Derivation of the Heston Model

by Fabrice Douglas Rouah
www.FRouah.com
www.Volopta.com

The stochastic volatility model of Heston [7] is one of the most popular equity option pricing models. This is due in part to the fact that the Heston model produces call prices that are in closed form, up to an integral that must evaluated numerically. In this Note we present a complete derivation of the Heston model.

1 Heston Dynamics

The Heston model assumes that the underlying, S_t, follows a Black-Scholes type stochastic process, but with a stochastic variance v_t that follows a Cox, Ingersoll, Ross process. Hence

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dW_{1,t} && (1) \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dW_{2,t} \\ E\left[dW_{1,t} dW_{2,t}\right] &= \rho dt. \end{aligned}$$

We will drop the time index and write $S = S_t, v = v_t, W_1 = W_{1,t}$, and $W_2 = W_{1,t}$ for notional convenience. The six parameters of the model are

μ the drift of the process for the stock
κ the mean reversion speed for the variance,
θ the mean reversion level for the variance,
σ the volatility of the variance,
ρ the correlation between the two Brownian motions W_1 and W_2, and
v_0 the initial (time zero) level of the variance.

1.1 Risk Neutral Parameters

To Come

2 Properties of the Variance Process

It is well-known (see CIR) that conditional on a realized value of v_s, the random variable $2c_t v_t$ follows a non-central chi-square distribution with $d = 4\kappa\theta/\sigma^2$ degrees of freedom and non-centrality parameter $2c_t v_s e^{-\kappa(t-s)}$, where

$$c_t = \frac{2\kappa}{\sigma^2\left(1 - e^{-\kappa(t-s)}\right)}. \qquad (2)$$

Consequently, the conditional mean and variance of v_t are, respectively,

$$\begin{aligned} m &= \theta + (v_s - \theta)\, e^{-\kappa(t-s)} \\ s^2 &= \frac{v_t \sigma^2 e^{-\kappa(t-s)}}{\kappa}\left(1 - e^{-\kappa(t-s)}\right) + \frac{\theta\sigma^2}{2\kappa}\left(1 - e^{-\kappa(t-s)}\right)^2 . \end{aligned} \tag{3}$$

3 The Heston PDE

In this section we explain how to derive the PDE from the Heston model. This derivation is a special case of a PDE for general stochastic volatility models which is described by Gatheral [9]. Form a portfolio consisting of one option $V = V(S, v, t)$, Δ units of the stock S, and ϕ units of another option $U = U(S, v, t)$ that is used to hedge the volatility. The portfolio has value

$$\Pi = V + \Delta S + \phi U$$

where $\Pi = \Pi_t$. Assuming the portfolio is self-financing, the change in portfolio value is

$$d\Pi = dV + \Delta dS + \phi dU.$$

3.1 Portfolio Dynamics

Apply Itō's Lemma to dV. We must differentiate with respect to the variables t, S, and v. Hence

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}dt + \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2}dt + \sigma v\rho S\frac{\partial^2 V}{\partial v \partial S}dt$$

since $(dS)^2 = vS^2(dW_1)^2 = vS^2 dt, (dv)^2 = \sigma^2 v dt$, and $dSdv = vS\sigma dW_1 dW_2 = vS\sigma\rho dt$. We have used the fact that $(dt)^2 = 0$ and $dW_1 dt = dW_2 dt = 0$. Applying Itō's Lemma to dU produces the identical result, but in U

$$dU = \frac{\partial U}{\partial t}dt + \frac{\partial U}{\partial S}dS + \frac{\partial U}{\partial v}dv + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2}dt + \frac{1}{2}\sigma^2 v\frac{\partial^2 U}{\partial v^2}dt + \sigma v\rho S\frac{\partial^2 U}{\partial v \partial S}dt.$$

Combining these two expressions, we can write the change in portfolio value, $d\Pi$, as

$$\begin{aligned} d\Pi &= dV + \Delta dS + \phi dU \\ &= \left\{\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \rho\sigma v S\frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2}\right\}dt + \\ &\quad \phi\left\{\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\sigma v S\frac{\partial^2 U}{\partial v \partial S} + \frac{1}{2}\sigma^2 v\frac{\partial^2 U}{\partial v^2}\right\}dt + \\ &\quad \left\{\frac{\partial V}{\partial S} + \phi\frac{\partial U}{\partial S} + \Delta\right\}dS + \left\{\frac{\partial V}{\partial v} + \phi\frac{\partial U}{\partial v}\right\}dv. \end{aligned} \tag{4}$$

3.2 The Riskless Portfolio

In order for the portfolio to be hedged against movements in the stock and against volatility, the last two terms in Equation (4) involving dS and dv must be zero. This implies that the hedge parameters must be

$$\begin{aligned} \phi &= -\frac{\frac{\partial V}{\partial v}}{\frac{\partial U}{\partial v}} \\ \Delta &= -\phi\frac{\partial U}{\partial S} - \frac{\partial V}{\partial S}. \end{aligned} \tag{5}$$

Moreover, the portfolio must earn the risk free rate, r. Hence

$$d\Pi = r\Pi dt = r\left(V + \Delta S + \phi U\right)dt.$$

Now with the values of ϕ and Δ from Equation (5) the change in value of the riskless portfolio is

$$\begin{aligned} d\Pi &= \left\{\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \rho\sigma vS\frac{\partial^2 V}{\partial v\partial S} + \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2}\right\}dt + \\ &\quad \phi\left\{\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\sigma vS\frac{\partial^2 U}{\partial v\partial S} + \frac{1}{2}\sigma^2 v\frac{\partial^2 U}{\partial v^2}\right\}dt \end{aligned}$$

which we write as

$$d\Pi = \left(A + \phi B\right)dt. \tag{6}$$

Hence we have

$$A + \phi B = r\left(V + \Delta S + \phi U\right).$$

Substituting for ϕ and re-arranging, produces the equality

$$\frac{A - rV + rS\frac{\partial V}{\partial S}}{\frac{\partial V}{\partial v}} = \frac{B - rU + rS\frac{\partial U}{\partial S}}{\frac{\partial U}{\partial v}} \tag{7}$$

which we exploit in the next section.

3.3 The PDE in Terms of the Price

The left-hand side of Equation (7) is a function of V only, and the right-hand side is a function of U only. This implies that both sides can be written as a function $f(S, v, t)$ of S, v, and t. Following Heston, specify this function as $f(S, v, t) = -\kappa(\theta - v) + \lambda(S, v, t)$, where $\lambda(S, v, t)$ is the price of volatility risk. Write the left-hand side of Equation (7) as

$$\frac{A - rV + rS\frac{\partial V}{\partial S}}{\frac{\partial V}{\partial v}} = -\kappa(\theta - v) + \lambda(S, v, t).$$

Substitute for A and rearrange to produce the Heston PDE expressed in terms of the price S

$$\begin{aligned} &\frac{\partial V}{\partial t}+\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma vS\frac{\partial^2 V}{\partial v\partial S}+\frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2} \\ &-rV+rS\frac{\partial V}{\partial S}+\left[\kappa(\theta-v)-\lambda(S,v,t)\right]\frac{\partial V}{\partial v}=0. \end{aligned} \tag{8}$$

This is Equation (6) of Heston [7]. The PDE in Equation (8) can be written

$$\frac{\partial V}{\partial t}+\mathcal{A}V-rV=0$$

where

$$\begin{aligned} \mathcal{A}&=rS\frac{\partial}{\partial S}+\frac{1}{2}vS^2\frac{\partial}{\partial S^2} \\ &+\left[\kappa(\theta-v)-\lambda(S,v,t)\right]\frac{\partial}{\partial v}+\frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2}+\rho\sigma vS\frac{\partial}{\partial v\partial S} \end{aligned} \tag{9}$$

is the generator of the Heston model. As explained by Lewis [14], the first line in Equation (9) is the generator of the Black-Scholes model, while the second line adds the corrections for stochastic volatility.

3.4 The PDE in Terms of the Log Price

Let $x=\ln S$ and express the PDE in terms of x, t and v instead of S, t, and v. This leads to a simpler form of the PDE. We need the following derivatives

$$\begin{aligned} \frac{\partial V}{\partial S} &= \frac{\partial V}{\partial x}\frac{\partial x}{\partial S}=\frac{\partial V}{\partial x}\frac{1}{S} \quad \text{by the chain rule.} \\ \frac{\partial^2 V}{\partial v\partial S} &= \frac{\partial}{\partial v}\left(\frac{\partial V}{\partial S}\right)=\frac{\partial}{\partial v}\left(\frac{1}{S}\frac{\partial V}{\partial x}\right)=\frac{1}{S}\frac{\partial^2 V}{\partial v\partial x} \quad \text{by the chain rule.} \\ \frac{\partial^2 V}{\partial S^2} &= \frac{\partial}{\partial S}\left(\frac{\partial V}{\partial S}\right)=\frac{\partial}{\partial S}\left(\frac{1}{S}\frac{\partial V}{\partial x}\right) \\ &= -\frac{1}{S^2}\frac{\partial V}{\partial x}+\frac{1}{S}\frac{\partial^2 V}{\partial S\partial x}=-\frac{1}{S^2}\frac{\partial V}{\partial x}+\frac{1}{S}\frac{\partial}{\partial x}\left(\frac{\partial V}{\partial S}\right) \\ &= -\frac{1}{S^2}\frac{\partial V}{\partial x}+\frac{1}{S^2}\frac{\partial^2 V}{\partial x^2} \quad \text{by the product rule.} \end{aligned}$$

Plug into the Heston PDE Equation (8). All the S terms cancel and we obtain the Heston PDE in terms of the log price $x=\ln S$

$$\begin{aligned} &\frac{\partial V}{\partial t}+\frac{1}{2}v\frac{\partial^2 V}{\partial x^2}+\left(r-\frac{1}{2}v\right)\frac{\partial V}{\partial x}+\rho\sigma v\frac{\partial^2 V}{\partial v\partial x}+ \\ &\frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2}-rV+\left[\kappa(\theta-v)-\lambda v\right]\frac{\partial V}{\partial v}=0 \end{aligned} \tag{10}$$

where, as in Heston, we have written the market price of risk to be a linear function of the volatility, so that $\lambda(S,v,t)=\lambda v$.

4 The Call Price

The call price is of the form

$$\begin{aligned} C_T(K) &= e^{-r\tau} E\left[(S_T - K)^+\right] \\ &= e^{-r\tau} E\left[S_T \mathbf{1}_{S_T>K}\right] - e^{-r\tau} K E\left[\mathbf{1}_{S_T>K}\right] \\ &= e^{x_t} P_1(x,v,\tau) - e^{-r\tau} K P_2(x,v,\tau). \end{aligned} \tag{11}$$

In this expression $P_j(x,v,\tau)$ each represent the probability of the call expiring in-the-money, conditional on the value $x_t = \ln S_t$ of the stock and on the value v_t of the volatility at time t, where $\tau = T - t$ is the time to expiration. Hence

$$P_j(x,v,\tau) = \Pr\left(\ln S_T > \ln K\right) \tag{12}$$

for $j = 1, 2$. These probabilities are obtained under different probability measures. In Equation (11), the expected value $E\left[\mathbf{1}_{S_T>K}\right]$ is the probability of the call expiring in-the-money under the original measure $\mathbb{Q}$ that makes W_1 and W_2 in Equation (1) Brownian motion. Hence

$$E^{\mathbb{Q}}\left[\mathbf{1}_{S_T>K}\right] = \mathbb{Q}\left(S_T > K\right) = \mathbb{Q}\left(\ln S_T > \ln K\right) = P_2\left(x,v,\tau\right).$$

Evaluating $e^{-r\tau} E\left[S_T \mathbf{1}_{S_T>K}\right]$ in Equation (11) requires changing the original measure $\mathbb{Q}$ to another measure $\mathbb{P}$. Define the Radon-Nikodym derivative

$$\mathbb{Z}_t = \frac{S_t / S_T}{B_t / B_T}$$

where $B_t = \exp\left(\int_0^t r\,du\right) = e^{rt}$. Then the second expectation is

$$\begin{aligned} e^{-r\tau} E^{\mathbb{Q}}\left[S_T \mathbf{1}_{S_T>K}\right] &= E^{\mathbb{Q}}\left[\frac{B_t}{B_T} S_T \mathbf{1}_{S_T>K}\right] \\ &= E^{\mathbb{P}}\left[\frac{B_t}{B_T} S_T \mathbf{1}_{S_T>K} \mathbb{Z}_t\right] \\ &= S_t E^{\mathbb{P}}\left[\mathbf{1}_{S_T>K}\right] \\ &= e^{x_t} \mathbb{P}\left(S_T > K\right) = e^{x_t} P_1\left(x,v,\tau\right). \end{aligned}$$

This implies that the call price in Equation (11) can be written in terms of both measures as

$$C_T(K) = S_t \mathbb{P}\left(S_T > K\right) - K e^{-r\tau} \mathbb{Q}\left(S_T > K\right).$$

In Section 6.1 we show a more general derivation of this. Now take derivatives of C

$$\begin{aligned} \frac{\partial C}{\partial t} &= e^x \frac{\partial P_1}{\partial t} - K r e^{-r\tau} P_2 - K e^{-r\tau} \frac{\partial P_2}{\partial t} \\ &= e^x \left[\frac{\partial P_1}{\partial t}\right] - K e^{-r\tau} \left[r P_2 + \frac{\partial P_2}{\partial t}\right], \end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\partial C}{\partial x} &= e^x P_1 + e^x \frac{\partial P_1}{\partial x} - e^{-r\tau} K \frac{\partial P_2}{\partial x} \\ &= e^x \left[P_1 + \frac{\partial P_1}{\partial x} \right] - K e^{-r\tau} \left[\frac{\partial P_2}{\partial x} \right],\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 C}{\partial x^2} &= e^x P_1 + 2e^x \frac{\partial P_1}{\partial x} + e^x \frac{\partial^2 P_1}{\partial x^2} - e^{-r\tau} K \frac{\partial^2 P_2}{\partial x^2} \\ &= e^x \left[P_1 + 2\frac{\partial P_1}{\partial x} + \frac{\partial^2 P_1}{\partial x^2} \right] - K e^{-r\tau} \left[\frac{\partial^2 P_2}{\partial x^2} \right],\end{aligned}$$

$$\frac{\partial C}{\partial v} = e^x \left[\frac{\partial P_1}{\partial v} \right] - K e^{-r\tau} \left[\frac{\partial P_2}{\partial v} \right],$$

$$\frac{\partial^2 C}{\partial v^2} = e^x \left[\frac{\partial^2 P_1}{\partial v^2} \right] - K e^{-r\tau} \left[\frac{\partial^2 P_2}{\partial v^2} \right],$$

and

$$\begin{aligned}\frac{\partial^2 C}{\partial x \partial v} &= e^x \frac{\partial P_1}{\partial v} + e^x \frac{\partial^2 P_1}{\partial v \partial x} - e^{-r\tau} K \frac{\partial^2 P_2}{\partial v \partial x} \\ &= e^x \left[\frac{\partial P_1}{\partial v} + \frac{\partial^2 P_1}{\partial v \partial x} \right] - K e^{-r\tau} \left[\frac{\partial^2 P_2}{\partial v \partial x} \right].\end{aligned}$$

Since in the first line of Equation (13) we actually need the derivative with respect to maturity $\tau = T - t$ rather than the derivative with respect to t, we actually need $\frac{\partial C}{\partial \tau} = \frac{\partial C}{\partial t}\frac{\partial t}{\partial \tau} = -\frac{\partial C}{\partial \tau}$. We use the derivatives from Equation (13) in the following section.

4.1 The PDE for P_1 and P_2

Since the call price C in Equation (11) is also an option, it also follows the PDE in Equation (10), which we write here in terms of C but using the time derivative with respect to τ rather than t

$$\begin{aligned}&-\frac{\partial C}{\partial \tau} + \frac{1}{2} v \frac{\partial^2 C}{\partial x^2} + \left(r - \frac{1}{2} v \right) \frac{\partial C}{\partial x} + \rho \sigma v \frac{\partial^2 C}{\partial v \partial x} + \\ &\frac{1}{2} \sigma^2 v \frac{\partial^2 C}{\partial v^2} - rC + \left[\kappa(\theta - v) - \lambda v \right] \frac{\partial C}{\partial v} = 0.\end{aligned} \tag{14}$$

To obtain the Heston PDE for P_1 and P_2, we note that the PDE in Equation (14) is independent of the terms of the call contract. Hence, by setting $K = 0$ and $S = 1$ in Equation (11) we obtain an option whose price is simply P_1. This option will also follow the PDE. Similarly, by setting $S = 0, K = 1$, and $r = 0$ in Equation (11) we obtain an option whose price is $-P_2$. Since $-P_2$ follows the PDE, so does P_2. This implies that from Equations (13) we can regroup terms common to P_1 and substitute into the PDE Equation (14) to obtain

$$\begin{aligned}-\frac{\partial P_1}{\partial \tau}+\frac{1}{2}v\left[P_1+2\frac{\partial P_1}{\partial x}+\frac{\partial^2 P_1}{\partial x^2}\right]+\left(r-\frac{1}{2}v\right)\left[P_1+\frac{\partial P_1}{\partial x}\right] \\ +\rho\sigma v\left[\frac{\partial P_1}{\partial v}+\frac{\partial^2 P_1}{\partial x\partial v}\right]+\frac{1}{2}\sigma^2 v\frac{\partial^2 P_1}{\partial v^2} \\ -r\left[e^x P_1-e^{-r\tau}KP_2\right]+\left[\kappa\left(\theta-v\right)-\lambda v\right]\frac{\partial P_1}{\partial v}=0.\end{aligned} \tag{15}$$

Set $K=0$ and $S=1$ so that the PDE for P_1 in Equation (15) becomes

$$\begin{aligned}-\frac{\partial P_1}{\partial \tau}+\left(r+\frac{1}{2}v\right)\frac{\partial P_1}{\partial x}+\frac{1}{2}v\frac{\partial^2 P_1}{\partial x^2}+\rho\sigma v\frac{\partial^2 P_1}{\partial x\partial v}+ \\ \left[\rho\sigma v+\kappa\left(\theta-v\right)-\lambda v\right]\frac{\partial P_1}{\partial v}+\frac{1}{2}\sigma^2 v\frac{\partial^2 P_1}{\partial v^2}=0.\end{aligned} \tag{16}$$

Similarly, in Equations (13), regroup terms common to P_2 and substitute into the PDE for P_2 in Equation (14)

$$\begin{aligned}\left[rP_2-\frac{\partial P_2}{\partial \tau}\right]+\frac{1}{2}v\frac{\partial^2 P_2}{\partial x^2}+\left(r-\frac{1}{2}v\right)\frac{\partial P_2}{\partial x}+\rho\sigma v\frac{\partial^2 P_2}{\partial v\partial x}+ \\ \frac{1}{2}\sigma^2 v\frac{\partial^2 P_2}{\partial v^2}-r\left[e^x P_1-e^{-r\tau}KP_2\right]+\left[\kappa\left(\theta-v\right)-\lambda v\right]\frac{\partial P_2}{\partial v}=0.\end{aligned} \tag{17}$$

Set $S=0, K=-1$, and $r=0$ so that the PDE for P_2 in Equation (17) becomes

$$\begin{aligned}-\frac{\partial P_2}{\partial \tau}+\frac{1}{2}v\frac{\partial^2 P_2}{\partial x^2}+\left(r-\frac{1}{2}v\right)\frac{\partial P_2}{\partial x}+\rho\sigma v\frac{\partial^2 P_2}{\partial v\partial x} \\ +\frac{1}{2}\sigma^2 v\frac{\partial^2 P_2}{\partial v^2}+\left[\kappa\left(\theta-v\right)-\lambda v\right]\frac{\partial P_2}{\partial v}=0.\end{aligned} \tag{18}$$

For notional convenience, we can combine Equations (16) and (18) into a single expression

$$\begin{aligned}-\frac{\partial P_j}{\partial \tau}+\rho\sigma v\frac{\partial^2 P_j}{\partial x\partial v}+\frac{1}{2}v\frac{\partial^2 P_j}{\partial x^2}+\frac{1}{2}v\sigma^2\frac{\partial^2 P_j}{\partial v^2} \\ +\left(r+u_j v\right)\frac{\partial P_j}{\partial x}+\left(a-b_j v\right)\frac{\partial P_j}{\partial v}=0\end{aligned} \tag{19}$$

for $j=1,2$ and where $u_1=\frac{1}{2}, u_2=-\frac{1}{2}, a=\kappa\theta, b_1=\kappa+\lambda-\rho\sigma$, and $b_2=\kappa+\lambda$. This is Equation (12) of Heston [7] but in terms of τ rather than t. That explains the minus sign in the first term of Equation (19) above.

4.2 Obtaining the Characteristic Functions

When the characteristic functions $f_1\left(\phi;x,v\right)$ and $f_2\left(\phi;x,v\right)$ corresponding to the in-the-money probabilities P_1 and P_2 are known, each probability can be recovered from its characteristic function via the inversion theorem

$$P_j = \Pr\left(\ln S_T > \ln K\right) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\ln K} f_j(\phi; x, v)}{i\phi}\right] d\phi. \tag{20}$$

At maturity, when $t = T$ so that $\tau = 0$, the probabilities are subject to the terminal condition

$$P_j(x, v, 0) = \mathbf{1}_{x>\ln K} \tag{21}$$

which simply means that at expiry, when $S_T > K$ the probability of the call being in-the-money is unity. Heston assumes that the characteristic functions are of the log-linear form

$$f_j(\phi; x, v) = \exp\left(C_j\left(\tau, \phi\right) + D_j\left(\tau, \phi\right) v + i\phi x\right) \tag{22}$$

where C_j and D_j are coefficients and $\tau = T - t$ is the time to maturity. The characteristic functions f_j will follow the PDE in Equation (19). This is a consequence of the Feynman-Kac theorem[1], which stipulates that if a function $f(x_t, t)$ of the SDE x_t satisfies the PDE $\frac{\partial f}{\partial t} - rf + \mathcal{A}f = 0$ with terminal condition $f(x_T, T)$, where $x_t = (\ln S_t, v_t)$ is the Heston bivariate diffusion and $\mathcal{A}$ is the Heston generator from Equation (9)

$$\begin{aligned}\mathcal{A} &= rS\frac{\partial}{\partial S} + \frac{1}{2}vS^2\frac{\partial}{\partial S^2} \\ &+ \left[\kappa(\theta - v) - \lambda(S, v, t)\right]\frac{\partial}{\partial v} + + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2} + \rho\sigma vS\frac{\partial}{\partial v\partial S}\end{aligned}$$

then the solution to f is

$$f(x_t, t) = E\left[\left. f(x_T, T)\right| x_t = x, v_t = v\right].$$

Using $f(x_T, T) = e^{i\phi x_T} = e^{i\phi\ln S_T}$ produces the solution

$$f(x_t, t) = E\left[\left. e^{i\phi\ln S_T}\right| \ln S_t = S, v_t = v\right]$$

which is the characteristic function for $\ln S_T$. Hence the PDE for the characteristic function is, from Equation (19)

$$\begin{aligned}-\frac{\partial f_j}{\partial \tau} + \rho\sigma v\frac{\partial^2 f_j}{\partial x\partial v} + \frac{1}{2}v\frac{\partial^2 f_j}{\partial x^2} + \frac{1}{2}v\sigma^2\frac{\partial^2 f_j}{\partial v^2} \\ + \left(r + u_j v\right)\frac{\partial f_j}{\partial x} + \left(a - b_j v\right)\frac{\partial f_j}{\partial v} = 0.\end{aligned} \tag{23}$$

---

[1] See the Note on www.FRouah.com for an explanation of the Feynman-Kac theorem.

To evaluate this PDE for the characteristic function we need the following derivatives

$$\begin{aligned}
\frac{\partial f_j}{\partial \tau} &= -\left(\frac{\partial C_j}{\partial \tau} + \frac{\partial D_j}{\partial \tau} v\right) f_j \\
\frac{\partial f_j}{\partial x} &= i\phi f_j \\
\frac{\partial^2 f_j}{\partial x^2} &= -\phi^2 f_j \\
\frac{\partial f_j}{\partial v} &= D_j f_j \\
\frac{\partial^2 f_j}{\partial v^2} &= D_j^2 f_j \\
\frac{\partial^2 f_j}{\partial v \partial x} &= i\phi D_j f_j
\end{aligned}$$

Substituting these derivatives in Equation (23) and drop the f_j terms to obtain

$$\begin{gathered}
-\left(\frac{\partial C_j}{\partial \tau} + v\frac{\partial D_j}{\partial \tau}\right) + \rho\sigma v i\phi D_j - \frac{1}{2}v\phi^2 + \frac{1}{2}v\sigma^2 D_j^2 \\
+ (r + u_j v)\, i\phi + (a - b_j v)\, D_j = 0
\end{gathered} \tag{24}$$

or equivalently

$$\begin{gathered}
v\left(-\frac{\partial D_j}{\partial \tau} + \rho\sigma i\phi D_j - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j i\phi - b_j D_j\right) \\
-\frac{\partial C_j}{\partial \tau} + ri\phi + aD_j = 0.
\end{gathered} \tag{25}$$

This produces two of differential equations

$$\begin{aligned}
\frac{\partial D_j}{\partial \tau} &= \rho\sigma i\phi D_j - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j i\phi - b_j D_j \\
\frac{\partial C_j}{\partial \tau} &= ri\phi + aD_j.
\end{aligned} \tag{26}$$

These are Equations (A7) in Heston [7]. Heston specifies the initial conditions $D_j(0, \phi) = 0$ and $C_j(0, \phi) = 0$. The first Equation in (26) is a Riccati equation in D_j while the second is an ODE for C_j that can solved using straightforward integration once D_j is obtained.

4.3 Solving the Heston Riccati Equation

4.3.1 Solution of the Riccati Equation

The Riccati equation for $y(t)$ with coefficients $P(t), Q(t)$ and $R(t)$

$$\frac{dy(t)}{dt} = P(t) + Q(t)y(t) + R(t)y(t)^2 \tag{27}$$

can be solved by considering the second-order ordinary differential equation for $w(t)$

$$w'' - \left[\frac{P'}{P} + Q \right] w' + PRw = 0 \tag{28}$$

which we write as $w'' + bw' + cw = 0$. The solution to Equation (27) is then

$$y(t) = -\frac{w'(t)}{w(t)} \frac{1}{R(t)}.$$

The ODE in Equation (28) can be solved via the auxiliary equation $r^2 + br + c = 0$ which has two solutions for given by

$$\alpha = \frac{-b + \sqrt{b^2 - 4ac}}{2} \text{ and } \beta = \frac{-b - \sqrt{b^2 - 4ac}}{2}.$$

The solution to the second-order ODE in Equation (28) is $w(t) = Ae^{\alpha t} + Be^{\beta t}$ where A and B are constants. The solution to the Riccati equation is therefore

$$y(t) = -\frac{A\alpha e^{\alpha t} + B\beta e^{\beta t}}{Ae^{\alpha t} + Be^{\beta t}} \frac{1}{R(t)}.$$

4.3.2 Solution of the Heston Riccati Equation

From Equation (26) the Heston Riccati equation is

$$\frac{\partial D_j}{\partial \tau} = P_j - Q_j D_j + RD_j^2 \tag{29}$$

where

$$\begin{aligned} P_j &= u_j i\phi - \frac{1}{2}\phi^2 \qquad (30) \\ Q_j &= b_j - \rho\sigma i\phi \\ R &= \frac{1}{2}\sigma^2. \end{aligned}$$

The corresponding second order ODE is

$$w'' + Q_j w' + P_j R = 0 \tag{31}$$

so that $D_j = -\frac{1}{R}\frac{w'}{w}$. The auxiliary equation is $r'' + Q_j r' + P_j R = 0$ which has roots

$$\alpha = \frac{-Q_j + \sqrt{Q_j^2 - 4P_j R}}{2} = \frac{-Q_j + d_j}{2}$$

and

$$\beta = \frac{-Q_j - \sqrt{Q^2 - 4P_j R}}{2} = \frac{-Q_j - d_j}{2},$$

where

$$\begin{aligned} d_j &= \alpha - \beta \\ &= \sqrt{Q_j^2 - 4P_jR} \\ &= \sqrt{\left(\rho\sigma i\phi - b_j\right)^2 - \sigma^2\left(2u_ji\phi - \phi^2\right)}. \end{aligned} \tag{32}$$

For notational simplicity we sometimes omit the j subscript for some of the variables. The solution to the Heston Riccati equation (29) is therefore

$$D_j = -\frac{1}{R}\frac{w'}{w} = -\frac{1}{R}\left(\frac{A\alpha e^{\alpha t} + B\beta e^{\beta t}}{Ae^{\alpha t} + Be^{\beta t}}\right) = -\frac{1}{R}\left(\frac{K\alpha e^{\alpha t} + \beta e^{\beta t}}{Ke^{\alpha t} + e^{\beta t}}\right) \tag{33}$$

where $K = \frac{A}{B}$. The initial condition $D_j(0,\phi) = 0$ implies that we set $t = 0$ in Equation (33), so that the numerator is $K\alpha + \beta = 0$ from which $K = -\frac{\beta}{\alpha}$. Hence, the solution for D_j becomes

$$\begin{aligned} D_j &= -\frac{\beta}{R}\left(\frac{-e^{\alpha\tau} + e^{\beta\tau}}{-g_je^{\alpha\tau} + e^{\beta\tau}}\right) \\ &= -\frac{\beta}{R}\left(\frac{1 - e^{d_j\tau}}{1 - g_je^{d_j\tau}}\right) \\ &= \frac{Q_j + d_j}{2R}\left(\frac{1 - e^{d_j\tau}}{1 - g_je^{d_j\tau}}\right) \end{aligned}$$

where

$$\begin{aligned} g_j &= -K \\ &= \frac{b_j - \rho\sigma i\phi + d_j}{b_j - \rho\sigma i\phi - d_j}. \end{aligned} \tag{34}$$

We can write the solution for D_j as

$$D_j = \frac{b_j - \rho\sigma i\phi + d_j}{\sigma^2}\left(\frac{1 - e^{d_j\tau}}{1 - g_je^{d_j\tau}}\right). \tag{35}$$

The solution for C_j is found by integrating the second equation in (26). Hence

$$C_j = \int_0^\tau r i\phi dy + a\left(\frac{Q_j + d_j}{2R}\right)\int_0^\tau \left(\frac{1 - e^{d_jy}}{1 - g_je^{d_jy}}\right) dy.$$

The first integral is $ri\phi\tau$ and the second integral can be found by substituting $x = e^{dy}$, from which $dx = de^{dy}dy$ and $y = \frac{1}{xd}dx$. We have

$$C_j = ri\phi\tau + \frac{a}{d_j}\left(\frac{Q_j + d_j}{\sigma^2}\right)\int_1^{e^{d\tau}} \left(\frac{1 - x}{1 - g_jx}\right)\frac{1}{x}dx. \tag{36}$$

The integral in Equation (36) is evaluated by partial fractions

$$\begin{aligned}\int_1^{e^{d\tau}} \frac{1-x}{x\,(1-gx)}dx &= \int_1^{e^{d\tau}} \left[\frac{1}{x} - \frac{1-g}{1-gx}\right] dx \\ &= \left[\ln x + \frac{1-g}{g}\ln\left(1-gx\right)\right]_{x=1}^{x=e^{d\tau}} \\ &= \left[d\tau + \frac{1-g}{g}\ln\left(\frac{1-ge^{d\tau}}{1-g}\right)\right].\end{aligned}$$

Substituting the integral back into (36), and substituting for d_j, Q_j, and g_j produces the solution for C_j

$$C_j = ri\phi\tau + \frac{a}{\sigma^2}\left[\left(b_j - \rho\sigma i\phi + d_j\right)\tau - 2\ln\left(\frac{1-g_je^{d_j\tau}}{1-g_j}\right)\right]. \tag{37}$$

This completes the original derivation of the Heston model.

4.4 Summary

Recall from Equation (11) that the call price is of the form

$$C_T\left(K\right) = e^{x_t}P_1(x,v,\tau) - e^{-r\tau}KP_2(x,v,\tau).$$

where the in-the-money probabilities P_1 and P_2 are

$$P_j = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\ln K}f_j(\phi;x,v)}{i\phi}\right]d\phi \tag{38}$$

with the characteristic functions f_1 and f_2

$$f_j(\phi;x,v) = \exp\left(C_j\left(\tau,\phi\right) + D_j\left(\tau,\phi\right)v + i\phi x\right).$$

To obtain the call price, we use the expressions for C_j and D_j in Equations (37) and (35) to obtain the two characteristic functions. We must then evaluate the integrals for P_1 and P_2 using a numerical integration algorithm.

4.5 Remark on the Characteristic Functions

It makes sense that two characteristic functions f_1 and f_2 are associated with the Heston model, because P_1 and P_2 are both in-the-money probabilities, but obtained under different measures. On the other hand, is also seems that a single characteristic function of S_T, the terminal log stock price, ought to exist, since there is only one security. Indeed, in many instances the probabilities P_1 and P_2 are written in terms of a single characteristic function $f\left(\phi\right) = f\left(\phi;x,v\right)$ as

$$P_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\ln K}f(\phi-i)}{i\phi f(-i)}\right]d\phi. \tag{39}$$

and

$$P_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re}\left[\frac{e^{-i\phi \ln K} f(\phi)}{i\phi}\right] d\phi. \tag{40}$$

In this section we show that these formulations are not contradictory. The "true" characteristic function is actually f_2, since it is associated with the probability measure $\mathbb{Q}$ that makes W_1 and W_2 in Equation (1) Brownian motion (see Section 4). Hence, using more compact notation, we can write $f(\phi) = f_2$, and we must show that

$$f_1(\phi) = \frac{f(\phi - i)}{f(-i)}. \tag{41}$$

Because of the log-linear form of the characteristic function, this is equivalent to showing that

$$\begin{aligned} C_2(\tau, \phi - i) - C_2(\tau, -i) &= C_1(\tau, \phi), \text{ and} \\ D_2(\tau, \phi - i) - D_2(\tau, -i) &= D_1(\tau, \phi). \end{aligned} \tag{42}$$

To simplify the task, we denote $d_j = d_j(\phi)$, $g_j = g_j(\phi)$,

$$\begin{aligned} G_j(\phi) &= \frac{1 - e^{d_j(\phi)\tau}}{1 - g_j(\phi)}, \text{ and} \\ H_j(\phi) &= \frac{1 - e^{d_j(\phi)\tau}}{1 - g_j(\phi) e^{d_j(\phi)\tau}} \end{aligned}$$

to emphasize the dependence of these quantities on ϕ. It is straightforward to show the following identities

$$\begin{aligned} d_2(\phi - i) &= d_1(\phi) \text{ and } d_2(-i) = \rho\sigma - \kappa, \\ g_2(\phi - i) &= g_1(\phi) \text{ and } g_2(-i) = 0, \\ G_2(\phi - i) &= G_1(\phi) \text{ and } G_2(-i) = 0, \\ H_2(\phi - i) &= H_1(\phi) \text{ and } H_2(-i) = 1 - e^{\rho\sigma - \kappa}. \end{aligned}$$

With these identities, it is easy to show that Equations (42) and (41) hold, and consequently, that the probabilities can be written written in terms of Equation (38) or equivalently in terms of (39) and (40). In Section (6) we present a formal proof of the relation (41).

5 The Little Heston Trap

Albrecher *et al.* [1] note that the solution to D_j in Equation (35) can be multiplied by $\frac{1}{g_j} e^{-d_j\tau}$ in the numerator and denominator, which leads to the equivalent form

$$\begin{aligned} D_j &= \frac{\frac{1}{g_j}(b_j - \rho\sigma i\phi + d_j)}{\sigma^2} \left(\frac{1 - e^{-d_j\tau}}{1 - \frac{1}{g_j} e^{-d_j\tau}}\right) \\ &= \frac{b_j - \rho\sigma i\phi - d_j}{\sigma^2} \left(\frac{1 - e^{-d_j\tau}}{1 - c_j e^{-d_j\tau}}\right) \end{aligned} \tag{43}$$

where, from Equation (34)

$$\begin{aligned} c_j &= \frac{1}{g_j} \\ &= \frac{b_j - \rho\sigma i\phi - d_j}{b_j - \rho\sigma i\phi + d_j}. \end{aligned} \tag{44}$$

The logarithm in the solution to C_j in Equation (37) can be written in the equivalent form

$$\begin{aligned} d_j\tau - 2\ln\left[\frac{1-g_je^{d_j\tau}}{1-g_j}\right] &= d_j\tau - 2\ln\left[\frac{1}{e^{-d_j\tau}}\left(\frac{e^{-d_j\tau}-g_j}{1-g_j}\right)\right] \\ &= d_j\tau - 2d_j\tau - 2\ln\left[\left(\frac{e^{-d_j\tau}-g_j}{1-g_j}\right)\right] \\ &= -d_j\tau - 2\ln\left[\left(\frac{1-c_je^{-d_j\tau}}{1-c_j}\right)\right]. \end{aligned} \tag{45}$$

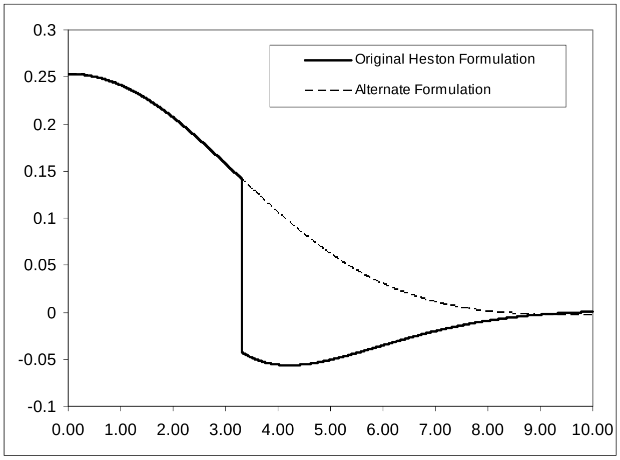
This implies that C_j can be written in the equivalent form

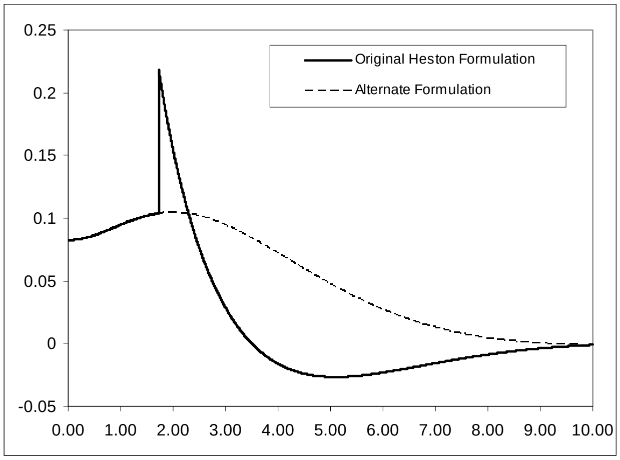
$$C_j = ri\phi\tau + \frac{a}{\sigma^2}\left[(b_j - \rho\sigma i\phi - d_j)\,\tau - 2\ln\left(\frac{1-c_je^{-d_j\tau}}{1-c_j}\right)\right]. \tag{46}$$

Implementing this formulation is very simple and involves only replacing C_j and D_j with slightly diffent forms. Albrecher *et al.* [1] explain that although Heston's original formulation and their formulation are identical, their formulation causes fewer numerical problems in the implementation of the model. We illustrate this by plotting the integrand

$$\mathrm{Re}\left[\frac{e^{-i\phi\ln K}f_j(\phi;x,v)}{i\phi}\right]$$

for both characteristic functions f_1 and f_2. We use the same parameter values as in [1], namely $\kappa = 1.5768$, $\sigma = 0.5751$, $\rho = -0.5711$, $\theta = 0.0398$, and $v_0 = 0.0175$. In addition, we use $S = K = 100$, and $r = 0.035$. These integrands appear in the Figures for ϕ ranging from 0 to 10.





Finally, the rotation algorithm of Jackael (REFERENCE) can be used to overcome the discontinuities brought on by the original Heston formulation. The algorithm, however, offers little benefit over the Albrecher *et al.* ([1]) formulation. Hence, we do not cover the rotation algorithm in this book, but refer interested readers to the article by Jackael ().

6 Bakshi and Madan Approach

Bakshi and Madan [8] present a general framework in which the characteristic function of the state-price density can be used to price most derivatives. In particular, the Heston model arises a special case of their approach. This is a more compact way of obtaining the call price, since only a (joint) characteristic function needs to be obtained–the PDE in Equation (19) is not needed.

6.1 Call Price

The time-t price $C_T(K)$ of a European call with strike K and maturity T is the expected discounted value under the risk neutral measure $\mathbb{Q}$

$$\begin{aligned} C_T(K) &= E_t^{\mathbb{Q}}\left[e^{-b(t,T)}\left(S_T - K\right)^+\right] \\ &= \int_{\varkappa} e^{-b(t,T)}\left(S_T - K\right) q(v)dv \end{aligned}$$

where $b(t,T) = \int_t^T r_u du$, $q(v)$ is the risk neutral joint density for $v = (b(t,T), S_T)$, S_T is the terminal stock price, and $\varkappa = \{S_T > K\}$ is the exercise region. Bakshi and Madan [8] show that the call price can be written

$$\begin{aligned} C_T(K) &= \int_{\varkappa} e^{-b(t,T)} S_T q(v)dv - K\int_{\varkappa} e^{-b(t,T)} q(v)dv \\ &= \int_{\Omega} e^{-b(t,T)} S_T q(v)dv \left(\frac{\int_{\varkappa} e^{-b(t,T)} S_T q(v)dv}{\int_{\Omega} e^{-b(t,T)} S_T q(v)dv}\right) \\ &\quad -K\int_{\Omega} e^{-b(t,T)} q(v)dv \left(\frac{\int_{\varkappa} e^{-b(t,T)} q(v)dv}{\int_{\Omega} e^{-b(t,T)} q(v)dv}\right) \\ &= G(t,T)\Pi_1 - KB(t,T)\Pi_2 \end{aligned}$$

where

$$G(t,T) = \int_{\Omega} e^{-b(t,T)} S_T q(v)dv = E_t^{\mathbb{Q}}\left[e^{-b(t,T)} S_T\right] \tag{47}$$

is the scaled forward price,

$$B(t,T) = \int_{\Omega} \exp\left(-\int_t^T r_u du\right) q(v)dv = E_t^{\mathbb{Q}}\left[e^{-b(t,T)}\right] \tag{48}$$

is the time-t price of a zero-coupon bond paying 1 at time T, and $\Omega = \{S_T > 0\}$ is the admissible region for the stock price at expiry. The Arrow-Debreu security Π_1 is

$$\begin{aligned} \Pi_1 &= \frac{\int_{\varkappa} e^{-b(t,T)} S_T q(v)dv}{\int_{\Omega} e^{-b(t,T)} S_T q(v)dv} = \int_{\varkappa}\left(\frac{e^{-b(t,T)} S_T}{G(t,T)}\right) q(v)dv \\ &= E_t^{\mathbb{Q}}\left[\frac{e^{-b(t,T)} S_T}{G(t,T)} \mathbf{1}_{\varkappa}\right] = E_t^{\mathbb{Q}^*}\left[\mathbf{1}_{\varkappa}\right], \end{aligned}$$

where $\mathbf{1}_{\varkappa}$ is the Heaviside function. The last expectation is written using $q^*(v)dv = \frac{e^{-b(t,T)} S_T}{G(t,T)} q(v)dv$, i.e., using the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = \frac{e^{-b(t,T)} S_T}{G(t,T)} = \frac{\exp\left(-\int_t^T r_u du\right) S_T}{\int_{\Omega} e^{-b(t,T)} S(T) q(v)dv}. \tag{49}$$

Similarly, the Arrow-Debreu security Π_2 is

$$\begin{aligned}
\Pi_2 &= \frac{\int_{\varkappa} e^{-b(t,T)} q(v) dv}{\int_{\Omega} e^{-b(t,T)} q(v) dv} = \int_{\varkappa} \left(\frac{e^{-b(t,T)} q(v) dv}{B(t,T)} \right) q(v) dv \\
&= E_t^{\mathbb{Q}} \left[\frac{e^{-b(t,T)}}{G(t,T)} \mathbf{1}_{\varkappa} \right] = E_t^{\mathbb{Q}^{**}} \left[\mathbf{1}_{\varkappa} \right],
\end{aligned}$$

which uses the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}^{**}}{d\mathbb{Q}} = \frac{e^{-b(t,T)}}{B(t,T)} = \frac{\exp\left(-\int_t^T r_u du\right)}{\int_{\Omega} e^{-b(t,T)} S_T q(v) dv}. \tag{50}$$

Since $\Pi_1 = E_t^{\mathbb{Q}^*} [\mathbf{1}_{\varkappa}]$ and $\Pi_2 = E_t^{\mathbb{Q}^{**}} [\mathbf{1}_{\varkappa}]$, each can be interpreted as the time-t probability that the option expires in-the-money, albeit under the different probability measures $\mathbb{Q}^*$ and $\mathbb{Q}^{**}$ respectively. Hence

$$\Pi_1 = \mathbb{Q}^* \left(S_T > K \right) \text{ and } \Pi_2 = \mathbb{Q}^{**} \left(S_T > K \right).$$

In the Heston model, these probabilities are P_1 and P_2 from Equation (12). The Heston probabilities require two separate valuation PDEs, regrouped in Equation (19), and two separate characteristic functions. In the following section we show that In Bakshi and Madan's approach only a single characteristic function is required.

6.2 Characteristic Function

The joint characteristic function for $v = (b(t,T), S_T)$ is

$$\begin{aligned}
\widetilde{f}(\phi, \varphi) &= E_t^{\mathbb{Q}} \left[e^{i\varphi b(t,T) + i\phi S_T} \right] \qquad (51) \\
&= \int_{\Omega} e^{i\varphi b(t,T) + i\phi S_T} q(v) dv.
\end{aligned}$$

By setting $\varphi = i$ in Equation (51) Bakshi and Madan [8] obtain the characteristic function of the state-price density $e^{-b(t,T)} q(v)$ as

$$\begin{aligned}
f(\phi) &= E_t^{\mathbb{Q}} \left[e^{-b(t,T) + i\phi S_T} \right] \qquad (52) \\
&= \int_{\Omega} e^{-b(t,T)} e^{i\phi S_T} q(v) dv.
\end{aligned}$$

Setting $\phi = 0$ in Equation (52) shows that $f(0) = B(t,T)$. Differentiating Equation (52) with respect to ϕ produces

$$f_{\phi}(\phi) = i \int_{\Omega} e^{-b(t,T)} e^{i\phi S(T)} S(T) q(v) dv. \tag{53}$$

Setting $\phi = 0$ in Equation (53) shows that $f_{\phi}(0) = iG(t,T)$. The characteristic function for S_T in terms of the probability Π_1 is the expected value of $e^{i\phi S(T)}$

under $\mathbb{Q}^*$ and is denoted $f_1(\phi)$. Use the Radon-Nikodym derivative (49) and write

$$\begin{aligned}
f_1(\phi) &= E_t^{\mathbb{Q}^*}\left[e^{i\phi S_T}\right] = E_t^{\mathbb{Q}}\left[\frac{d\mathbb{Q}^*}{d\mathbb{Q}}e^{i\phi S_T}\right] \qquad (54)\\
&= \int_\Omega \frac{e^{-b(t,T)}S_T}{G(t,T)}e^{i\phi S_T}q(v)dv\\
&= \frac{\int_\Omega e^{-b(t,T)}S_T e^{i\phi S_T}q(v)dv}{\int_\Omega e^{-b(t,T)}S_T q(v)dv}\\
&= \frac{f_\phi(\phi)}{f_\phi(0)} = \frac{f_\phi(\phi)}{iG(t,T)}.
\end{aligned}$$

Similarly, the characteristic function for S_T in terms of the probability Π_2 is the expected value of $e^{i\phi S(T)}$ under $\mathbb{Q}^{**}$ and is denoted $f_2(\phi)$. Use the Radon-Nikodym derivative (50) and write

$$\begin{aligned}
f_2(\phi) &= E_t^{\mathbb{Q}^{**}}\left[e^{i\phi S_T}\right] = E_t^{\mathbb{Q}}\left[\frac{d\mathbb{Q}^{**}}{d\mathbb{Q}}e^{i\phi S_T}\right] \qquad (55)\\
&= \int_\Omega \frac{e^{-b(t,T)}}{B(t,T)}e^{i\phi S_T}q(v)dv\\
&= \frac{\int_\Omega e^{-b(t,T)}e^{i\phi S_T}q(v)dv}{\int_\Omega e^{-b(t,T)}q(v)dv}\\
&= \frac{f(\phi)}{f(0)} = \frac{f(\phi)}{B(t,T)}.
\end{aligned}$$

Hence we have the important result that the two characteristic functions f_1 and f_2 can each be expressed in terms of the characteristic function f of the state-price density, its derivative, the bond price $B(t,T)$, and the forward price $G(t,T)$, as

$$f_1(\phi) = \frac{f_\phi(\phi)}{iG(t,T)} \text{ and } f_2(\phi) = \frac{f(\phi)}{B(t,T)}.$$

6.3 Characteristic Function for the Heston Model

Since Π_1 and Π_2 are analogous to the in-the-money probabilities P_1 and P_2, respectively, in Equation (12), by the inversion theorem they can each be expressed in terms of their characteristic functions f_1 and f_2 as

$$\Pi_j = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \text{Re}\left[\frac{e^{-i\phi K}f_j(\phi)}{i\phi}\right]d\phi. \qquad (56)$$

The probabilities P_1 and P_2 are for the log price $X_T = \ln S_T$ and the log strike price $\ln K$. To translate them into Π_1 and Π_2 we must substitute $\ln K$ into Equation (56) and express the derivation in Section 6 in terms of $\ln S_T$ instead of S_T. In this case the inversion theorem produces

$$\Pi_j = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \ln K} f_j(\phi)}{i\phi}\right] d\phi \tag{57}$$

which is identical to Equation (20). The scaled forward price $G(t,T)$ in Equation (47) becomes

$$G(t,T) = \int_\Omega e^{-b(t,T)} \ln S_T q(v) dv \tag{58}$$

and the characteristic function $f(\phi)$ in Equation (52) becomes

$$f(\phi) = \int_\Omega e^{-b(t,T)} e^{i\phi \ln S_T} q(v) dv \tag{59}$$

with derivative

$$f_\phi(\phi) = i\int_\Omega e^{-b(t,T)} e^{i\phi \ln S_T} \ln S_T q(v) dv. \tag{60}$$

Recall from Equations (54) and (55) that the two characteristic functions are $f_1(\phi) = \frac{f_\phi(\phi;t,T)}{iG(t,T)}$ and $f_2(\phi) = \frac{f(\phi;t,T)}{B(t,T)}$, where now we use the updated form of the characteristic function in Equation (59). Substitute for f_2 in Equation (57) to obtain the second survival probability

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \ln K} f(\phi)}{i\phi B(t,T)}\right] d\phi$$

or, when interest rates are fixed

$$\Pi_2 = \frac{1}{2} + \frac{e^{r(T-t)}}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \ln K} f(\phi)}{i\phi}\right] d\phi.$$

Substitute $\phi = -i$ in Equation (59) and compare with Equation (58) to obtain $G(t,T) = f(-i)$. Next, note that from Equations (59) and (60) we have the relationship $f_\phi(\phi;t,T) = if(\phi - i)$. This implies that the first characteristic function can be written $f_1(\phi) = \frac{f(\phi-i)}{f(-i)}$. Hence the first survival probability is

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \ln K} f(\phi - i)}{i\phi f(-i)}\right] d\phi.$$

6.4 Comparison with Heston's Approach

To evaluate the in-the-money probabilities P_1 and P_2 via the inversion theorem, two characteristic functions f_1 and f_2 for the logarithm of the terminal stock price are required because the probabilities P_1 and P_2 are obtained under two different measures. Under Bakshi and Madan's approach, a *single* characteristic function, namely $\widetilde{f}(\phi,\varphi)$ in Equation 51, is obtained. This is the *joint* characteristic function for the bivariate vector $v = (b(t,T), S_T)$. From $\widetilde{f}$ they obtain $f(\phi)$ in Equation (52), the characteristic function of the state-price density. Bakshi and Madan then show that f_1 and f_2 can be expressed solely in terms of $f(\phi)$ and its derivatives. Under Heston's approach, f_1 and f_2 must obtained separately.

7 Derivation Using the Fourier Transform

In this section we show how to derive the Heston call price using the Fourier transform. We follow the derivation described by Gatheral [9].

7.1 Fourier Transform and its Derivative

There are several definitions of the Fourier transform $\widehat{f}$ of a function f. The one used by Gatheral [9] is

$$\widehat{f}(k) = \int_{-\infty}^{\infty} e^{-ikx} f(x) dx.$$

The derivative of $\widehat{f}$ is found by considering the Fourier transform of the derivative of f, denoted $\widehat{f'}$

$$\widehat{f'}(k) = \int_{-\infty}^{\infty} e^{-ikx} f'(x) dx.$$

Perform integration by parts, using $u' = f'$ and $v = e^{-ikx}$

$$\begin{aligned} \widehat{f'}(k) &= e^{-ikx} f(x)\big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} ike^{-ikx} f(x) dx \\ &= ik\widehat{f}(k) \end{aligned}$$

which assumes that $\lim_{x \to \pm\infty} e^{-ikx} f(x) = 0$. Continuing in this way shows that differentiation of the Fourier transform corresponds to multiplication by ik so that

$$\widehat{f^{(n)}}(k) = (ik)^n \widehat{f}(k).$$

The inverse transform is

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} \widehat{f}(k) dk.$$

7.2 The Heston Model by the Fourier Transform

We follow the derivation in Gatheral [9]. He explains that the Heston model can be simplified by the change of variable $x_t = \ln \frac{F_{t,T}}{K}$ where K is the strike price of the option and $F_{t,T} = S_t e^{\mu(T-t)}$ is the forward price. Applying Itō's Lemma to the first equation in (1) shows that x_t follows the diffusion

$$dx = -\frac{1}{2} v dt + \sqrt{v} dW_1.$$

Under the assumptions used by Gatheral the market price of volatility risk is assumed to be zero so that $\lambda(S, v, t) = 0$. With x_t defined as $\ln \frac{F_{t,T}}{K}$ instead of $\ln S_t$, the terminal condition in Equation (21) becomes

$$P_j(x, v, 0) = \mathbf{1}_{x>0}. \tag{61}$$

With these modifications the PDE in Equation (19) becomes

$$\begin{gathered}-\frac{\partial P_j}{\partial\tau}+\rho\sigma v\frac{\partial^2 P_j}{\partial x\partial v}+\frac{1}{2}v\frac{\partial^2 P_j}{\partial x^2}+\frac{1}{2}v\sigma^2\frac{\partial^2 P_j}{\partial v^2}\\+u_j v\frac{\partial P_j}{\partial x}+(a-b_j v)\frac{\partial P_j}{\partial v}=0\end{gathered} \tag{62}$$

where the parameters are redefined as $u_1=\frac{1}{2}, u_2=-\frac{1}{2}, a=\kappa\theta, b_1=\kappa-\rho\sigma$, and $b_2=\kappa$ since it is assumed that $\lambda=0$. Now consider the Fourier transform $\widehat{P}_j$ of the probabilities P_j

$$\widehat{P}_j(k,v,\tau)=\int_{-\infty}^{\infty}e^{-ikx}P_j(x,v,\tau)dx.$$

Using the terminal condition (61) we have

$$\widehat{P}_j(k,v,0)=\int_0^{\infty}e^{-ikx}dx=-\frac{1}{ik}e^{-ikx}\Big|_{x=0}^{x=\infty}=\frac{1}{ik}.$$

Remembering that differentiation of the Fourier transform $\widehat{P}_j$ with respect to x corresponds to multiplication by ik, the PDE in Equation (62) for $\widehat{P}_j$ is

$$\begin{gathered}-\frac{\partial\widehat{P}_j}{\partial\tau}+\rho\sigma vik\frac{\partial\widehat{P}_j}{\partial v}-\frac{1}{2}vk^2\widehat{P}_j+\frac{1}{2}v\sigma^2\frac{\partial^2\widehat{P}_j}{\partial v^2}\\+u_jvik\widehat{P}_j+(a-b_jv)\frac{\partial\widehat{P}_j}{\partial v}=0.\end{gathered} \tag{63}$$

Re-arranging, we obtain

$$v\left\{\alpha\widehat{P}_j-\beta\frac{\partial\widehat{P}_j}{\partial v}+\gamma\frac{\partial^2\widehat{P}}{\partial v^2}\right\}+a\frac{\partial\widehat{P}_j}{\partial v}-\frac{\partial\widehat{P}_j}{\partial\tau}=0 \tag{64}$$

where

$$\begin{aligned}\alpha &= u_jik-\frac{k^2}{2}\\ \beta &= b_j-\rho\sigma ik\\ \gamma &= \frac{\sigma^2}{2}.\end{aligned}$$

Note that these coefficients are identical to the coefficients P, Q, and R in Equation (30) that define the Riccati equation in Section 4.3.2, except that k replaces ϕ. Now the ansatz is that the $\widehat{P}_j$ are of the form

$$\begin{aligned}\widehat{P}_j(k,v,\tau) &= \exp\left[C_j(k,\tau)\theta+D_j(k,\tau)v\right]\widehat{P}_j(u,v,0)\\ &= \frac{1}{ik}\exp\left[C_j(k,\tau)\theta+D_j(k,\tau)v\right].\end{aligned} \tag{65}$$

Take derivatives of Equation (65)

$$\begin{aligned}\frac{\partial \widehat{P}_j}{\partial \tau} &= \left[\frac{\partial C_j}{\partial \tau}\theta + \frac{\partial D_j}{\partial \tau}v\right]\widehat{P}_j \\ \frac{\partial \widehat{P}_j}{\partial v} &= D_j \widehat{P}_j \\ \frac{\partial^2 \widehat{P}_j}{\partial v^2} &= D_j^2 \widehat{P}_j .\end{aligned}$$

Substitute these derivatives back into the PDE in Equation (64) and dropping the $\widehat{P}_j$ terms produces

$$v\left\{\alpha - \beta D_j + \gamma D_j^2\right\} + aD_j - \left[\frac{\partial C_j}{\partial \tau}\theta + \frac{\partial D_j}{\partial \tau}v\right] = 0$$

which implies that the following two equations must be satisfied, which appear as Equation (2.11) in Gatheral [9]

$$\begin{aligned}\frac{\partial C_j}{\partial \tau} &= \kappa D_j \qquad (66)\\ \frac{\partial D}{\partial \tau} &= \alpha - \beta D_j + \gamma D_j^2 \\ &= \gamma\left(D - r_+\right)\left(D - r_-\right).\end{aligned}$$

The first equation obtains since $a = \kappa\theta$. The second equation is identical to the Riccati equation in (29). Define, as in Gatheral [9]

$$\begin{aligned}r_\pm &= \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha} = \\ &= \frac{b_j - \rho\sigma ik \pm d}{\sigma^2}\end{aligned}$$

where

$$d = \sqrt{\left(\rho\sigma ik - b_j\right)^2 - \sigma^2\left(2u_j ik - k^2\right)}.$$

This expression for d is identical to that in Equation (32), with ϕ replaced by k. Then g_2 in Equation (44) can be written as $g_2 = \frac{r_-}{r_+}$ with ϕ replaced by k,. The solution for D_j in Equation (43) is then

$$D_j = r_-\left(\frac{1 - e^{-d\tau}}{1 - g_2 e^{-d\tau}}\right)$$

since γ cancels out of Equation (66). The solution for C_j from Equation (46), remembering that $r = 0$ in this derivation, is

$$\begin{aligned}C_j &= \frac{\kappa}{\sigma^2}\left[\left(b_j - \rho\sigma i\phi - d\right)\tau - 2\ln\left(\frac{1 - g_2 e^{-d\tau}}{1 - g_2}\right)\right] \\ &= \kappa\left[r_-\tau - \frac{2}{\sigma^2}\ln\left(\frac{1 - g_2 e^{-d\tau}}{1 - g_2}\right)\right]\end{aligned}$$

since θ cancels out. These expressions for D_j and C_j are Equations (2.12) in Gatheral [9].

7.3 Alternate Derivation Using the Fourier Transform

We use the following form of the Fourier transform

$$\widehat{f}(k) = \int_{-\infty}^{\infty} e^{ikx} f(x) dx$$

which is more consistent with the fact that the characteristic function of a random variable is the Fourier transform of its probability density.

7.3.1 The Convolution Theorem

The convolution of two function $f(x)$ and $g(x)$ is another function, denoted $(f \star g)(x)$ and defined as

$$(f \star g)(x) = \int_{-\infty}^{\infty} f(u)\, g(x-u)\, du. \tag{67}$$

The Convolution theorem states that the Fourier Transform of $(f \star g)(x)$ is the product of the Fourier transforms of $\widehat{f}$ and $\widehat{g}$, namely that

$$\widehat{(f \star g)}(k) = \widehat{f}(k)\widehat{g}(k). \tag{68}$$

Indeed

$$\begin{aligned}
\widehat{(f \star g)}(k) &= \int_{-\infty}^{\infty} e^{ikx} (f \star g)(x)\, dx \\
&= \int_{-\infty}^{\infty} e^{ikx} \left[\int_{-\infty}^{\infty} f(u)\, g(x-u)\, du \right] dx
\end{aligned}$$

Use the transformation $\tau = x - u$ and reverse the order of integration

$$\begin{aligned}
\widehat{(f \star g)}(k) &= \int_{-\infty}^{\infty} e^{ik(\tau+u)} \left[\int_{-\infty}^{\infty} f(u)\, g(\tau)\, du \right] d\tau \\
&= \int_{-\infty}^{\infty} e^{iku} f(u) \left[\int_{-\infty}^{\infty} e^{ik\tau} g(\tau)\, d\tau \right] du \\
&= \widehat{f}(k)\, \widehat{g}(k).
\end{aligned}$$

7.3.2 Plancharel and Parseval Identities

The Plancharel identity states that the product of f and $\overline{g}$ (the complex conjugate of g) can be expressed in terms of their Fourier transforms

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\, \overline{g(x)} dx &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(k)\, \overline{\widehat{g}}(k)\, dk \qquad (69) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(k)\, \widehat{g}(-k)\, dk
\end{aligned}$$

since for a complex number $g = |g| e^{i\theta} = |g| (\cos\theta + i \sin\theta)$, its complex conjugate is $\overline{g} = |g| (\cos\theta - i \sin\theta) = |g| e^{-i\theta}$. To prove this identity, use the fact that Dirac's delta function $\delta(x)$ can be expressed in integral form as

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\pm ix\phi} d\phi \tag{70}$$

The Parserval identity arises as a special case of the Plancharel identity, by setting $g = f$ in Equation (69)

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \widehat{f}(k) \right|^2 dk \tag{71}$$

since the absolute value of a complex number $z = x + iy$ is defined as $|z| = \sqrt{x^2 + y^2}$

8 The Fundamental Transform

To come. Alan Lewis.

9 Heston Greeks

To come.

10 Numerical Integration Schemes

To come

10.1 Newton-Coates Formulas

We evaluate three different formulas. The mid-point rule, the trapezoidal rule, and Simpson's rule. The mid-

10.2 Quadratures

To come

11 Fast Fourier Transform

This is proposed by Carr and Madan ([5]) as a way to speed up the computation of option prices. Recall the call price is

$$\begin{aligned} C_T(k) &= e^{-r\tau} \int_k^{\infty} \left(e^{x_T} - e^k \right) q_T(x) \, dx \\ &= e^{x_t} \mathbb{P}\left(S_T > e^k \right) - K e^{-r\tau} \mathbb{Q}\left(S_T > e^k \right) \end{aligned}$$

where $k = \ln K$, $x = \ln S_T$, $\tau = T - t$ and where $\mathbb{P}(S_T > K)$ and $\mathbb{Q}(S_T > K)$ are found by inverting the characteristic function. This inversion obviously requires that the characteristic function be integrable. Since

$$\begin{aligned}
\lim_{k \to -\infty} C_T(k) &= \lim_{k \to -\infty} e^{-r\tau} \int_k^\infty \left(e^x - e^k\right) q_T(x)\, dx \\
&= e^{-r\tau} E^{\mathbb{Q}}\left[e^x\right] \\
&= S_t
\end{aligned}$$

which is not zero, $C_T(k)$ is not integrable. Carr and Madan solve this by defining the modified call price $c_T(k)$ as

$$c_T(k) = e^{\alpha k} C_T(k)$$

which includes the dampening factor $e^{\alpha k}$ on $C_T(k)$. Since

$$\begin{aligned}
\lim_{k \to -\infty} c_T(k) &= \lim_{k \to -\infty} e^{-r\tau} \int_k^\infty \left(e^{\alpha k + x} - e^{(\alpha+1)k}\right) q_T(x)\, dx \\
&= e^{-rT} E^{\mathbb{Q}}[0]
\end{aligned}$$

which *is* zero, $c_T(k)$ *is* integrable. The idea is to find the Fourier transform $\psi_T(v)$ for $c_T(k)$, and obtain the Fourier transform for $C_T(k)$ in terms of $\psi_T(v)$ and α. We thus have

$$\begin{aligned}
\psi_T(v) &= \int_{-\infty}^\infty e^{ivk} c_T(k)\, dk \\
&= \int_{-\infty}^\infty e^{ivk} e^{\alpha k} C_T(k)\, dk \\
&= e^{-r\tau} \int_{-\infty}^\infty e^{(\alpha+iv)k} \left[\int_k^\infty \left(e^x - e^k\right) q_T(x)\, dx\right] dk.
\end{aligned}$$

The area of integration $-\infty < k < \infty$ and $k < x < \infty$ is equivalent to $-\infty < x < \infty$ and $-\infty < k < x$, so we can write$_j$

$$\begin{aligned}
\psi_T(v) &= e^{-r\tau} \int_{-\infty}^\infty q_T(x) \left[\int_{-\infty}^x \left(e^{(\alpha+iv)k+x} - e^{(\alpha+iv+1)k}\right) dk\right] dx \\
&= e^{-r\tau} \int_{-\infty}^\infty q_T(x) \left[\frac{e^{(\alpha+iv+1)x}}{\alpha^2 + \alpha - v^2 + iv(2\alpha+1)}\right] dx \\
&= \frac{e^{-r\tau} \phi_T(v - (\alpha+1)i)}{\alpha^2 + \alpha - v^2 + iv(2\alpha+1)}
\end{aligned}$$

since

$$\phi_T(u) = E^{\mathbf{Q}}\left[e^{iux}\right] = \int_{-\infty}^\infty e^{iux} q_T(x) dx$$

is the characteristic function of $x = \ln S_T$. This call price is therefore

$$\begin{aligned} C_T(k) &= \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \psi_T(v)\, dv \qquad (72) \\ &= \frac{e^{-\alpha k}}{\pi} \int_{0}^{\infty} \text{Re}\left[e^{-ivk} \psi_T(v) \right] dv. \end{aligned}$$

This last equality holds because while the integrand $e^{-ivk}\psi_T(v)$ is a complex number, the call price $C_T(k)$ is a real number. This implies that we can ignore the imaginary part of the integrand, and work only with the real part, which is even-valued.

11.1 Fast Fourier Transform

The discrete Fourier transform maps a vector of points $\mathbf{x} = (x_1, \cdots, x_N)$ to another vector of points $\widehat{\mathbf{x}} = (\widehat{x}_1, \cdots, \widehat{x}_N)$ via the relation

$$\widehat{x}_k = \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} x_j \quad \text{for } k = 1, \cdots, N. \qquad (73)$$

Computing these sums independently of one another would require N^2 steps. The fast Fourier transform (FFT) computes these sums simultaneously, which requires $N \log_2 N$ steps. Recall from Equation (72) that for a given value of the log strike k, the call price is

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_{0}^{\infty} \text{Re}\left[e^{-ivk} \psi_T(v) \right] dv. \qquad (74)$$

The objective is to discretize the expression for $C_T(x)$ and express it in terms of (73). Note that the inverse FFT of $\widehat{\mathbf{x}} = (\widehat{x}_1, \cdots, \widehat{x}_N)$ is the vector $\mathbf{x} = (x_1, \cdots, x_N)$ defined as

$$x_k = \frac{1}{N} \sum_{j=1}^{N} e^{i\frac{2\pi}{N}(j-1)(k-1)} \widehat{x}_j \quad \text{for } k = 1, \cdots, N.$$

The FFT and inverse FFT are intuitive discrete analogues of their continuous counterparts. We can denote the FFT that maps $\mathbf{x}$ to $\widehat{\mathbf{x}}$ by $\widehat{\mathbf{x}} = D(\mathbf{x})$ and the inverse FFT by $\mathbf{x} = D^{-1}(\widehat{\mathbf{x}})$.

11.1.1 Discretization of the Integration Domain

We can approximate Equation (74) by the trapezoidal rule over the truncated range of integration $[0, a]$ for v, using N equidistant points

$$\{v_j = (j-1)\eta\}_{j=1}^{N}$$

where η is the increment. The trapezoidal rule approximates $C_T(k)$ as

$$\begin{aligned} C_T(k) &\approx \frac{e^{-\alpha k}}{\pi} \operatorname{Re}\left[\frac{1}{2} e^{-iv_1 k} \psi_T(v_1) + e^{-iv_2 k} \psi_T(v_2) + \cdots \right. \\ &\quad \left. + e^{-iv_{N-1} k} \psi_T(v_{N-1}) + \frac{1}{2} e^{-iv_N k} \psi_T(v_N) \right] \eta \\ &= \frac{\eta e^{-\alpha k}}{\pi} \sum_{j=1}^{N} \operatorname{Re}\left[e^{-iv_j k} \psi_T(v_j)\right] w_j \end{aligned} \tag{75}$$

where $w_1 = w_N = \frac{1}{2}$ and $w_j = 1$ for $j = 2, \cdots, N-1$ are weights. If Simpson's rule is used instead of the trapezoidal rule, we set $w_1 = w_N = \frac{1}{3}$ and $w_j = \frac{1}{3}\left(3 + (-1)^j\right)$.

11.1.2 Discretization of the Strike Range

We are interested in strikes around the money, so we need to define the discretization range of the log strikes so that it is centered about the log spot price, $\log S_t$. A good choice is to use N equidistant points

$$\{k_u = -b + (u-1)\lambda + \log S_t\}_{u=1}^{N}$$

where λ is the increment and $b = \frac{N\lambda}{2}$. This produces log strikes over the range $[\log S_t - b, \log S_t + b - \lambda]$. For a log strike value k_u on the grid we can write Equation (75) as

$$C_T(k_u) \approx \frac{\eta e^{-\alpha k_u}}{\pi} \sum_{j=1}^{N} \operatorname{Re}\left[e^{-iv_j k_u} \psi_T(v_j)\right] w_j.$$

Substitute for v_j and k_u

$$\begin{aligned} C_T(k_u) &\approx \frac{\eta e^{-\alpha k_u}}{\pi} \sum_{j=1}^{N} \operatorname{Re}\left[e^{-i(j-1)\eta[-b+(u-1)\lambda+\log S_t]} \psi_T(v_j) w_j\right] \\ &= \frac{\eta e^{-\alpha k_u}}{\pi} \sum_{j=1}^{N} \operatorname{Re}\left[e^{-i\lambda\eta(j-1)(u-1)} e^{i(b-\log S_t)v_j} \psi_T(v_j) w_j\right]. \end{aligned} \tag{76}$$

In order to express Equation (76) in terms of the FFT (73), we must have the following constraint on the increments η and λ

$$\lambda\eta = \frac{2\pi}{N}.$$

This is an important limitation of the FFT, since it implies a trade off between the grid sizes. For a fixed N, choosing a fine grid for the integration range will produce a coarse grid for the log strikes range, and vice-versa. The only way to increase the granularity of both grids is to increase N, and consequently, the computation time. Defining

$$x_j = e^{i(b-\log S_t)v_j}\psi_T(v_j)\,w_j \tag{77}$$

and $\widehat{x}_u = C_T(k_u)$ the call price is expressed as the FFT

$$\widehat{x}_u = \frac{\eta e^{-\alpha k_u}}{\pi}\sum_{j=1}^{N}\mathrm{Re}\left[e^{-i\frac{2\pi}{N}(j-1)(u-1)}x_j\right] \tag{78}$$

for $u = 1, \cdots, N$.

11.2 Fractional Fast Fourier Transform

This method was proposed by Chourdakis ([6]) for option pricing. It relaxes the restrictive constraint on the increments $\lambda\eta = \frac{2\pi}{N}$. The fast fractional Fourier transform (FRFT) replaces $\frac{1}{N}$ in the exponent to Equation (78) with a general term

$$\widehat{x}_u = \frac{\eta e^{-\alpha k_u}}{\pi}\sum_{j=1}^{N}e^{-i2\pi\beta(j-1)(u-1)}x_j.$$

The FRFT arises as the special case $\beta = \frac{1}{N}$. We can denote the FRFT that maps $\mathbf{x}$ to $\widehat{\mathbf{x}}$ with the parameter β by $\widehat{\mathbf{x}} = D(\mathbf{x}, \beta)$ Chourdakis ([6]) describes how the FRFT can be used for option pricing. First, the define the vectors $\mathbf{y}$ and $\mathbf{z}$ each of dimension $2N$ as

$$\begin{aligned}\mathbf{y} &= \left(\left\{e^{-i\pi(j-1)^2\beta}x_j\right\}_{j=1}^{N}, \{0\}_{j=1}^{N}\right)\\ \mathbf{z} &= \left(\left\{e^{i\pi(j-1)^2\beta}\right\}_{j=1}^{N}, \left\{e^{i\pi(N-j+1)^2\beta}\right\}_{j=1}^{N}\right).\end{aligned}$$

Next, take the FFT of $\mathbf{y}$ and $\mathbf{z}$ to obtain the vectors $\widehat{\mathbf{y}} = D(\mathbf{y})$ and $\widehat{\mathbf{z}} = D(\mathbf{z})$, and take their element by element product, which produces the vector $\widehat{\mathbf{h}}$ of dimension $2N$

$$\widehat{\mathbf{h}} = \widehat{\mathbf{y}} \odot \widehat{\mathbf{z}} = \{y_j z_j\}_{j=1}^{2N}.$$

Now take the inverse FFT of $\widehat{\mathbf{h}}$ to produce the vector $\mathbf{h} = D^{-1}\left(\widehat{\mathbf{h}}\right)$ of dimension $2N$, and multiply element-by-element the first N elements of $\mathbf{h}$ by $\left\{e^{-i\pi(j-1)^2\beta}\right\}_{j=1}^{N}$. Hence, we can write the fractional FFT in compact form as

$$\begin{aligned}\widehat{\mathbf{x}} &= D(\mathbf{x}, \beta)\\ &= \left\{e^{-i\pi(j-1)^2\beta}\right\}_{j=1}^{N} \odot D^{-1}\left(D(\mathbf{y}) \odot D(\mathbf{z})\right).\end{aligned}$$

Note that, similar to the FFT, the FRFT takes the N-vector $\mathbf{x}$ and maps it to the N-vector $\widehat{\mathbf{x}}$. The FRFT, however, uses the intermediate $2N$-vectors

$\widehat{\mathbf{y}}$ and $\widehat{\mathbf{z}}$, which requires that two FFT be computed in the intermediate steps. Neverthess, the increase in computational time required by the two intermediate FFT is usually offset by the increase in accuracy

11.2.1 The FFT and FRFT for the Heston Model

Both the FFT and FRFT require the points $x_j = e^{ibv_j}\psi_T(v_j)\, w_j$ defined in Equation (77), where $\psi_T(\phi)$ is the characteristic function $f_2(\phi; x, v)$ defined in Equation (22)

$$\begin{aligned}\psi_T(\phi) &= \exp(i\phi x) \\ &\quad\times \exp\left(ri\phi\tau + \frac{\kappa\theta}{\sigma^2}\left[(\kappa - \rho\sigma i\phi + d) - 2\ln\left(\frac{1 - ge^{d\tau}}{1 - g}\right)\right]\right) \\ &\quad\times \exp\left(\frac{\kappa - \rho\sigma i\phi + d}{\sigma^2}\left[\frac{1 - e^{d\tau}}{1 - ge^{d\tau}}\right] v_0\right)\end{aligned}$$

where $x = \ln S_t$ the log spot price, v_0 is the initial variance, $\tau = T - t$, the time to maturity, and where

$$\begin{aligned} d &= \sqrt{(\rho\sigma i\phi - \kappa)^2 - \sigma^2\left(-i\phi - \phi^2\right)} \\ g &= \frac{\kappa - \rho\sigma i\phi - d}{\kappa - \rho\sigma i\phi + d}.\end{aligned}$$

If we use the formulation by Albrecher *et al.* [1], the characteristic function is

$$\begin{aligned}\psi_T(\phi) &= \exp(i\phi x) \\ &\quad\times \exp\left(ri\phi\tau + \frac{\kappa\theta}{\sigma^2}\left[(\kappa - \rho\sigma i\phi - d) - 2\ln\left(\frac{1 - ce^{-d\tau}}{1 - c}\right)\right]\right) \\ &\quad\times \exp\left(\frac{\kappa - \rho\sigma i\phi - d}{\sigma^2}\left[\frac{1 - e^{-d\tau}}{1 - ce^{-d\tau}}\right] v_0\right)\end{aligned}$$

where $c = 1/g$.

12 Effect of the Heston Parameters

12.1 Heston Terminal Spot Price S_T

If we generate distributions of the stock price at maturity, S_T, under the Heston model, we will get a distribution that is more skewed, and that has higher kurtosis, than under Black-Scholes.

12.1.1 Effect of ρ on Skewness

If we choose $\rho > 0$ the distribution of S_T will be more positively skewed than the distribution of S_T generated under the Black-Scholes model, and if we choose

$\rho < 0$ the distribution will be more negatively skewed. In the following figures we use the parameter settings in Heston [7] to generate 10,000 stock price paths. The parameter settings are $v_0 = 0.01, \kappa = 2, \theta = 0.01, \lambda = 0$, and $\sigma = 0.1$. We generate Heston prices under two scenarios: $\rho = 0.5$,. and $\rho = -0.5$. The volatilities used in the Black Scholes model are those matched by Heston [7], so that when $\rho = 0.5$ we use $\sigma_{BS} = 0.100409$, and when $\rho = -0.5$, we use $\sigma_{BS} = 0.099561$. In Equation (1) we use Milstein discretization of the Heston variance process, and Euler discretization for the stock price. Figure 1 clearly shows that $\rho = 0.5$ produces a distribution of S_T that is positively skewed.

Figure 1: Distribution of S_T under Black-Scholes and Heston, positive correlation

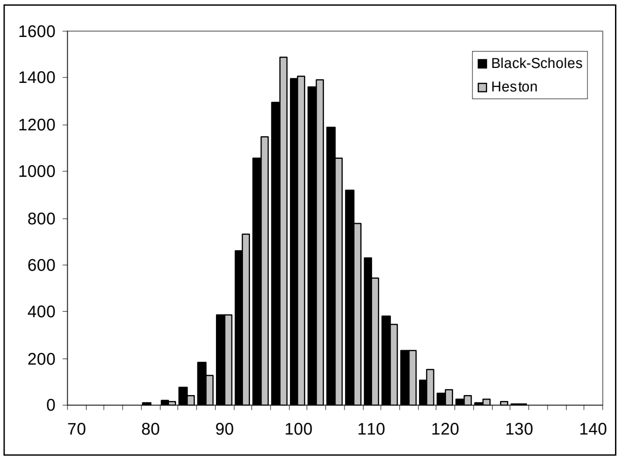


Figure 2 shows that $\rho = -0.5$ produces a distribution of S_T that is negatively skewed.

12.2 Heston Implied Volatility

Another feature of the Heston model is that implied volatilities backed out from option prices generated by the model will show a smile or skew. The shape of the skew is driven by the values of the parameters. The correlation parameter ρ determines the direction of skew, with $\rho > 0$ corresponding to a positive skew, and $\rho < 0$ corresponding to a negative skews. This is illustrated by generating call prices using $S = 100, r = 0.05, T-t = 0.25, \kappa = 2, \theta = 0.01, \lambda = 0, v_0 = 0.01$. We use two values for the volatility of variance, $\sigma_1 = 0.3$ and $\sigma_2 = 0.5$. Figure 3 plots the implied volatility for $\rho = -0.5$. It clearly shows a negative skew for the implied volatilities.

Figure 2: Distribution of S_T under Black-Scholes and Heston, negative correlation

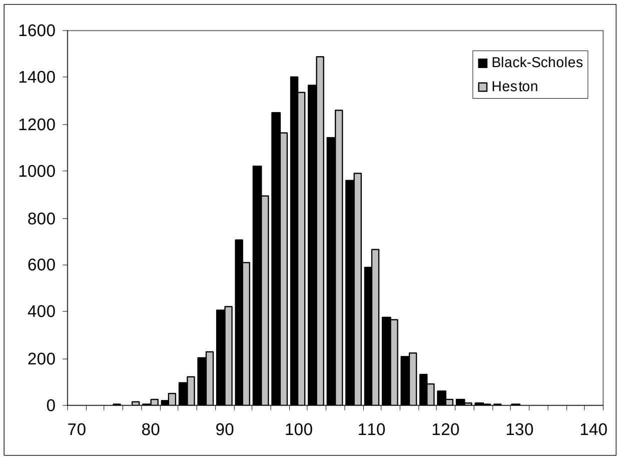


Figure 3:

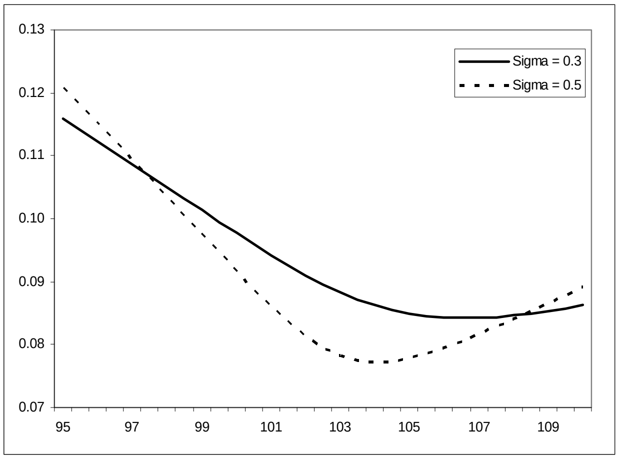


Figure 3. Effect of negative correlation on implied volatility from Heston call prices

Figure 4 plots the implied volatility for $\rho = 0.5$, and shows a positive skew.

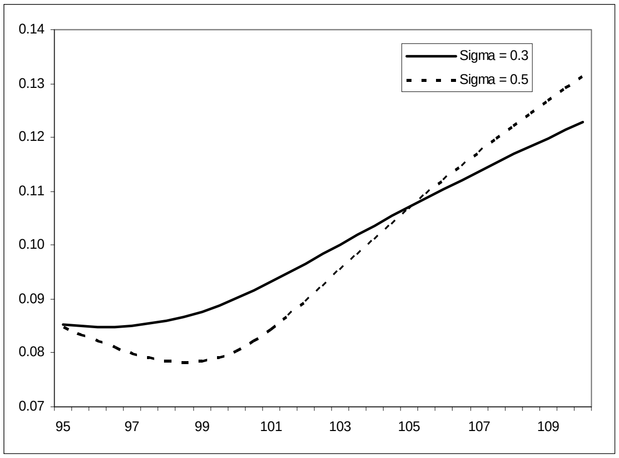


Figure 4. Effect of positive correlation on implied volatility from Heston call prices

Finally, Figure 5 plots the implied volatility for $\rho = 0$. The implied volatilities are more symmetric.

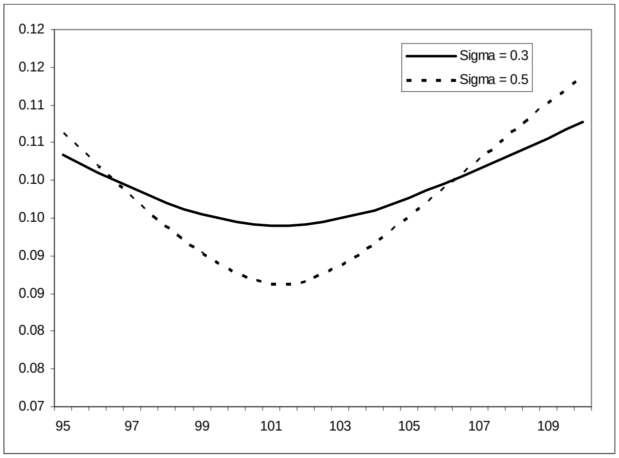


Figure 5. Effect of zero correlation on implied volatility from Heston call prices

12.3 Heston Call Prices

The results illustrated in Figures 1 and 2 suggest that because of the skew in S_T produced by the Heston model relative to Black-Scholes, option prices from the Heston model should differ in a way that is sensible. Call prices in the Heston model are affected by the correlation parameter, ρ, and by the volatility of variance parameter, σ.

12.3.1 Effect of Correlation on Skewness

When $\rho > 0$, the skew in S_T is positive, so more weight is assigned to the right tail of the distribution. This implies that deep out-of-the-money calls from the Heston model should be more expensive than those produced by the Black-Scholes model. Since less weight is assigned to the left tail, deep in-the-money calls from the Heston model should be less expensive than those produced by Black-Scholes. Similarly, when $\rho < 0$ the skew in S_T is negative, so more weight is assigned to the left tail, and less in the right tail. This implies that deep in-of-the-money calls from the Heston model should be less expensive than those produced by the Black-Scholes model, and deep out-of-the-money calls should be more expensive. Table 1 summarizes these observations.

Table 1. Price Comparisons

	$\rho > 0$	$\rho < 0$
OTM Calls	Heston > BS	Heston < BS
ITM Calls	Heston < BS	Heston > BS

Figure 6 reproduces Figure in Heston [7]. It plots the difference between call prices from both models (Heston price minus Black-Scholes price), as the spot price varies from \$70 to \$140. We use the same parameter settings as in Section 12.1, and the same "matched" Black-Scholes volatilities. Figure 6 indicates that the summary in Table 1 is correct.

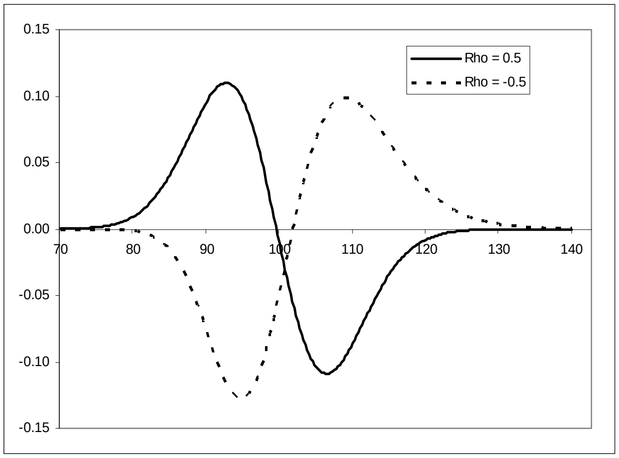


Figure 6. Effect of correlation on Heston prices relative to Black-Scholes

The left portion of Figure 6 correspond to low stock prices and out-of-the-money calls, and the right portion to high stock prices and in-the-money calls. Clearly, when correlation is positive (solid line), the difference is positive in the OTM call region. Heston OTM calls are more expensive than Black-Scholes OTM calls, due to the thickness in the right tail of the distribution of S_T generated by the Heston model. When correlation is negative (dotted line), the difference is positive in the ITM call region. Heston ITM calls are more expensive than Black-Scholes ITM calls, due to the thickness in the left tail of the distribution of S_T generated by the Heston model. Similar arguments can be made for negative differences.

12.3.2 Effect of Volatility of Variance on Kurtosis

The effect of increasing σ is to increase kurtosis. This makes sense, since a high volatility of variance will increase the range of terminal stock price values. This is illustrated in Figure 7, which compares the difference between the Heston and Black-Scholes call prices when $\rho = 0$, and when $\sigma = 0.2$ and $\sigma = 0.3$. It indicates that Heston prices are higher than Black-Scholes prices in ITM and OTM regions, and lower in the at-the-money region. These two observations are consistent with thicker tails of the distribution of S_T generated by the Heston model. Not surprisingly, the difference is more pronounced when σ is higher.

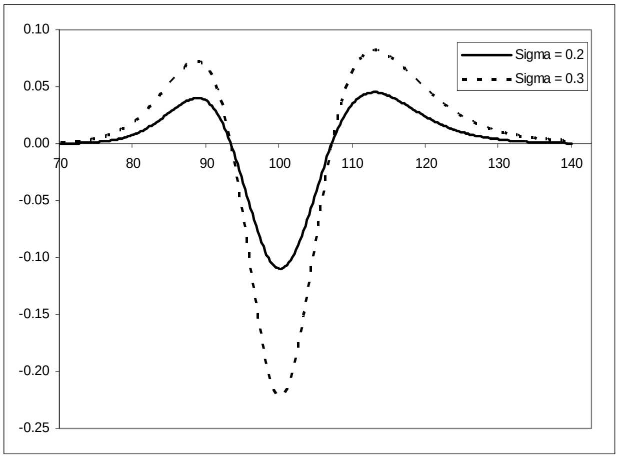


Figure 7. Effect of volatility of variance on Heston prices relative to Black-Scholes

13 Simulation in the Heston Model

To obtain the price of a European option using simulation in the Heston model, we simulate the bivariate process (S_t, v_t) in Equation (1), generate M stock price paths from $t = 0$ to $t = T$, retain the last stock price S_T and obtain the value of the European option, take the average over all stock price paths and discount back to time zero. Hence, for example, the call price is

$$C = e^{-r\tau} \frac{1}{M} \sum_{i=1}^{M} \max\left(0, S_{T,i} - K\right),$$

where $S_{T,i}$ is the terminal stock price generated by the ith stock price path, $i = 1, 2, \cdots, M$. This requires that we have estimates of the parameters $v_0, \rho, \kappa, \theta$, and σ. There are two issues that arise when simulating the bivariate process (S_t, v_t). The first is the slow speed of convergence. The second, more serious issue, is that since v_t follows a CIR process, many simulation schemes–including the Euler and Milstein schemes–will generate negative values for v_t. The simplest way to deal with this second issue is to override negative values. There are at least two ways to do this

1. In the *full truncation scheme*, a negative value for v_t is overidden with zero. Hence, v_t is replaced by $(v_t)^+ = \max(0, v_t)$ everywhere in the discretization of v_t.

2. In the *reflection scheme*, a negative value for v_t is overridden with $-v_t$. Hence, v_t is replaced by $|v_t|$ everywhere in the discretization.

Another way to deal with negative simulated values of v_t is to devise simulation schemes for v_t that do not produce negative values. Much of the research on simulating the CIR variance process in the Heston model is devoted to this approach. Alternatively, the process for $\ln v_t$ can be simulated, and then exponentiated. Finally, the stock price S_t or the log-stock price $x_t = \ln S_t$ can be simulated.

All of the simulation schemes contain the same basic steps. First, two correlated standard normal random variables Z_s and Z_v are generated by generating two standard normal independent random variables Z_1 and Z_2, and by defining $Z_v = Z_1$ and $Z_s = \rho Z_v + \sqrt{1-\rho^2} Z_2$. We then have $E[Z_v] = E[Z_s] = 0$ and $E[Z_v Z_s] = \rho E\left[Z_1^2\right] + \sqrt{1-\rho^2} E[Z_1 Z_2] = \rho$, as required. We then set $dW_{1,t} = \sqrt{dt} Z_s$ and $dW_{2,t} = \sqrt{dt} Z_v$, which produces correlated Brownian motion The basic steps of the simulation schemes are

Step 0. Initialize S_0 to the spot price (or x_0 to the log spot price), and initialize v_0 to the current variance parameter.

Step 1. Generate two dependent normal random variables Z_v and Z_s. Set $W_{1,t} = \sqrt{dt} Z_s$ and $W_{2,t} = \sqrt{dt} Z_v$.

Step 2. Given S_t (or x_t) and v_t, first simulate the updated value v_{t+1},

Step 3. Given v_{t+1} simulate the updated value S_{t+1} (or x_{t+1}).

The following subsections describe some common discretization schemes for (S_t, v_t). We assume that the time grid is discretized as $0 = t_1 < t_2 < \cdots < t_m = T$, where the time increments are equally spaced with width dt.

13.1 Euler Scheme

The simplest way to discretize the process in Equation (1) is to use Euler discretization on the risk-neutral process (using $\mu = r$, the risk free rate), as described in Gatheral [9] and derived for general processes in Glasserman [10].

13.1.1 Process for v_t

The SDE for v_t in (1) in integral form is

$$v_{t+dt} = v_t + \int_t^{t+dt} \kappa\left(\theta - v_u\right) du + \int_t^{t+dt} \sigma\sqrt{v_u} dW_u. \tag{79}$$

The Euler discretization approximates the first integral as

$$\int_t^{t+dt} \kappa\left(\theta - v_u\right) du \approx \kappa\left(\theta - v_t\right) dt.$$

The right hand side involves $(\theta - v_t)$ rather than $(\theta - v_{t+dt})$ since at time t we don't know the value of v_{t+dt}. The second integral is approximated as

$$\int_t^{t+dt} \sigma\sqrt{v_u} dW_{2,u} \approx \sigma\sqrt{v_t}\left(W_{t+dt} - W_t\right) = \sigma\sqrt{v_t}\sqrt{dt}Z_v$$

where Z_v is a standard normal random variable. This leaves us with

$$v_{t+dt} = v_t + \kappa\left(\theta - v_t\right) dt + \sigma\sqrt{v_t}\sqrt{dt}Z_v.$$

13.1.2 Process for S_t

In a similar fashion, the SDE for S_t in (1) (with μ replaced by r) is written in integral form as

$$S_{t+dt} = S_t + r\int_t^{t+dt} S_u du + \int_t^{t+dt} \sqrt{v_u} S_u dW_u.$$

The Euler discretization approximates the first integral as

$$\int_t^{t+dt} S_u du \approx S_t dt$$

and the second as

$$\begin{aligned} \int_t^{t+dt} \sqrt{v_u} S_u dW_{1,u} &\approx \sqrt{v_{t+dt}} S_t \left(W_{t+dt} - W_t\right) \qquad (81) \\ &= \sqrt{v_{t+dt}} S_t \sqrt{dt} Z_s \end{aligned}$$

where Z_s is a standard normal random variable that has correlation ρ with Z_v. Note that we can write $\sqrt{v_{t+dt}}$ on the right hand side of (81) because in the process for S_t the value v_{t+dt} is assumed to be known. This leaves us with

$$S_{t+dt} = S_t + rS_t dt + \sqrt{v_{t+dt}} S_t \sqrt{dt} Z_s.$$

13.1.3 Process for (S_t, v_t)

Given simulated values for $v_t$and S_t we first obtain v_{t+dt} under the full truncation scheme from

$$v_{t+dt} = v_t + \kappa\left(\theta - v_t\right)dt + \sigma\sqrt{\left(v_t\right)^+}\sqrt{dt}Z_v$$

and we obtain S_{t+dt} from

$$S_{t+dt} = S_t + rS_t dt + \sqrt{\left(v_{t+dt}\right)^+}S_t\sqrt{dt}Z_s.$$

To use the reflection scheme, replace $()^+$ with $||$ in both equations.

13.2 Milstein Scheme

This scheme is described in Glasserman [10] and in Kloeden and Platen [13] for general processes, and in Kahl and Jackel [12] for stochastic volatility models.

13.2.1 Process for v_t

The SDE for v_t is

$$dv_t = \kappa\left(\theta - v_t\right)dt + \sigma\sqrt{v_t}dW_t = a_t dt + b_t dW_t$$

where, for notational simplicity, $a_t = a\left(v_t\right)$ and $b_t = b\left(v_t\right)$. In integral form

$$v_{t+dt} = v_t + \int_t^{t+dt} a_s ds + \int_t^{t+dt} b_s dW_s. \tag{82}$$

Apply Ito's Lemma to the coefficients a_s and b_s. These will follow the process

$$\begin{aligned} da_s &= \left(a_s a'_s + \frac{1}{2}b_s a''_s\right)ds + \left(b_s a'_s\right)dW_s \\ db_s &= \left(a_s b'_s + \frac{1}{2}b_s^2 b''_u\right)du + \left(b_u b'_u\right)dW_u. \end{aligned}$$

In integral form (with $t < s < t + dt$)

$$\begin{aligned} a_s &= a_t + \int_t^s \left(a_u a'_u + \frac{1}{2}b_u a''_u\right)du + \int_t^s b_u a'_u dW_u \\ b_s &= b_t + \int_t^s \left(a_u b'_u + \frac{1}{2}b_u^2 b''_u\right)du + \int_t^s b_u b'_u dW_u \end{aligned}$$

Substitute for a_s and b_s into Equation (82)

$$\begin{aligned} v_{t+dt} &= v_t + \int_t^{t+dt}\left[a_t + \int_t^s \left(a_u a'_u + \frac{1}{2}b_u a''_u\right)du + \int_t^s b_u a'_u dW_u\right]ds \\ &\quad + \int_t^{t+dt}\left[b_t + \int_t^s \left(a_u b'_u + \frac{1}{2}b_u^2 b''_u\right)du + \int_t^s b_u b'_u dW_u\right]dW_s. \end{aligned}$$

The terms higher than order one are $dsdu = \mathcal{O}\left((dt)^2\right)$ and $dsdW_u = \mathcal{O}\left((dt)^{3/2}\right)$ and are ignored. The terms involving $dW_u dW_s$ are retained since $dW_u dW_s = \mathcal{O}(dt)$ is of order one. This leaves us with

$$v_{t+dt} = v_t + a_t \int_t^{t+dt} ds + b_t \int_t^{t+dt} dW_s + \int_t^{t+dt} \int_t^s b_u b'_u dW_u dW_s. \tag{83}$$

Apply Euler discretization to the last term in (83)

$$\begin{aligned} \int_t^{t+dt} \int_t^s b_u b'_u dW_u dW_s &\approx b_t b'_t \int_t^{t+dt} \int_t^s dW_u dW_s \\ &= b_t b'_t \int_t^{t+dt} \left[W_s - W_t\right] dW_s \\ &= b_t b'_t \left[\int_t^{t+dt} W_s dW_s - W_t W_{t+dt} + W_t^2 \right] \end{aligned}$$

Now define $dY_t = W_t dW_t$. Using Ito's Lemma, it is easy to show[2] that Y_t has solution $Y_t = \frac{1}{2}W_t^2 - \frac{1}{2}t$ so that

$$\int_t^{t+dt} W_s dW_s = Y_{t+dt} - Y_t = \frac{1}{2}W_{t+dt}^2 - \frac{1}{2}W_t^2 - \frac{1}{2}dt.$$

Substitute back to obtain

$$\begin{aligned} \int_t^{t+dt} \int_t^s b_u b'_u dW_u dW_s &\approx \frac{1}{2} b_t b'_t \left[\left(W_{t+dt} - W_t\right)^2 - dt \right] \\ &= .\frac{1}{2} b_t b'_t \left[\Delta W_t^2 - dt \right]. \end{aligned}$$

where $\Delta W_t = W_{t+dt} - W_t$. Hence from (83) the general form of Milstein discretization is

$$v_{t+dt} = v_t + a_t dt + b_t \Delta W_t + \frac{1}{2} b_t b'_t \left(\Delta W_t^2 - dt \right). \tag{84}$$

To obtain the Milstein discretization of the Heston model, substitute for $a_t = \kappa(\theta - v_t)\, dt$ and $b_t = \sigma\sqrt{v_t}$ and use the fact that ΔW_t is equal in distribution to $\sqrt{dt} Z_v$, where Z_v is standard normal. This produces

$$v_{t+dt} = v_t + \kappa(\theta - v_t)\, dt + \sigma\sqrt{v_t}\sqrt{dt} Z_v + \frac{1}{4}\sigma^2 dt \left(Z_v^2 - 1\right) \tag{85}$$

which can be written

$$v_{t+dt} = \left(\sqrt{v_t} + \frac{1}{2}\sigma\sqrt{dt} Z_v \right)^2 + \kappa(\theta - v_t)\, dt - \frac{1}{4}\sigma^2 dt.$$

---

[2] Indeed, $\frac{\partial Y}{\partial t} = -\frac{1}{2}$, $\frac{\partial Y}{\partial W} = W$, and $\frac{\partial^2 Y}{\partial W^2} = 1$, so that $dY_t = \left(-\frac{1}{2} + 0 + \frac{1}{2} \cdot 1 \cdot 1\right) dt + (W_t \cdot 1)\, dW_t = W_t dW_t$.

This last equation is also Equation (2.18) of Gatheral [9]. Milstein discretization of the variance process produces far fewer negative values for the variance than Euler discretization. Nevertheless, the full truncation scheme or the reflection scheme must be applied to (85) as well.

13.2.2 Process for $\ln S_t$

Gatheral [9] advocates the discretization of the log-stock process. By Ito's Lemma $\ln S_t$ follows the process

$$d\ln S_t = \left(r - \frac{1}{2}v_t\right)dt + \sqrt{v_t}dW_{1,t}. \tag{86}$$

He explains that the discretization of $\ln S_t$ rather than S_t means that there are no higher corrections to be brought to the Euler discretization. Hence $\ln S_t$ can be discretized as

$$\ln S_{t+dt} = \ln S_t + \left(r - \frac{1}{2}v_t\right)dt + \sqrt{v_t}\sqrt{dt}dZ_s.$$

In later sections we show a refinement to this scheme.

13.3 Implicit Milstein Scheme

In Equation (84), the drift coefficient a_t and the volatility coefficient b_t are both functions of v_t. In the Milstein *drift-implicit scheme*, or simply *implicit scheme*, the drift coefficient is a function of v_{t+dt}. Hence under this scheme the drift coefficient is known only implicitly, and not explicitly, as is the case when it depends on v_t. Under the *Ito version* of this scheme Equation (84) becomes

$$v_{t+dt} = v_t + a_{t+dt}dt + b_t\Delta W_t + \frac{1}{2}b_t b_t'\left(\Delta W_t^2 - dt\right)$$

where $a_{t+dt} = a\left(v_{t+dt}\right)$. See Kloeden and Platen [13]. It is also possible to interpolate between explicit and implicit Milstein schemes as follows

$$v_{t+dt} = v_t + \left[\alpha a_{t+dt} + (1-\alpha)a_t\right]dt + b_t\Delta W_t + \frac{1}{2}b_t b_t'\left(\Delta W_t^2 - dt\right)$$

where $\alpha \in [0,1]$. The explicit Mistein scheme corresponds to $\alpha = 0$, and the implicit Milstein scheme to $\alpha = 1$. Under the Heston model, the implicit scheme involves replacing $\kappa\left(\theta - v_t\right)dt$ in Equation (85) with $\kappa\left(\theta - v_{t+dt}\right)dt$. Bringing the term $kv_{t+dt}dt$ over to the left-hand side of the resulting equation and dividing by $1 + \kappa dt$ produces

$$v_{t+dt} = \frac{v_t + \kappa\theta dt + \sigma\sqrt{v_t}\sqrt{dt}Z_v + \frac{1}{4}\sigma^2 dt\left(Z_v^2 - 1\right)}{1 + \kappa dt}. \tag{87}$$

13.4 Balanced Implicit Milstein Scheme

This is presented by Platen and Heath [15] as

$$v_{t+dt} = v_t + a_t dt + b_t \Delta W_t + (v_t - v_{t+dt})\, C(v_t)$$

where

$$C(v_t) = c^0(v_t)\, dt + c^1(v_t)\, |\Delta W_t|$$

with c^0 and c^1 suitably chosen positive real valued and bounded functions. Kahl and Jackel [12] suggest $c^0(v_t) = \kappa$ and $c^1(v_t) = \sigma\sqrt{v_t}$. Their balanced implied Milstein scheme for the Heston model is therefore

$$\begin{aligned} v_{t+dt} &= v_t + \kappa(\theta - v_t)\, dt + \sigma\sqrt{v_t}\sqrt{dt} Z_v + (v_t - v_{t+dt})\, C(v_t) \\ &= \frac{v_t(1 + C(v_t)) + \kappa(\theta - v_t)\, dt + \sigma\sqrt{v_t}\sqrt{dt} Z_v}{C(v_t)} \end{aligned}$$

where ΔW_t is approximated as $\sqrt{dt} Z_v$ so that $C(v_t) = \left(\kappa dt + \sigma\sqrt{v_t}\sqrt{dt}\,|Z_v|\right)$. Under this scheme, simulated values of v_t are guaranteed to be positive. Unfortunately, as shown by Kahl and Jackel [12], the convergence of this scheme can be very poor. For details, see [11], Equations (6.25) through (6.28)

13.5 Pathwise Approximation

One method for discretization of the variance is the pathwise approximation illustrated in Kahl and Jackel [12]. Its convergence is fast, especially for small values of σ. The discretization scheme is given by

$$v_{t+dt} = v_t + \left(\kappa\left(\tilde{\theta} - v_t\right) + \sigma\beta_n\sqrt{v_t}\right)\left(1 + \frac{\sigma\beta_n - 2\kappa\sqrt{v_t}}{4\sqrt{v_t}} dt\right) dt$$

where $\tilde{\theta} = \theta - \frac{\sigma^2}{4\kappa}$ and where $\beta_n = Z_v/\sqrt{dt}$.

13.6 Kahl-Jackel Scheme

This was suggested by Kahl and Jackel [12]. It involves simulating v_t with an implicit Milstein scheme from (87), and simulating $\ln S_t$ with their IJK discretization

$$\begin{aligned} v_{t+dt} &= \frac{v_t + \kappa\theta dt + \sigma\sqrt{v_t}\sqrt{dt} Z_v + \frac{1}{4}\sigma^2 dt\left(Z_v^2 - 1\right)}{1 + \kappa dt} \qquad (88) \\ \ln S_{t+dt} &= \ln S_t + \left(r - \frac{v_t + v_{t+dt}}{4}\right) dt + \rho\sqrt{v_t} Z_v \sqrt{dt} \\ &\quad + \frac{1}{2}\left(\sqrt{v_t} + \sqrt{v_{t+1}}\right)(Z_s + \rho Z_v)\sqrt{dt} + \frac{\rho\sigma dt}{2}\left(Z_v^2 - 1\right). \end{aligned}$$

Since this can also produce negative variances, in Equation (88) the full truncation scheme or the reflection scheme should be used. See their paper for details of the derivation.

13.7 QE Scheme

Recall from Section 2 that the value of v_{t+dt} conditional on a realized value v_t follows a non-central chi-square distribution. Andersen (reference) suggests sampling from an approximation to the distribution, depending on whether the non-centrality parameter, $2c_{t+dt}v_{t+dt}e^{-\kappa dt}$, is large or small. Since the non-centrality parameter is proporational to v_t, large or small values of the parameter correspond to large or small values of v_t. The algorithm switches back and forth between two different approximations to the non-central chi-square distribution, depending on the magnitude of v_t.

1. For moderate or high values of v_t, a non-central chi-square random variable can be approximated by a power function applied to a standard normal variable Z_v

   $$v_{t+dt} = a\left(b + Z_v\right)^2 \tag{89}$$

   where a and b are determined by moment-matching using the mean and variance of the CIR process described in Section 2.

2. For low values of v_t, the non-central chi-square density can be approximated by a weighted average of a term involving Dirac's delta function δ and a term involving $e^{-\beta x}$

   $$\Pr\left(v_{t+dt} \in [x, x+dx]\right) = \left(p\delta\left(0\right) + \beta\left(1-p\right)e^{-\beta x}\right). \tag{90}$$

   where p and β are also determined by moment-matching. Note that $0 \le p \le 1$. Integrating Equation (90) and inverting produces the inverse distribution function

   $$\Psi^{-1}\left(u\right) = \left\{ \begin{array}{ll} 0 & \text{for } 0 \le u \le p \\ \beta^{-1}\ln\frac{1-p}{1-u} & \text{for } p < u \le 1 \end{array} \right. . \tag{91}$$

   The sampling scheme for low values of v_s is therefore

   $$v_t = \Psi^{-1}\left(U_v\right) \tag{92}$$

   where U_v is a uniform random number. The QE sampling schemes are defined by Equations (89) and (92).

13.7.1 Moment-Matching

Recall that the mean and variance of the CIR process are

$$\begin{aligned} m &= \theta + \left(v_{t+dt} - \theta\right)e^{-\kappa dt}, \\ s^2 &= \frac{v_{t+dt}\sigma^2 e^{-\kappa dt}}{\kappa}\left(1 - e^{-\kappa dt}\right) + \frac{\theta\sigma^2}{2\kappa}\left(1 - e^{-\kappa dt}\right)^2. \end{aligned} \tag{93}$$

1. For moderate or high values of v_s, since $v_{t+dt} = a\left(b + Z_v\right)^2$ from (89), we have that $E\left[v_{t+dt}\right] = a\left(1 + b^2\right)$ and $\text{Var}[v_{t+dt}] = 2a^2\left(1 + 2b^2\right)$. Equating these to m and s^2 respectively and solving for a and b produces

   $$\begin{aligned} b &= \left(\frac{2}{\psi} - 1 + \sqrt{\frac{2}{\psi}\left(\frac{2}{\psi} - 1\right)}\right)^{\frac{1}{2}} \\ a &= \frac{m}{1 + b^2}, \end{aligned} \tag{94}$$

   where $\psi = s^2/m^2$. Note that b is only defined when $\psi \leq 2$.

2. For low values of v_s, the mean and variance of v_t are found by integrating Equation (90).directly, which produces $E\left[v_{t+dt}\right] = (1 - p)/\beta$ and $\text{Var}[v_{t+dt}] = \left(1 - p^2\right)/\beta^2$. Again, equating these to m and s^2 and solving for p and β produces

   $$p = \frac{\psi - 1}{\psi + 1} \quad \text{and} \quad \beta = \frac{1 - p}{m}. \tag{95}$$

   Note that the condition that $0 \leq p \leq 1$ in Equation (90) requires that $\psi \geq 1$.

The value of ψ stipulates which approximation to use. Indeed, for $\psi \leq 2$ we can use the first approximation, and for $\psi \geq 1$ we can use the second approximation. This implies that a critical level $\psi_c \in [1, 2]$ ought to be defined as a threshold. Andersen uses $\psi_c = 1.5$

The scheme can be summarized as follows, using a cutoff value of $1 \leq \psi_c \leq 2$.

1. Given v_t, obtain m and s^2 from Equation (93) using estimates of θ and κ, and obtain $\psi = \frac{s^2}{m^2}$.

2. Draw a uniform random number U_v.

   if $\psi < \psi_c$ compute a and b from Equation (94), and $Z_v = \Phi^{-1}\left(U_v\right)$. Define $v_{t+dt} = a\left(b + Z_v\right)^2$

   if $\psi > \psi_c$ compute β and p from Equation (95). Define $v_{t+dt} = \Psi^{-1}\left(U_v\right)$ from Equation (91).

In the same paper, Andersen describes the Truncated Gaussian (TG) scheme, which uses the approximation $v_{t+dt} = \left(\mu + \sigma Z_v\right)^+$ for μ and σ constants. The performance of TG scheme is inferior that to the QE scheme, however, and Andersen recommends that the QE scheme be the default choice.

13.7.2 Process for $\ln S_t$

Andersen proposes a discretization scheme for x_t that overcomes the problem of "leaky" correlation brought on by a Euler discretization of x_t. The integral form of the process for v_t is

$$v_{t+dt} = v_t + \kappa\theta dt - \int_t^{t+dt} v_u du + \sigma \int_t^{t+dt} \sqrt{v_u} dW_{2,u}.$$

Rearranging terms produces

$$\int_t^{t+dt} \sqrt{v_u} dW_{2,u} = \frac{1}{\sigma}\left(v_{t+dt} - v_t - \kappa\theta dt + \kappa \int_t^{t+dt} \sqrt{v_u} du \right). \tag{96}$$

Now applying a Cholesky decomposition to the process for $\ln S_t$ means that it can be written in the integral form

$$\ln S_{t+dt} = \ln S_t + rdt - \frac{1}{2}\int_t^{t+dt} v_u du + \int_t^{t+dt} \sqrt{v_u}\left(\rho dW_{2,u} + \sqrt{1-\rho^2} dW_{1,u} \right).$$

Now substitute Equation (96) to obtain

$$\begin{aligned} \ln S_{t+dt} &= \ln S_t + rdt + \frac{\rho}{\sigma}\left(v_{t+dt} - v_t - \kappa\theta dt \right) \\ &\quad + \left(\frac{\kappa\rho}{\sigma} - \frac{1}{2} \right) \int_t^{t+dt} v_u du + \sqrt{1-\rho^2} \int_t^{t+dt} \sqrt{v_u} dW_{1,u}. \end{aligned}$$

Andersen uses the approximations

$$\begin{aligned} \int_t^{t+dt} v_u du &\approx dt\left(\gamma_1 v_t + \gamma_2 v_{t+dt} \right) \\ \int_t^{t+dt} \sqrt{v_u} dW_u &\approx Z_v \sqrt{dt} \sqrt{\gamma_1 v_t + \gamma_2 v_{t+dt}} \end{aligned}$$

to arrive at

$$\ln S_{t+dt} = \ln S_t + rdt + K_0 + K_1 v_t + K_2 v_{t+dt} + \sqrt{K_3 v_t + K_4 v_{t+dt}} Z_v \tag{97}$$

where

$$\begin{aligned} K_0 &= -\frac{\kappa\rho\theta}{\sigma} dt \\ K_1 &= \left(\frac{\kappa\rho}{\sigma} - \frac{1}{2} \right) \gamma_1 dt - \frac{\rho}{\sigma} \\ K_2 &= \left(\frac{\kappa\rho}{\sigma} - \frac{1}{2} \right) \gamma_1 dt + \frac{\rho}{\sigma} \\ K_3 &= \left(1 - \rho^2 \right) \gamma_1 dt \\ K_4 &= \left(1 - \rho^2 \right) \gamma_2 dt. \end{aligned}$$

The constants γ_1 and γ_2 are arbitrary. Setting $\gamma_1 = 1$ and $\gamma_2 = 0$ produces an Euler-type scheme, while $\gamma_1 = \gamma_2 = \frac{1}{2}$ produces a central discretization. With these values, the algorithm to generate a value of S_{t+dt}, given S_t,v_t, and v_{t+dt} is evident.

13.7.3 Martingale Correlation

Under the risk neutral measure $\mathbb{Q}$, the discounted asset price will be a martingale in continuous time. On the other hand, the discretizated stock price

$$S_{t+dt} = S_t \exp\left(rdt + K_0 + K_1 v_t + K_2 v_{t+dt} + \sqrt{K_3 v_t + K_4 v_{t+dt}} Z_v\right) \tag{98}$$

will not be a martingale. The resulting bias is minor, since the drift away from the martingale can be controlled by reducing the size of the time increment, dt. Nonetheless, Andersen shows that the martingale property can be satisfied simply by replacing K_0 with K_0^*. Incorporating the martingale correction into the QE algorithm is straightforward. Andersen shows that

$$K_0^* = \begin{cases} \frac{Ab^2 a}{1-2Aa} + \frac{1}{2}\ln(1-2Aa) - \left(K_1 + \frac{1}{2}K_3\right) v_t & \text{when } \psi \leq \psi_c \\ \\ -\ln\left(p + \frac{\beta(1-p)}{\beta - A}\right) - \left(K_1 + \frac{1}{2}K_3\right) v_t & \text{when } \psi > \psi_c \end{cases}$$

where $A = K_2 + \frac{1}{2}K_4$, and where a and b are defined in Equation (94), and p and β are from Equation (95). The martingale correction holds provided that $A < \frac{1}{2a}$ when $\psi \leq \psi_c$, and provided that $A < \beta$ in the case $\psi > \psi_c$.

13.8 Exact Simulation Broadie and Kaya

To come.

13.9 Moment Matching Scheme

Andersen and Brotherton-Ratcliffe [3] propose a moment-matched discretization scheme that preserves positivity of the variance. The scheme produces a variance that is distributed as lognormal, so a natural choice of parameterization is one that matches the first two moments of the discretized process to the lognormal moments. This produces a discretization of the form

$$v_{t+dt} = \left(e^{-\kappa dt} v_t + \left(1 - e^{-\kappa dt}\right)\theta\right) \exp\left(-\frac{1}{2}\Gamma_t^2 dt + \Gamma_t \sqrt{dt} Z_v\right)$$

where

$$\Gamma_t = \frac{1}{dt} \ln\left(1 + \frac{\frac{1}{2\kappa}\sigma^2 v_t \left(1 - e^{-2\kappa dt}\right)}{\left(e^{-\kappa dt} v_t + \left(1 - e^{-\kappa dt}\right)\theta\right)^2}\right).$$

14 Variance Swap in the Heston Model

Recall from Equation (1) that the volatility of the Heston model is driven by the CIR process

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dZ_t$$

where $Z_t = Z_t^{(2)}$. It is straightforward to show that the expected value of v_t, given v_s $(s < t)$ is

$$\begin{aligned} E\left[\left. v_t \right| v_s\right] &= v_s e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right) \\ &= \theta + (v_s - \theta)\, e^{-\kappa(t-s)}. \end{aligned}$$

See, for example, Brigo and Mercurio [4]. As explained by Gatheral [9], a variance swap requires an estimate of the future variance over the $(0, T)$ time period, namely of the total variance $w_T = \int_0^T v_t dt$. A fair estimate of w_T is its conditional expected value $E\left[\left. w_T \right| v_0\right]$. This is given by

$$\begin{aligned} E\left[\left. w_T \right| v_0\right] &= E\left[\left. \int_0^T v_t dt \right| v_0\right] \\ &= \int_0^T E\left[\left. v_t \right| v_0\right] dt \\ &= \int_0^T \left[\theta + (v_0 - \theta)\, e^{-\kappa t}\right] dt \\ &= \theta T + \frac{1 - e^{-\kappa T}}{\kappa}\left(v_0 - \theta\right). \end{aligned}$$

Since v_T represents the total variance over $(0, T)$, it must be scaled by T in order to represent a fair estimate of annual variance (assuming that T is expressed in years.) Hence the strike variance for a variance swap is given by

$$\frac{1}{T} E\left[\left. w_T \right| v_0\right] = \theta + \frac{1 - e^{-\kappa T}}{\kappa T}\left(v_0 - \theta\right).$$

This is the expression on page 138 of Gatheral [9].

15 Parameter Estimation

To come.

References

[1] Albrecher, H., Mayer, P, Schoutens, W, and J. Tistaert (2006). "The Little Heston Trap." Working Paper, Radon Institute, Austrian Academy of Sciences.

[2]

[3]

[4] Brigo, D., and F. Mercurio (2006). *Interest Rate Models - Theory and Practice: With Smile, Inflation, and Credit.* Second Edition. New York, NY: Springer.

[5] Carr, P., and D. Madan (1999). "Option Valuation Using the Fast Fourier Transform." Journal of Computational Finance, Vol. 2, No. 4, pp. 61-73.

[6] Chourdakis, K. (2005). "Option Pricing Using Fractional FFT." Journal of Computational Finance, Vol. 8, No 2., pp. 1-18.

[7] Heston, S.L. (1993). "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *Review of Financial Studies*, Vol. 6, pp 327-343.

[8] Bakshi, G., and D. Madan (2000). "Spanning and Derivative-Security Valuation." *Journal of Financial Economics*, Vol. 55, pp 205-238.

[9] Gatheral, J. (2006) *The Volatility Surface: A Practitioner's Guide*. New York, NY: John Wiley & Sons.

[10] Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. New York, NY: Springer.

[11] Kahl, C. (2008). *Modeling and Simulation of Stochastic Volatility in Finance*. Published by Dissertation.com.

[12] Kahl, C., and P. Jackel (2006). Fast Strong Approximation Monte-Carlo Schemes for Stochastic Volatility Models. *Quantitative Finance*, Vol. 6, No. 6.

[13] Kloeden, P.E., and E. Platen (1992). *Numerical Solution of Stochastic Differential Equations*. New York, NY: Springer.

[14] Lewis, A.L.(2000). *Option Valuation Under Stochastic Volatility: With Mathematica Code*. Finance Press.

[15] Platen, E., and D. Heath (2009). *A Benchmark Approach to Quantitative Finance, Volume 13*. New York, NY: Springer.