx volatility matrix for $\alpha_0 = 10\%$ and $\alpha_0 = 30\%$, while $a_0 = 6\%$ is given. We see that (given two quite different values

---

[13]Alternatively one could select K_1 (resp. K_3) as the strike corresponding to 25% (resp. 75%) of the ATM option's delta. This choice is common practice in FX markets.

of the initial volvol coefficient) the model is capable of efficiently calibrating the whole volatility surface since all errors are within the bid ask spread (typically around 1%). However, when α_0 is set to 10%, the resulting term structure of $\alpha(t)$ is much more irregular than in the second case with $\alpha_0 = 30\%$. In particular, a jump in $\alpha(t)$ exists between the 6 month and the 1Y smile. When this happens, adjusting λ can help in partially reducing the oscillations of $\alpha(t)$, but to completely remove the jumps it is usually necessary to adjust α_0, as Table 5 demonstrates. Interestingly, the term structure of $a(t)$ is very regular and smooth in all cases. In addition, the choice of $\xi(t)$ ensures that the jumps gradually vanish in approaching T^* (condition C1), so that the transition between the two regimes is smooth (condition C2).

Statistical robustness of the calibrated solutions is addressed next. Table 6 shows the output of a calibration with $\alpha_0 = 10\%$ after a shock of 1% has been applied uniformly across the whole volatility matrix. This is the order of magnitude of the shocks occurring between two consecutive days in the market. In order to determine whether our algorithm is robust, we keep all parameters at the same level before the shock (in particular we do not revaluate $a_0, \rho_0, \alpha_0, \lambda$ and ω). Results show that the calibration accuracy is unaffected by the shock and, more importantly, that the new set of calibrated coefficients is extremely close to the old one. This clearly demonstrates that the optimal solution is stable and that meaningful risk-management is possible is this framework.

In Fig 5 we plot the calibrated volatility surface. A comparison with Fig.1b (the original market volatility surface) shows that the calibration errors are always very small respect to the market bid/ask spread.

As an important final remark, we note that the calibrated correlation term structure $\rho(t)$ tends to converge to -1 at large maturities. This clearly indicates that the market-implied skewness is larger than the one predicted by a SVJD model and is in contrasts with correlation estimations based on historical data.[14] For instance, Eraker *et al.* (2000) report that ρ varies typically in $[-0.4, -0.5]$ for the S&P 500 and in $[-0.3, -0.4]$ for the Nasdaq 100 based on statistical estimations. To take into account these features, dynamics Eq.(1) must be generalized. From a statistical point of view there is strong evidence of presence of jumps in volatility (Eraker *et al.* (2000)). Alternatively, these effects could be accounted for by an extension of the present model to include more complex forms of local volatility (Hagan *et al.* (2002)).

---

[14] Although the tests presented here refer to the EuroStoxx 50, the same conclusion applies to other indices, including S&P 500 and FTSE 100.

7　Conclusions

In this paper we have introduced a market-implied calibration technique that can be used for certain classes of stochastic volatility jump diffusion models. In particular, we focused on a model within the linear-quadratic class since generalizations to include stochastic interest rates and multi-currency markets are viable in this setting. We have demonstrated that calibration of the entire volatility surface is possible in this framework and we have studied both precision and stability of the algorithm. Our empirical study indicates, at the same time, that the algorithmic complexity associated to the calibration of more general SVJD models might represent a major problem in the applications. Further theoretical and numerical developments in this direction are left to future research.

References

[1] Aït-SahaliaY. and R. Kimmel (2004) "Maximum likelihood estimation od stochastic volatility models", Working Paper, Princeton University.

[2] Andersen L. and Andreasen J. (2000), "Jump-Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Option Pricing" , *Review of Derivatives Research,* 4, 231-262.

[3] Andersen T., Benzoni, L. and Lund, J., (2002), "An Empirical Investigation of continuous-time equity return models", *Journal of Finance,* 57, 1239-1284.

[4] Backus D., Foresi S., Li K. and Wu L. (1997) "Accounting for Biases in Black Scholes", Working Paper, Stern School of Business, New York University.

[5] Bakshi G., Cao C. and Chen Z. (1997) "Empirical Performance of Alternative Option Pricing models", *Journal of Finance,* 52, 2003 - 2049.

[6] Bates, D. (1996), "Jumps and Stochastic Volatility: Exchange Rates Processes Implicit in Deutsche Mark Options", *Review of Financial Studies,* 9, 69-107.

[7] Bates D.S. (2000), "The crash of '87: was it expected ? The evidence from option markets", *Journal of Finance*, 46, 1009-1044.

[8] Chernov M., Gallant R., Ghysels E. and Tauchen G, (1999) "A New Class of Stochastic Volatility Models with Jumps: Theory and Estimation", Working Paper, Pennsylvania State University.

[9] Chernov M., Gallant R., Ghysels E. and Tauchen G (2003), "Alternative Models for Stock Price Dynamics", *Journal of Econometrics*, 116.

[10] Cont. R., Tankov. P., (2004), "Non-parametric calibration of jump-diffusion option pricing models", *Journal of Computational Finance*, 7, 1-43.

[11] Craine R, Lochstoer L.A. and Syrtveit K., "Estimation of a Stochastic-Volatility Jump-Diffusion model", Working Paper, University of California at Berkeley (2000).

[12] Deelstra, G., Ezzine, A. and Janssen J. (2003) "Option Valuation in a Non-Affine Stochastic Volatility Jump-Diffusion Model", Working Paper, University of Bruxelles.

[13] Detlefsen K., (2005), "Hedging Exotic Options in Stochastic Volatility and Jump Diffusion Models", Working Paper, Humboldt University Berlin.

[14] Derman, E., and Kani, I., (1994), "Riding on a Smile", *Risk Magazine,* February, 32-39.

[15] Di Graziano G., and Galluccio, S., (2005), "Evaluating Tracking Hedging Errors for General Processes: Theory and Experiments", Working Paper, BNP Paribas and University of Cambridge.

[16] Duffie D., Pan J. and Singleton K. (2000) "Transform Analysis and Asset Pricing for Affine Jump-Diffusions", *Econometrica,* 68, 1343 - 1376.

[17] Dumas B., Fleming J. and Whaley R.E. (1997), "Implied Volatility Functions: Empirical Tests", *Journal of Finance,* 53, 2059 - 2106.

[18] Dupire, B. (1994), "Pricing with a Smile", *Risk Magazine,* January, 18-20.

[19] Eraker B., Johannes M. and Polson M., (2003) "The Impact of Jumps in Volatility and Returns", *Journal of Finance,* 58, 1269 - 1300.

[20] Fengler M., (2005) "Arbitrage-free smoothing of the implied volatility surface", Working Paper, Humboldt University Berlin.

[21] Galluccio, S., and Le Cam, Y., (2005), "Modelling Hybrids with Jumps and Stochastic Volatility", Working Paper, BNP Paribas and University of Evry.

[22] Hagan, P., Lesniewski, A., Kumar, D., Woodward, D., (2002) "Managing Smile Risk", *Wilmott Magazine*, September, 84-108.

[23] Heston S., (1993) "Closed-form Solutions for Options with Stochastic Volatility, with Applications to Bond and Currency Options", *Review of Financial Studies*, 6, 327-343.

[24] Hull, J, and White A., (1987), "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance,* 42, 281- 300.

[25] Jones C.S., (2003) "The Dynamics of Stochastic Volatility: Evidence from Underlying and Option Markets", *Journal of Econometrics* 116, 181-224.

[26] Merton, R., Theory of Rational Option Pricing, (1973), *Bell Journal of Economics and Management Science,* 4, 141-183.

[27] Medvedev, and Scaillet, O., (2005), Working Paper, University of Geneva and FAME.

[28] Pan J., (2002), "The Jump Risk Premia Implicit in Options: Evidence from an Integrated Time Series Study", *Journal of Financial Econometrics,* 63, 3-50.

[29] Peng, C. and Scaillet O. (2004), "Linear-Quadratic Jump-Diffusion Modelling with Application to Stochastic Volatility", Working Paper, University of Geneva and FAME.

25

[30] Piazzesi, M., (2003), "An Econometric Model of the Yield Curve with Macroeconomic Jump Effects", NBER Working paper 8246.

[31] Rebonato, R., (2000), Volatility and Correlation, John Wiley and Sons, NewYork.

[32] Stein E.M. and Stein J.C., (1991), "Stock Price Distributions with Stochastic Volatility: An Analytic Approach", *Review of Financial Studies*, 4, 727-752.

[33] Zhang, J.E., and Xiang, Y., (2005), "Implied Volatility Smirk", Working Paper, University of Hong Kong.

A   Appendix

The proof is similar to Duffie et al. (2000) and Peng and Scaillet (2004). We start by identifying the coefficients of the dynamics Eq.(4). In a more compact notation we have $dX_t = \zeta(X_t,t)dt + \Sigma(X_t,t)dW_t + dZ_t$. The triplet of characteristics of this semimartingale are quadratic functions of the state variables since, by direct inspection,

$$\zeta(x,t) = K_0(t) + K_1(t)x + \sum_{k,l=1,2} K_2^{(k,l)}(t)x^{(k)}x^{(l)}$$

with

$$K_0(t) = \begin{pmatrix} r - d_t - \mu_t \\ \lambda_\eta a(t) \end{pmatrix}, \quad K_1(t) = \begin{pmatrix} 0 & 0 \\ 0 & -\lambda_\eta \end{pmatrix},$$

$$K_2^{(2,2)}(t) = 0 \quad \text{if } (k,l) \neq (2,2), \quad \text{and} \quad K_2^{(2,2)}(t) = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Similarly, the quadratic variation reads as

$$\Sigma(x,t)\Sigma(x,t)^T = H_0(t) + \sum_{k=1,2} H_1^{(k)}(t)x^{(k)} + \sum_{k,l=1,2} H_2^{(k,l)}(t)x^{(k)}x^{(l)}$$

since

$$\Sigma(x,t)\Sigma(x,t)^t = \begin{pmatrix} x_2^2 & \rho_1\alpha_t x_2 \\ \rho_1\alpha_t x_2 & (\rho_1\alpha_t^2 + \rho_2\alpha_t^2) \end{pmatrix}$$

and

$$H_0 = \begin{pmatrix} 0 & 0 \\ 0 & (\rho_1^2\alpha_t^2 + \rho_2^2\alpha_t^2) \end{pmatrix}, \quad H_1^{(2)}(t) = \begin{pmatrix} 0 & \rho_1\alpha_t \\ \rho_1\alpha_t & 0 \end{pmatrix}, \quad H_1^{(2,2)}(t) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$H_1^{(1)}(t) = H_1^{(3)}(t) = H_2^{(1,1)}(t) = H_2^{(1,2)}(t) = H_2^{(2,1)}(t) = 0.$$

We remark that $e^{-\int_0^t r_s ds}\psi(u, X_t, t, T)$ is a $\mathbb{P}^*$-martingale. Equivalently, the process $h(t, X_t) = e^{Y_t + \gamma(t,T) + \beta(t,T)\cdot X_t + \delta(t,T)(X_t^2)^2}$ is a $\mathbb{P}^*$-martingale where $Y_s = -\int_0^t r_s ds$ is a deterministic process. Since the predictable finite variation process of this semi martingale must be equal to zero, application of Itô formula to $h(t, X_t)$ allows identifying the drift term which results into the following equation

$$\begin{aligned} 0 &= \partial_t f(t,x,y) + \sum_{i=1,2} \partial_{x_i} h(t,x,y)\zeta_i(t,x) - r_t.h(t,x,y) \\ &\quad + \frac{1}{2} \sum_{i,j=1,2} (\Sigma\Sigma^t)_{i,j}(t,x)\partial_{i,j}^2 h(t,x,y). + \mathcal{A}h(t,x,y), \end{aligned}$$

where ζ is the drift vector and Σ the volatility matrix of the diffusion. $\mathcal{A}$ is the infinitesimal generator of the jump process, i.e.,

$$\mathcal{A}f(t, X_1, X_2) = \xi(t, X_1, X_2) \int_{\mathbb{R}} [f(t, X_1 + u, X_2) - f(t, X_1, X_2)] \, d\mathbb{G}(u),$$

Because

$$\mathcal{A}h(t, X_1, X_2) = \left(\xi^0 + \xi^1 X_1 + \xi^2 X_2 + \xi^3 X_2^2 \right) (\mathcal{L}(\beta_1(t, T)) - 1) h(t, X_1, X_2),$$

after some algebra, we finally get

$$
\begin{aligned}
0 \ = \ & \partial_t \gamma + \partial_t \beta_1(t) X_1 + \partial_t \beta_2(t) X_2 + \partial_t \delta(t) X_2^2 \\
& + \beta_1(t) \left(r_t - X_2^2/2 - d_t - \mu_t \right) + \lambda(\beta_2(t) + 2\delta(t)X_2)(a(t) - X_2) \\
& + \frac{1}{2} \left(\beta_1(t)^2 \left(X_2 \right)^2 + 2\beta_1(t)(\beta_2(t) + 2\delta(t)X_2)\rho_1 X_2 \alpha_t + \left(2\delta(t) + (\beta_2(t) + 2\delta X_2)^2 \right) \alpha_t^2 \right) \\
& + \left(\xi^0 + \xi^1 X_1 + \xi^2 X_2 + \xi^3 X_2^2 \right) (\mathcal{L}(\beta_1(t, T)) - 1)
\end{aligned}
$$

For a completely generic choice of X_1 and X_2 this expression is a second order polynomial in X and is identically equal to zero if and only if all its coefficients are identically zero, which provides the four ODE's.

B   Appendix

We consider a generic time interval $[T^{i-1}, T^i)$ where all equation coefficients are supposed to be constant. To solve the system of Riccati ODE's, a precise order must be followed. In this appendix we will omit specifying the time dependency of some variables to lighten notation.

- **First equation.** Solution subject to the final condition $\beta_1(T^i) = u_i^{(1)}$ is immediate, and reads

$$\beta_1(t) = u_i^{(1)}. \tag{21}$$

- **Third equation.** The equation satisfied by $\delta(t)$ is a second-order Riccati equation with terminal condition $\delta(T^i) = u_i^{(3)}$

$$\frac{\partial \delta(t)}{\partial t} \ = \ \tfrac{1}{2} \left(u_1^i - (u_1^i)^2 \right) + 2 \left(\lambda - \alpha_i \rho_1 u_1^i \right) \delta(t) - 2\alpha^2 \delta(t)^2,$$

and we have used $\beta_1(t)$ given Eq.(21). After a little algebra, we can rewrite the equations as

$$
\begin{aligned}
\frac{\partial \delta(t)}{\partial t} \ &= \ 2 \left(\frac{1}{\alpha_i^2} \left(B_i^2 + \alpha_i^2 A_i \right) - \left(\alpha_i \delta(t) + \frac{B_i}{\alpha_i} \right)^2 \right) \\
&= \ -2 \left(\alpha_i \delta(t) + \frac{B_i}{\alpha_i} + \frac{\Gamma_i}{\alpha_i} \right) \left(\alpha_i \delta(t) + \frac{B_i}{\alpha_i} - \frac{\Gamma^i}{\alpha_i} \right),
\end{aligned}
$$

28

or, equivalently, by separating the variables

$$4\Gamma^i dt = \alpha_i d\delta(t) \left[\frac{1}{\left(\alpha_i\delta(t) + \frac{B_i}{\alpha_i} + \frac{\Gamma_i}{\alpha_i}\right)} - \frac{1}{\left(\alpha_i\delta(t) + \frac{B_i}{\alpha_i} - \frac{\Gamma_i}{\alpha_i}\right)} \right]$$

and the solution, given the above final condition $u_i^{(3)}$ is therefore

$$\delta(t) = -\frac{1}{\alpha_i^2} \left((B_i + \Gamma_i) + \frac{2\Gamma_i C_i}{e^{4\Gamma_i(T_i - t)} - C_i} \right). \tag{22}$$

- **Second equation.** This equation is linear and its solution (with terminal condition $\beta_2(T^i) = u_i^{(2)}$ is lengthy but straightforward. We have

$$\frac{\partial\beta_2(t)}{\partial t} = -2\lambda a_i\delta - \left(\rho_1\alpha_i\beta_1 + 2\alpha_i^2\delta\right)\beta_2$$

Introducing a new set of functions

$$U(t) = -2\lambda a_i\delta, \quad V(t) = -\left(\rho_1\alpha_i\beta_1 + 2\alpha_i^2\delta\right)$$

This equation becomes $\partial\beta_2(t)/\partial t = U(t) + V(t)\beta_2(t)$ so that, formally

$$\beta_2(t) = u_i^{(2)}e^{-\int_t^{T^i}V(s)ds} - e^{-\int_t^{T^i}V(s)ds}\int_t^{T^i}U(x)e^{\int_x^{T^i}V(s)ds}dx.$$

After a number of algebraic manipulations it is possible to solve all integrals explicitly, and we finally obtain

$$\beta_2(t) = M(t)\left(u_i^{(2)} - K(t)\right). \tag{23}$$

with $M(t)$ and $K(t)$ defined in the main text.

- **Fifth equation** The equation to solve reads as

$$\frac{\partial\gamma}{\partial t} = -\Phi^0(\beta_1) + (d + \mu - r)\beta_1 - \lambda_\eta a_i\beta_2 - \alpha_i^2\left(\delta + \beta_2^2/2\right)$$

with terminal condition $\gamma(T^i) = u_i^{(4)}$. Once again, the solution is straightforward but requires some lengthy algebra. We point out that integrals of $\beta_2(t)$ and $\beta_2(t)^2$ can be both expressed in terms of hypergeometric functions but the solution is rather involved. For this reason, we found more convenient to present the solution in integral form. In the applications, both integrals can be easily evaluated numerically through a simple Gaussian quadrature.

29

C   Appendix

If Z_t satisfies the SDE $dZ_t = \left(r - \eta^2/2 - \kappa_t\right)dt + \eta dB_t + dN_t$ where the jump Poisson process is defined as in the text, its characteristic function is given by $\varphi_t(\theta) = \mathbb{E}\left[\exp\left(i\theta Z_t\right)\right] = \exp\left[t\psi(\theta)\right]$, since Z is a Lévy process and, in particular, its increments are stationary and independent (Jacod and Shiryaev (1987)). Function $\psi(\theta)$ is known as characteristic exponent and can be computed as follows. By recalling that $N_t = \sum_{j=1}^{n_t} Y_j$, and that Y_j are i.i.d. random variables, we get (by the laws of conditional expectation)

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(i\theta Z_t\right)\right] &= \exp\left[i\left(r - \frac{\eta^2}{2} - \kappa_t\right)\theta t - \frac{\eta^2}{2}\theta^2 t\right]\mathbb{E}\left[\exp\left(i\theta N_t\right)\right]\\
&= \exp\left[i\left(r - \frac{\eta^2}{2} - \kappa_t\right)\theta t - \frac{\eta^2}{2}\theta^2 t\right]\sum_{k=0}^{\infty}\mathbb{E}\left[\exp\left(i\theta N_t\right)|N_t = k\right]\mathbb{P}\left(N_t = k\right)\\
&= \exp\left[i\left(r - \frac{\eta^2}{2} - \kappa_t\right)\theta t - \frac{\eta^2}{2}\theta^2 t\right]\exp\left[\xi t\int_{\mathbb{R}}\left(e^{i\theta x} - 1\right)n(x)dx\right], \quad (24)
\end{aligned}
$$

where we have used the property $\mathbb{P}\left(N_t = k\right) = \frac{(\xi t)^k}{k!}\exp\left(-\xi t\right)$ and $\mathbb{E}\left[\exp\left(i\theta N_t\right)|N_t = k\right] = \mathbb{E}\left[\exp i\theta Y\right]^k$. Here, $n(x)$ is the Gaussian distribution with mean q and variance v^2. Thus,

$$
\begin{aligned}
\psi(\theta) &= i\left(r - \frac{\eta^2}{2} - \kappa_t\right)\theta - \frac{\eta^2}{2}\theta^2 + \xi\int_{\mathbb{R}}\left(e^{i\theta x} - 1\right)n(x)dx\\
&= i\left(r - \frac{\eta^2}{2} - \kappa_t\right)\theta - \frac{\eta^2}{2}\theta^2 + \xi\left[\exp\left(i\theta q - \frac{\theta^2 v^2}{2}\right) - 1\right].
\end{aligned}
$$

From the knowledge of $\varphi_t(\theta)$, the centered moment $K_l = \mathbb{E}\left[\left(Z_t - \mathbb{E}\left(Z_t\right)\right)^l\right]$ of order l can be evaluated from $K_l = \frac{1}{i^l}\frac{d}{d\theta^l}\left[\ln\varphi_t(\theta)\right]\big|_{\theta=0} = \frac{t}{i^l}\frac{d}{d\theta^l}\left[\psi(\theta)\right]\big|_{\theta=0}$. Simple algebra then provides the result.

Table 1
Robustness of the minimization algorithm for jumps calibration

Table 1a

		Initial jumps standard deviation v(0)						
		2%	**5%**	**15%**	**20%**	**30%**	**50%**	**80%**
	-80%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%
	-50%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%
	-20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%
Initial jumps	**0%**	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%
average q(0)	**20%**	7.56%	7.56%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%
	50%	7.56%	7.56%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%
	80%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%	-7.20%

Table 1b

		Initial jumps standard deviation v(0)						
		2%	**5%**	**15%**	**20%**	**30%**	**50%**	**80%**
	-80%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%
	-50%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%
	-20%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%
Initial jumps	**0%**	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%
average q(0)	**20%**	0.00%	0.00%	7.95%	7.95%	7.95%	7.95%	7.95%
	50%	0.00%	0.00%	7.95%	7.95%	7.95%	7.95%	7.95%
	80%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%	7.95%

Robustness of the minimization algorithm for jumps. For a given row and column, Table 1a (resp. 1b) provides the optimal
 jumps average (resp. standard deviation) that is obtained by initializing a Levenberg-Marquardt algorithm at the given couple (q(0),v(0))
We see that the algorithm converges to the optimal couple (q = -7.2%, v = 7.95%) for a wide range of
initialization parameters. Coefficients used are: delta =1, a(0)= 2%, rho(0)=0%, alpha(0)= 10%, omega = -0.3

Table 2

Minimum of the objective function at varying omega parameter

Omega	-0.100	-0.200	-0.500	-0.700	-1.000	-1.500	-1.800	-2.000
Least squares error	0.006%	0.006%	0.006%	0.006%	0.006%	0.006%	0.006%	0.006%

Precision of the minimization algorithm at varying omega. The error refers to the minimum of the objective function
with weights equal to 1. Other parameters are: delta =1, a(0)= 2%, rho(0)=0%, alpha(0)= 10%

Table 3

Differences between market and model implied volatility for different values of the initial volvol (alpha(0))

	alpha(0) = 10%			alpha(0) = 30%			alpha(0) = 50%		
	-1 st. dev.	ATM	+1 st. dev.	-1 st. dev.	ATM	+1 st. dev.	-1 st. dev.	ATM	+1 st. dev.
1Y	5.73%	3.17%	2.71%	2.13%	1.13%	-0.09%	**-1.59%**	**-0.90%**	**-0.37%**
2Y	9.40%	6.74%	5.36%	5.43%	4.45%	2.22%	2.23%	1.55%	0.52%
3Y	11.62%	8.81%	6.89%	7.76%	6.75%	3.89%	3.78%	3.11%	2.52%
5Y	13.85%	11.32%	9.01%	10.47%	9.67%	6.46%	7.63%	6.74%	5.28%

The difference between market and model implied volatility at varying time to expiry and for different values of
the initial volvol parameter. The model is calibrated to smiles up to 6M expiry only, then all coefficients are set to 0.
Parameters used are a(0) = 11%, rho = -40%, Omega = -0.3, delta = 0.14.
The 1Y model-generated smile with alpha(0) = 50% (in bold) is much higher than the market one.
It is then impossible to calibrate the market by constraining the volatility mean reversion level to take positive values

Table 4

Calibration errors and associated calibrated parameters with alpha(0) = 10%

		alpha(0) = 10%				a	alpha	rho	csi	q	v
		-1 st. dev.	ATM	+1 st. dev.							
1M	Mod. Vol.	20.06%	15.06%	14.81%							
	Market Vol.	19.85%	15.94%	14.81%		6.00%	10.00%	-60.00%	120.0%	-6.15%	14.25%
	Error	0.21%	-0.88%	0.00%							
3M	Mod. Vol.	18.96%	16.47%	15.01%							
	Market Vol.	18.91%	16.71%	15.45%		6.00%	10.00%	-60.00%	95.0%	-6.15%	14.25%
	Error	0.05%	-0.24%	-0.44%							
6M	Mod. Vol.	18.61%	17.07%	15.78%							
	Market Vol.	18.85%	17.19%	16.02%		6.00%	10.00%	-60.00%	76.0%	-6.15%	14.25%
	Error	-0.24%	-0.12%	-0.24%							
1Y	Mod. Vol.	21.84%	18.10%	14.76%							
	Market Vol.	22.49%	17.78%	15.19%		6.46%	39.35%	-95.57%	0.0%	0.00%	0.00%
	Error	-0.65%	0.32%	-0.43%							
2Y	Mod. Vol.	22.62%	18.45%	15.28%							
	Market Vol.	22.91%	18.22%	15.44%		6.10%	14.00%	-67.45%	0.0%	0.00%	0.00%
	Error	-0.29%	0.23%	-0.16%							
3Y	Mod. Vol.	22.91%	18.60%	15.27%							
	Market Vol.	23.36%	18.44%	15.46%		5.57%	25.10%	-98.83%	0.0%	0.00%	0.00%
	Error	-0.45%	0.16%	-0.19%							
5Y	Mod. Vol.	23.25%	19.06%	15.79%							
	Market Vol.	23.38%	18.86%	15.81%		5.87%	23.24%	-99.60%	0.0%	0.00%	0.00%
	Error	-0.13%	0.20%	-0.02%							

Comparison between market and model generated implied volatilities on three different strikes after calibration .

Up to smiles of 6M expiry, all SV coefficients are kept to a constant level and smiles are calibrated with jumps only

For the remaining part jumps are switched off and market is calibrated through stochastic volatility

Alpha(0) is fixed at 10%. Other model parameters are: lambda = 0.6, a(0) = 0.06, rho(0) = -0.6, delta = 0.1, volatility initial condition =0.11.

Errors are in absolute units

Table 5

Calibration errors and associated calibrated parameters with alpha(0) = 30%

		alpha(0) = 30%			a	alpha	rho	csi	q	v
		-1 st. dev.	ATM	+1 st. dev.						
1M	Mod. Vol.	19.99%	15.46%	14.94%						
	Market Vol.	19.85%	15.94%	14.81%	6.00%	30.00%	-60.00%	133.2%	-2.75%	9.34%
	Error	0.14%	-0.48%	0.13%						
3M	Mod. Vol.	19.10%	16.79%	15.23%						
	Market Vol.	18.91%	16.71%	15.45%	6.00%	30.00%	-60.00%	99.3%	-2.75%	9.34%
	Error	0.19%	0.08%	-0.22%						
6M	Mod. Vol.	19.46%	17.95%	16.70%						
	Market Vol.	18.85%	17.19%	16.02%	6.00%	30.00%	-60.00%	75.5%	-2.75%	9.34%
	Error	0.61%	0.76%	0.68%						
1Y	Mod. Vol.	22.13%	17.98%	15.17%						
	Market Vol.	22.49%	17.78%	15.19%	8.43%	22.23%	-88.34%	0.0%	0.00%	0.00%
	Error	-0.36%	0.20%	-0.02%						
2Y	Mod. Vol.	22.64%	18.51%	15.75%						
	Market Vol.	22.91%	18.22%	15.44%	9.31%	21.58%	-78.45%	0.0%	0.00%	0.00%
	Error	-0.27%	0.29%	0.31%						
3Y	Mod. Vol.	23.02%	18.64%	15.56%						
	Market Vol.	23.36%	18.44%	15.46%	6.37%	20.40%	-96.34%	0.0%	0.00%	0.00%
	Error	-0.34%	0.19%	0.10%						
5Y	Mod. Vol.	23.16%	19.09%	16.09%						
	Market Vol.	23.38%	18.86%	15.81%	5.89%	23.45%	-98.48%	0.0%	0.00%	0.00%
	Error	-0.22%	0.23%	0.28%						

Comparison between market and model generated implied volatilities on three different strikes after calibration .
Up to smiles of 6M expiry, all SV coefficients are kept to a constant level and smiles are calibrated with jumps only
For the remaining part jumps are switched off and market is calibrated through stochastic volatility
Alpha(0) is fixed at 30%. Other model parameters are: lambda = 0.6, a(0) = 0.06, rho(0) = -0.6, delta = 0.4, volatility initial condition =0.11.
Errors are in absolute units

Table 6

Calibration errors and associated calibrated parameters with alpha(0) = 10% after a 1% shock in the market volatility

		alpha(0) = 10%				a	alpha	rho	csi	q	v
		-1 st. dev.	ATM	+1 st. dev.							
1M	Mod. Vol.	20.73%	16.15%	16.08%							
	Market Vol.	20.85%	16.94%	15.81%		6.00%	10.00%	-60.00%	120.0%	-6.02%	15.87%
	Error	-0.12%	-0.79%	0.27%							
3M	Mod. Vol.	20.16%	17.80%	16.55%							
	Market Vol.	19.91%	17.71%	16.45%		6.00%	10.00%	-60.00%	95.0%	-6.02%	15.87%
	Error	0.25%	0.09%	0.10%							
6M	Mod. Vol.	20.12%	18.64%	17.48%							
	Market Vol.	19.85%	18.19%	17.02%		6.00%	10.00%	-60.00%	76.0%	-6.02%	15.87%
	Error	0.27%	0.45%	0.46%							
1Y	Mod. Vol.	22.74%	19.09%	16.01%							
	Market Vol.	23.49%	18.78%	16.19%		6.13%	40.12%	-95.00%	0.0%	0.00%	0.00%
	Error	-0.75%	0.31%	-0.18%							
2Y	Mod. Vol.	23.11%	19.06%	16.02%							
	Market Vol.	23.91%	19.22%	16.44%		6.03%	13.45%	-80.00%	0.0%	0.00%	0.00%
	Error	-0.80%	-0.16%	-0.42%							
3Y	Mod. Vol.	23.88%	19.69%	16.41%							
	Market Vol.	24.36%	19.44%	16.46%		5.87%	29.54%	-98.00%	0.0%	0.00%	0.00%
	Error	-0.48%	0.25%	-0.05%							
5Y	Mod. Vol.	24.19%	20.02%	16.77%							
	Market Vol.	24.38%	19.86%	16.81%		5.65%	23.12%	-99.00%	0.0%	0.00%	0.00%
	Error	-0.19%	0.16%	-0.04%							

Comparison between market and model generated implied volatilities on three different strikes after calibration .

Market volatility surface has been artificially shocked by 1%

Up to smiles of 6M expiry, all SV coefficients are kept to a constant level and smiles are calibrated with jumps only

For the remaining part jumps are switched off and market is calibrated through stochastic volatility

Alpha(0) is fixed at 10%. Other model parameters are: lambda = 0.6, a(0) = 0.06, rho(0) = -0.6, delta = 0.06, volatility initial condition =0.12.

Errors are in absolute units
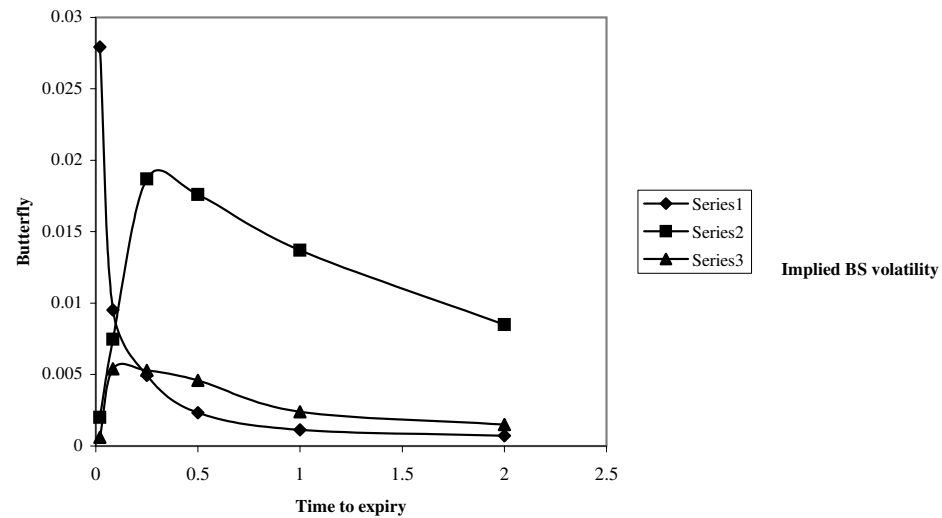
EuroStoxx 50 Implied Volatility

Fig 1a
Comparison of butterfly prices between market and the pure SV model
Series 1: Market butterfly on Feb 2nd, 2004
Series 2: Model butterfly with a= 0.1, rho =0, alpha =0.1, lambda = 5%
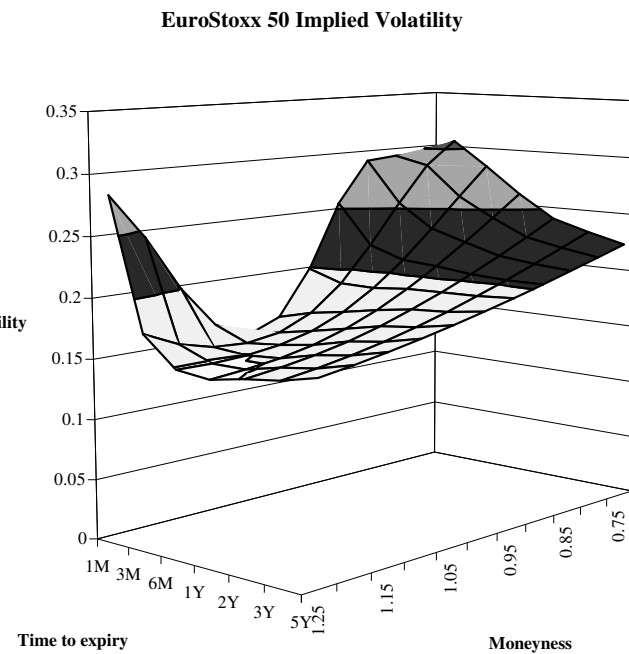Series 3: Model butterfly with a= 0.1, rho =0, alpha =0.1, lambda = 60%

Fig 1b
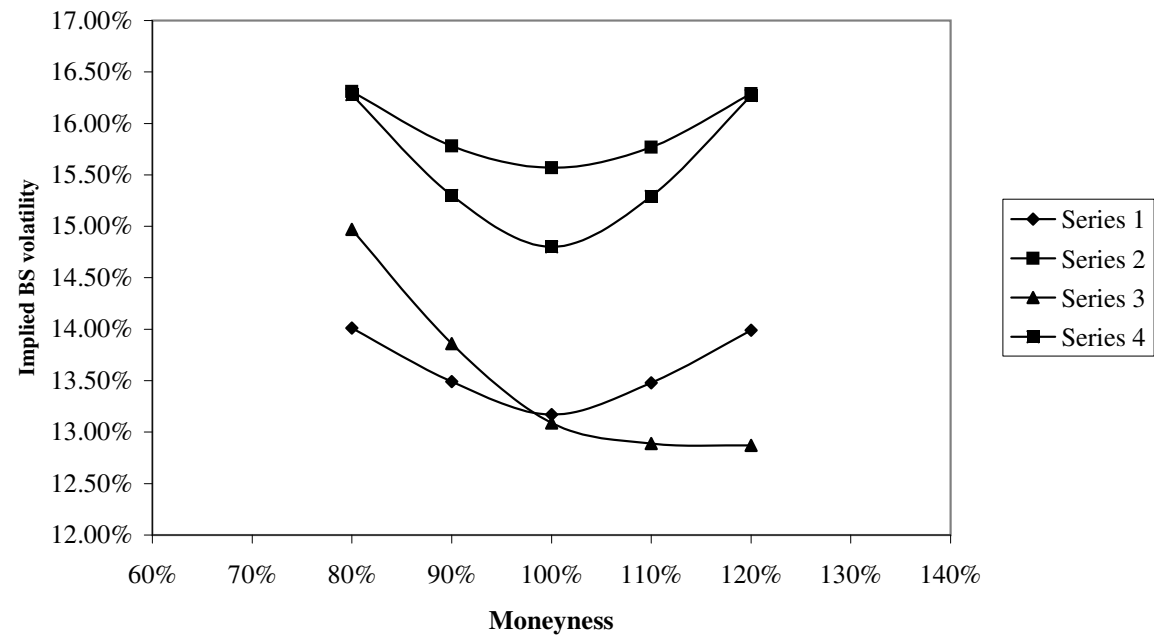The EuroStoxx implied volatility matrix
on Feb, 2nd 2004

Fig 1

Fig 2

The picture shows and compares the impact of a shock in different SV parameters on the smile. Jumps are set to zero

Volatility initial condition eta(0) = 0.11.

Series 1: Base scenario. Parameters: a = 0.1, rho = 0, alpha =0.1

Series 2: Shock in the volatility mean reversion level a. Parameters: a = 0.15, rho = 0, alpha =0.1

Series 3: Shock in the correlation rho. Parameters: a = 0.1, rho = -0.2, alpha =0.1

Series 4: Shock in the volvol alpha. Parameters: a = 0.1, rho = 0, alpha =0.15
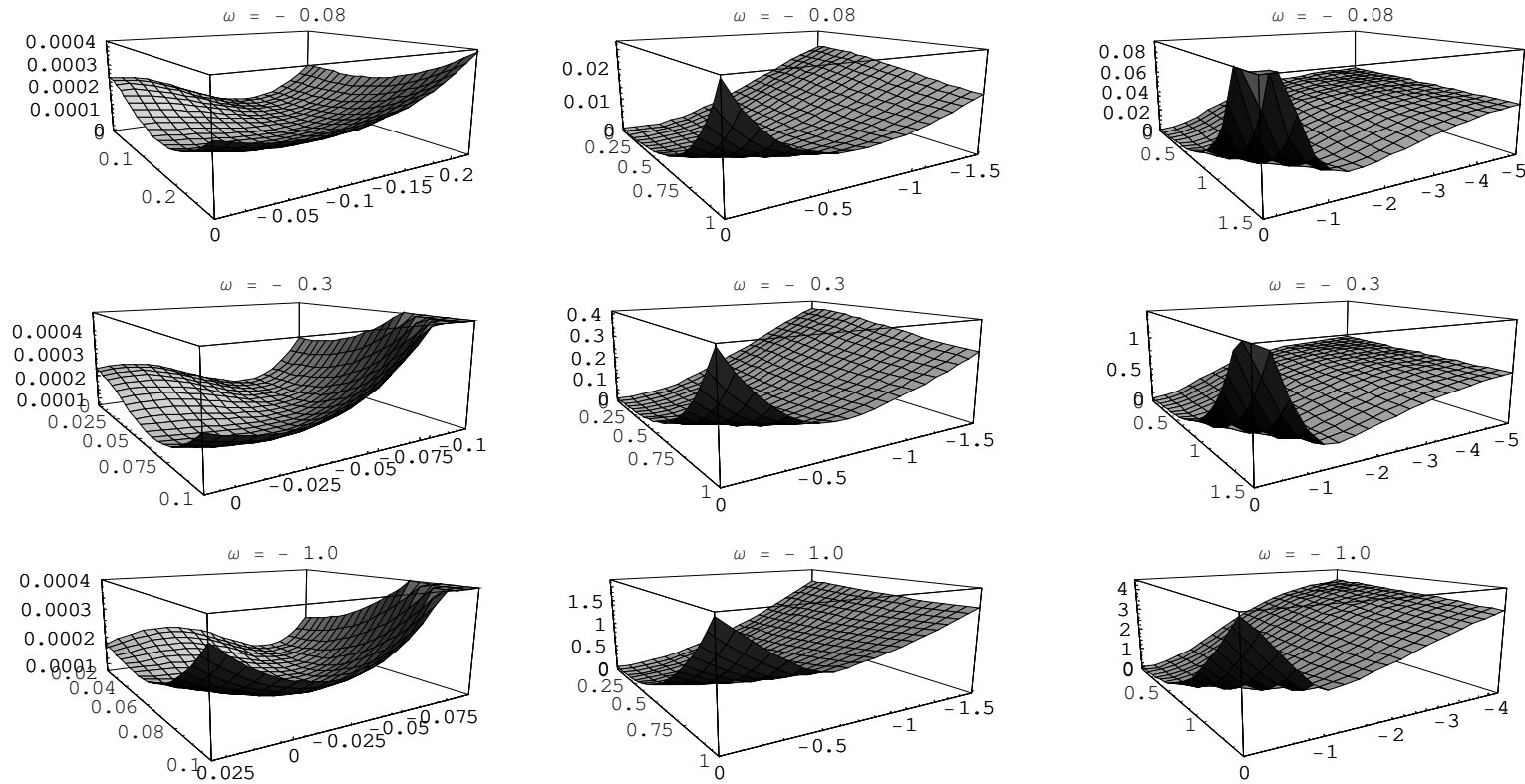
Fig. 3

Plot of the objective function associated to the optimization problem Eq.(14) for different values of
the omega parameter. For a given omega, the surface is plotted at three different scales
 Delta is fixed to 1 in this example.
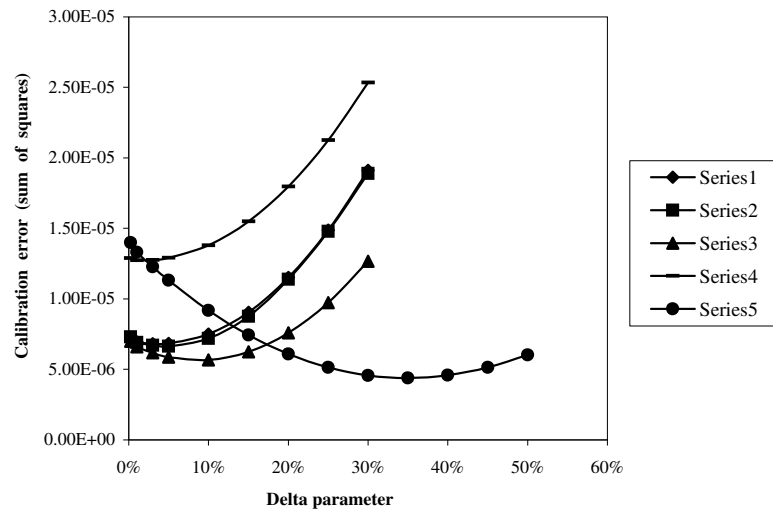Other coefficients are: a(0)= 2%, rho(0)=0%, alpha(0)= 10%.

Fig 4a

Minimum of the objectve function as a function of delta

Series 1:   a(0)=7%, rho(0) =-40%, alpha(0) =10%, omega = -0.3
Series 2:   a(0)=7%, rho(0) =-40%, alpha(0) =10%, omega = -0.6
Series 3:   a(0)=13%, rho(0) =-40%, alpha(0) =10%, omega = -0.3
Series 4:   a(0)=7%, rho(0) =-99%, alpha(0) =10%, omega = -0.3
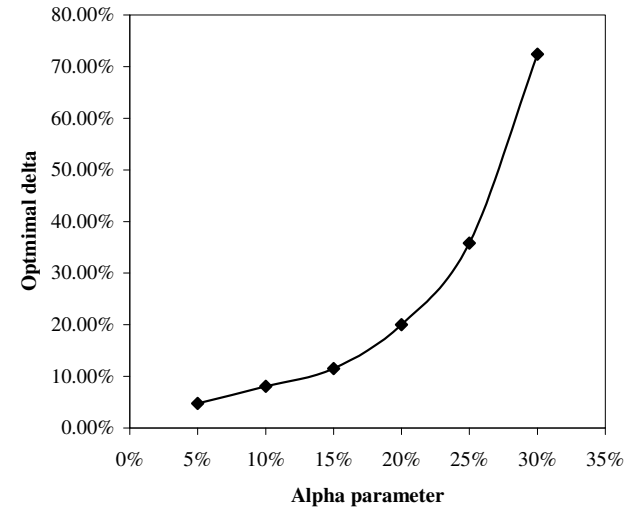Series 5:   a(0)=7%, rho(0) =-40%, alpha(0) =25%, omega = -0.3

Fig4b

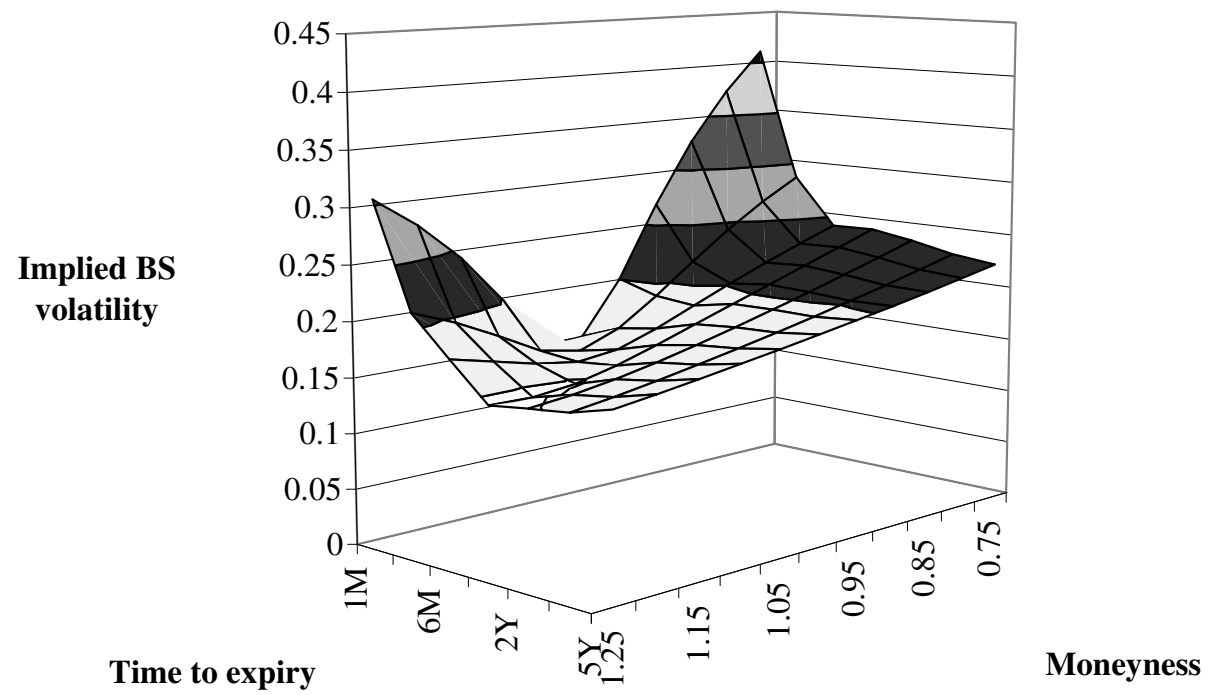Dependency of optimal delta from alpha(0)

Fig 4

Model implied volatility

Fig 5
Calibrated   volatility   surface