OPTION PRICING UNDER VARIANCE-GAMMA DYNAMICS, WHEN THE PARAMETERS ARE STOCHASTIC

KYRIAKOS CHOURDAKIS

Abstract. In this paper we extend option pricing under Variance-Gamma (VG) dynamics, by assuming that the parameters of the VG process are stochastic. We present a methodology that allows one to compute option prices, under virtually any set of diffusive dynamics for the parameters of the VG process. We show that this extension can accommodate for a plethora of implied volatility skews and term structures. An example of a VG process with time varying skewness is presented in detail, to illustrate the methodology. We find that such an extension provides enough flexibility to allow for complex volatility term structure dynamics, such as skew reversals. An example, based on a set of observed SP500 option prices, exhibits the benefits of this approach.

1. Introduction

Over the last thirty years, a vast number of pricing models have been proposed as an alternative to the classic Black-Scholes (BS) approach, driven by the well documented biases of the BS formula. Moving into the implied volatility space, observed option prices exhibit well defined patterns across moneyness and maturity. In some markets, such as the currency markets, a convex implied volatility smile is observed across moneyness, while in others, such as the stock index options markets, the pattern resembles more a downward sloping skew. Across maturities, the term structure of implied volatilities can be downward or upward sloping, similar to the term structure of interest rates. In order to accommodate for these biases, some of the BS assumptions have been relaxed over the years.

There are two key assumptions that need to be made, in order to price in the BS world: that asset prices follow continuous sample paths, and that there is a single source of uncertainty. Under these assumptions, a continuously rebalanced portfolio can be used to perfectly hedge an options position, thus determining a unique price for the option. Relaxing the assumption of a unique source of uncertainty leads to the *stochastic volatility* family of

models, where the volatility parameter follows a separate diffusion; relaxing the assumption of continuous sample paths leads to the *jump diffusion* models, where infrequent jumps are added to the standard BS diffusion. In both cases the market is incomplete, and option prices are not determined in a unique way.

It has been recently recognized that, although the presence of jumps is pivotal, allowing for continuous sample paths in not necessary to price option contracts. The family of pure jump models, where the underlying asset is subject to (perhaps infinitely dense) jumps, suffices to capture stylized facts of asset returns, such as the negative skewness and excess kurtosis. The theory of Lévy processes links such specifications to the characteristic functions of their respective densities, and provides the probabilistic framework that offers option prices. Broadly speaking, a Lévy log-price will exhibit a sequence of independent and identically distributed increments, and offers a more general framework than the lognormal assumption of BS.[1] More importantly, though, assuming a jump process breaks the link between the statistical properties of the underlying asset and the risk neutral properties, since a replicating hedge cannot be constructed. Therefore, the risk-neutral parameters under Lévy dynamics can be separated from the 'objective' set of parameters of the underlying process.

Although Lévy processes are very versatile, from a distribution point of view, they fail to capture the time series properties of asset returns. In fact, since the densities assumed are i.i.d., a Lévy process cannot generate the time-varying higher moments, or the volatility clusters that are typical of financial time series. This paper addresses this issue and presents a class of Lévy processes that exhibit stochastic volatilities, or other higher moments.

2. The Variance-Gamma process with stochastic parameters

2.1. **The Variance-Gamma process.** The Variance-Gamma (VG) process is defined as a Brownian motion with drift θ and volatility σ, evaluated at a random Gamma time

$$X_{VG}(t;\theta,\sigma,\nu) = \theta G(t;\nu) + \sigma W(G(t;\nu)) \tag{2.1}$$

The Gamma process $G(t;\nu)$ has (normalized) mean rate equal to one, and variance rate ν, and is assumed independent of the Brownian motion W. It

[1] In this context, the BS analysis becomes a special case.

can be shown that the characteristic function of X is given by

$$E\exp\{\mathtt{i}uX_{VG}(t)\} = \left(1 - \mathtt{i}u\theta\nu + \frac{1}{2}u^2\sigma^2\nu\right)^{-\frac{t}{\nu}} \tag{2.2}$$

Broadly speaking, out of the three VG parameters, σ controls for the volatility, θ is responsible for skewness and ν generates kurtosis. Therefore, the VG density can take a variety of shapes, and can therefore accommodate for a wide range of implied volatility smiles. Assuming that the log-asset price follows the VG process described above, Madan, Carr and Chang [12] show that option prices can be computed using Fourier inversion.

The major drawback of the VG approach is that (being a Lévy process) it can only exhibit i.i.d. increments. Therefore, although almost any shape of the volatility skew can be generated for a given maturity, the VG model fails to capture the term structure of implied volatilities. In addition, VG fails to conform with the well documented time-variation of higher moments of asset returns, volatility clustering and volatility persistence. In order to capture these features, CGMY [2] have recently proposed 25 variations that can generate time-varying volatilities; the common feature of them being that they are all Lévy processes subordinated to a stochastic clock. The higher (or lower) the rate of time changes induced by this clock, the higher (or lower) the resulting volatility. More recently, and motivated by evidence of time varying skews in the currency options markets, Carr and Wu [4] propose a specification with two stochastic clocks, one for the upward moves and one for the downward ones.

2.2. **Introducing stochastic parameters.** In this paper we focus on option pricing when the parameters of (2.1) themselves are random, following a process of the general form

$$Y(t) = \int_0^t \alpha(Y(s))ds + \int_0^t \beta(Y(s))dZ(s) \tag{2.3}$$

In the context of the VG model that we utilize, Y can stand for any one of the parameters σ, θ or ν. As an example, we will focus on the time variation of the parameter θ that controls the skewness. This choice is motivated by the evidence presented in Carr and Wu on exactly this salient feature of currency options markets. The same approach can be used to treat any other parameter as stochastic, or perhaps even all of them, by introducing additional SDEs. The treatment for these cases follows in a straightforward fashion. To ensure that skewness can be positive as well as negative, and to

force stationarity, we will assume that the parameter θ follows an Ohrstein-Uhlenbeck (OU) process

$$\theta(t) = \int_0^t \kappa\left(\bar{\theta} - \theta(s)\right) ds + \int_0^t \phi dZ(s) \tag{2.4}$$

The Brownian motion Z is considered independent of all processes encountered so far. Again, the choice of an OU process is arbitrary; the approximation methodology can handle almost all functional forms for $\alpha(\cdot)$ and $\beta(\cdot)$. It is also straightforward to extend this approach to other infinitely divisible processes, such as the NIG or the CGMY process.

3. Approximating the diffusion

As Kushner [10, 8] shows, the diffusion (2.3) can be approximated arbitrarily well using a carefully selected Markov chain that lives on a finite grid. The Markov chain has to satisfy the 'local consistency' requirements, namely that it exhibits the same instantaneous drift and volatility as the diffusion in question.[2] In essence, the log-price process is approximated by a sufficiently dense regime switching model; Chourdakis [6] shows how closed form characteristic functions for such regime switching models can be used to approximate option prices, and how correlations between the state process and the prices process can be introduced.

3.1. **The 'local consistency' concept.** Consider a grid $\Upsilon^h = \{Y_0^h, \cdots, Y_{N^h}^h\}$, indexed by $h > 0$. Assume that as $h \to 0$, $|Y_{j+1}^h - Y_j^h| \to 0$ for all $j = 1, \cdots, N^h$, or in other words that the grid becomes progressively denser. Also assume that $N^h \to \infty$, and that the grid tends to cover the support of the diffusion (2.3). We denote the approximating chain with $Y_{MC}^h(t) \in \Upsilon$, and the corresponding rate matrix with $\mathbf{Q}^h = [q_{i,j}^h]$. Also let $Y(t) = Y_{MC}^h(t)$. It can be shown, for example Aït-Sahalia [1], that allowing switches to the neighboring states Y_{j-1}^h and Y_{j+1}^h is sufficient to ensure local consistency. The local consistency conditions stipulate that, for $\delta > 0$

$$\begin{aligned} E_t^h(Y_{MC}(t+\delta) - Y_{MC}(t)) &= \alpha(Y(t))\delta + o(\delta) \\ E_t^h(Y_{MC}(t+\delta) - Y_{MC}(t))^2 &= \beta^2(Y(t))\delta + o(\delta) \end{aligned} \tag{3.1}$$

The expectations operator E^h is applied with respect to the rate matrix $\mathbf{Q}^h$. Kushner shows that given local consistency, the Markov chains indexed by h will converge weakly to the diffusion 2.3, as $h \to 0$.

[2] Kushner and DiMasi [11] generalize this approach for jump diffusions.

Given a grid, equations (3.1) naturally lead to a system that can be solved to retrieve the approximating rate matrix. For convenience we drop the index h from the rest of the exposition, keeping in mind though that the results will only hold asymptotically. Over a time interval δ, we can either remain at Y_j, move up by d_U, or move down by d_D (to the points $Y_{j\pm1}$). The local consistency conditions can be restated as

$$\begin{aligned} -q_{j-1,j}d_D\delta + q_{j+1,j}d_U\delta &= \alpha(Y_j)\delta + o(\delta) \\ q_{j-1,j}d_D^2\delta + q_{j+1,j}d_U^2\delta &= \beta^2(Y_j)\delta + o(\delta) \end{aligned}$$

The above relationships can lead to the approximations schemes. Passing to the limit $\delta \to 0$, we can solve the above system to retrieve the rate matrix elements, for all $j = 2, \cdots, N-1$, as

$$\begin{aligned} q_{j-1,j} &= \tfrac{1}{d_D(d_U+d_D)}\left(\beta^2(Y_j) - d_U\alpha(Y_j)\right) \\ q_{j+1,j} &= \tfrac{1}{d_U(d_U+d_D)}\left(\beta^2(Y_j) + d_D\alpha(Y_j)\right) \\ q_{j,j} &= -q_{j+1,j} - q_{j+1,j} \end{aligned} \tag{3.2}$$

For $\mathbf{Q}$ to be a valid rate matrix, we have to ensure that the off-diagonal elements are non-negative. For this reason, when the equations (3.2) above produce invalid elements, one can replace them with a scheme that always remains valid

$$\begin{aligned} q_{j-1,j} &= \tfrac{1}{d_D(d_U+d_D)}\left(\beta^2(Y_j) + (d_D+d_U)\alpha^-(Y_j)\right) \\ q_{j+1,j} &= \tfrac{1}{d_U(d_U+d_D)}\left(\beta^2(Y_j) + (d_D+d_U)\alpha^+(Y_j)\right) \\ q_{j,j} &= -q_{j+1,j} - q_{j+1,j} \end{aligned} \tag{3.3}$$

The superscripts $\alpha\pm(\cdot)$ denote the positive and negative parts of $\alpha(\cdot)$. Both schemes match the instantaneous drift, but only the first one matches the instantaneous volatility too.[3] We now turn in issues related to the grid construction.

3.2. **The grid.** The grid of the Markov chain has to be carefully chosen, keeping in mind that: (i) it has to be sufficiently dense, in order for the local consistency conditions to apply, and (ii) it has to be sufficiently wide, in order for the effect of truncating the support of the diffusion to be negligible. In this paper, we will use the unconditional density to create the grid: for the OU process the unconditional density is a normal, with mean $\bar{\theta}$ and variance $\frac{\phi^2}{2\theta}$. We can construct an M-point grid by setting Y_j equal to

[3]It is straightforward to verify that the instantaneous volatility of approximations (3.3) is $\beta^2(Y_j) + (d_D + d_U)|\alpha(Y_j)|$. This will asymptotically converge to the diffusion volatility $\beta(\cdot)$ as the grid becomes finer.

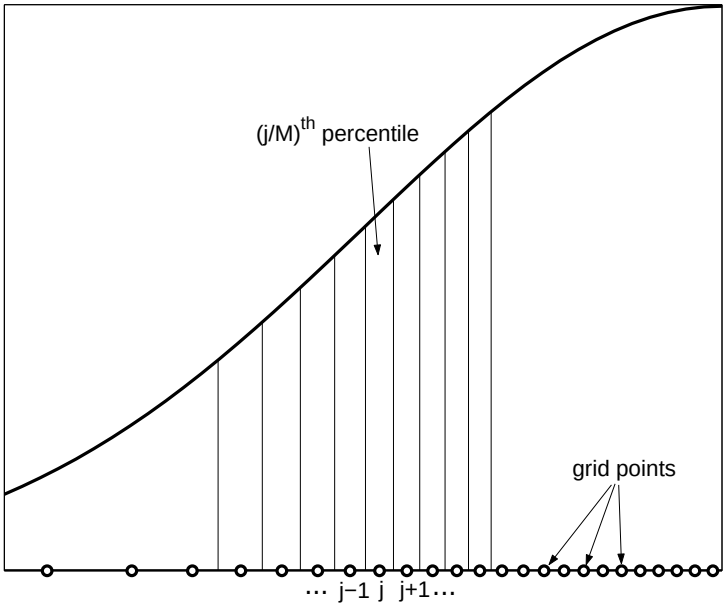


FIGURE 1. Grid construction. Setting the grid points at the $\left(\frac{1}{2M}+(j-1)\frac{1}{M}\right)$-percentile of the unconditional distribution of the diffusion.

the $\left(\frac{1}{2M}+(j-1)\frac{1}{M}\right)$-percentile of the unconditional distribution. Figure 1 illustrates this construction. This approach offers two major benefits: (i) the grid extends and covers the diffusion support as M increases, and (ii) the grid is tighter where it matters most, that is to say, where the diffusion is more likely to take values in the future. Of course other constructions are also possible; for example, if more extreme values are required, the symmetric regularized Beta function $B_{SR}(\cdot,\eta)$ can be used to 'spread out' the percentiles generated above.[4] Figure 2 gives examples of $B_{SR}(\cdot,\eta)$ for different values of the parameter η.

It is assumed that the grid has be selected in such a way, that the behavior of the chain on the boundaries is not important. Generally speaking, this will be the case if the probability of reaching the boundaries (in the relevant time frame) is negligible. These probabilities can be actually examined, since the time-t transition density of the Markov chain is given by the matrix

[4]The symmetric regularized Beta function maps the interval $[0,1]\to[0,1]$ in a sigmoid way, maintaining $\frac{1}{2}\to\frac{1}{2}$. It is given by $B_{SR}(\cdot,\eta)=\frac{B(\cdot,\eta,\eta)}{B(\eta,\eta)}$, where the numerator and denominator are the incomplete and complete Beta functions, respectively. It can also be computed as the cumulative density function of the Beta distribution, with both shape parameters equal to η. Setting $\eta=1$ results in a linear (in probability) grid, while values of $\eta>1$ spread out the grid points towards more extreme values.

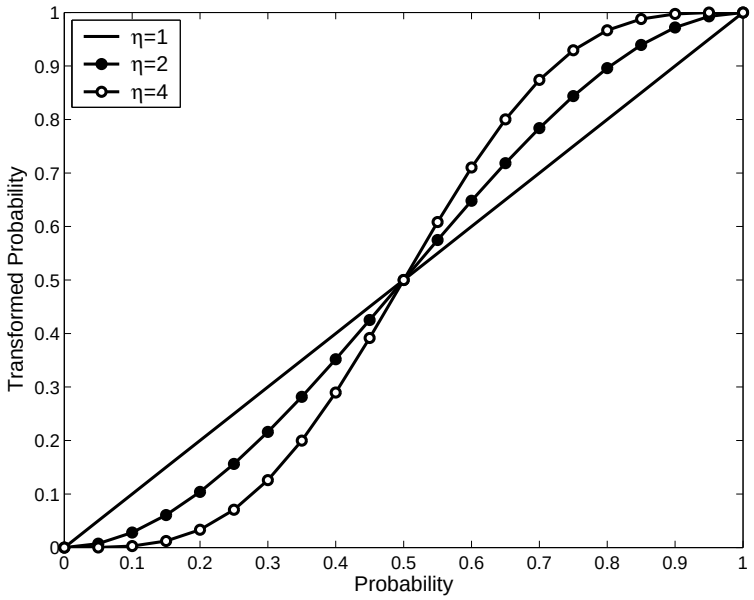


FIGURE 2. The symmetric regularized Beta function, $B_{SR}(\cdot, \eta)$, for different values of η. This function can be used to 'spread out' the percentiles that lead to the grid construction.

exponential $\mathbf{P}(t) = \exp\{t\mathbf{Q}\}$. If this is the case, the boundaries can be made absorbent or reflective, without any significant impact on the approximation procedure.

3.3. **The characteristic function.** As Madan, Carr and Chang [12] show, the VG asset price process under risk neutrality has to satisfy (after normalizing $S(0) = 1$)

$$S(t) = \exp\{(r + \mu)t + X(t; \sigma, \theta, \nu)\}, \text{ with}$$
$$\mu = \frac{1}{\nu} \log\left\{1 - \theta\nu - \frac{1}{2}\sigma^2\nu\right\}$$

MCC show that the process of the log-price is infinitely divisible, with characteristic function given by

$$f(u) = \exp\{t\Phi(u)\}, \text{ with}$$
$$\Phi(u) = (r + \mu)\mathtt{i} - \frac{1}{\nu} \log\left\{1 - \mathtt{i}u\theta\nu - \frac{1}{2}u^2\sigma^2\nu\right\}$$

By definition, $\Phi(\cdot)$ is the characteristic exponent of the infinitely divisible Lévy process.

Chourdakis [5, 6] shows that if some model parameters are considered stochastic, the (time-t) characteristic function of the log-asset price $s(t) = \log S(t)$ can be approximated by the characteristic function of the discrete state process. This function, can in turn be computed in terms of a matrix exponential, which is based on the rate matrix of the approximating chain, and the conditional characteristic exponents.[5] The general form of the characteristic function is

$$f_{MC}(u) = \iota' \cdot \exp\{t\mathbf{\Psi}(u)\} \cdot \mathbf{y}_0, \text{ where} \tag{3.4}$$

$$\mathbf{\Psi}(u) = [\psi_{i,j}(u)] = \begin{cases} q_{i,i} + \Phi(u|Y_i) & \text{, if} j = i \\ q_{i,i\pm1} & \text{, if} j = i \pm 1 \\ 0 & \text{, otherwise} \end{cases}$$

Here, $\psi_{i,j}$ denotes the typical element of matrix $\mathbf{\Psi}$, and $\Phi(u|Y_i)$ denotes the conditional characteristic exponent. ι is a vector of ones, and $\mathbf{y}_0$ is the initial (time-0) distribution of the parameter Y. For example, if $Y(0) = Y_k$, then $\mathbf{y}_0$ would be a vector of zeros, with the k-th element equal to one.

3.4. **Example: time varying skewness.** To illustrate the construction of the approximating characteristic function, we now turn to equation (2.4), and show how the approximation methodology can be applied. All numerical illustrations assume the parameter values $\{\kappa, \bar{\theta}, \phi\} = \{5.00, -0.30, 0.50\}$ for the OU diffusion. In this setting, the skewness parameter of the VG process is time-varying, and follows a mean reverting process. These parameters imply that the unconditional density for the skewness parameter θ is a normal, with mean $\bar{\theta} = -0.30$, and variance $\frac{\phi^2}{2\kappa} \approx 0.16^2$. The rest of the VG parameters, which are kept constant, assume the values $\{\sigma, \nu\} = \{0.30, 1.00\}$.

We initially create a 41-point grid of the skewness parameters using the following procedure: Initially, we construct a linear set of probabilities, $\left(\frac{1}{2\times 41} + (j-1)\frac{1}{41}\right)$, for $j = 1, \cdots, 41$. Then, we spread out these probabilities, using function $B_{SR}(\cdot, \eta)$, for $\eta = 2$, as shown in figure 2. Finally, we invert these transformed probabilities, in order to retrieve the grid. We will denote this grid with $\mathbf{\Theta} = \{\theta_1, \cdots, \theta_{41}\}$. For this inversion we use the unconditional density of the state process, which is a normal with mean -0.30

[5]The setting here is slightly simpler than the one in [6], since the correlation between the processes is zero. The characteristic function of the regime switching model is given in [6] and Elliott and Osakwe [9]. Correlations between the processes can be easily introduced, by allowing state switches to be accompanied by jumps in the asset process. See [6] for details.

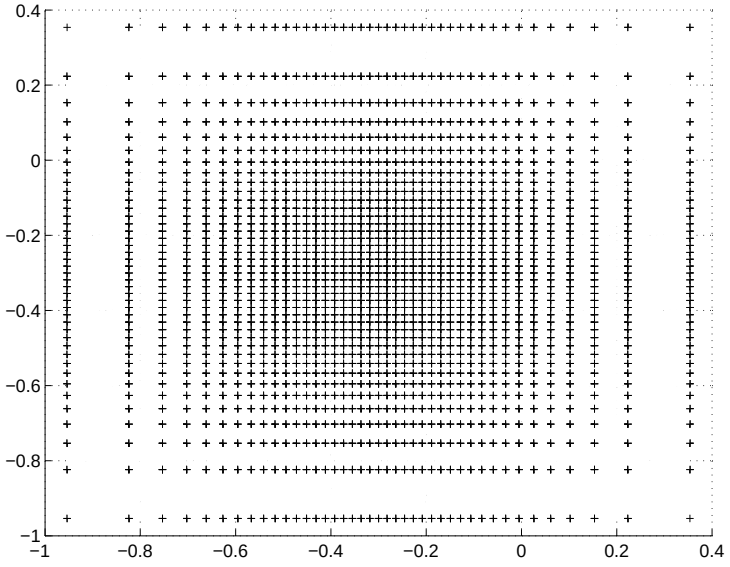


FIGURE 3. The grid that approximates the skewness parameter process $\theta(t)$. The figure presents all possible transitions across the discretized state-space.

and variance 0.16^2. Figure 3 shows the grid, and presents the possible transitions on the discrete grid. By construction, the grid is centered at the long run value of the skewness parameter, $\bar{\theta} = -0.30$, while the values range from -0.95 to $+0.35$. It is important to note that this implies that although the long run skew is negative, it is possible to temporarily experience a positive skew.

The approximating Markov chain can be easily constructed, using equations (3.2) or (3.3). The extreme states have been made reflective, by setting the corresponding rate matrix elements to a 'large value', equal to twice the maximum of all other rate elements.[6] A standard result in the theory of continuous time Markov chains postulates that the matrix exponential $\mathbf{P}(t) = \exp\{t\mathbf{Q}\}$ gives the time-t transition density of the chain. Figure 4 shows such densities for a forecasting period of one month, based on three different initial values $\theta(0)$. Of course, in this case these densities are normal, since $\theta(t)$ follows an OU process, but in general the transition densities need not be standard ones.

[6]Note that the approximating chain can be pre-computed, and then used to price contracts of different maturities.

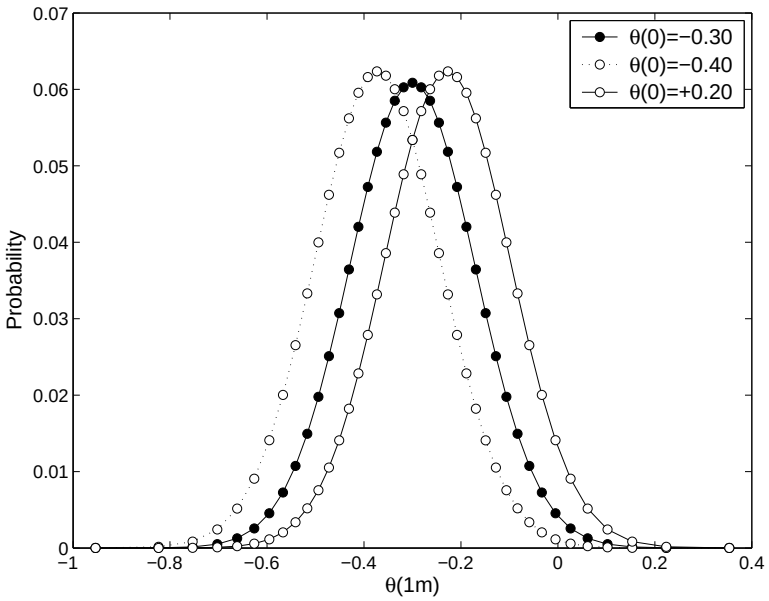


FIGURE 4. One month transition densities for the skewness parameter $\theta(1m)$, conditional on three initial values $\theta(0)$.

Matrix $\mathbf{\Psi}(u)$ can be computed for any value of u, based on the (41×41) rate matrix $\mathbf{Q}$, and the 41 conditional characteristic exponents. The typical (i,j)-th element, for $i,j = 1, \cdots, 41$ will be given by

$$\begin{cases} q_{i,i} + (r+\mu_i)\mathtt{i} - \frac{1}{\nu}\log\left\{1 - \mathtt{i}u\theta_i\nu - \frac{1}{2}u^2\sigma^2\nu\right\} & \text{, if} j = i \\ q_{i,i\pm1} & \text{, if} j = i \pm 1 \\ 0 & \text{, otherwise} \end{cases}$$

for $\mu_i = \frac{1}{\nu}\log\left\{1 - \theta_i\nu - \frac{1}{2}\sigma^2\nu\right\}$. Then, one can apply equation (3.4) to retrieve the approximating characteristic function.

The prices of options can be recovered from the characteristic functions, using the Fast Fourier Transform methodology of Carr and Madan [3], or the more efficient Fractional variant of Chourdakis [7].

4. Option prices and the volatility term structure

This section investigates the ability of the stochastic skewness specification (2.4) to generate varying volatility term structures. The numerical values of the previous sections will be assumed for all parameters. The interest rate and the dividend yields are assumed equal to zero.

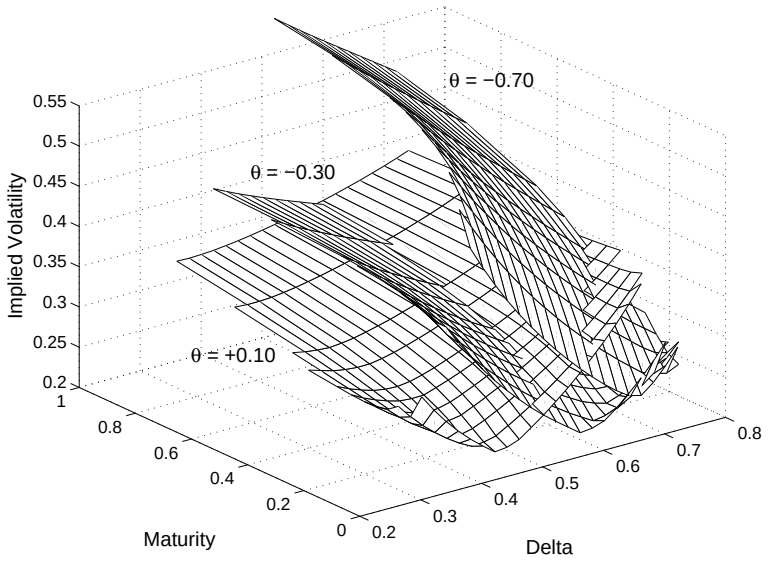
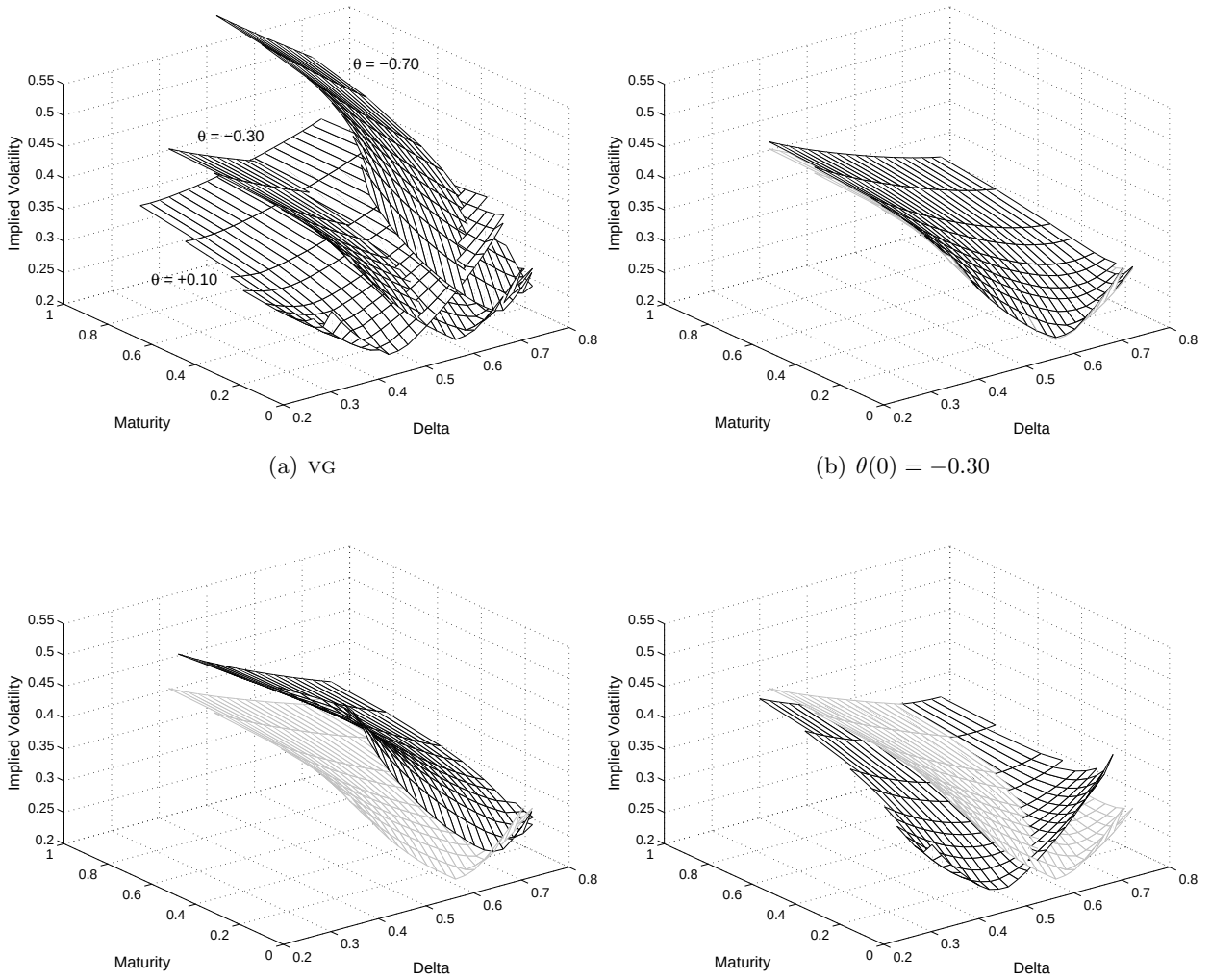


FIGURE 5. Volatility skews based on the VG model, with constant parameters. Three skews, for $\theta = -0.70$, $\theta = -0.30$ and $\theta = +0.10$ are presented.

All volatility surfaces are constructed in the Delta-Maturity space. Initially, option prices were computed for a range of strike prices, ranging from far-in- to far-out-of-the-money. Typically, 256 to 512 option prices were computed for each maturity, and a grid of 40 was selected to work with. Then, these option prices were inverted, in order to retrieve the implied volatilities. Based on these implied volatility values, the corresponding BS Deltas were computed, and the implied volatilities were interpolated using these Deltas, over a constant Delta grid, that extends from 20% to 80%, in increments of 2%. This procedure was repeated for a set of 11 maturities, ranging from 0.04 years (approximately two weeks), up to one year.

To carry out this comparative study, we will use the VG model with constant parameters, as our benchmark. Since we focus on the effect of the skewness controlling parameter θ, we present in figure 5 three volatility skews, based on three different values for θ. One can easily verify the relatively slow-varying nature of the volatility term structure, where the skew flattens-out in a predictable fashion.

Figure (6) presents the volatility skews for the VG model with stochastic θ. In order to assist comparisons, the constant VG skews are also replicated.



(a) VG (b) $\theta(0) = -0.30$

(c) $\theta(0) = -0.70$ (d) $\theta(0) = +0.10$

FIGURE 6. Volatility skews based on the stochastic version of the VG model. Three skews, for three different initial values $\theta(0)$ are presented. For comparisons, the constant-parameter VG skews are also given. In each graph, the grey surface gives the 'long run' implied volatilities.

Three different skews are presented, for three different initial values $\theta(0) = -0.30$, -0.70 and $+0.10$. The VG surface, for $\theta = -0.30$, is also presented in grey. Intuitively, this serves as the 'unconditional' or 'long run' set of implied volatilities.

It is apparent from these graphs, that imposing a stochastic structure on the VG parameters can offer a richer family of volatility term structures. The resulting graphs are also very intuitive: the initial value $\theta(0)$ controls the short end of the skew, while the long run value of $\bar{\theta}$ will control the long end. In particular, for short maturities, the stochastic VG skews resemble the short skews of their constant parameter counterparts, while for longer maturities all skews appear to converge towards the 'long run' surface of $\theta = -0.30$. One interesting scenario is offered in figure 6(d), which corresponds to the initial value $\theta(0) = +0.10$. here we observe a volatility skew that is reversed across maturity: for short maturities the skew is positive, while as the maturity increases the volatility skew becomes negative.

We now turn into an application of the above methodology, an investigate how well a stochastic Variance-Gamma model explains a set of observed option prices.

5. Example: Fitting the model to SP500 option prices

This section investigates how the stochastic parameter model can provide a better fit for European option prices on the SP500 index. We use all three alternatives to model the stochastic behavior of the model parameters, namely we fit a model with stochastic volatility (parameter σ), stochastic skewness (parameter θ), and stochastic kurtosis (parameter ν). In the case of stochastic skewness, the parameter θ follows the standard OU process discussed in the previous section. In the cases of stochastic σ and ν, since the parameters have to be restricted in positive domain, a log-OU process is assumed.

It it well known that the VG model can provide the flexibility to fit any set of option prices, *for a constant maturity*. The three parameters control over not only the volatility, but also the skewness and kurtosis of the risk neutral distribution. This is illustrated in figure 7: the fit of the time values is near-perfect, but each maturity needs a different set of parameter values to achieve it. Overall, one can attribute this feature to parameter variations through time. In fact, one can easily observe the monotonic behavior of these parameters.

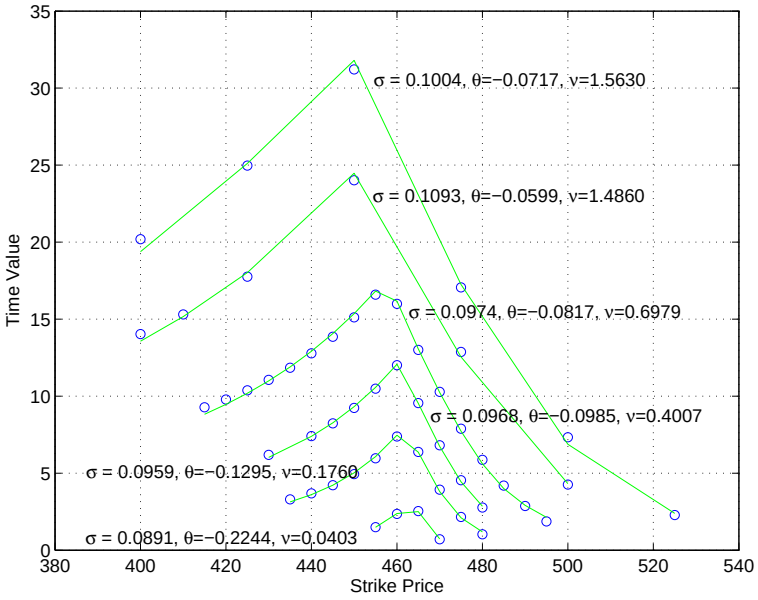


FIGURE 7. The VG model fitted to a set of observed option prices. Note that a new model is fitted for each maturity, and a different set of parameters is retrieved.

days to mat	σ	θ	ν	vol	skew	kurt
9	0.0891	-0.2244	0.0403	0.0998	-0.2533	3.1651
37	0.0959	-0.1295	0.1760	0.1102	-0.5701	3.7534
72	0.0968	-0.0985	0.4007	0.1151	-0.9278	4.8037
128	0.0974	-0.0817	0.6979	0.1189	-1.2804	6.2457
219	0.1093	-0.0599	1.4860	0.1314	-1.8225	9.7848
310	0.1004	-0.0717	1.5630	0.1346	-2.1286	10.9262

TABLE 1. The calibrated parameters for the VG processes, based on different maturities. Also presented are the 'implied' volatility, skewness and kurtosis, that are retrieved from these parameters.

Table 1 presents the estimates, together with the 'implied' volatility, skewness and kurtosis.[7]

The same set of prices was fitted using the stochastic parameter VG process. The resulting parameter values are reported in table 2, together with the corresponding root-mean-square-error (RMSE).

[7]Simply differentiating the characteristic function allows us to compute the moments. In particular, $Var = \sigma^2 + \nu\theta^2$, $Skew = \theta\nu(3\sigma^2 + 2\nu\theta^2)(\sigma^2 + \nu\theta^2)^{-3/2}$, and $Kurt = 3 + 6\nu - 3\nu\sigma^4(\sigma^2 + \nu\theta^2)^{-2}$.

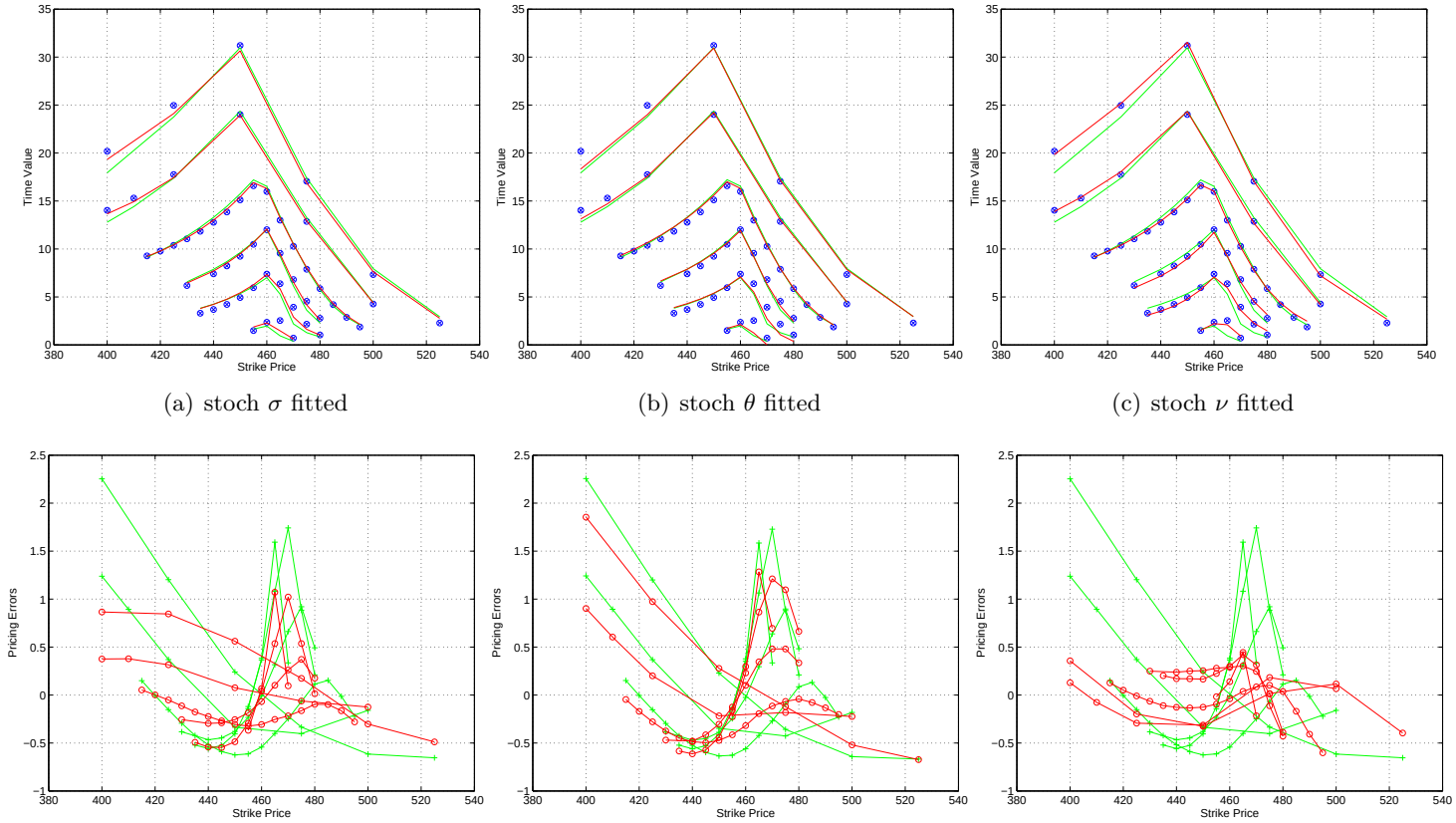
Model	σ	θ	ν	α	β	initial	RMSE
VG	0.0974	-0.0878	0.7320				0.672
stoch σ	0.0864	-0.0978	0.4642	20.0595	1.0376	0.1237	0.387
stoch θ	0.0095	-0.1286	0.3935	11.6281	0.7616	-0.1763	0.566
stoch ν	0.1035	-0.0745	2.2697	4.4861	4.4100	0.0995	0.241

TABLE 2. The calibrated parameters for a set of stochastic VG processes.

One can recognize that in terms of the RMSE, augmenting the model with time varying parameters can dramatically improve the fit. The RMSE is reduced by nearly two thirds, when a model that exhibits stochastic $\nu(t)$ is considered. Of course, pinpointing the value of the latent parameters α and β based on option prices alone is a tough exercise, since options do not use nor reveal all the information content of these parameters. The objective function with respect to these parameters is very flat, allowing a wide set of choices that produce similar values for the RMSE. Augmenting the data with basic exotic contracts, such as forward starting options, can lead to more accurate estimates of these parameters.

REFERENCES

1. Yacine Aït-Sahalia, *Telling from discrete data whether the underlying continuous-time model is a diffusion.*, The Journal of Finance **57** (2002), 2075–2112.
2. Peter Carr, Hélyette Geman, Dilip Madan, and Marc Yor, *Stochastic volatility for Lévy processes*, Mathematical Finance **13** (2003), no. 3, 345–382.
3. Peter Carr and Dilip Madan, *Option valuation using the Fast Fourier Transform.*, Journal of Computational Finance **3** (1999), 463–520.
4. Peter Carr and Liuren Wu, *Stochastic skew models for FX options.*, Tech. report, 2004.
5. Kyriakos Chourdakis, *Stochastic volatility and jumps driven by continuous time Markov chains.*, Tech. Report 430, Queen Mary, University of London, 2000.
6. ———, *Non-affine option pricing.*, Journal of Derivatives **11** (2004), no. 3, 10–25.
7. ———, *Option pricing using the Fractional FFT.*, Journal of Computational Finance (2005).
8. Paul G. Dupuis and Harold J. Kushner, *Numerical methods for stochastic control problems in continuous time.*, Applications of Mathematics, vol. 24, Springer Verlag, New York, N.Y., 2001.
9. Robert Elliott and Carlton-James Osakwe, *Option pricing for pure jump processes with markov switching compensators.*, Unpublished manuscript.
10. Harold J. Kushner, *Numerical methods for stochastic control problems in continuous time.*, SIAM Journal of Control and Optimization **28** (1990), no. 5, 999–1048.



(a) stoch σ fitted (b) stoch θ fitted (c) stoch ν fitted

(d) stoch σ errors (e) stoch θ errors (f) stoch ν errors

FIGURE 8. The pricing performance of the various stochastic parameter variants of the VG process. The green lines represent the performance of the VG model fitted simultaneously to all prices.

11. Harold J. Kushner and Giovanni DiMasi, *Approximations of functionals and optimal control problems on jump diffusion processes.*, Journal of Mathematical Analysis and Applications **63** (1978), 772–800.
12. Dilip Madan, Peter Carr, and Eric Chang, *The variance gamma process and option pricing.*, European Finance Review **2** (1998), 79–105.

AFIS Department, University of Canterbury, Aotearoa - New Zealand, url: http://www.theponytail.net/

E-mail address: kyriakos.chourdakis@canterbury.ac.nz