Abstract

"One model fits all" has been taken to heart by 90's market participants, hungering for ever more sophisticated models to capture the informational content of option prices using heavy numerical artillery, even to price European vanilla options before considering more complex derivatives. This led to a boom in models that could capture the market-implied volatility (IV) surface across strikes and maturities, while still inheriting some tractability that makes the models implementable and computationally fast. Among the many derived models, the family of exponential Lévy models provided a way to model jumps in the asset price process and consequently improved the capturing of the short-term curvature of the market surface. While shortcomings of the exponential Lévy models were present - such as independent increments and fast decay of skew and kurtosis across time to maturity - it spun off a combination of stochastic volatility (SV) models and jump processes, which in theory, should be able to capture the short-term curvature while still allowing for the leverage effect and dependence of increments in long-term smiles (an indirect consequence of volatility clustering). We provide a survey for two models in the class of infinite activity (IA) exponential Lévy models namely, the Variance-Gamma (VG) model and the CGMY model, to analyze how their simplistic nature fare against more complex models such as the Heston SV model and Bates stochastic volatility with jumps (SVJ) model. We provide a thorough theoretical introduction and calibrate each model across strikes and maturities. The conclusion is twofold: First, the exponential Lévy models under analysis have a hard time calibrating across maturities due to a combination of the floating smile property and the decay in skew and kurtosis, thus leading to an underpricing of long-term OTM options. For short-term options, they over-compensate the skew, therefore leading to an overpricing of the short-term options. Secondly, the Heston and Bates model capture the market surface considerably better due to the increased complexity and the incorporated stylized properties of the asset returns, described above.

1 Introduction

In the 90's there was a special emphasis on models that could capture the 'cross-sectional' behavior of the implied volatility surface, while still providing some tractability, making them implementable and computationally fast. One of the families under interest was the exponential Lévy models, which provided a greater fit of the short-term skew as opposed to SV based models. In the three-parameter case for the VG model, the extra parameters provide for modelling of the excess kurtosis and skewness of the risk-neutral return distribution, in turn giving us the desired effects. In essence, the CGMY model provides the same setup with an added parameter Y that accommodates for behaviors represented by pure jumps and diffusions thus allowing its arrival rates and variations to be either finite or infinite. However, as is empirical often evident, the exponential Lévy models have some statistical drawbacks, where one of them is the assumption of independent increments that leads to a lack of observed stylized facts of asset returns such as volatility clustering[1], the leverage effect and the floating smile property. In turn, we compare the Lévy models with the SV model of (Heston, 1993) and the SVJ model of (Bates, 1996), who incorporates both of the lacking properties. The Bates model has increased flexibility due to an added Gaussian jump specification, thus being capable of capturing more of the short-term skew than any pure diffusion-based models. While the SV and SVJ models clearly have the advantage, the purpose of the survey is to observe how 'standard' exponential Lévy models fare against these models with a focus on their drawbacks. Thus we will mainly concentrate on the exponential Lévy models. The article is organized as follows: Section 2 introduces the general theory of the asset price process for SV/SVJ type models, exponential Lévy models and furthermore provides for thorough derivations of their corresponding characteristic functions needed to recover option prices using Fourier methods. The section ends with a brief theoretical introduction to the drawbacks of exponential Lévy models. Section 3 deals with the calibration to market data and the assessment of the fit. Section 4 provides the conclusion and a short discussion on how to alleviate some of the lagging properties of exponential Lévy models by incorporating time-changed underlying dynamics.

2 General theory

2.1 The dynamics of the asset price process

We introduce two different specifications for the asset price process, in which the first nests the SVJ model and SV model, and the latter nests any exponential Lévy model omitting closed-form characteristic function. This is because SVJ type models are not Lévy processes since the process is path-dependent in the drift and diffusion component and thus violates the definition of a Lévy process, therefore, leading to two different specifications[2]. In turn, the same applies for the SV model[3].

2.1.1 Asset dynamics for SV and SVJ type models

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ be a filtered probability space. Throughout the analysis, we will only work under the risk-neutral measure $\mathbb{Q}$ unless otherwise specified. Moreover, we make the assumption of a deterministic risk-free interest rate $(r_t)_{t\geq 0}$ and constant

[1] Large jumps happens independent over time.

[2] see (Gatheral, 2011) p. 57.

[3] Section 2 and Section 3 are inspired from (Kokholm, 2016) (albeit with constant intensity), (Madan, Carr, & Chang, 1998) and (Tankov, 2003) (Ch. 7 and 15).

dividend yield q. Assume that the solution to the stock-price process at time t is given as

$$S_t^* = S_0 e^{\int_0^t r_s\, ds - qt + X_t}. \tag{1}$$

Then transforming (1) by $S_t = \ln(S_t^*)$ the dynamics of the corresponding transformed process can then be described as:

$$dS_t = (r_t - q)\, dt + dX_t, \tag{2}$$

with X_t following the risk-neutral dynamics of:

$$dX_t = \left(-\frac{1}{2}V_t - \lambda \int_{-\infty}^{\infty} [e^{\upsilon} - 1]\Pi(d\upsilon)\right) dt + \sqrt{V_t} dW_t^1 + dC_t \tag{3}$$

$$dV_t = \kappa\left[\bar{V} - V_t\right] dt + \sigma\sqrt{V_t} dW_t^2 \tag{4}$$

$$d\langle W^1, W^2\rangle_t = \rho\, dt \tag{5}$$

$$J_t \sim \mathcal{N}\left(\alpha, \delta^2\right) \tag{6}$$

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda\, dt \\ 0 & \text{with probability } 1 - \lambda\, dt \end{cases}. \tag{7}$$

Here, $(W_t^1)_{t\geq 0}$ and $(W_t^2)_{t\geq 0}$ are two correlated Brownian motions with correlation $\rho \in [-1, 1]$, $\Pi(d\upsilon)$ is the moment generating function of the jump size J_1 with density described above, $C_t = \sum_{i=1}^{N_t} J_i$ is a compound Poisson process with intensity $\lambda > 0$ and J_t as the jump size (conditioned on an occurrence of a jump) which is independent and identically distributed normal variables with parameters α and $\delta \geq 0$[4]. Moreover, J_t are assumed to be independent of the Brownian components. Finally κ, $\bar{V}$ and $\sigma \geq 0$ are respectively the speed of adjustment, long-run mean and the volatility of the variance process. While the fully specified asset price process is a SVJ model (Bates, 1996), it furthermore nests the Heston SV model formulated in (Heston, 1993) by setting $\lambda = 0$.

From standard option pricing theory, we know that analytical pricing formulas arise, when integrating the payoff against the density of the stock price process. However, closed-form solutions for the densities of both processes are not available. Thus, we exploit the fact that they both have (semi) closed-form characteristic functions, which provide the necessary condition to use Fourier pricing methods and recover the European option prices. The characteristic function at time T for the process $(X_t)_{t\geq 0}$, can be described as

$$\phi_T(u) = \mathbb{E}\left[e^{iuX_T}\right]. \tag{8}$$

Due to the assumed independence between the Brownian components and J_t, we can further split the characteristic function into two parts:

$$\phi_T(u) = \mathbb{E}\left[e^{iu\left(-\frac{1}{2}\int_0^T V_s\, ds + \int_0^T \sqrt{V_s}\, dW_s^1\right)}\right] \cdot \mathbb{E}\left[e^{iu\left(-\lambda\int_{-\infty}^{\infty}(e^{\upsilon}-1)\Pi(d\upsilon) + \sum_{i=1}^{N_T} J_i\right)}\right] = \phi_T^{SV}(u) \cdot \phi_T^{J}(u). \tag{9}$$

This implies that the characteristic function of the SVJ model can be decomposed into the characteristic function for the Heston model and the characteristic function of a compound Poisson process (CPP) with Gaussian jumps under S_t.

2.1.2 Asset dynamics for exponential Lévy models

Let everything be given as above. Then for $(S_t^*)_{t\geq 0}$ being an exponential Lévy process, it follows that the risk-neutral dynamics of $S_t = \ln(S_t^*)$ are given by

$$dS_t = (r_t - q - \kappa^L(1))dt + dX_t, \tag{10}$$

for $(X_t)_{t\geq 0}$ being the Lévy process and $\mathbb{E}\left[e^{X_1}\right] < \infty$. Moreover, $\kappa^L(1)$ is the cumulant generating function $\kappa^L(1) = \ln(\mathbb{E}\left[e^{X_1}\right])$ and it further follows that $\phi_T(u) = e^{T\kappa^L(iu)}$ for $\phi_T(\cdot)$ being a time T characteristic function. The above results stems from a change of measure from $\mathbb{P}$ to $\mathbb{Q}$ under the Esscher transform, which is used when a diffusion component is absent (and thus we cannot change the drift). Under the Esscher transform the Radon-Nikodym derivative change of measure is found to be

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \frac{e^{\theta X_t}}{\mathbb{E}\left[e^{\theta X_t}\right]} = e^{\theta X_t - \kappa^L(\theta)t} \qquad \theta \in \mathbb{R}, \tag{11}$$

[4]Under the transformation of $S_t = \ln(S_t^*)$, J_t will be normal distributed implying that under the untransformed process S_t^*, J_t is log-normal distributed.

where $\mathbb{Q}$ is the Esscher martingale transform for the exponential Lévy process e^{X_t}. Indeed, it is a martingale since,

$$\begin{aligned}\mathbb{E}\left[e^{\theta X_t - \kappa^L(\theta)t}\middle|\mathcal{F}_s\right] &= e^{-\kappa^L(\theta)t}e^{\theta X_s}\mathbb{E}\left[e^{\theta(X_t-X_s)}\middle|\mathcal{F}_s\right]\\ &= e^{-\kappa^L(\theta)t}e^{\theta X_s}\mathbb{E}\left[e^{\theta X_{t-s}}\middle|\mathcal{F}_s\right]\\ &= e^{-\kappa^L(\theta)t}e^{\theta X_s}e^{(t-s)\kappa^L(\theta)}\\ &= e^{\theta X_s - \kappa^L(\theta)s},\end{aligned}$$

for $s < t$. Then for $(S_t)_{t\geq 0}$ we have that,

$$\begin{aligned}e^{S_t} &= e^{\ln(S_0)+\int_0^t r_s\,ds - qt - \kappa^L(1)t + X_t}\\ \frac{e^{\ln(S_t^*)}}{e^{\ln(S_0^*)+\int_0^t r_s\,ds - qt}} &= e^{-\kappa^L(1)t+X_t}\\ e^{\ln\left(\frac{S_t^*}{S_0^*}\right)-\int_0^t r_s\,ds+qt} &= e^{-\kappa^L(1)t+X_t} \qquad \text{for } \theta = 1,\end{aligned}$$

which makes the discounted price process a martingale thus validifying the risk-neutral dynamics of the log-price process described in (10). In conclusion the characteristic function of the standardized log-price then becomes

$$\phi_T(u) = \mathbb{E}\left[e^{iu(S_T - S_0 - \int_0^T r_s\,ds + qT)}\right] = \mathbb{E}\left[e^{iuX_T - iu\kappa^L(1)T}\right] = e^{T\kappa^L(iu) - iuT\kappa^L(1)}. \tag{12}$$

In both the VG and the CGMY model, the characteristic exponents $\kappa(iu)$ are known in closed-form and as a consequence, we can apply Fourier pricing methods on (12) to recover European option prices.

2.2 The characteristic function of the Heston model & the Bates model

Setting $\lambda = 0$ we observe that the dynamics in (3) reduces to:

$$dX_t = -\frac{1}{2}V_t\,dt + \sqrt{V_t}dW_t^1 \tag{13}$$

$$dV_t = \kappa\left[\bar{V} - V_t\right]dt + \sigma\sqrt{V_t}dW_t^2 \tag{14}$$

$$d\langle W^1, W^2\rangle_t = \rho\,dt, \tag{15}$$

where $(W_t^i)_{t\geq 0}$ for $i = 1, 2$ are two correlated Brownian motions with correlation ρ, as described above. We recover the characteristic function $\phi_T^{SV}(u)$ following (Tankov, 2003) and consider the function:

$$M_t = \mathbb{E}\left[e^{iuX_T}\,\middle|\,\mathcal{F}_t\right] \tag{16}$$

with X_T being the log price at time T. Now from the 2-dimensional Itô's formula on $M_t = f(X_t, V_t)$ we deduce that:

$$\begin{aligned}dM_t &= \frac{\partial M_t}{\partial t}dt + \frac{\partial M_t}{\partial X_t}dX_t + \frac{\partial M_t}{\partial V_t}dV_t + \frac{1}{2}\frac{\partial^2 M_t}{\partial X_t^2}d[X,X]_t + \frac{1}{2}\frac{\partial^2 M_t}{\partial V_t^2}d[V,V]_t + \frac{\partial M_t^2}{\partial X_t \partial V_t}d[X,Y]_t\\ &= \left(\frac{\partial M_t}{\partial t} - \frac{1}{2}\frac{\partial M_t}{\partial X_t}V_t + \frac{\partial M_t}{\partial V_t}[\kappa(\bar{V} - V_t)] + \frac{1}{2}\frac{\partial^2 M_t}{\partial X_t^2}V_t + \frac{1}{2}\frac{\partial^2 M_t}{\partial V_t^2}\sigma^2 V_t + \frac{\partial^2 M_t}{\partial X_t \partial V_t}\sigma\rho V_t\right)dt + \frac{\partial M_t}{\partial X_t}\sqrt{V_t}dW_t^1 + \frac{\partial M_t}{\partial V_t}\sigma\sqrt{V_t}dW_t^2\end{aligned}$$

where the drift-term needs to equal zero in order for M_t to be a martingale

$$\frac{\partial M_t}{\partial t} - \frac{1}{2}\frac{\partial M_t}{\partial X_t}V_t + \frac{\partial M_t}{\partial V_t}[\kappa(\bar{V} - V_t)] + \frac{1}{2}\frac{\partial^2 M_t}{\partial X_t^2}V_t + \frac{1}{2}\frac{\partial^2 M_t}{\partial V_t^2}\sigma^2 V_t + \frac{\partial^2 M_t}{\partial X_t \partial V_t}\sigma\rho V_t = 0. \tag{17}$$

Together with the terminal condition $\mathbb{E}\left[e^{iuX_T}\,\middle|\,\mathcal{F}_T\right] = e^{iuX_T}$, the equation (17) allows us to compute the characteristic function of the log-price. We make the assumption that the solution is an log-affine function on the form:

$$M(\tau, V_t, X_t) = e^{iuX_t + C(\tau,iu) + D(\tau,iu)V_t}, \tag{18}$$

with $\tau = T - t$. Inserting the derivatives from (18) in the drift-term and reducing the equation yields:

$$\begin{aligned}0 &= \left(-\frac{\partial C}{\partial \tau} - V_t\frac{\partial D}{\partial \tau}\right) - \frac{1}{2}iuV_t + D(\tau)\kappa\bar{V} - D(\tau)\kappa V_t + \frac{1}{2}(iu)^2V_t + \frac{1}{2}D(\tau)^2\sigma^2V_t + iuD(\tau)\sigma\rho V_t \\ &= V_t\left(-\frac{\partial D}{\partial \tau} - \frac{1}{2}iu - D(\tau)\kappa + \frac{1}{2}(iu)^2 + \frac{1}{2}D(\tau)^2\sigma^2 + iuD(\tau)\sigma\rho\right) + \left(-\frac{\partial C}{\partial \tau} + D(\tau)\kappa\bar{V}\right),\end{aligned}$$

which again can be rearranged such that the functions $C(\cdot,\cdot)$ and $D(\cdot,\cdot)$ are the solution to two Riccati ordinary differential equations,

$$\begin{aligned}\frac{\partial C}{\partial \tau} &= D(\tau)\kappa\bar{V} \\ \frac{\partial D}{\partial \tau} &= \frac{1}{2}\left((iu)^2 - iu\right) - D(\tau)\cdot(\kappa - iu\sigma\rho) + \frac{1}{2}D(\tau)^2\sigma^2,\end{aligned}$$

with $D(0,iu) = 0$, $C(0,iu) = 0$ and terminal condition described above. The solutions to the Riccati ordinary differential equations are described in (Schoutens, Simons, & Tistaert, 2003) and are given by

$$C(\tau,iu) = \frac{\kappa\bar{V}}{\sigma^2}\left[(\xi(iu) - d(iu))\,\tau - 2\log\left(\frac{1 - g(iu)e^{-d(iu)\tau}}{1 - g(iu)}\right)\right] \tag{19}$$

$$D(\tau,iu) = \frac{1}{\sigma^2}\left(\xi(iu) - d(iu)\right)\frac{1 - e^{-d(iu)\tau}}{1 - g(iu)e^{-d(iu)\tau}}, \tag{20}$$

with

$$\xi(iu) = \kappa - \rho\sigma ui, \qquad d(iu) = \sqrt{(-\xi(iu))^2 - \sigma^2(-iu - u^2)}, \qquad g(iu) = \frac{\xi(iu) - d(iu)}{\xi(iu) + d(iu)}.$$

In conclusion the characteristic function for the Heston model when $t = 0$ is now given in (18) and as a formality can be rewritten to:

$$\phi_t^{SV}(u,T) = e^{iuX_t + C(\tau,iu) + D(\tau,iu)V_t}. \tag{21}$$

Now defining the stock price process under $\mathbb{Q}$ as a CPP with log-normal jumps:

$$S_t^* = S_0 e^{\left(-\lambda\int_{-\infty}^{\infty}(e^v - 1)\Pi(dv)\right)t + \sum_{i=1}^{N_t} J_i} \tag{22}$$

and thus,

$$dS_t = \left(-\lambda\int_{-\infty}^{\infty}(e^v - 1)\Pi(dv)\right)dt + dC_t \tag{23}$$

for J_i being log-normal under S_t^* and normal under S_t and $\lambda\int_{-\infty}^{\infty}(e^v - 1)\Pi(dv)$ being the drift correction term chosen such that S_t^* is a martingale under $\mathbb{Q}$. As described above, the correction term corresponds to the cumulant generating function of a CPP evaluated in 1, ie. $\kappa^J(1) = \log\left(\mathbb{E}\left[e^{L_1}\right]\right)$ and thus the characteristic function of the standardized log-price under the risk-neutral measure becomes:

$$\phi_T^J(u) = \mathbb{E}\left[e^{iuS_T}\right] = e^{-iuT\kappa^J(1) + T\kappa^J(iu)}, \tag{24}$$

where $\kappa^J(iu)$ is the characteristic exponent. From the fact that J_i is Gaussian under the log-price, then the characteristic exponent becomes[5]:

$$\begin{aligned}\kappa^J(iu) &= \lambda\int_{-\infty}^{\infty}(e^{iuv} - 1)\Pi(dv) \\ &= \lambda\left(e^{iu\alpha - \frac{u^2\delta^2}{2}} - 1\right),\end{aligned}$$

The corresponding cumulant generating function is given by[6]

$$\kappa^J(1) = \lambda\left(e^{\alpha - \frac{\delta^2}{2}} - 1\right), \tag{25}$$

[5] In both (Duffie, Pan, & Singleton, 2000) and (Gatheral, 2011) (p. 66), they use the mean α, while in the original article of (Bates, 1996) he used $\ln(1+\alpha) - \frac{1}{2}\delta^2$, as the mean parameter. Calibrating the Bates model in our situation under the two specifications of the mean parameter makes no difference whatsoever. Therefore we chose to generalize it and followed (Gatheral, 2011).

[6] Another intuitive formulation of the c.g.f is, $\kappa^J(1) = \ln(M_{N_1}(1)) = \lambda(M_{J_1}(1) - 1) = \lambda(e^{\alpha - \frac{\delta^2}{2}} - 1)$. Now substitute 1 with iu and the same applies for $\kappa^J(iu)$.

which finally yields the characteristic function for the Lévy process with log-normal jumps,

$$\phi_T^J(u) = e^{-iuT\lambda\left(e^{\alpha-\frac{\delta^2}{2}}-1\right)+T\lambda\left(e^{iu\alpha-\frac{u^2\delta^2}{2}}-1\right)}. \tag{26}$$

The characteristic function for the Bates SVJ model can now be recovered using (9).

2.3 The characteristic function of the CGMY and the VG model

The CGMY model is a generalization of the VG model, incorporating another parameter Y into the Lévy density which accommodates behaviors represented by pure jumps and diffusions thus allowing its arrival rates and variations to be either finite or infinite. For instance, $Y < 0$ allows for finite activity, $0 \leq Y < 1$ allows for infinite activity with finite variation, while $1 \leq Y < 2$ allows infinite activity with infinite variation. Moreover, as described in (Carr, Geman, Madan, & Yor, 2002) we restrict $C, G, M > 0$ and $Y < 2$, which is induced by the requirement that Lévy densities integrate x^2 in the neighborhood of 0. The parameter C can be thought of as a measure of overall activity, while G and M are measures of skewness. In the original article of (Carr et al., 2002) it is further noted that special cases arise when $Y = -1$ and $Y = 0$ giving us the Kou jump diffusion model and the VG model, respectively. Let $(X_t)_{t\geq 0}$ be the Lévy process corresponding to the CGMY model. Then we know that the characteristic function can be described by the Lévy-Khintchine theorem with characteristic triplet $(0, 0, \Psi(dx))$ under $(X_t)_{t\geq 0}$

$$\phi_T(u) = \mathbb{E}\left[e^{iuX_T}\right] = e^{T\int_{-\infty}^{\infty}(e^{iux}-1)\Psi(dx)} = e^{T\int_{-\infty}^{\infty}(e^{iux}-1)\mu(x)\,dx}, \tag{27}$$

and $\Psi(dx) = \mu(x)\,dx$ is the Lévy measure that has Lévy density $\mu(x)$. The Lévy density of the CGMY model is defined as:

$$\mu_{CGMY}(x) = \begin{cases} C\frac{e^{-M|x|}}{|x|^{1+Y}} & \text{for } x > 0 \\ C\frac{e^{-G|x|}}{|x|^{1+Y}} & \text{for } x < 0 \end{cases} = C\frac{e^{-M|x|}}{|x|^{1+Y}}\mathbb{1}_{\{x>0\}} + C\frac{e^{-G|x|}}{|x|^{1+Y}}\mathbb{1}_{\{x<0\}}, \tag{28}$$

and can be inserted in the characteristic exponent as the first step to recover the characteristic function of the risk-neutral process,

$$\begin{aligned}
\kappa^L(iu) &= \int_{-\infty}^{\infty}(e^{iux}-1)\left(C\frac{e^{-M|x|}}{|x|^{1+Y}}\mathbb{1}_{\{x>0\}} + C\frac{e^{-G|x|}}{|x|^{1+Y}}\mathbb{1}_{\{x<0\}}\right)dx \\
&= \int_{-\infty}^{0}(e^{iux}-1)C\frac{e^{-G|x|}}{-x^{1+Y}}\,dx + \int_{0}^{\infty}(e^{iux}-1)C\frac{e^{-M|x|}}{x^{1+Y}}\,dx \\
&= \left(\int_{-\infty}^{0}C\frac{e^{iux+Gx}}{-x^{1+Y}}\,dx - \int_{-\infty}^{0}C\frac{e^{Gx}}{-x^{1+Y}}\,dx\right) + \left(\int_{0}^{\infty}C\frac{e^{iux-Mx}}{x^{1+Y}}\,dx - \int_{0}^{\infty}C\frac{e^{-Mx}}{x^{1+Y}}\,dx\right),
\end{aligned}$$

where we in the third equation have used that $\int_{-\infty}^{0}|x|\,dx = \int_{-\infty}^{0}-x\,dx$ (due to $|x|$ being symmetric). Now doing a change of variable for the first integral $\omega(u,x) = -(iu+G)x$ with $\frac{d\omega}{dx} = -(iu+G) \Leftrightarrow dx = \frac{d\omega}{-(iu+G)}$ we further get

$$\begin{aligned}
\int_{-\infty}^{0}Ce^{-\omega(u,x)}\frac{1}{\left(\frac{\omega(u,x)}{(iu+G)}\right)^{1+Y}}\frac{d\omega}{-(iu+G)} &= \int_{-\infty}^{0}-Ce^{-\omega(u,x)}\frac{1}{\left(\frac{\omega(u,x)}{(iu+G)}\right)^{1+Y}}\frac{d\omega}{(iu+G)} \\
&= \int_{0}^{\infty}Ce^{-\omega(u,x)}\omega(u,x)^{-1-Y}(iu+G)^{1+Y}(iu+G)^{-1}d\omega \\
&= \int_{0}^{\infty}Ce^{-\omega(u,x)}\omega(u,x)^{-1-Y}(iu+G)^{Y}d\omega.
\end{aligned}$$

Following up with a corresponding change of variable $\omega(0,x) = -Gx$ for the second integral and in the second bracket $\omega_2(u,x) = (M-iu)x$ and $\omega_2(0,x) = Mx$, the same derivation follows. In conclusion, we get four integrals omitting analytical solutions, since integrating with respect to ω yields a Gamma function with parameter $-Y$ [7]:

$$\begin{aligned}
\cdots &= \left(\int_{0}^{\infty}(G+iu)^{Y}w(u,x)^{-1-Y}Ce^{-w(u,x)}\,dw - \int_{0}^{\infty}CG^{Y}w(0,x)^{-1-Y}e^{-w(0,x)}dw\right) \\
&+ \left(\int_{0}^{\infty}C(M-iu)^{Y}w_2(u,x)^{-1-Y}e^{-w_2(u,x)}\,dw - \int_{0}^{\infty}CM^{Y}w_2(0,x)^{-1-Y}e^{-w_2(0,x)}\,dw\right) \\
&= \left(C\Gamma(-Y)\left[(G+iu)^{Y}-G^{Y}\right]\right) + \left(C\Gamma(-Y)\left[(M-iu)^{Y}-M^{Y}\right]\right) \\
&= C\Gamma(-Y)\left[(M-iu)^{Y}-M^{Y}+(G+iu)^{Y}-G^{Y}\right].
\end{aligned}$$

[7] That is, $\int_0^{\infty}e^{-\omega(u,x)}\omega(u,x)^{-1-Y} = \Gamma(-Y)$.

Therefore the time T characteristic function for the CGMY exponential Lévy model becomes:

$$\phi_T^{\text{CGMY}}(u) = e^{T\kappa^L(iu)} = e^{TC\cdot\Gamma(-Y)\left[(M-iu)^Y - M^Y + (G+iu)^Y - G^Y\right]}. \tag{29}$$

In its original setting of (Madan et al., 1998), the VG model is described as a subordinated Brownian motion with subordination component being Gamma distributed with unit mean rate[8]. Let $V_T \sim \Gamma(\frac{T}{\nu}, \nu)$ be Gamma distributed then the VG process is:

$$X_t = \theta V_t + \sigma W_{V_t}, \tag{30}$$

where V and W are independent, $\nu \geq 0$ is the variance rate of the Gamma time-change and θ, $\sigma \geq 0$ denotes the corresponding drift and volatility for the Brownian motion. Furthermore, (Tankov, 2003) establishes that the Brownian subordination again is a Lévy process (Theorem 4.2). Thus the time T characteristic function of $(X_t)_{t\geq 0}$ can be derived as:

$$\begin{aligned}
\phi_T^{VG}(u) &= \mathbb{E}\left[e^{iuX_T}\right] \\
&= \mathbb{E}\left[e^{iu\theta V_t + \sigma W_{V_t}}\right] \\
&= \mathbb{E}\left[e^{iu\theta V_T}\right] \cdot \mathbb{E}\left[e^{iu\sigma W_T}\middle| V_T\right] \\
&= (1-\nu\theta iu)^{-\frac{T}{\nu}} \cdot \mathbb{E}\left(\mathbb{E}\left[e^{iu\sigma W_T}\middle| V_T\right]\right) \\
&= (1-\nu\theta iu)^{-\frac{T}{\nu}} \cdot \mathbb{E}\left[e^{-\frac{\sigma^2u^2V_T}{2}}\right] \\
&= (1-\nu\theta iu)^{-\frac{T}{\nu}} \cdot M_{V_T}\left(-\frac{\sigma^2u^2}{2}\right) \\
&= (1-\nu\theta iu)^{-\frac{T}{\nu}} \cdot \left(1+\nu\frac{\sigma^2u^2}{2}\right)^{-\frac{T}{\nu}} \\
&= \left(1-\nu\theta iu+\nu\frac{\sigma^2u^2}{2}\right)^{-\frac{T}{\nu}}
\end{aligned}$$

using the fact that $\theta V_T \sim \Gamma(\frac{T}{\nu}, \theta\nu)$, the tower rule and the moment generating function for the Gamma distribution. Now, taking the logarithm on both sides of the time T characteristic function we can recover $T\kappa^L(iu)$,

$$T\kappa^L(iu) = -\frac{T}{\nu}\ln\left(1-\nu\theta iu+\nu\frac{\sigma^2u^2}{2}\right). \tag{31}$$

In the last equation of $\phi_T^{VG}(u)$, we can factorize the bracket into two parts by rewriting it as

$$\left(\frac{1}{1-\nu\theta iu+\nu\frac{\sigma^2u^2}{2}}\right) = \left(\frac{1}{1-i\eta_p u}\right)\left(\frac{1}{1+i\eta_n u}\right), \tag{32}$$

with η_p, η_n being given by

$$\begin{aligned}
\eta_p &= \sqrt{\frac{\theta^2\nu^2}{4}+\frac{\sigma^2\nu}{2}}+\frac{\theta\nu}{2} \\
\eta_n &= \sqrt{\frac{\theta^2\nu^2}{4}+\frac{\sigma^2\nu}{2}}-\frac{\theta\nu}{2}
\end{aligned}$$

where the parameters in the CGMY model now arises from[9]

$$C = \frac{1}{\nu}, \qquad G = \frac{1}{\eta_n}, \qquad M = \frac{1}{\eta_p}, \qquad Y = 0.$$

With the new parameterization the VG model can now be written as the difference between two independent Gamma processes,

$$X_t^{VG} = \Gamma_t(C,M) - \Gamma_t(C,G). \tag{33}$$

In conclusion, we observe that both the CGMY and the VG model have closed-form characteristic functions. However, we will stick to the parameterization of (Madan et al., 1998). For both of the exponential Lévy models we can now recover the time T characteristic function of the standardized log-price using (12), and thus price European call options using Fourier methods.

[8] That is, for $V_T \sim \Gamma(aT, \nu)$ and $\mu = a\nu = 1$ being the mean rate then $a = \frac{1}{\nu}$.

[9] see (Carr et al., 2002) p. 5.

2.4 Smile asymptotics with a focus on exponential Lévy models

In this section we give a brief introduction to some of the drawbacks related to exponential Lévy models, starting with the floating smile property:

Per definition the implied volatility can be recovered by equating the Black-Scholes (BS) call prices with the Model call prices

$$C_{BS}(t;T,S_t,K,\sigma) = C_{Model}(t;K,T), \tag{34}$$

and then inverting the BS formula to solve for the implied volatility. If we let the implied volatility of an exponential Lévy model $(X_t)_{t\geq 0}$ be given by $I_t(\tau,k)$, for $k = \ln(\frac{K}{S_0 e^{\int_0^T rs\,ds - qT}})$ then (Tankov, 2003) provides us with proposition (11.1) described below.

Proposition 1. *Let $(X_t)_{t\geq 0}$ be given by a risk-neutral exponential Lévy process. Then the implied volatility for a given moneyness k and time to maturity $\tau = T - t$ does not depend on time t,*

$$I_t(\tau,k) = I_0(\tau,k), \qquad \text{for all } \ t \geq 0. \tag{35}$$

This is called the floating smile property.

The floating smile property for exponential Lévy models comes from the independent increments property. The intuitive interpretation of this drawback is that even though the underlying moves in time t, the smile uniquely depends on moneyness and time to maturity, making it constant in calendar time and thus the surface does not depend on the underlying at all. In conclusion, if exponential Lévy models were to give the correct picture of real markets, then the calibrated surface today, will be the same as the predicted surface tomorrow[10]. This can be compared to the flat implied volatility surface of the BS model with a height of σ across time to maturity, just with the exception that exponential Lévy models exhibit skew and excess kurtosis in the risk-neutral return distribution, thus giving the surface some curvature. This is, however, rejected by data: The volatility surface is not constant in time, In particular, the short-term skew tends to become more pronounced in times of crises/market crushes while it flattens in more quiet periods. Moreover, the change in the level of the surface tends to change depending on the economic situation (higher in crisis, lower in calm). Therefore the exponential Lévy models simply cannot capture the dynamics of the surface - which moves rapidly for short-term maturities. This can be considered one of the main drawbacks of the exponential Lévy models (see (Derman, 1999)). From the SVJ/SV type models the situation is different: While the smile (in moneyness) still does not depend on the current level of the underlying, it depends on the current level of the instantaneous variance $(V_t)_{t\geq 0}$. Therefore as the variance moves in time, the same happens to the smile thus changing level, skew and kurtosis depending on where the variance goes.

Another drawback comes from the fast decay of skew and kurtosis. While this is in line with empirics, the exponential Lévy models simply decays too fast. Assume that the Lévy model $(X_t)_{t\geq 0}$ on $\mathbb{R}$ with characteristic triplet (A,ν,γ) has finite n-th absolute moment and define

$$\kappa_n = \left.\frac{d^n \kappa(u)}{du^n}\right|_{u=0}, \tag{36}$$

where $\kappa(u) = \ln(\mathbb{E}\left[e^{uX_1}\right])$ is the cumulant generating function. Then the moments can be explicitly computed from its characteristic function. Moreover, it can be shown that[11]:

$$\begin{aligned}
\mathbb{E}[X_t] &= t\left(\gamma + \int_{|x|\geq 1} x\nu(dx)\right) \\
\kappa_2(X_t) &= t\left(A + \int_{-\infty}^{\infty} x^2\,\nu(dx)\right) \\
&\vdots \\
\kappa_n(X_t) &= t\int_{-\infty}^{\infty} x^n\,\nu(dx), \qquad \text{for } n \geq 3.
\end{aligned}$$

Thus the skewness and excess kurtosis of a Lévy process with infinitely divisible distributions can be found to be,

$$\text{Skew}(X_t) = \frac{1}{t^{1/2}}\frac{\kappa_3}{\kappa_2^{3/2}}, \qquad \text{Kurt}(X_t) = \frac{1}{t}\frac{\kappa_4}{\kappa_2^2}, \qquad \text{for } \ \kappa_3 = \kappa_4 \neq 0 \text{ and } \ \kappa_2 > 0. \tag{37}$$

[10] In retrospect as described in (Derman, 1999), the current level of at-the-money volatility should remain unchanged as the underlying moves, implying that an option 10% OTM after the index moves, should still have the same implied volatility as before the move.

[11] See (Tankov, 2003) p. 103.

Therefore the skew decays at a rate of $t^{-1/2}$ and the kurtosis with a rate of t^{-1}, implying that for options with long maturities the skew and excess kurtosis tends to zero thus smoothening out the implied volatility surface and when $t \to 0$ the skewness and kurtosis explode. As described in (Tankov, 2003) proposition (3.13) if $(X_t)_{t\geq 0}$ is the log-price of an exponential Lévy process then its 4-th cumulant $\kappa_4(X_t)$ is always proportional to t which is one of the reasons the kurtosis always decays with t^{-1} no matter the model. In contrary - for the cost of increased dimensionality[12] - if we, for the simplicity of the argument, let $(X_t)_{t\geq 0}$ be a SV type model given by $X_t = \int_0^t \sigma_s \, dW_s$ and assuming that σ_s is deterministic and independent of the Wiener process, then the 4-th cumulant for this model is $\kappa_4(X_t) = 3 \cdot \mathbb{V}ar\left(\int_0^t \sigma_s^2 \, ds\right)$, which gives many different dependence types on t. Especially if we let $\sigma_t = V_t$ be the solution to the dynamics described in (3), it paves the way for models containing the leverage effect and volatility clustering, which in turn increase the flexibility of the long-term smile dynamics.

To end the section, we note that for exponential Lévy models with finite second moment, the Central Limit theorem shows that for T being large then $\frac{X_T - \mathbb{E}[X_T]}{\sqrt{T}}$ becomes approximately standard Gaussian. As argued in (Rogers & Tehranchi, 2010) they emphasize that the flattening of the implied volatility smile is a universal property of all martingale models, also in models where the return distribution have infinite second moment. Therefore, for models with finite second moment, prices will be closer to BS prices with the implied volatility smile becoming flat for $T \to \infty$. While this effect is shown in the market, it is much more apparent in exponential Lévy models.

3 Empirical results

3.1 Preliminary data cleaning & recovery of interest rate and dividends

We utilize option data from the S&P 500 in the period of 2017/12/13, which contains data for maturities ranging from one month to two years. We filter the data by using the mid-price as a "fair" option price and remove the quotes where the bid-price is zero for both call and put options, since it reflects a non-existing demand for these specific options. Under the empirical analysis, we will only work with annualized maturities,

$$T \in \mathcal{T} = \{37/365,\ 65/365,\ 93/365,\ 184/365,\ 282/365,\ 373/365\} \tag{38}$$

and with a moneyness defined in the range of $0.77 \leq \frac{K}{S} \leq 1.28$. Furthermore, we will only work with call options, since the surface of put options can be recovered using put-call parity. We follow in the style of (Carr & Madan, 2005) filtering out options that violate the no-arbitrage bounds by excluding static arbitrage formed by spread strategies[13]. Let $C(S_0, K_i, T)$ denote the mid quote for a call of strike K_i, $i = 1, \ldots, \infty$ and maturity $T \in \mathcal{T}$, then the call needs to satisfy the bounds:

$$S_0 > C(S_0, K_i, T) \geq \max\left(S_0 e^{-qT} - K e^{-\int_0^T r_s \, ds}, 0\right) \tag{39}$$

$$\frac{C(S_0, K_{i-1}, T) - C(S_0, K_i, T)}{K_i - K_{i-1}} \in \left[0,\ e^{-\int_0^T r_s \, ds}\right] \tag{40}$$

$$\frac{C(S_0, K_{i-1}, T) - C(S_0, K_i, T)}{K_i - K_{i-1}} - \frac{C(S_0, K_i, T) - C(S_0, K_{i+1}, T)}{K_{i+1} - K_i} \geq 0 \tag{41}$$

$$C(S_0, K_i, T_{j+1}) - C(S_0, K_i, T_j) \geq 0, \qquad \{T_{j+1}, T_j\} \in \mathcal{T} \quad \text{and} \quad T_{j+1} > T_j > 0, \tag{42}$$

where $0 \leq K_{i-1} < K_i < K_{i+1}$ and the last three bounds are derived from the cost of a vertical spread, butterfly spread and a calendar spread, respectively. From the fundamental theorem of asset pricing we know that absence of arbitrage is equivalent to finding an equivalent martingale measure $\mathbb{Q}$ in which the model framework is a martingale. Therefore any model under $\mathbb{Q}$[14], will only improve by removing the options violating the no-arbitrage bounds reducing the noise in the data and conclusively provide a marginally better fit. Moreover, (Carr, Geman, Madan, & Yor, 2003) establish the equivalence of no *static* arbitrage opportunities with the existence of a Markov martingale $\mathbb{Q}$, in some filtration for exponential Lévy processes.

We have recovered the interest rates and dividend yield across the first 6 maturities using 9 pairs of call options with strikes,

$$K \in \mathcal{K} = \{2550,\ 2575,\ 2600,\ 2625,\ 2650,\ 2675,\ 2700,\ 2725,\ 2750\}.$$

From these call options we compute the corresponding put options using put-call parity

$$P(S_0, K, T) = C(S_0, K, T) - S_0 e^{-qT} - K e^{-\int_0^T r_s \, ds} \qquad \text{for } T \in \mathcal{T}, \tag{43}$$

[12] In stochastic volatility models the asset price process is no longer a Markov process: the evolution of the price process is determined not only by its value but also by the level of volatility. To regain a Markov process one must consider the two-dimensional process (S_t, σ_t), as seen when deriving the Heston model.

[13] Static arbitrage, hereby the definition of a costless position in the underlying that only depends on time and asset dimension and will at a future time provide a positive return (with positive probability) without taking any losses.

[14] or an equivalent martingale measure.

and choose $\int_0^{T_j} r_s\, ds$ for all $T_j \in \mathcal{T}$ and q for $j = 1,\dots,6$ and $T_j > T_{j-1}$ such that the put prices derived from the put-call parity matches the observed market prices[15]. The interest rate and dividend yield can now be found by minimizing the sum of squared error:

$$\min_{r_j,q} \sum_{j=1}^{6}\sum_{i=1}^{9}\Big(Put_{PC}(K_i,T_j,r_{T_j},q) - Put_{Market}(K_i,T_j)\Big)^2. \tag{44}$$

In Table 1 we have described the estimates for the interest rate and the dividend yield across the six maturities. Furthermore, we have provided a graphical illustration in the appendix of the interest rates.

Rates/Maturities	T_1	T_2	T_3	T_4	T_5	T_6
Interest rate (%)	3.3492	2.3736	2.1370	1.9830	1.9979	1.9872
Dividend yield (%)			1.5935			

Table 1: The estimated interest rates and dividend yield found by minimizing equation (44) under 9 call options around the ATM call option with strike, $K = 2660$.

3.2 Calibration procedure

The calibration procedure involves minimizing the sum of squared error between the model implied volatility and the market implied volatility,

$$\theta^* = \min_{\theta} \sum_{j\in\mathcal{T}}\sum_{i\in K_{T_j}} \left(\sigma^{IV}_{Model}(S_0,T_j,K_i,r_{T_j},q,C^{ij}_{Model}(\theta)) - \sigma^{IV}_{Market}(S0,K_i,T_j,r_{T_j},q,C^{ij}_{Market})\right)^2. \tag{45}$$

where θ^* is the optimal parameters for the specified model that minimizes the sum of squared error, K_{T_j} denotes the strike space at maturity T_j and $C^{ij}_{Model}(\theta)$ are the model call prices for maturity T_j at strike K_i for the parameters θ. The above method provides for a significantly better fit than the classical procedure of calibrating to the call surface by minimizing the sum of squared error between the corresponding call prices. Of course, the above objective function minimizes the distance directly and thus have an advantage against the latter method which minimizes the implied volatility surface indirectly as a consequence of calibrating to said call price surface. As seen in Table 2 there is a significant improvement in minimizing the sum of squared errors between the implied volatilities, since it contributes to a lower mean absolute implied volatility error between the model IV and market IV.

Model/MAIE	*IV calibration* ($\times 10^{-3}$)	*Call calibration* ($\times 10^{-3}$)
VG	12.7972	20.7523
CGMY	11.7670	19.1640
Heston	3.9053	4.7479
Bates	3.2030	4.1130

Table 2: A comparison of the mean absolute implied volatility error (MAE^{IV}) between the model implied volatilties and the market implied volatilities for the two calibration procedures (This is just $MAPE^{IV}$ in eq. (54) without normalizing with the market implied volatility). Understandably, since the IV calibration procedure involves minimizing this distance directly, it will be far superior in contrast to the call calibration, which recover the model IV's by inverting the call prices. Call calibration is defined as the classical procedure of minimizing the sum of squared differences between the model and market call prices.

Of course, the model implied volatilities depends on the model call prices, which are computed by the use of Fourier pricing methods. In the traditional setup, option prices are computed by integrating the payoff over the density function for the log-returns on the underlying. However, closed-form solutions of the densities are not available which motivates the use of Fourier pricing methods[16]. For all the models considered in this analysis, the corresponding characteristic function of the standardized log-price is available in closed or semi closed-form, which allow us to obtain call prices using the Lewis approach (Lewis, 2001) described in the next proposition.

[15] We essentially recover the interest rates on the form of $\int_0^{T_j} r_s\, ds$, but since this is the only formulation we work with (involving the interest rates) it is justified.

[16] With the exception of the VG model, since the VG model omits a closed-form density and therefore also an analytical option pricing formula, see (Madan et al., 1998). However we will only work with Fourier pricing methods to ensure a comparable framework for all models.

Proposition 2. *The price of a European call on the asset $(S_t)_{t\geq 0}$ at strike K under interest rates $(r_t)_{t\geq 0}$ and the constant dividend yield q at maturity T can be computed using the formula*

$$C(S_0,K,T) = Se^{-qT} - \frac{Ke^{-\int_0^T r_s\,ds}}{2\pi}\int_{iz_2-\infty}^{iz_2+\infty} e^{-izk}\phi_T(-z)\frac{dz}{z^2-iz}, \quad z_2\in(0,1),$$

where $k = \log(S_0/K) + \int_0^T r_s\,ds - qT$. For $z_2 = \frac{1}{2}$ the formula reduces to

$$C(S_0,K,T) = S_0e^{-qT} - \frac{\sqrt{S_0K}e^{-\frac{1}{2}\left(\int_0^T r_s\,ds+qT\right)}}{\pi}\int_0^\infty \Re\mathfrak{e}\left[e^{iuk}\phi_T^0\left(u-\frac{i}{2}\right)\right]\frac{du}{u^2+\frac{1}{4}}, \tag{46}$$

with $\phi_T^0(\cdot)$ being the standardized characteristic function of $X_T = \ln\left(\frac{S_T}{S_0}\right) - \int_0^T r_s\,ds + qT$, which is found by solving the stock price process $S_T = S_0e^{\int_0^T r_s\,ds - qT + X_T}$.

Proof. We start by deriving the payoff transform of a covered call, which is used in the proof of the Lewis integral. Let $x = \ln(S_T)$ and $K > 0$, then the payoff of a covered call is $\omega(x) = \min(e^x, K)$ with corresponding generalized Fourier transform $\hat{\omega}(z) = \mathscr{F}[\omega(x)]$ for $z\in\mathbb{C}$, which can be found to be:

$$\begin{aligned}
\hat{\omega}(z) &= \int_{\mathbb{C}} e^{izx}\omega(x)\,dx \\
&= \int_{\mathbb{C}} e^{izx}\min(e^x,K)\,dx \\
&= \int_{-\infty}^{\ln(K)} e^{izx}e^x\,dx + K\int_{\ln(K)}^{\infty} e^{izx}\,dx \\
&= \left[\frac{e^{(iz+1)x}}{iz+1}\right]_{x=-\infty}^{x=\ln(K)} + \left[K\frac{e^{izx}}{iz}\right]_{x=\ln(K)}^{x=\infty}
\end{aligned}$$

Before continuing we derive the strip of regularity for the call transform[17]. Decomposing $z = z_r + iz_i$ we can rewrite the first component as,

$$\frac{e^{izx}e^x}{iz+1} = \frac{e^{i(z_r+iz_i)x}e^x}{i(z_r+iz_i)+1} = \frac{e^{iz_rx-z_ix}e^x}{iz_r-z_i+1} = \frac{e^{(1-z_i)x}}{iz_r-z_i+1}\cdot[\cos(z_rx)+i\sin(z_rx)],$$

where we have used Eulers rule for complex numbers. Evaluating this in the lower limit of $x=-\infty$ it is observed that $\Im\mathfrak{m}(z) < 1$ otherwise $e^{(1-z_i)x} = \infty$ and the entire component is undefined. Therefore let $\Im\mathfrak{m}(z) < 1$, then evaluating the first component gives us,

$$\frac{e^{(iz+1)x}}{iz+1}\Big|_{x=-\infty}^{x=\ln(K)} = \frac{e^{(iz+1)\ln(K)}}{iz+1} - \lim_{x\to-\infty}\frac{e^{(iz+1)x}}{iz+1} = \frac{e^{(iz+1)\ln(K)}}{iz+1} - 0 = \frac{e^{(iz+1)\ln(K)}}{iz+1}. \tag{47}$$

The same idea can be applied to the second term, which decomposes to

$$K\frac{e^{izx}}{iz} = K\frac{e^{iz_rx}e^{-z_ix}}{iz} = K\frac{e^{-z_ix}}{iz}\cdot[\cos(z_rx)+i\sin(z_rx)],$$

and is well-defined as long as $\Im\mathfrak{m}(z) > 0$. Evaluating the second component we observe that,

$$K\frac{e^{izx}}{iz}\Big|_{x=\ln(K)}^{x=\infty} = \lim_{x\to\infty}K\frac{e^{izx}}{iz} - K\frac{e^{iz\ln(K)}}{iz} = -K\frac{e^{iz\ln(K)}}{iz}. \tag{48}$$

Conclusively we get:

$$\hat{\omega}(z) = \frac{e^{(iz+1)\ln(K)}}{iz+1} - K\frac{e^{iz\ln(K)}}{iz} = \frac{K^{(iz+1)}}{iz+1} - K\frac{K^{iz}}{iz} = \frac{K^{iz+1}iz - K^{iz+1}(iz+1)}{(iz+1)iz} = \frac{K^{(iz+1)}}{z^2-iz}, \qquad \text{for } 0<\Im\mathfrak{m}(z)<1. \tag{49}$$

Now under the risk-neutral measure $\mathbb{Q}$, we know that the price of a call at time 0 is given by:

$$C(S_0,K,T) = e^{-\int_0^T r_s\,ds}\mathbb{E}^{\mathbb{Q}}[\omega(x)], \tag{50}$$

[17] that is, the place where the payoff transform is well behaved.

with $\omega(x) = (S_T - K)^+ = S_T - \min(S_T, K)$, which then reduces (50) to

$$\begin{aligned}
C(S_0, K, T) &= e^{-\int_0^T r_s\, ds}\mathbb{E}^{\mathbb{Q}}[S_T - \min(S_T, K)] \\
&= e^{-\int_0^T r_s\, ds}\mathbb{E}^{\mathbb{Q}}[S_T \mid \mathcal{F}_0] - e^{-\int_0^T r_s\, ds}\mathbb{E}^{\mathbb{Q}}[\min(S_T, K)] \\
&= e^{-\int_0^T r_s\, ds} S_0 e^{\int_0^T r_s\, ds - qT + 0} - \frac{e^{-\int_0^T r_s\, ds}}{2\pi}\mathbb{E}^{\mathbb{Q}}\left[\int_{iz_i-\infty}^{iz_i+\infty} e^{-izx}\hat{\omega}(z)\, dz\right] \\
&= S_0 e^{-qT} - \frac{e^{-\int_0^T r_s\, ds}}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty}\mathbb{E}^{\mathbb{Q}}\left[e^{-izx}\right]\hat{\omega}(z)\, dz \\
&= S_0 e^{-qT} - \frac{e^{-\int_0^T r_s\, ds}}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty}\phi_T(-z)\hat{\omega}(z)\, dz
\end{aligned}$$

Normalizing the characteristic function by $X_T = \ln(\frac{S_T}{S_0 e^{\int_0^T r_s\, ds - qT}})$ we further get

$$\begin{aligned}
\phi_T(-z) &= \mathbb{E}^{\mathbb{Q}}\left[e^{-iz\ln(S_T)}\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[e^{-iz\ln\left(\frac{S_T}{S_0 e^{\int_0^T r_s\, ds - qT}}\right)}\right] \cdot e^{-iz\left[\ln(S_0) + \int_0^T r_s\, ds - qT\right]} \\
&= \mathbb{E}^{\mathbb{Q}}\left[e^{-iz\ln(X_T)}\right] \cdot e^{-iz\left[\ln(S_0) + \int_0^T r_s\, ds - qT\right]} \\
&= \phi_T^0(-z) \cdot e^{-iz\left[\ln(S_0) + \int_0^T r_s\, ds - qT\right]},
\end{aligned}$$

and thus inserting the standardized characteristic function yields

$$\begin{aligned}
\cdots &= S_0 e^{-qT} - \frac{e^{-\int_0^T r_s\, ds}}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty}\phi_T^0(-z) \cdot e^{-iz\left[\ln(S_0) + \int_0^T r_s\, ds - qT\right]}\frac{K^{iz+1}}{z^2 - iz}\, dz \\
&= S_0 e^{-qT} - \frac{e^{-\int_0^T r_s\, ds} K}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty}\phi_T^0(-z) \cdot e^{-iz\left[\ln(\frac{S_0}{K}) + \int_0^T r_s\, ds - qT\right]}\frac{1}{z^2 - iz}\, dz.
\end{aligned}$$

Now let $k = \ln(\frac{S_0}{K}) + \int_0^T r_s\, ds - qT$, choose $z_i = \frac{1}{2}$ and substitute $z = u + \frac{1}{2}i$:

$$\begin{aligned}
\cdots &= S_0 e^{-qT} - \frac{e^{-\int_0^T r_s\, ds} K}{2\pi}\int_{-\infty}^{\infty}\phi_T^0\left(-u - \frac{1}{2}i\right)e^{-i(u+\frac{1}{2}i)k}\frac{1}{(u+\frac{1}{2}i)^2 - i(u+\frac{1}{2}i)}du \\
&= S_0 e^{-qT} - \frac{e^{-\int_0^T r_s\, ds + \frac{1}{2}k} K}{2\pi}\int_{-\infty}^{\infty}\phi_T^0\left(-u - \frac{1}{2}i\right)e^{-iuk}\frac{1}{u^2 + \frac{1}{4}}du \\
&= S_0 e^{-qT} - \frac{e^{-\frac{1}{2}\left[\int_0^T r_s\, ds + qT\right]}\sqrt{S_0 K}}{2\pi}\int_{-\infty}^{\infty}\phi_T^0\left(-u - \frac{1}{2}i\right)e^{-iuk}\frac{1}{u^2 + \frac{1}{4}}du.
\end{aligned}$$

Using the symmetry property of Fourier integrals (which is a consequence of $\mathfrak{Re}(x) = \frac{x+\bar{x}}{2}$ and $\phi(-x) = \bar{\phi}(x)$. Eg. see (Schmelzle, 2010) formula (3.13)), we finally obtain:

$$C(S_0, K, T) = S_0 e^{-qT} - \frac{e^{-\frac{1}{2}\left[\int_0^T r_s\, ds + qT\right]}\sqrt{S_0 K}}{\pi}\int_0^{\infty}\mathfrak{Re}\left[\phi_T^0\left(u - \frac{1}{2}i\right)e^{iuk}\right]\frac{1}{u^2 + \frac{1}{4}}du.$$

As argued in (Schmelzle, 2010) doing a sign change in the characteristic function does not change the value of the integral as long as the sign in the corresponding phase factor (e^{-iuk}) is changed as well. □

From the above pricing formula, we can derive an equation to efficiently compute the term structure of the ATM volatility skew. Recovering the implied volatility can be done as described in Section 2.4. Under the Lewis approach, call prices for both models

can be recovered by (46) where the BS price can be subtracted on both sides of the equation thus giving us,

$$S_0 e^{-qT} - \frac{\sqrt{S_0 K} e^{-\frac{1}{2}\left(\int_0^T r_s\, ds + qT\right)}}{\pi} \int_0^\infty \Re\mathfrak{e}\left[e^{iuk}\left(\phi_{Model,T}\left(u - \frac{i}{2}\right) - \phi^0_{BS,T}\left(u - \frac{i}{2}\right)\right)\right] \frac{du}{u^2 + \frac{1}{4}} = 0$$

$$S_0 e^{-qT} - \frac{\sqrt{S_0 K} e^{-\frac{1}{2}\left(\int_0^T r_s\, ds + qT\right)}}{\pi} \int_0^\infty \Re\mathfrak{e}\left[e^{iuk}\left(\phi_{Model,T}\left(u - \frac{i}{2}\right) - e^{-\frac{1}{2}(u^2 + \frac{1}{4}\sigma^2_{BS,T}T)}\right)\right] \frac{du}{u^2 + \frac{1}{4}} = 0$$

Since the $\phi(\cdot)$ does not depend on k then (Gatheral, 2011) shows that differentiating with respect to k and evaluating at $k = 0$ yields the ATM implied volatility skew:

$$\left.\frac{\partial \sigma^{IV}(k,T)}{\partial k}\right|_{k=0} = -e^{\frac{\sigma^2_{BS}T}{8}} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{T}} \int_0^\infty \frac{u\, \Im\mathfrak{m}\left[\phi_{Model,T}(u - \frac{i}{2})\right]}{u^2 + \frac{1}{4}}\, du. \tag{51}$$

3.3 Assessment of the calibration procedure

The calibration procedure was done by simultaneously fitting the implied volatility surface across strikes and maturities for all models. In more complex models such as the SVJ model, one could argue that we could have produced a better fit by calibrating the Heston dynamics, then the jump process and conclusively use these calibrated parameters as initial values for a final recalibration of the model. This is however not done in order to establish a fair framework for all models. To assess the performance of the calibration, we utilize different accuracy measures involving prices and implied volatilities. We use the mean squared error for both call prices (MSE^{price}) and implied volatilities (MSE^{IV}) together with the mean absolute percentage error of prices ($MAPE^{price}$) and implied volatilities ($MAPE^{IV}$). The motivation for using both accuracy measures is that mean squared error is scale-dependent, and thus are more sensitive to outliers. Mean absolute percentage error is less senitive and also provides us with a percentage deviation from the true data, which is easier to interpret. Lastly, we do not want to have any preferential treatment to the mean squared error, therefore including another accuracy measure. For ease of notation we have removed S_0, r, q from the implied volatility measures and thus all measures are described below:

$$MSE^{price} = \frac{1}{\#options} \sum_{j \in \mathcal{T}} \sum_{i \in K_{T_j}} \left(C^{ij}_{Market} - C^{ij}_{Model}(\theta)\right)^2, \qquad MAPE^{price} = \frac{1}{\#options} \sum_{j \in \mathcal{T}} \sum_{i \in K_{T_j}} \left|\frac{C^{ij}_{Market} - C^{ij}_{Model}(\theta)}{C^{ij}_{Market}}\right| \tag{52}$$

$$MSE^{IV} = \frac{1}{\#options} \sum_{j \in \mathcal{T}} \sum_{i \in K_{T_j}} \left(\sigma^{IV}_{Market}(T_j, K_i, C^{ij}_{Market}) - \sigma^{IV}_{Model}(T_j, K_i, C^{ij}_{Model})\right)^2 \tag{53}$$

$$MAPE^{IV} = \frac{1}{\#options} \sum_{j \in \mathcal{T}} \sum_{i \in K_{T_j}} \left|\frac{\sigma^{IV}_{Market}(T_j, K_i, C^{ij}_{Market}) - \sigma^{IV}_{Model}(T_j, K_i, C^{ij}_{Model})}{\sigma^{IV}_{Market}(T_j, K_i, C^{ij}_{Market})}\right|. \tag{54}$$

	θ	ν	C	G	M	Y	V_0	κ	$\bar{V}$	σ	ρ	α	δ	λ
VG	-0.0986	0.8050								0.0932				
CGMY			0.0160	1.0438	33.4980	1.2993								
Heston model							0.0029	5.1602	0.0342	1.6581	-0.7239			
Bates model							0.0010	4.3905	0.0375	1.9875	-0.7952	-0.0027	0.0152	8.3718

Table 3: Shows the parameters for the calibrated models. While σ in Heston and Bates model is the vol of vol parameter, it is the variance of the diffusion component for the VG model. Moreover, it is argued by (Lewis, 2001) that the VG model is a good stock market model if $x + y \geq 0$ and $y - x \leq -1$ with $x = \sqrt{2/(\nu\sigma^2) + \theta^2/\sigma^4}$ and $y = \theta/\sigma^2$. In our case we have $x + y = 1.8118$ and $y - x = -2.1161$.

	MSE^{price}	$MAPE^{price}$	MSE^{IV} $(\times 10^{-4})$	$MAPE^{IV}$
VG	62.2926	0.1323	2.4746	0.0836
CGMY	63.0292	0.0930	1.9243	0.0735
Heston model	1.8131	0.0593	0.2137	0.0278
Bates model	2.6771	0.0505	0.1440	0.0223

Table 4: The corresponding accuracy measures for the calibrated models. The accuracy measures from left to right are, the mean squared pricing error (MSE^{price}), mean absolute percentage pricing error ($MAPE^{price}$), mean squared implied volatility error (MSE^{IV}) and the mean absolute implied volatility percentage error ($MAPE^{IV}$).

From Table 3 we observe the calibrated parameters for each model and from Table 4 the accuracy measures for each model. The tables can be summarized as follows. First, we see that the skewness of the Gamma distribution in the VG model is $\theta < 0$, therefore giving rise to a negatively skewed implied volatility surface. Furthermore, ν is quite large thus raising the likelihood of large jumps, which in turn raises the tail probabilities and excess kurtosis as described in (Madan et al., 1998). Unlike the geometric Brownian motion model, that lacks a kurtosis parameter (ν), in the VG model this property of higher prices for tail events (thus higher IV), is reflected in the increased (risk-neutral) kurtosis of the process with volatilities typically being unaffected[18]. We find similar results with the CGMY model, since $G < M$ then the left tail of the distribution for $(X_t)_{t\geq 0}$ is heavier than the right tail, which is in line with the risk-neutral distribution implied from option prices. This also gives us a negatively skewed jump distribution which again gives rise to skew in the smile. It is furthermore observed that the 'Y' parameter in the CGMY model is $1 < Y < 2$ implying that the CGMY model exhibits infinite activity together with infinite variation. The infinite variation just tells us that the Lévy measure around [-1,1] is infinite, ie. $\int_{\{|x|\leq 1\}} |x|\, \Psi(dx) = \infty$, and thus the Lévy process exhibits infinitely many small jumps. However, as seen by the performance evaluation in Table 4 the extra parameter does not seem to add a noticeable improvement over the VG model, which is also noted in Figure 2 and furthermore seen in examples analyzed in current literature (see eg. (Fiorani, 2004) and (Carr et al., 2002)).

Secondly, we see that Heston and Bates capture the leverage effect since $\rho < 0$, but both models still need a high vol of vol parameter σ to fit the short-term skew, thus violating the Feller condition. It is empirically evident that Heston and Bates violate the Feller condition since it implies a higher mean reversion and reduces the variance of the stochastic volatility, which leads to a decrease of the convexity of the smile. In order to increase the smile convexity, we need to increase the vol of vol parameter σ, but cannot do so, since it is bounded from above. In continuity, a violated Feller condition does not impose any problem when the variance hits zero, since the drift-term ($dV_t = \kappa \bar{V}\, dt$) would instantly make it positive. However, when violating the Feller condition, we do observe a less pronounced improvement by allowing for jumps, which is also argued in (Kokholm & Stisen, 2015). Thus the penalty for imposing the Feller condition is most severe for the SV model because it constrains the parameter space (especially the vol of vol parameter), whereas the more flexible SVJ model can compensate for the Feller condition constraint with the added jumps[19]. When that is said, the Gaussian jump specification in the Bates model do extend the curvature of the short-term implied volatility (as seen in Figure 4), enabling it to provide a greater short-term fit.

[18] The same empirical studies was done in (Madan et al., 1998), which shows the same outcome.
[19] Again see (Kokholm & Stisen, 2015).

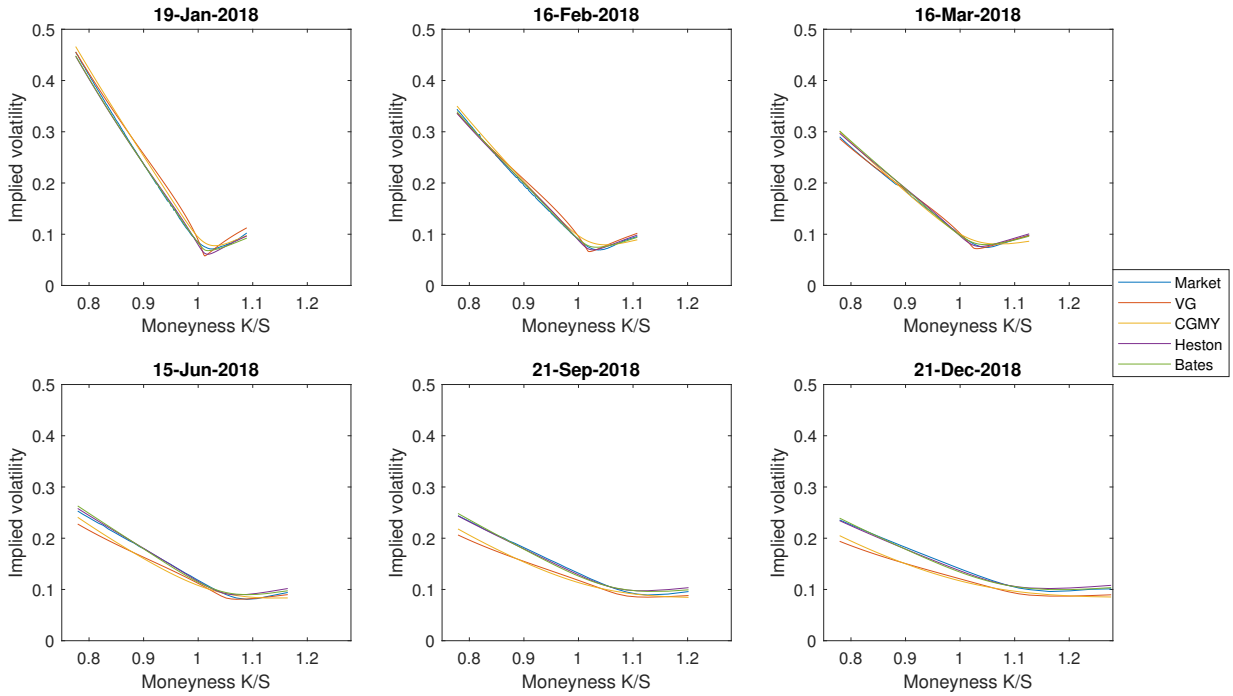


Figure 1: Fit of the implied volatility smiles for all calibrated models. The rapid decay in skewness for the exponential Lévy models is evident here by observing the fit for the long maturities.

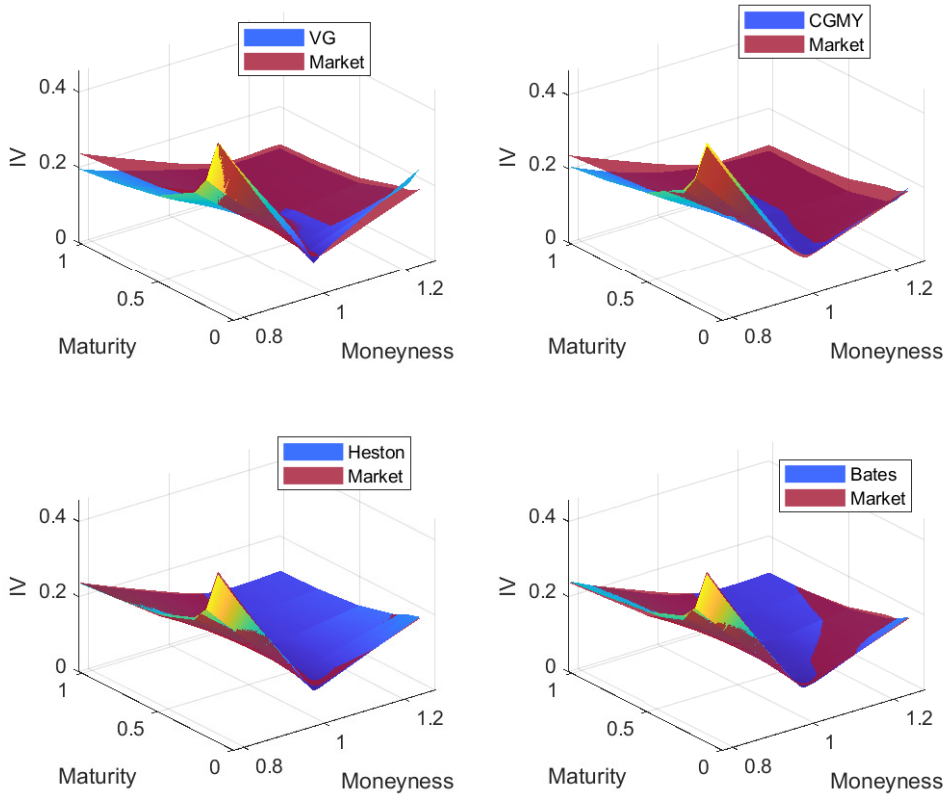


Figure 2: The fit of the implied volatility surface for the models against the market surface. To ensure an equivalent observation set across maturities for market calls, we utilized linear interpolation and extrapolation. The model surfaces was calculated by inverting the call prices found using Fourier pricing methods with corresponding calibrated parameters.

From Figure 1, Figure 2 and the corresponding error surfaces provided in Figure 4 in the appendix, we observe as is empirically often evident that the VG and the CGMY model are not able to capture the long-term implied volatility consistently with the market. Since the BS call prices are increasing in sigma, we see that the VG and CGMY model under-prices long maturity options due to a combination of the floating smile property and the decay in the curvature when $T \to \infty$. Moreover, it seems that in our case, the VG model overprices short-term options away from ATM (aka. OTM and ITM), whereas the CGMY model has a hard time capturing the short-term OTM calls. As observed for the Heston and Bates model they fare considerably better, as expected. However, both models fall a bit short of fully capturing the short-term upward sloping curvature of the OTM options. This increase in implied volatility can be due to market participants speculating in short-term price jumps/increases and are thus willing to pay more for OTM short-term options. For the Heston model the price increase has to be captured through the diffusion component, where it is almost able to generate enough curvature (at a cost of violating the feller condition) to capture the dynamics of the short-term options. However, as seen from the error surface, it cannot completely capture the short-term skew for the left-wing of the implied volatility surface.

For the Bates model, the Gaussian jump specification in the asset log-price process improves the short-term maturity behavior of implied volatility without deteriorating long-term smiles. Indeed, the Bates model is able to generate an acceptable amount of skew reducing the surface error even further, but still not able to completely fit the short-term left wing skew. While Heston under-prices long-term ITM calls it seems that both models capture the market surface quite well with minor inconsistencies in contrast to the VG and CGMY model[20]. Lastly, the left wing curvature is reasoned by the fact that OTM puts functions as market crash insurance for larger portfolios, thus helping institutional investors minimizing the risk from rare-events, which in turn drives up the option prices (and thus implied volatility).

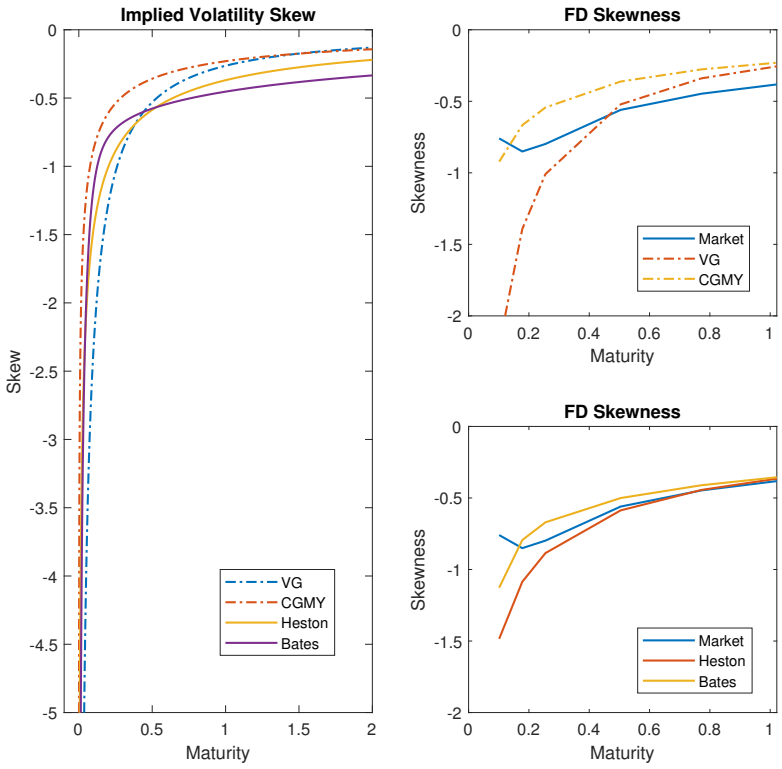


Figure 3: The graph on the left side, shows the calculated ATM implied volatility skew from the equation in (51), where the two other figures compares the ATM implied volatility skew for the models against the market ATM skew computed using Central Finite Difference on $\frac{\partial \sigma^{IV}(k,T)}{\partial k}\big|_{k=0}$ for $k = \ln(\frac{K}{S_0})$.

Figure 3 shows how the skew behaves for each model when increasing time to maturity (on the left) and then compared to the market skew, calculated using central finite difference (on the right). Indeed, the exponential Lévy models do provide a high short-term skew, which decays faster than the SVJ/SV models. However, it seems that the skew for both models decays too fast, which is the fine reason why the models are not able to capture the long-term IV market skew. While the skew for the Lévy

[20] Please be aware that linear extrapolation has been used for short-term OTM IV for all models which was from a moneyness level at 1.09 till 1.25. This of course, have a clear saying in the feasibility of the fit at this level, but nevertheless the upward sloping right wing for short-term OTM options is in agreement with empirics.

models do not align with the market skew across maturity, the slow decaying long-term skew for the SVJ/SV model obtains a good fit against the market skew. The reason lies with the correlation parameter ρ since it is important to realize that - due to the symmetry of Brownian increments - the asymmetry of the (risk-neutral) distribution - therefore, the implied volatility skew - is completely determined by this correlation coefficient which plays a fundamental role for generating smiles and skews. Thus a correlation parameter $\rho < 0$ includes the leverage effect and (somewhat) 'explains' the downward sloping volatility skew. In conclusion the overall long-term skew should be in-between the skew of the Heston model and Bates model[21].

4 Conclusion and further research

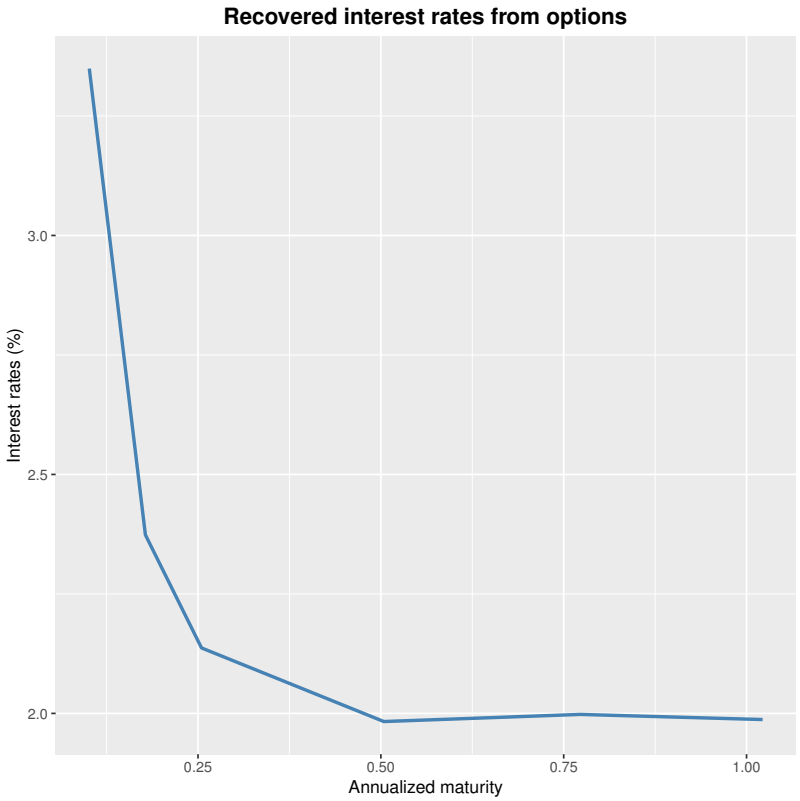
We constructed a survey for the performance of two ordinary exponential Lévy models against two well-known SV/SVJ-type models to observe how they fare against complex models. Through the extensive theoretical introduction we establish ways to price both families of models using characteristic functions and Fourier pricing methods. Moreover, we establish a theoretical base of reasons why the exponential Lévy models fail to fully capture the market surface (besides the lack of parameters), which is further strengthened in the empirical analysis. In order to provide the best fit possible we removed static arbitrage strategies as done in (Carr & Madan, 2005), and calibrated directly to the implied volatilities instead of the classical method of calibrating to the call prices. We conclude that due to the fast decay in skew and kurtosis for any exponential Lévy model whatsoever, it will not be able to fit the long-term implied volatility surface of the market. Of course, in hindsight to making this assignment it could have been more feasible to test whether the VG model with stochastic volatility induced via time-change - thus eliminating moment decays through leverage effect - provided a better fit to the surface and conclusively whether it could contest the benchmark models. Moreover, due to the simplicity of both exponential Lévy models[22] they are inconsistent in capturing the short-term market skew when trying to accommodate the entire term-structure of the surface.

[21] In our particular coincidence.

[22] In order to alleviate confusion "simplicity" in the amount of parameters to fit the market surface: The VG and CGMY model have 3 and 4 parameters to calibrate versus 5 and 8 for the Heston and Bates model respectively.

Appendix A Interest rate graph

The graph below depicts the recovered interest rates across the first 6 maturities.



Appendix B Error surfaces

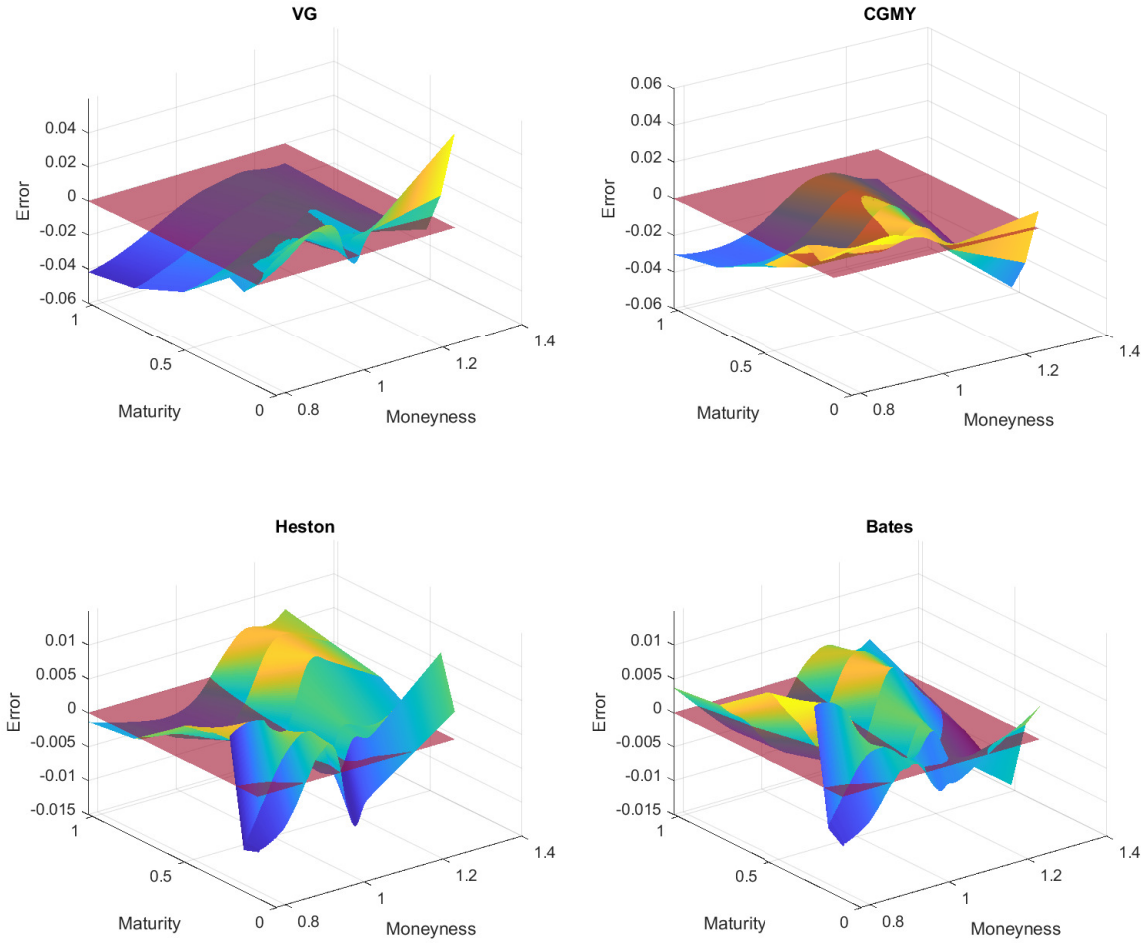


Figure 4: The associated error surfaces for all models, calculated using $Error = \sigma^{IV}_{Model}(\cdot) - \sigma^{IV}_{Market}(\cdot)$ with a flat transparent surface at zero for a better overview. For the Heston and Bates model the lowest error observation for the short-term left wing (at deep ITM) is -0.0102 and -0.006920 respectively.

References

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, *9*(1), 69–107.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *The Journal of Business*, *75*(2), 305–332.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for lévy processes. *Mathematical finance*, *13*(3), 345–382.

Carr, P., & Madan, D. B. (2005). A note on sufficient conditions for no arbitrage. *Finance Research Letters*, *2*(3), 125–130.

Derman, E. (1999). Regimes of volatility. *Risk*, 55–59.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, *68*(6), 1343–1376.

Fiorani, F. (2004). Option pricing under the variance gamma process. *Available at SSRN 1411741*.

Gatheral, J. (2011). *The volatility surface: a practitioner's guide* (Vol. 357). John Wiley & Sons.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, *6*(2), 327–343.

Kokholm, T. (2016). Pricing and hedging of derivatives in contagious markets. *Journal of Banking & Finance*, *66*, 19–34.

Kokholm, T., & Stisen, M. (2015). Joint pricing of vix and spx options with stochastic volatility and jump models. *The Journal of Risk Finance*.

Lewis, A. L. (2001). A simple option formula for general jump-diffusion and other exponential lévy processes. *Available at SSRN 282110*.

Madan, D. B., Carr, P. P., & Chang, E. C. (1998). The variance gamma process and option pricing. *Review of Finance*, *2*(1), 79–105.

Rogers, L., & Tehranchi, M. (2010). Can the implied volatility surface move by parallel shifts? *Finance and Stochastics*, *14*(2), 235–248.

Schmelzle, M. (2010). Option pricing formulae using fourier transform: Theory and application. *Preprint, http://pfadintegral. com*.

Schoutens, W., Simons, E., & Tistaert, J. (2003). A perfect calibration! now what? *The best of Wilmott*, 281.

Tankov, P. (2003). *Financial modelling with jump processes*. Chapman and Hall/CRC.