CHAPTER 1

The Volatility Problem

Suppose we use the standard deviation of possible future returns on a stock as a measure of its volatility. Is it reasonable to take that volatility as a constant over time? I think not.

—Fischer Black

INTRODUCTION

It is widely accepted today that an assumption of a constant volatility fails to explain the existence of the volatility smile as well as the leptokurtic character (fat tails) of the stock distribution. The Fischer Black quote, made shortly after the famous constant-volatility Black-Scholes model was developed, proves the point.

In this chapter, we will start by describing the concept of Brownian Motion for the Stock Price Return, as well as the concept of historic volatility.

We will then discuss the derivatives market and the ideas of hedging and risk neutrality. We will briefly describe the Black-Scholes Partial Derivatives Equation (PDE) in this section.

Next, we will talk about jumps and level-dependent volatility models. We will first mention the jump-diffusion process and introduce the concept of leverage. We will then refer to two popular level-dependent approaches: the Constant Elasticity Variance (CEV) model and the Bensoussan-Crouhy-Galai (BCG) model.

At this point, we will mention local volatility models developed in the recent past by Dupire and Derman-Kani and we will discuss their stability.

Following this, we will tackle the subject of stochastic volatility where we will mention a few popular models such as the Square-Root model and GARCH.

We will then talk about the Pricing PDE under stochastic volatility and the risk-neutral version of it. For this we will need to introduce the concept of Market Price of Risk.

The Generalized Fourier Transform is the subject of the following section. This technique was used by Alan Lewis extensively for solving stochastic volatility problems.

Next, we will discuss the Mixing Solution, both in a correlated and non-correlated case. We will mention its link to the Fundamental Transform and its usefulness for Monte-Carlo-based methods.

We will then describe the Long-Term Asymptotic case, where we get closed-form approximations for many popular methods such as the Square-Root model.

We will finally talk about pure-jump models such as Variance Gamma and VGSA.

THE STOCK MARKET

The Stock Price Process

The relationship between the stock market and the mathematical concept of Brownian Motion goes back to Bachelier [19]. A Brownian Motion corresponds to a process the increments of which are independent stationary normal random variables. Given that a Brownian Motion can take negative values, it cannot be used for the stock price. Instead, Samuelson [222] suggested to use this process to represent the *return* of the stock price, which will make the stock price a Geometric (or exponential) Brownian Motion.

In other words, the stock price S follows a log-normal process[1]

$$dS_t = \mu S_t dt + \sigma S_t dB_t \tag{1.1}$$

where dB_t is a Brownian Motion process, μ the instantaneous expected total return of the stock (possibly adjusted by a dividend yield), and σ the instantaneous standard deviation of stock price returns, called the *volatility* in financial markets.

Using Ito's lemma,[2] we also have

$$d\ln(S_t) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dB_t. \tag{1.2}$$

[1]For an introduction to Stochastic Processes see Karatzas [175] or Oksendal [207].
[2]See for example Hull [153].

The stock return μ could easily become time dependent without changing any of our arguments. For simplicity, we will often refer to it as μ even if we mean μ_t. This remark holds for other quantities such as r_t the interest-rate, or q_t the dividend-yield.

The equation (1.1) represents a continuous process. We can either take this as an approximation to the real discrete tick by tick stock movements, or consider it the real unobservable dynamics of the stock price, in which case the discrete prices constitute a *sample* from this continuous ideal process. Either way, the use of a continuous equation makes the pricing of financial instruments more analytically tractable.

The discrete equivalent of (1.2) is

$$\ln S_{t+\Delta t} = \ln S_t + \left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}B_t \tag{1.3}$$

where B_t is a sequence of independent normal random variables with zero mean and variance of 1.

Historic Volatility

This suggests a first simple way to estimate the volatility σ, namely the *historic volatility*. Considering $S_1, \ldots, S_N$ a sequence of known historic daily stock close prices and calling $R_n = \ln(S_{n+1}/S_n)$ the stock price return between two days and $\bar{R} = \frac{1}{N}\sum_{n=0}^{N-1} R_n$ the mean return, the historic volatility would be the annualized standard deviation of the returns, namely

$$\sigma_{hist} = \sqrt{\frac{252}{N-1}\sum_{n=0}^{N-1}(R_n - \bar{R})^2}. \tag{1.4}$$

Because we work with annualized quantities, and we are using daily stock close prices, we needed the factor 252, supposing that there are approximately 252 business days in a year.[3]

Note that N, the number of observations, can be more or less than one year; hence when talking about a historic volatility, it is important to know what time horizon we are considering. We can indeed have three-month historic volatility or three-year historic volatility. Needless to say, taking too few prices would give an inaccurate estimation. Similarly, the begin and end date of the observations matter. It is preferable to take the end date as close as possible to today, so that we include more recent observations.

[3] Clearly the observation frequency does not have to be daily.

An alternative was suggested by Parkinson [210] where instead of daily close prices, we use the high and the low prices of the stock on that day, and $R_n = \ln(S_n^{high}/S_n^{low})$.

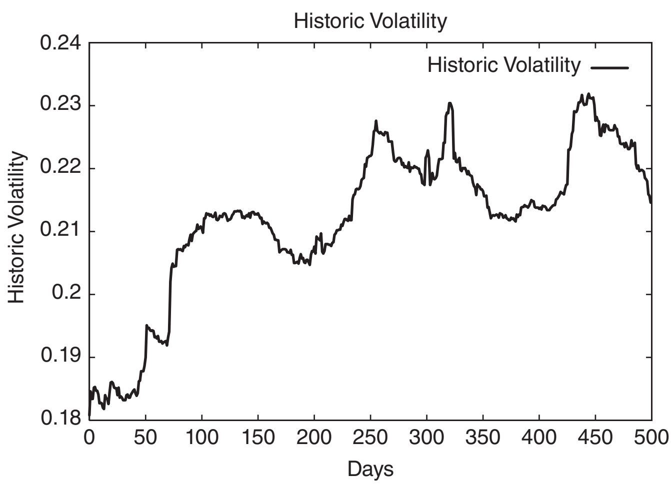
The volatility would then be

$$\sigma_{parkinson} = \sqrt{\frac{252}{N-1}\frac{1}{4\ln(2)}\sum_{n=0}^{N-1}(R_n - \bar{R})^2}.$$

This second moment estimation derived by Parkinson is based on the fact that the range R_n of the asset follows a *Feller* distribution.

Plotting for instance the one-year rolling[4] historic volatility (1.4) of S&P 500 Stock Index, it is easily seen that this quantity is *not* constant over time. This observation was made as early as the 1960s by many financial mathematicians and followers of the *Chaos Theory*. We therefore need time-varying volatility models.

One natural extension of the constant volatility approach is to make σ_t a deterministic function of time. This is equivalent to giving the volatility a term structure, by analogy with interest rates.



SPX Historic Rolling Volatility from January 3, 2000, to December 31, 2001. As we can see, the volatility is clearly non-constant.

[4]By *rolling* we mean that the one-year interval slides within the total observation period.

THE DERIVATIVES MARKET

Until now we only mentioned the stock price movements independently from the derivatives market. We now are going to include the financial derivatives (specialty options) prices as well. These instruments became very popular and as liquid as the stocks themselves after Balck and Scholes introduced their risk-neutral pricing formula in [40].

The Black-Scholes Approach

The Black-Scholes approach makes a number of reasonable assumptions about markets being frictionless and uses the log-normal model for the stock price movements. It also supposes a constant or deterministically time dependent stock drift and volatility. Under these conditions they prove that it is possible to hedge a position in a contingent claim dynamically by taking an offsetting position in the underlying stock and hence become *immune* to the stock movements. This risk neutrality is possible because, as they show, we can replicate the financial derivative (for instance an option) by taking positions in cash and the underlying security. This condition of possibility of replication is called *market completeness*.

In this situation everything happens as if we were replacing the stock drift μ_t with the risk-free rate of interest r_t in (1.1), or $r_t - q_t$ if there is a dividend-yield q_t. The contingent claim $f(S,t)$ having a payoff $G(S_T)$ will satisfy the famous Black-Scholes equation

$$rf = \frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}. \tag{1.5}$$

Indeed, the Hedged Portfolio $\Pi = f - \frac{\partial f}{\partial S}S$ is immune to the stock random movements and according to Ito's lemma verifies

$$d\Pi = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}\right) dt$$

which must also be equal to $r\Pi dt$ or else there would be possibility of Riskless Arbitrage.[5]

Note that this equation is closely related to the Feynman-Kac equation satisfied by $F(S,t) = \mathbf{E}_t(h(S_T))$ for any function h under the risk-neutral

[5]For a detailed discussion see Hull [153].

measure. $F(S,t)$ must be a martingale[6] under this measure and therefore must be driftless, which implies $dF = \sigma S \frac{\partial F}{\partial S} dB_t$ and

$$0 = \frac{\partial F}{\partial t} + (r-q)S\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}.$$

This would indeed be a different way to reach the same Black-Scholes equation, by using $f(S,t) = \exp(-rt)F(S,t)$, as was done for instance in Shreve [229].

Let us insist again on the fact that the real drift of the stock price does not appear in the previous equation, which makes the volatility σ_t the only unobservable quantity.

As we said, the volatility could be a deterministic function of time without changing the earlier argument, in which case all we need to do is to replace σ^2 with $\frac{1}{t}\int_0^t \sigma_s^2 ds$ and keep everything else the same.

For calls and puts, where the payoffs $G(S_T)$ are respectively $MAX(0, S_T - K)$ and $MAX(0, K - S_T)$ where K is the strike price and T the maturity of the option, the Black Scholes Partial Derivatives Equation is solvable and gives the celebrated Black Scholes formula

$$call_t = S_t e^{-q(T-t)}\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{1.6}$$

and

$$put_t = -S_t e^{-q(T-t)}\Phi(-d_1) + Ke^{-r(T-t)}\Phi(-d_2) \tag{1.7}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{u^2}{2}} du$ is the Cumulative Standard Normal function and $d_1 = d_2 + \sigma\sqrt{T-t}$ and $d_2 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r - q - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$.

Note that using the well-known symmetry property for normal distributions $\Phi(-x) = 1 - \Phi(x)$ in the above formulae, we could reach the *Put-Call Parity* relationship

$$call_t - put_t = S_t e^{-q(T-t)} - Ke^{-r(T-t)} \tag{1.8}$$

that we can also rearrange as

$$S_t e^{-q(T-t)} - call_t = Ke^{-r(T-t)} - put_t.$$

The left-hand side of the last equation is called a *covered call* and is equivalent to a short position in a put combined with a bond.

[6]For an explanation see Shreve [229] or Karatzas [175].

The Cox Ross Rubinstein Approach

Later, Cox, Ross, and Rubinstein [71] developed a simplified approach using the Binomial Law to reach the same pricing formulae. The approach commonly referred to as the Binomial Tree uses a tree of recombining spot prices where at a given time-step n we have $n + 1$ possible $S[n][j]$ spot prices, with $0 \leq j \leq n$.

Calling p the upward transition probability and $1 - p$ the downward transition probability, S the stock price today, and $S_u = uS$ and $S_d = dS$ upper and lower possible future spot prices, we can write the expectation equation[7]

$$\mathbf{E}[S] = puS + (1 - p)dS = e^{r\Delta t}S$$

which immediately gives us

$$p = \frac{a - d}{u - d}$$

with $a = \exp(r\Delta t)$.

We can also write the variance equation

$$Var[S] = pu^2S^2 + (1 - p)d^2S^2 - e^{2r\Delta t}S^2 \approx \sigma^2 S^2 \Delta t$$

which after choosing a centering condition such as $ud = 1$, will provide us with $u = \exp(\sigma\sqrt{\Delta t})$ and $d = \exp(-\sigma\sqrt{\Delta t})$.

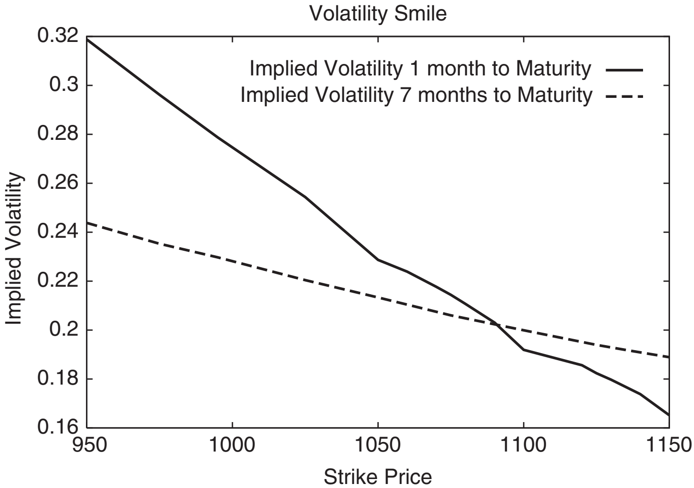
Using these values for u, d, and p, we can build the tree; and using the final payoff, we can calculate the option price by backward induction.[8] We can also build this tree by applying an Explicit Finite Difference scheme to the PDE (1.5) as was done in Wilmott [250]. An important advantage of the tree method is that it can be applied to American Options (with early exercise) as well.

It is possible to deduce the *implied* volatility of call and put options by solving a reverse Black-Scholes equation: that is, find the volatility that would equate the Black-Scholes price to the market price of the option.

This is a good way to see how derivatives markets *perceive* the underlying volatility. It is easy to see that if we change the maturity and strike prices of options (and keep everything else fixed), the implied volatility will *not* be constant. It will have a linear skew and a convex form as the strike price

[7] The expectation equation is written under the risk-neutral probability.
[8] For an in-depth discussion on Binomial Trees see Cox [72].

changes. This famous "smile" cannot be explained by simple time dependence, hence the necessity of introducing new models.[9]



SPX volatility smile on February 12, 2002, with Index = 1107.5 USD, one month and seven months to maturity. The negative skewness is clearly visible. Note how the smile becomes *flatter* as time to maturity increases.

JUMP DIFFUSION AND LEVEL-DEPENDENT VOLATILITY

In addition to the volatility smile observable from the implied volatilities of the options, there is evidence that the assumption of a pure normal distribution (also called *pure diffusion*) for the stock return is not accurate. Indeed "fat tails" have been observed away from the mean of the stock return. This phenomenon is called *leptokurticity* and could be explained in many different ways.

Jump Diffusion

Some try to explain the smile and the leptokurticity by changing the underlying stock distribution from a diffusion process to a jump-diffusion process.

[9] It is interesting to note that this smile phenomenon was practically nonexistent prior to the 1987 stock market crash. Many researchers therefore believe that the markets have learned to factor in a crash possibility, which creates the volatility smile.

A jump diffusion is *not* a level-dependent volatility process. However, we are mentioning it in this section to demonstrate the importance of the *leverage effect*.

Merton [200] was first to actually introduce jumps in the stock distribution. Kou [180] recently used the same idea to explain both the existence of fat tails and the volatility smile.

The stock price will follow a modified stochastic process under this assumption. If we add to the Brownian Motion dB_t a Poisson (jump) process[10] dq with an intensity[11] λ, then calling $k = \mathbf{E}(Y - 1)$ with $Y - 1$ the random variable percentage change in the stock price, we will have

$$dS_t = (\mu - \lambda k)S_t dt + \sigma S_t dB_t + S_t dq \tag{1.9}$$

or equivalently

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2} - \lambda k\right)t + \sigma B_t\right] Y_n$$

where $Y_0 = 1$ and $Y_n = \prod_{j=1}^{n} Y_j$ with Y_j's independently identically distributed random variables and n a Poisson random variable with a parameter λt.

It is worth noting that for the special case where the jump corresponds to total ruin or *default*, we have $k = -1$, which will give us

$$dS_t = (\mu + \lambda)S_t dt + \sigma S_t dB_t + S_t dq \tag{1.10}$$

and

$$S_t = S_0 \exp\left[\left(\mu + \lambda - \frac{\sigma^2}{2}\right)t + \sigma B_t\right] Y_n.$$

Given that in this case $\mathbf{E}(Y_n) = \mathbf{E}(Y_n^2) = e^{-\lambda t}$ it is fairly easy to see that in the risk-neutral world

$$\mathbf{E}(S_t) = S_0 e^{rt}$$

exactly as in the pure diffusion case, but

$$Var(S_t) = S_0^2 e^{2rt}(e^{(\sigma^2+\lambda)t} - 1) \approx S_0^2(\sigma^2 + \lambda)t \tag{1.11}$$

unlike the pure diffusion case where $Var(S_t) \approx S_0^2\sigma^2 t$.

---

[10]See for instance Karatzas [175].

[11]The intensity could be interpreted as the mean number of jumps per time unit.

Proof

Indeed

$$\mathbf{E}(S_t) = S_0 \exp((r+\lambda)t) \exp\left(-\frac{\sigma^2}{2}t\right) \mathbf{E}[\exp(\sigma B_t)]\mathbf{E}(Y_n)$$
$$= S_0 \exp((r+\lambda)t) \exp\left(-\frac{\sigma^2}{2}t\right) \exp\left(\frac{\sigma^2}{2}t\right) \exp(-\lambda t) = S_0 \exp(rt)$$

and

$$\mathbf{E}(S_t^2) = S_0^2 \exp(2(r+\lambda)t) \exp(-\sigma^2 t)\mathbf{E}[\exp(2\sigma B_t)]\mathbf{E}(Y_n^2)$$
$$= S_0^2 \exp(2(r+\lambda)t) \exp(-\sigma^2 t) \exp\left(\frac{(2\sigma)^2}{2}\right) \exp(-\lambda t)$$
$$= S_0^2 \exp((2r+\lambda)t) \exp(\sigma^2 t)$$

and as usual

$$Var(S_t) = \mathbf{E}(S_t^2) - \mathbf{E}^2(S_t)$$

(QED)

Link to Credit Spread Note that for a zero-coupon risky bond Z with no recovery, a credit spread C and a face value X paid at time t we have

$$Z = e^{-(r+C)t}X = e^{-\lambda t}(e^{-rt}X) + (1 - e^{-\lambda t})(0);$$

consequently, $\lambda = C$, and using (1.8) we can write

$$\tilde{\sigma}^2(C) = \sigma^2 + C$$

where σ is the fixed (pure diffusion) volatility and $\tilde{\sigma}$ is the modified jump diffusion volatility. This equation relates the volatility and *leverage*, a concept we will see later in level-dependent models as well.

Also, we could see that everything happens as if we were using the Black-Scholes pricing equation but with a modified "interest rate," which is $r + C$. Indeed, the Hedged Portfolio $\Pi = f - \frac{\partial f}{\partial S}S$ now satisfies

$$d\Pi = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}\right) dt$$

under the no-default case, which occurs with a probability of $e^{-\lambda dt} \approx 1 - \lambda dt$ and

$$d\Pi = -\Pi$$

under the default case, which occurs with a probability of $1 - e^{-\lambda dt} \approx \lambda dt$.

We therefore have

$$\mathbf{E}(d\Pi) = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} - \lambda\Pi\right) dt$$

and using a diversification argument we can always say that $\mathbf{E}(d\Pi) = r\Pi dt$, which provides us with

$$(r+\lambda)f = \frac{\partial f}{\partial t} + (r+\lambda)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} \tag{1.12}$$

which again is the Black-Scholes PDE with a "risky rate."

A generalization of the jump-diffusion process would be the use of the Levy process, a stochastic process with independent and stationary increments. Both the Brownian Motion and the Poisson process are included in this category. For a description, see Matacz [196].

Level-Dependent Volatility

Many assume that the smile and the fat tails are due to the level dependence of the volatility. The idea would be to make σ_t level dependent or a function of the spot itself; we would therefore have

$$dS_t = \mu_t S_t dt + \sigma(S,t) S_t dB_t. \tag{1.13}$$

Note that to be exact, a level-dependent volatility is a function of the spot price alone. When the volatility is a function of the spot price *and* time, it is referred to as local volatility, which we shall discuss further.

The Constant Elasticity Variance Approach One of the very first attempts to use this approach was the Constant Elasticity Variance (CEV) method realized by Cox [69] and [70]. In this method, we would suppose an equation of the type

$$\sigma(S,t) = CS_t^{\gamma} \tag{1.14}$$

where C and γ are parameters to be calibrated either from the stock price returns themselves or from the option prices and their implied volatilities. The CEV method was recently analyzed by Jones [173] in a paper where he uses two γ exponents.

This level-dependent volatility represents an important feature that is observed in options markets as well as in the underlying prices: the negative correlation between the stock price and the volatility, also called the *leverage effect*.

The Bensoussan Crouhy Galai Approach Bensoussan, Crouhy, and Galai (BCG) [34] try to find the level dependence of the volatility differently from Cox and Ross. Indeed, in the CEV model, Cox and Ross *first* suppose that $\sigma(S, t)$ has a certain exponential form and only then try to calibrate the model parameters to the market. BCG, on the other hand, try to deduce the functional form of $\sigma(S, t)$ by using a firm structure model.

The idea of firm structure is not new and goes back to Merton [199] where he considers that the Firm Assets follow a log-normal process

$$dV = \mu_V V dt + \sigma_V V dB_t \tag{1.15}$$

where μ_V and σ_V are the assets return and volatility. One important point is that σ_V is considered *constant*.

Merton then argues that the equity S of the firm could be considered a call option on the assets of the firm with a strike price K equal to the face value of the firm liabilities and an expiration T equal to the average liability maturity.

Using Ito's lemma, it is fairly easy to see that

$$dS = \mu S dt + \sigma(S, t) S dB_t = \left(\frac{\partial S}{\partial t} + \mu_V V \frac{\partial S}{\partial V} + \frac{1}{2}\sigma_V^2 V^2 \frac{\partial^2 S}{\partial V^2} \right) dt + \sigma_V V \frac{\partial S}{\partial V} dB_t \tag{1.16}$$

which immediately provides us with

$$\sigma(S, t) = \sigma_V \frac{V}{S} \frac{\partial S}{\partial V} \tag{1.17}$$

which is an implicit functional form for $\sigma(S, t)$.

BCG then eliminate the asset term in the above functional form and end up with a nonlinear PDE

$$\frac{\partial \sigma}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \sigma}{\partial S^2} + (r + \sigma^2) S \frac{\partial \sigma}{\partial S} = 0. \tag{1.18}$$

This PDE gives the dependence of σ on S and t.

Proof

A quick sketch of the proof is as follows: S being a contingent-claim on V we have the risk-neutral Black-Scholes PDE

$$\frac{\partial S}{\partial t} + rV \frac{\partial S}{\partial V} + \frac{1}{2}\sigma_V^2 V^2 \frac{\partial^2 S}{\partial V^2} = rS$$

and using $\frac{\partial S}{\partial V} = 1 / \frac{\partial V}{\partial S}$ as well as $\frac{\partial S}{\partial t} = -\frac{\partial S}{\partial V}\frac{\partial V}{\partial t}$ and $\frac{\partial^2 S}{\partial V^2} = -\frac{\partial^2 V}{\partial S^2} \Big/ \left(\frac{\partial V}{\partial S}\right)^3$ we have the reciprocal Black-Scholes equation

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV.$$

Now posing $\Psi(S,t) = \ln V(S,t)$, we have $\frac{\partial V}{\partial t} = V\frac{\partial \Psi}{\partial t}$ as well as $\frac{\partial V}{\partial S} = V\frac{\partial \Psi}{\partial S}$ and $\frac{\partial^2 V}{\partial S^2} = V\left(\frac{\partial^2 \Psi}{\partial S^2} + \left(\frac{\partial \Psi}{\partial S}\right)^2\right)$ and will have the new PDE

$$r = \frac{\partial \Psi}{\partial t} + rS\frac{\partial \Psi}{\partial S} + \frac{1}{2}\sigma^2 S^2\left(\frac{\partial^2 \Psi}{\partial S^2} + \left(\frac{\partial \Psi}{\partial S}\right)^2\right)$$

and the equation

$$\sigma = \sigma_V \Big/ \left(S\frac{\partial \Psi}{\partial S}\right).$$

This last identity implies $\frac{\partial \Psi}{\partial S} = \frac{\sigma_V}{S\sigma}$ as well as $\frac{\partial^2 \Psi}{\partial S^2} = \frac{-\sigma_V\left(\sigma + S\frac{\partial \sigma}{\partial S}\right)}{S^2\sigma^2}$ and therefore the PDE becomes

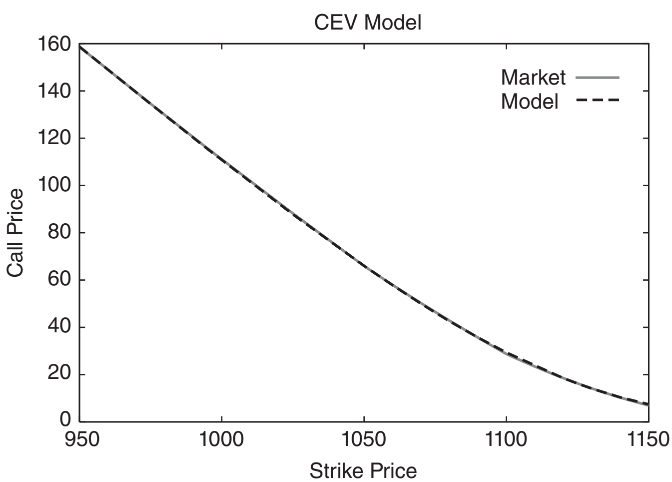
$$r = \frac{\partial \Psi}{\partial t} + r\sigma_V/\sigma + \frac{1}{2}\left(\sigma_V^2 - \sigma_V\left(\sigma + S\frac{\partial \sigma}{\partial S}\right)\right).$$

Taking the derivative with respect to S and using $\frac{\partial^2 \Psi}{\partial S \partial t} = -\frac{\sigma_V}{S\sigma^2}\frac{\partial \sigma}{\partial t}$, we get the final PDE

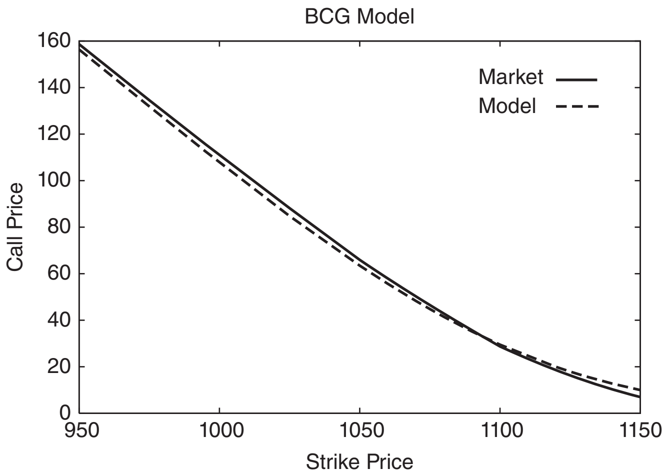
$$\frac{\partial \sigma}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 \sigma}{\partial S^2} + (r + \sigma^2)S\frac{\partial \sigma}{\partial S} = 0$$

as previously stated. *(QED)*

We therefore have an implicit functional form for $\sigma(S,t)$ and just like for the CEV case, we need to calibrate the parameters to the market data.



CEV model for SPX on February 12, 2002, with Index = 1107.5 USD, one month to maturity. The smile is fitted well, but the model assumes a perfect (negative) correlation between the stock and the volatility.



BCG model for SPX on February 12, 2002, with Index = 1107.5 USD, one month to maturity. The smile is fitted well.

LOCAL VOLATILITY

In the early 1990s, Dupire [94] on the one hand, and Derman & Kani [79] on the other, developed a concept called *local volatility* where the volatility smile was retrieved from the option prices.

The Dupire Approach

The Breeden & Litzenberger Identity This approach uses the options prices to get the implied distribution for the underlying stock. To do this we can write

$$V(S_0, K, T) = call(S_0, K, T) = e^{-rT} \int_0^{+\infty} (S - K)^+ p(S_0, S, T) dS \qquad (1.19)$$

where S_0 is the stock price at time $t = 0$ and K the strike price of the call, and $p(S_0, S, T)$ the *unknown* transition density for the stock price. As usual, $x^+ = MAX(x, 0)$.

Using the equation (1.19) and differentiating with respect to K twice, we get the Breeden & Litzenberger [47] implied distribution

$$p(S_0, K, T) = e^{rT} \frac{\partial^2 V}{\partial K^2} \qquad (1.20)$$

Proof

The proof is straightforward if we write

$$e^{rT}V(S_0, K, T) = \int_K^{+\infty} Sp(S_0, S, T)dS - K\int_K^{+\infty} p(S_0, S, T)dS$$

and take the first derivative

$$e^{rT}\frac{\partial V}{\partial K} = -Kp(S_0, K, T) + Kp(S_0, K, T) - \int_K^{+\infty} p(S_0, S, T)dS$$

and the second derivative in the same manner. *(QED)*

The Dupire Identity Now according to the Fokker-Planck (or forward Kolmogorov) equation[12] for this density, we have

$$\frac{\partial p}{\partial T} = \frac{1}{2}\frac{\partial^2(\sigma^2(S,t)S^2p)}{\partial S^2} - r\frac{\partial(Sp)}{\partial S}$$

and therefore after a little rearrangement

$$\frac{\partial V}{\partial T} = \frac{1}{2}\sigma^2K^2\frac{\partial^2 V}{\partial K^2} - rK\frac{\partial V}{\partial K}$$

which provides us with the local volatility formula

$$\sigma^2(K, T) = \frac{\frac{\partial V}{\partial T} + rK\frac{\partial V}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 V}{\partial K^2}}. \tag{1.21}$$

Proof

For a quick proof of this, let us use the zero interest rates case (the general case could be done similarly). We would then have

$$p(S_0, K, T) = \frac{\partial^2 V}{\partial K^2}$$

as well as Fokker-Planck

$$\frac{\partial p}{\partial T} = \frac{1}{2}\frac{\partial^2(\sigma^2(S,t)S^2p)}{\partial S^2}.$$

[12] See for example Wilmott [249] for an explanation on Fokker-Planck equation.

Now

$$\frac{\partial V}{\partial T} = \int_0^{+\infty} (S_T - K)^+ \frac{\partial p}{\partial T} dS_T$$
$$= \int_0^{+\infty} (S_T - K)^+ \frac{1}{2} \frac{\partial^2(\sigma^2(S,T)S^2 p)}{\partial S^2} dS_T$$

and integrating by parts twice and using the fact that

$$\frac{\partial^2 (S_T - K)^+}{\partial K^2} = \delta(S_T - K)$$

with $\delta(.)$ the Dirac function, we will have

$$\frac{\partial V}{\partial T} = \frac{1}{2}\sigma^2(K,T)K^2 p(S_0, K, T) = \frac{1}{2}K^2\sigma^2(K,T)\frac{\partial^2 V}{\partial K^2}$$

as stated. *(QED)*

It is also possible to use the implied volatility σ_{BS} from the Black-Scholes formula (1.5) and express the above local volatility in terms of σ_{BS} instead of V. For a detailed discussion, we could refer to Wilmott [249].

Local Volatility vs. Instantaneous Volatility Clearly the local volatility is related to the instantaneous variance v_t, as Gatheral [118] shows. The relationship could be written as

$$\sigma^2(K,T) = \mathbf{E}[v_T | S_T = K] \tag{1.22}$$

That is, local variance is the risk-neutral expectation of the instantaneous variance conditional on the final stock price being equal to the strike price.[13]

Proof

Let us show the above identity for the case of zero interest rates.[14]

As mentioned above, we have

$$\sigma^2(K,T) = \frac{\frac{\partial V}{\partial T}}{\frac{1}{2}K^2 \frac{\partial^2 V}{\partial K^2}}.$$

[13]Note that this is independent from the process for v_t, meaning that *any* stochastic volatility model satisfies this property, which is an attractive feature of local volatility models.

[14]For the case of non-zero rates we need to work with the forward price instead of the stock price.

On the other hand using the call payoff $V(S_0, K, t = T) = \mathbf{E}[(S_T - K)^+]$, we have

$$\frac{\partial V}{\partial K} = \mathbf{E}[H(S_T - K)]$$

with $H(.)$ the Heaviside function and

$$\frac{\partial^2 V}{\partial K^2} = \mathbf{E}[\delta(S_T - K)]$$

with $\delta(.)$ the Dirac function.

Therefore, the Ito lemma at $t = T$ would provide

$$d(S_T - K)^+ = H(S_T - K)dS_T + \frac{1}{2}v_T S_T^2 \delta(S_T - K)dT.$$

Using the fact that the forward price (here with zero interest rates, the stock price) is a martingale under the risk-neutral measure

$$dV = d\mathbf{E}[(S_T - K)^+] = \frac{1}{2}\mathbf{E}[v_T S_T^2 \delta(S_T - K)]dT,$$

now we have

$$\begin{aligned}\mathbf{E}[v_T S_T^2 \delta(S_T - K)] &= \mathbf{E}[v_T | S_T = K]K^2 \mathbf{E}[\delta(S_T - K)] \\ &= \mathbf{E}[v_T | S_T = K]K^2 \frac{\partial^2 V}{\partial K^2}.\end{aligned}$$

Putting all this together

$$\frac{\partial V}{\partial T} = \frac{1}{2}K^2 \frac{\partial^2 V}{\partial K^2}\mathbf{E}[v_T | S_T = K]$$

and by the above expression of $\sigma^2(K, T)$, we will have

$$\sigma^2(K, T) = \mathbf{E}[v_T | S_T = K]$$

as claimed. *(QED)*

The Derman Kani Approach

The Derman Kani technique is very similar to the Dupire approach, except it uses the Binomial (or Trinomial) Tree framework instead of the continuous one.

Using the Binomial Tree notations, their upward transition probability p_i from the spot s_i at time t_n to the upper node S_{i+1} at the following time-step t_{n+1} is obtained from the usual

$$p_i = \frac{F_i - S_i}{S_{i+1} - S_i} \tag{1.23}$$

where F_i is the stock forward price known from the market, and S_i the lower spot at the step t_{n+1}.

In addition, we have for a call expiring at time-step t_{n+1}

$$C(K, t_{n+1}) = e^{-r\Delta t}\sum_{j=1}^{n}[\lambda_j p_j + \lambda_{j+1}(1 - p_{j+1})]MAX(S_{j+1} - K, 0)$$

where λ_j's are the known Arrow-Debreu prices corresponding to the discounted probability of getting to the point s_j at time t_n from S_0, the initial stock price. These probabilities could easily be derived iteratively.

This allows us after some calculation to obtain S_{i+1} as a function of s_i and S_i, namely

$$S_{i+1} = \frac{S_i[e^{r\Delta t}C(s_i, K, t_{n+1}) - \Sigma] - \lambda_i s_i(F_i - S_i)}{[e^{r\Delta t}C(s_i, K, t_{n+1}) - \Sigma] - \lambda_i(F_i - S_i)}$$

where the term Σ represents the sum $\sum_{j=i+1}^{n}\lambda_j(F_j - s_i)$. This means that after choosing the usual centering condition for the Binomial Tree

$$s_i^2 = S_i S_{i+1},$$

we have all the elements to build the tree and deduce the implied distribution from the Arrow-Debreu prices.

Stability Issues

The local volatility models are very elegant and theoretically sound; however, they present in practice many stability issues. They are *Ill-Posed Inversion* problems and are extremely sensitive to the input data.[15] This might introduce arbitrage opportunities and in some cases negative probabilities or variances. Derman and Kani suggest overwriting techniques to avoid such problems.

---

[15] See Tavella [237] or Avellaneda [17].

Andersen [14] tries to improve this issue by using an Implicit Finite Difference method. However, he recognizes that the negative variance problem could still happen.

One way to make the results smoother is to use a constrained optimization.

In other words, when trying to fit theoretical results C_{theo} to the market prices C_{mrkt}, instead of minimizing

$$\sum_{j=1}^{N} (C_{theo}(K_j) - C_{mrkt}(K_j))^2$$

we could minimize

$$\lambda \frac{\partial \sigma}{\partial t} + \sum_{j=1}^{N} (C_{theo}(K_j) - C_{mrkt}(K_j))^2$$

where λ is a constraint parameter, which could also be interpreted as a Lagrange multiplier.

However, this is an artificial way to smoothen the results and the real issue remains that once again, we have an inversion problem that is inherently unstable.

What is more, local volatility models imply that future implied volatility smiles will be flat relative to today's, which is another limitation.[16] As we will see in the following section, stochastic volatility models offer more time-homogeneous volatility smiles.

An alternative approach suggested in [17] would be to choose a prior risk-neutral distribution for the asset (based on a subjective view) and then minimize the relative Entropy distance between the desired surface and this prior distribution. This approach uses the Kullback-Leibler distance (which we will discuss in the context of MLE) and performs the minimization via Dynamic Programming [37] on a tree.

Calibration Frequency

One of the most attractive features of local-vol models is their ability to match plain-vanilla puts and calls *exactly*. This will avoid arbitrage situations, or worse, market manipulations by traders to create "phantom" profits.

[16] See Gatheral [119].

As explained in Hull [154], these arbitrage-free models were developed by researchers with a Single Calibration (SC) methodology assumption. However, traders use them with a Continual Recalibration (CR) strategy in practice. Indeed, if they used the SC version of the model, significant errors would be introduced from one week to the following, as shown by Dumas et al. [93].

However, once this CR version is used, there is no guarantee that the no-arbitrage property of the original SC model is preserved. Indeed, the Dupire equation determines the marginal stock distribution at different points in time, but not the joint distribution of these stock prices. Therefore, a path-dependent option could very well be mispriced, and the more path-dependent this option, the greater the mispricing.

Hull [154] takes the example of a Bet Option, a Compound Option, and a Barrier Option. The Bet Option depends on the distribution of the stock at one point in time and therefore is correctly priced with a Continually Recalibrated Local-Vol model. The Compound Option has some path dependency, hence a certain amount of mispricing compared to a stochastic volatility (SV) model. Finally, the Barrier Option has a strong degree of path dependency and will introduce large errors.

Note that this is due to the discrete nature of the data. Indeed, the maturities we have are limited. If we had all possible maturities in a continuous way, the joint distribution would be determined completely.

Also, when interpolating in time, it is customary to interpolate upon the true variance $t\sigma_t^2$ rather than the volatility σ_t given the equation

$$T_2\sigma^2(T_2) = T_1\sigma^2(T_1) + (T_2 - T_1)\sigma^2(T_1, T_2).$$

Interpolating upon the true variance will provide smoother results as shown by Jackel [159].

Proof

Indeed, calling for $0 \leq T_1 \leq T_2$ the spot return variances

$$Var(0, T_2) = T_2\sigma^2(T_2)$$
$$Var(0, T_1) = T_1\sigma^2(T_1)$$

for a Brownian Motion, we have independent increments and therefore a forward variance $Var(T_1, T_2)$ such that

$$Var(0, T_1) + Var(T_1, T_2) = Var(0, T_2)$$

which demonstrates the point. *(QED)*

STOCHASTIC VOLATILITY

Unlike non-parametric local volatility models, parametric stochastic volatility (SV) models define a specific stochastic differential equation for the unobservable instantaneous variance. As we shall see, the previously defined CEV model could be considered a special case of these models.

Stochastic Volatility Processes

The idea would be to use a different stochastic process for σ altogether. Making the volatility a deterministic function of the spot is a special "degenerate" two-factor, a natural generalization of which would precisely be to have two stochastic processes with a non-perfect correlation.[17]

Several different stochastic processes have been suggested for the volatility. One popular one is the *Ornstein-Uhlenbeck* (OU) process:

$$d\sigma_t = -\alpha\sigma_t dt + \beta dZ_t \tag{1.24}$$

where α and β are two parameters; remembering the stock equation

$$dS_t = \mu_t S_t dt + \sigma_t S_t dB_t,$$

there is a (usually negative) correlation ρ between dZ_t and dB_t which can in turn be time or level dependent.

Heston [141] and Stein [234] were among those who suggested the use of this process. Using Ito's lemma, we can see that the stock-return variance $v_t = \sigma_t^2$ satisfies a *Square-Root* or Cox-Ingersoll-Ross (CIR) process

$$dv_t = (\omega - \theta v_t)dt + \xi\sqrt{v_t}dZ_t \tag{1.25}$$

with $\omega = \beta^2$, $\theta = 2\alpha$ and $\xi = 2\beta$.

Note that the OU process has a closed-form solution

$$\sigma_t = \sigma_0 e^{-\alpha t} + \beta\int_0^t e^{-\alpha(t-s)}dZ_s$$

which means that σ_t follows in law $\Phi(\sigma_0 e^{-\alpha t}, \frac{\beta^2}{2\alpha}(1 - e^{-2\alpha t}))$, with Φ again the normal distribution. This was discussed in Fouque [109] and Shreve [229].

[17]Note that here, the *instantaneous* volatility is stochastic. Recent work by researchers such as Schonbucher supposes a Stochastic Implied-Volatility process, which is a completely different approach. See for instance [224].

On the other hand, Avellaneda *et al.* [18] use the concept of *uncertain volatility* for pricing and hedging derivative securities. They make the volatility switch between two extreme values based on the convexity of the derivative contract and obtain a nonlinear *Black-Scholes-Barenblatt* equation, which they solve on a grid.

Heston and Nandi [144] show that this process corresponds to a special case of the General Auto-Regressive Conditional Heteroskedasticity (GARCH) model that we will discuss further in the next section.

Another popular process is the GARCH(1,1) process, where we would have

$$dv_t = (\omega - \theta v_t)dt + \xi v_t dZ_t. \tag{1.26}$$

GARCH and Diffusion Limits

The most elementary GARCH process called GARCH(1,1) was developed originally in the field of econometrics by Engle [99] and Bollerslev [42] in a *discrete* framework. The stock discrete equation (1.3) could be rewritten by taking $\Delta t = 1$ and $v_n = \sigma_n^2$ as

$$\ln S_{n+1} = \ln S_n + \left(\mu - \frac{1}{2}v_{n+1}\right) + \sqrt{v_{n+1}}B_{n+1} \tag{1.27}$$

and calling the mean adjusted return

$$u_n = \ln\left(\frac{S_n}{S_{n-1}}\right) - \left(\mu - \frac{1}{2}v_n\right) = \sqrt{v_n}B_n. \tag{1.28}$$

The variance process in GARCH(1,1) is supposed to be

$$v_{n+1} = \omega_0 + \beta v_n + \alpha u_n^2 = \omega_0 + \beta v_n + \alpha v_n B_n^2 \tag{1.29}$$

where α and β are weight parameters and ω_0 a parameter related to the long-term variance.[18]

Nelson [204] shows that as the time interval length decreases and becomes infinitesimal, equation (1.29) becomes precisely the previously cited equation (1.26). To be more accurate, there is a *weak convergence* of the discrete GARCH process to the continuous diffusion limit.[19] For a GARCH(1,1) continuous diffusion, the correlation between dZ_t and dB_t is zero.

---

[18]It is worth mentioning that as explained in [105] a GARCH(1,1) model could be rewritten as an Auto-Regressive Moving Average model of first order ARMA(1,1) and therefore an Auto-Regressive model of infinite order AR($+\infty$).

GARCH is therefore a parsimonious model that can fit the data with only a few parameters. Fitting the same data with an ARCH or AR model would require a much larger number of parameters. This feature makes the GARCH model very attractive.

[19]For an explanation on weak convergence see for example Varadhan [241].

It might appear surprising that even if the GARCH(1,1) process has only *one* source of randomness, namely B_n, the continuous version has two independent Brownian Motions. This is understandable if we consider B_n a standard normal random variable and $A_n = B_n^2 - 1$ another random variable. It is fairly easy to see that A_n and B_n are uncorrelated even if A_n is a function of B_n. As we go toward the continuous version, we can use Donsker's theorem,[20] by letting the time interval $\Delta t \to 0$, to prove that we end up with two uncorrelated and therefore independent Brownian Motions. This is a limitation of the GARCH(1,1) model, hence the introduction of the Nonlinear Asymmetric GARCH (NGARCH) model.

Duan [88] attempts to explain the volatility smile by using the NGARCH process expressed by

$$v_{n+1} = \omega_0 + \beta v_n + \alpha(u_n - c\sqrt{v_n})^2 \tag{1.30}$$

where c is a parameter to be determined.

The NGARCH process was first introduced by Engle [102]. The continuous counterpart of NGARCH is the same equation (1.26), except unlike the equation resulting from GARCH(1,1) there *is* a non-zero correlation between the stock process and the volatility process. This correlation is created precisely because of the parameter c that was introduced, and is once again called the leverage effect. The parameter c is sometimes referred to as the leverage parameter.

We can find the following relationships between the discrete process and the continuous diffusion limit parameters by letting the time interval become infinitesimal

$$\omega = \frac{\omega_0}{dt^2}$$

$$\theta = \frac{1 - \alpha(1 + c^2) - \beta}{dt}$$

$$\xi = \alpha\sqrt{\frac{\kappa - 1 + 4c^2}{dt}}$$

and the correlation between dB_t and dZ_t

$$\rho = \frac{-2c}{\sqrt{\kappa - 1 + 4c^2}}$$

[20]For a discussion on Donsker's theorem, similar to the central limit theorem, see for instance Whitt [247].

where κ represents the Pearson kurtosis[21] of the mean adjusted returns (u_n). As we can see, the sign of the correlation ρ is determined by the parameter c.

Proof

A quick proof of the convergence to diffusion limit could be outlined as follows. Let us assume that $c = 0$ for simplicity, we therefore are dealing with the GARCH(1,1) case.

As we saw

$$v_{n+1} = \omega_0 + \beta v_n + \alpha v_n B_n^2$$

therefore

$$v_{n+1} - v_n = \omega_0 + \beta v_n - v_n + \alpha v_n - \alpha v_n + \alpha v_n B_n^2$$

or

$$v_{n+1} - v_n = \omega_0 - (1 - \alpha - \beta)v_n + \alpha v_n(B_n^2 - 1).$$

Now, allowing the time-step Δt to become variable and posing $Z_n = (B_n^2 - 1)/\sqrt{\kappa - 1}$

$$v_{n+\Delta t} - v_n = \omega\Delta t^2 - \theta\Delta t v_n + \xi v_n\sqrt{\Delta t}Z_n$$

and annualizing v_n by posing $v_t = v_n/\Delta t$ we shall have

$$v_{t+\Delta t} - v_t = \omega\Delta t - \theta\Delta t v_t + \xi v_t\sqrt{\Delta t}Z_n$$

and as $\Delta t \to 0$ we get

$$dv_t = (\omega - \theta v_t)dt + \xi v_t dZ_t$$

as claimed. *(QED)*

Note that the discrete GARCH version of the Square-Root process (1.15) is

$$v_{n+1} = \omega_0 + \beta v_n + \alpha(B_n - c\sqrt{v_n})^2 \tag{1.31}$$

as Heston and Nandi show[22] in [144].

Also, note that having a diffusion process $dv_t = b(v_t)dt + a(v_t)dZ_t$ we can apply an Euler approximation[23] to discretize and obtain a Monte-Carlo process such as $v_{n+1} - v_n = b(v_n)\Delta t + a(v_n)\sqrt{\Delta t}Z_n$. It is important to note that if we use a GARCH process and go to the continuous diffusion limit, and

[21]The kurtosis corresponds to the fourth moment. The Pearson kurtosis for a normal distribution is equal to 3.

[22]For a detailed discussion on the convergence of different GARCH models toward their diffusion limits, also see Duan [90].

[23]See for instance Jones [173].

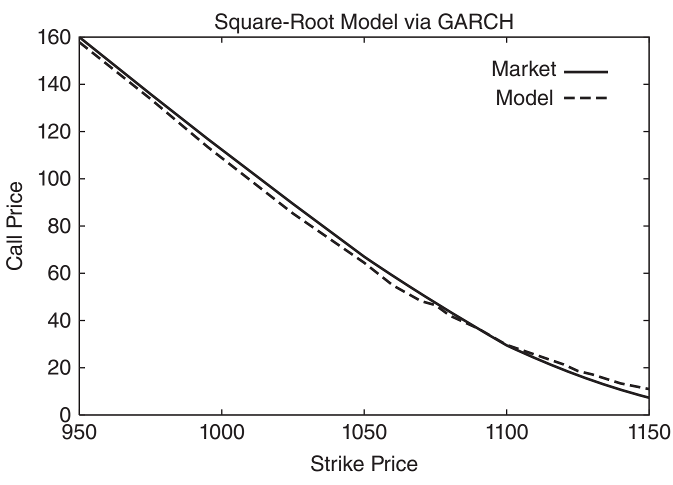
then apply an Euler Approximation, we will *not necessarily* find the original GARCH process again. Indeed, there are many different ways to discretize the continuous diffusion limit and the GARCH process corresponds to one special way.

In particular, if we use (1.31) and allow $\Delta t \to 0$ to get to the continuous diffusion limit, we shall obtain (1.25). As we will see later in the section on *Mixing Solutions*, we can then apply a discretization to this process and obtain a Monte-Carlo simulation

$$v_{n+1} = v_n + (\omega - \theta v_n)\Delta t + \xi\sqrt{v_n}\sqrt{\Delta t}Z_n$$

which is again different from (1.31) but obviously has to be consistent in terms of pricing. However, we should know which method we are working with from the very beginning to perform our calibration on the corresponding specific process.

Corradi [66] explains this in the following manner: The discrete GARCH model could converge either toward a two-factor continuous limit if one chooses the Nelson parameterization, or could very well converge to a one-factor diffusion limit if one chooses another parameterization. What is more, an appropriate Euler discretization of the one-factor continuous model will provide a GARCH discrete process, while as previously mentioned the discretization of the two-factor diffusion model provides a two-factor discrete process. This distinction is fundamental and could explain why GARCH and SV behave differently in terms of parameter estimation.



GARCH Monte-Carlo simulation with the Square-Root model for SPX on February 12, 2002, with Index = 1107.5 USD, one month to maturity. Powell Optimization method was used for Least-Square Calibration.

THE PRICING PDE UNDER STOCHASTIC VOLATILITY

A very important issue to underline here is that, because of the unhedgeable second source of randomness, the concept of Market Completeness is lost. We can no longer have a straightforward risk neutral pricing. This is where the *market price of risk* will come into consideration.

The Market Price of Volatility Risk

Indeed, taking a more general form for the variance process

$$dv_t = b(v_t)dt + a(v_t)dZ_t \tag{1.32}$$

as we previously said, using the Black-Scholes risk-neutrality argument, the equation (1.1) could be replaced with

$$dS_t = (r_t - q_t)S_t dt + \sigma_t S_t dB_t. \tag{1.33}$$

This is equivalent to changing the probability measure from the real one to the *risk-neutral* one.[24] We therefore need to use (1.33) together with the risk adjusted volatility process

$$dv_t = \tilde{b}(v_t)dt + a(v_t)dZ_t \tag{1.34}$$

where

$$\tilde{b}(v_t) = b(v_t) - \lambda a(v_t)$$

with λ the market price of volatility risk. This quantity is closely related to the market price of risk for the stock $\lambda_e = (\mu - r)/\sigma$. Indeed, as Hobson [147] and Lewis [185] both show, we have

$$\lambda = \rho\lambda_e + \sqrt{1-\rho^2}\lambda^* \tag{1.35}$$

where λ^* is the market price of risk associated with $dB_t - \rho dZ_t$, which can also be regarded as the market price of risk for the hedged portfolio.

The passage from equation (1.32) to (1.34) and the introduction of the market price of volatility risk could also be explained by the Girsanov theorem as was done for instance in Fouque [109].

[24] See Hull [153] or Shreve [229] for more detail.

It is important to underline the difference between the real and the risk-neutral measures here. If we use historic stock prices together with the real stock-return drift μ to estimate the process parameters, we will obtain the real volatility drift $b(v)$. An alternative method would be to estimate $\tilde{b}(v)$ by using current option prices and performing a least square estimation. These calibration methods will be discussed in more detail in the following chapters.

The risk-neutral version for a discrete NGARCH model would also involve the market price of risk and instead of the usual

$$\ln S_{n+1} = \ln S_n + \left(\mu - \frac{1}{2}v_{n+1}\right) + \sqrt{v_{n+1}}B_{n+1}$$
$$v_{n+1} = \omega_0 + \beta v_n + \alpha v_n (B_n - c)^2,$$

we would have

$$\ln S_{n+1} = \ln S_n + \left(r - \frac{1}{2}v_{n+1}\right) + \sqrt{v_{n+1}}\tilde{B}_{n+1} \tag{1.36}$$
$$v_{n+1} = \omega_0 + \beta v_n + \alpha v_n (\tilde{B}_n - c - \lambda_e)^2$$

where $\tilde{B}_n = B_n + \lambda_e$ which could be regarded as the discrete version of the Girsanov theorem. Note that the market price of risk for the stock λ_e is *not* separable from the leverage parameter c in this formulation.

Duan shows in [89] and [91] that the risk-neutral GARCH system (1.36) will indeed converge toward the continuous risk-neutral GARCH

$$dS_t = S_t r dt + S_t\sqrt{v_t}dB_t$$
$$dv_t = (\omega - \tilde{\theta} v_t)dt + \xi v_t dZ_t$$

as we expected.

The Two-Factor PDE

From here, writing a two-factor PDE for a derivative security f becomes a simple application of the two-dimensional Ito's lemma. The PDE will be[25]

$$rf = \frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + \frac{1}{2}vS^2\frac{\partial^2 f}{\partial S^2} + \tilde{b}(v)\frac{\partial f}{\partial v} + \frac{1}{2}a^2(v)\frac{\partial^2 f}{\partial v^2} + \rho a(v)\sqrt{v}S\frac{\partial^2 f}{\partial S \partial v} \tag{1.37}$$

[25]For a proof of the derivation see Wilmott [249] or Lewis [185].

Therefore, it is possible, after calibration, to apply a Finite Difference method[26] to the above PDE to price the derivative $f(S, t, v)$. An alternative would be to use directly the stochastic processes for dS_t and dv_t and apply a two-factor Monte-Carlo simulation. Later in the chapter we will also mention other possible methods such as the Mixing Solution or Asymptotic Approximations.

Other possible approaches for incomplete markets and stochastic volatility assumption include *Super-Replication* and *Local Risk Minimization*.[27]

The Super-Replication strategy is the cheapest self-financing strategy with a terminal value no less than the payoff of the derivative contract. This technique was primarily developed by El-Karoui and Quenez in [96].

Local Risk Minimization involves a partial hedging of the risk. The risk is reduced to an "intrinsic component" by taking an offsetting position in the underlying security as usual. This method was developed by Follmer and Sondermann in [107].

THE GENERALIZED FOURIER TRANSFORM

The Transform Technique

One useful technique to apply to the PDE (1.37) is the *Generalized Fourier Transform*.[28] First, we can use the variable $x = \ln S$ in which case, using Ito's lemma, (1.37) could be rewritten as

$$rf = \frac{\partial f}{\partial t} + \left(r - q - \frac{1}{2}v\right)\frac{\partial f}{\partial x} + \frac{1}{2}v\frac{\partial^2 f}{\partial x^2} + \tilde{b}(v)\frac{\partial f}{\partial v} + \frac{1}{2}a^2(v)\frac{\partial^2 f}{\partial v^2} + \rho a(v)\sqrt{v}\frac{\partial^2 f}{\partial x \partial v} \tag{1.38}$$

Calling

$$\hat{f}(k, v, t) = \int_{-\infty}^{+\infty} e^{ikx} f(x, v, t)dx \tag{1.39}$$

where k is a *complex* number,[29] $\hat{f}$ will be defined in a complex *strip* where the imaginary part of k is between two real numbers α and β.

[26]See for instance Tavella [238] or Wilmott [249] for a discussion on Finite Difference Methods.
[27]For a discussion on both these techniques see Frey [112].
[28]See Lewis [185] for a detailed discussion on this technique.
[29]As usual we note $i = \sqrt{-1}$.

Once $\hat{f}$ is suitably defined, meaning that $k_i = \mathcal{I}(k)$ (the imaginary part of k) is within the appropriate strip, we can write the Inverse Fourier Transform

$$f(x, v, t) = \frac{1}{2\pi} \int_{ik_i-\infty}^{ik_i+\infty} e^{-ikx} \hat{f}(k, v, t) dk \tag{1.40}$$

where we are integrating for a *fixed* k_i parallel to the real axis.

Each derivative satisfying (1.37) or equivalently (1.38) has a known payoff $G(S_T)$ at maturity. For instance, as we said before, a call option has a payoff $MAX(0, S_T - K)$ where K is the call strike price. It is easy to see that for $k_i > 1$ the Fourier Transform of a call option exists and the payoff transform is

$$-\frac{K^{ik+1}}{k^2 - ik} \tag{1.41}$$

Proof

Indeed, we can write

$$\begin{aligned}
\int_{-\infty}^{+\infty} e^{ikx}(e^x - K)^+ dx &= \int_{\ln K}^{+\infty} e^{ikx}(e^x - K) dx \\
&= 0 - \left(\frac{K^{ik+1}}{ik+1} - K\frac{K^{ik}}{ik} \right) \\
&= -K^{ik+1} \left(\frac{1}{ik+1} - \frac{1}{ik} \right) = -K^{ik+1} \frac{1}{k^2 - ik}
\end{aligned}$$

as stated. *(QED)*

The same could be applied to a put option or other derivative securities. In particular, a covered call (stock minus call) having a payoff $MIN(S_T, K)$ will have a transform for $0 < k_i < 1$ equal to

$$\frac{K^{ik+1}}{k^2 - ik} \tag{1.42}$$

Applying the transform to the PDE (1.38) and introducing $\tau = T - t$ and

$$\hat{h}(k, v, \tau) = e^{(r+ik(r-q))\tau} \hat{f}(k, v, \tau) \tag{1.43}$$

and posing[30] $c(k) = \frac{1}{2}(k^2 - ik)$, we get the new PDE equation

$$\frac{\partial \hat{h}}{\partial \tau} = \frac{1}{2} a^2(v) \frac{\partial^2 \hat{h}}{\partial v^2} + (\tilde{b}(v) - ik\rho(v)a(v)\sqrt{v}) \frac{\partial \hat{h}}{\partial v} - c(k)v\hat{h}. \tag{1.44}$$

[30] We are following Lewis [185] notations.

Lewis calls the *Fundamental Transform* a function $\hat{H}(k, v, \tau)$ satisfying the PDE (1.44) and satisfying the initial condition $\hat{H}(k, v, \tau = 0) = 1$.

If we know this Fundamental Transform, we can then multiply it by the derivative security's payoff transform and then divide it by $e^{(r+ik(r-q))\tau}$ and apply the inverse Fourier technique by keeping k_i in an appropriate strip and finally get the derivative as a function of $x = \ln S$.

Special Cases

There are cases where the Fundamental Transform is known. The case of a constant (or deterministic) volatility is the most elementary one. Indeed, using (1.44) together with $dv_t = 0$ we can easily find

$$\hat{H}(k, v, \tau) = e^{-c(k)v\tau}$$

which is analytic in k over the entire complex plane. Using the call payoff transform (1.41), we can rederive the Black-Scholes equation. The same can be done if we have a deterministic volatility $dv_t = b(v_t)dt$ by using the function $Y(v, t)$ where $dY = b(Y)dt$.

The Square-Root model (1.25) is another important case where $\hat{H}(k, v, \tau)$ is known and analytic. We have for this process

$$dv_t = (\omega - \theta v_t)dt + \xi\sqrt{v_t}dZ_t$$

or under the risk-neutral measure

$$dv_t = (\omega - \tilde{\theta} v_t)dt + \xi\sqrt{v_t}dZ_t$$

with $\tilde{\theta} = (1-\gamma)\rho\xi + \sqrt{\theta^2 - \gamma(1-\gamma)\xi^2}$ where $\gamma \leq 1$ represents the risk-aversion factor.

For the Fundamental Transform, we get

$$\hat{H}(k, v, \tau) = \exp[f_1(t) + f_2(t)v] \tag{1.45}$$

with

$$t = \frac{1}{2}\xi^2\tau \quad \tilde{\omega} = \frac{2}{\xi^2}\omega \quad \tilde{c} = \frac{2}{\xi^2}c(k) \quad \text{and}$$

$$f_1(t) = \left[tg - \ln\left(\frac{1 - he^{td}}{1 - h}\right)\right]\tilde{\omega}$$

$$f_2(t) = \left[\frac{1 - e^{td}}{1 - he^{td}}\right]g$$

where

$$d = \sqrt{\bar{\theta}^2 + 4\tilde{c}} \qquad g = \frac{1}{2}(\bar{\theta} + d) \qquad h = \frac{\bar{\theta} + d}{\bar{\theta} - d} \qquad \text{and}$$

$$\bar{\theta} = \frac{2}{\xi^2}[(1 - \gamma + ik)\rho\xi + \sqrt{\theta^2 - \gamma(1 - \gamma)\xi^2}]$$

This transform has a cumbersome expression, but it can be seen that it is analytic in k and therefore always exists. For a proof refer to Lewis [185].

The Inversion of the Fourier Transform for the Square-Root (Heston) model is a very popular and powerful approach. It is appealing due to its robustness and speed.

The following example is based on SPX options as of March 9, 2004, expiring in one to eight years from the calibration date.

SPX implied surface as of 03/09/2004. T is the maturity and $M = K/S$ the inverse of the moneyness.

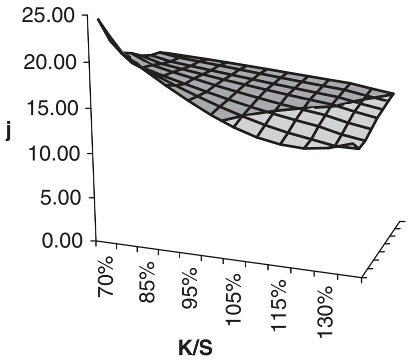
T / M	0.70	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.30
1.000	24.61	21.29	19.73	18.21	16.81	15.51	14.43	13.61	13.12	12.94	13.23
2.000	21.94	18.73	18.68	17.65	16.69	15.79	14.98	14.26	13.67	13.22	12.75
3.000	20.16	18.69	17.96	17.28	16.61	15.97	15.39	14.86	14.38	13.96	13.30
4.000	19.64	18.48	17.87	17.33	16.78	16.26	15.78	15.33	14.92	14.53	13.93
5.000	18.89	18.12	17.70	17.29	16.88	16.50	16.13	15.77	15.42	15.11	14.54
6.000	18.46	17.90	17.56	17.23	16.90	16.57	16.25	15.94	15.64	15.35	14.83
7.000	18.32	17.86	17.59	17.30	17.00	16.71	16.43	16.15	15.88	15.62	15.15
8.000	17.73	17.54	17.37	17.17	16.95	16.72	16.50	16.27	16.04	15.82	15.40

Heston prices fitted to the March 9, 2004, surface.

T / M	0.70	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.30
1.000	30.67	21.44	17.09	13.01	9.33	6.18	3.72	2.03	1.03	0.50	0.13
2.000	31.60	22.98	18.98	15.25	11.87	8.89	6.37	4.35	2.83	1.78	0.66
3.000	32.31	24.18	20.44	16.98	13.82	11.00	8.55	6.47	4.77	3.43	1.66
4.000	33.21	25.48	21.93	18.66	15.63	12.91	10.50	8.39	6.61	5.10	2.93
5.000	33.87	26.54	23.20	20.09	17.22	14.63	12.30	10.21	8.39	6.82	4.36
6.000	34.56	27.55	24.34	21.36	18.60	16.08	13.79	11.73	9.89	8.26	5.64
7.000	35.35	28.61	25.52	22.64	19.96	17.49	15.24	13.19	11.35	9.70	6.97
8.000	35.77	29.34	26.39	23.64	21.07	18.69	16.51	14.51	12.68	11.04	8.24

As we shall see further, the optimal Heston parameter-set to fit this surface could be found via a Least Square Estimation approach, and for the index at $S = 1156.86$ USD we find the optimal parameters $\hat{v}_0 = 0.1940$ and

$$\hat{\Psi} = (\hat{\omega}, \hat{\theta}, \hat{\xi}, \hat{\rho}) = (0.052042332, 1.8408, 0.4710, -0.4677).$$



SPX implied surface as of March 9, 2004. We can observe the negative skewness as well as the flattening of the slope with maturity.

THE MIXING SOLUTION

The Romano Touzi Approach

The idea of *Mixing Solutions* was probably presented for the first time by Hull and White [156] for a zero correlation case. Later, Romano and Touzi [220] generalized this approach for a correlated case.

The basic idea is to *separate* the random processes of the stock and the volatility, integrate the stock process conditionally on a given volatility, and finally end up with a one-factor problem.

Let us recall the two processes we had:

$$dS_t = (r_t - q_t)S_t dt + \sigma_t S_t dB_t$$

and

$$dv_t = \tilde{b}(v_t)dt + a(v_t)dZ_t$$

under a risk-neutral measure.

Given a correlation ρ_t between dB_t and dZ_t, we can introduce the Brownian Motion dW_t independent of dZ_t and write the usual Cholesky[31] factorization:

$$dB_t = \rho_t dZ_t + \sqrt{1-\rho_t^2}\,dW_t$$

We can then introduce the same $X_t = \ln S_t$ and write the new system of equations:

$$dX_t = (r-q)dt + dY_t - \frac{1}{2}(1-\rho_t^2)\sigma_t^2 dt + \sqrt{1-\rho_t^2}\,\sigma_t dW_t \tag{1.46}$$
$$dY_t = -\frac{1}{2}\rho_t^2\sigma_t^2 dt + \rho_t\sigma_t dZ_t$$
$$dv_t = \tilde{b}_t dt + a_t dZ_t$$

where once again, the two Brownian Motions are independent.

It is now possible to integrate the stock process for a given volatility and end up with an expectation on the volatility process only. We can think of (1.46) as the limit of a discrete process, while the time step $\Delta t \to 0$.

For a derivative security $f(S_0, v_0, T)$ with a payoff[32] $G(S_T)$, using the bivariate normal density for two uncorrelated variables, we can write

$$\begin{aligned} f(S_0, v_0, T) &= e^{-rT}\mathbf{E}_0[G(S_T)] \\ &= e^{-rT}\lim_{\Delta t\to 0}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} G(S_T) \\ &\quad\times\prod_{t=0}^{T-\Delta t}\exp\left[-\frac{1}{2}(Z_t^2+W_t^2)\right]\frac{dZ_t dW_t}{2\pi}. \end{aligned} \tag{1.47}$$

Now if we know how to integrate this over dW_t for a given volatility and we know the result $f^*(S, v, T)$ (for instance, for a European call option, we know the Black-Scholes formula (1.6), there are many other cases where we have closed-form solutions), then we can introduce the auxiliary variables[33]

$$S^{eff} = S_0 e^{Y_T} = S_0 \exp\left(-\frac{1}{2}\int_0^T \rho_t^2\sigma_t^2 dt + \int_0^T \rho_t\sigma_t dZ_t\right) \tag{1.48}$$

[31] See for example Press [214].
[32] The payoff should *not* depend on the volatility process.
[33] Again, all notations are taken from Lewis [185].

and

$$v^{eff} = \frac{1}{T}\int_0^T (1-\rho_t^2)\sigma_t^2 dt \tag{1.49}$$

and as Romano and Touzi prove in [220] we will have

$$f(S_0, v_0, T) = \mathbf{E}_0[f^*(S^{eff}, v^{eff}, T)] \tag{1.50}$$

where this last expectation is being taken on dZ_t only.

Note that in the zero correlation case discussed by Hull and White [156] we have $S^{eff} = S_0$ and $v^{eff} = v_T = \frac{1}{T}\int_0^T \sigma_t^2 dt$, which makes the expression (1.50) a natural weighted average.

A One-Factor Monte-Carlo Technique

As Lewis suggests, this will enable us to run a single-factor Monte-Carlo simulation on the dZ_t and apply the known closed-form for each simulated path. The method does suppose however that the payoff $G(S_T)$ does *not* depend on the volatility.

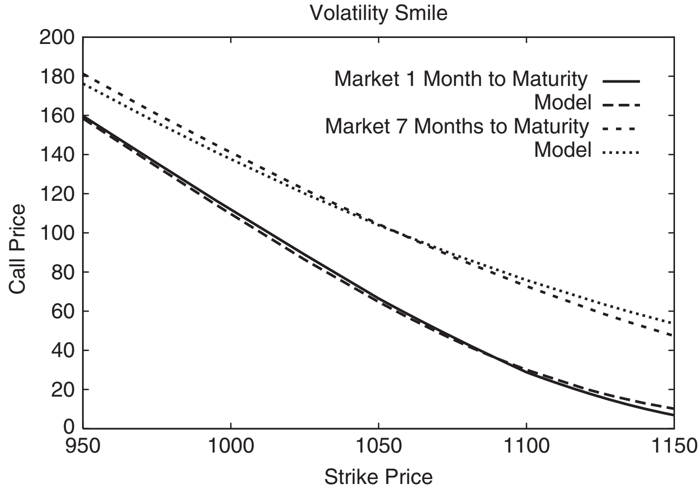
Indeed, going back to (1.46) we can do a simulation on Y_t and v_t using the random sequence of (Z_t), then after one path is generated, we can calculate $S^{eff} = S_0 \exp(Y_T)$ and $v^{eff} = \frac{1}{T}\sum_{t=0}^{T-\Delta t}(1-\rho_t^2)v_t\Delta t$ and then apply the known closed-form (e.g., Black-Scholes for a call or put) with S^{eff} and v^{eff}. Repeating this procedure for a large number of times and averaging over the paths, as we usually do in Monte-Carlo methods, we will have $f(S_0, v_0, T)$.

This will give us a way to calibrate the model parameters to the market data. For instance using the Square-Root model

$$dv_t = (\omega - \theta v_t)dt + \xi\sqrt{v_t}dZ_t,$$

we can estimate ω, θ, ξ, and ρ from the market prices via a least-square estimation applied to theoretical prices obtained from the above Monte-Carlo method. We can either use a single calibration and suppose we have time-independent parameters, or perform one calibration per maturity. The single calibration method is known to provide a bad fit, hence the idea of adding jumps to the stochastic volatility process as described by Matytsin [197]. However, this method will introduce new parameters for calibration.[34]

[34]Eraker et al. [103] claim that a model containing jumps in the return *and* the volatility process will fit the options and the underlying data well, and will have no misspecification left.



Mixing Monte-Carlo simulation with the Square-Root model for SPX on February 12, 2002, with Index = 1107.5 USD, one month and seven months to maturity. Powell Optimization method was used for Least-Square Calibration. As we can see, both maturities are fitted fairly well.

THE LONG-TERM ASYMPTOTIC CASE

In this section, we will discuss the case where the contract time to maturity is very large, that is, $t \to \infty$. We will focus on the special case of a Square-Root process, since this is the model we will use in many cases.

The Deterministic Case

We shall start with the case of deterministic volatility and use that for the more general case of the stochastic volatility.

We know that under the Square-Root model the variance follows

$$dv_t = (\omega - \theta v_t)dt + \xi\sqrt{v_t}dZ_t.$$

As an *approximation*, we can drop the stochastic term and obtain

$$\frac{dv_t}{dt} = \omega - \theta v_t$$

which is an ordinary differential equation providing us immediately with

$$v_t = \frac{\omega}{\theta} + \left(v - \frac{\omega}{\theta}\right)e^{-\theta t} \tag{1.51}$$

where v is the initial variance for $t = 0$.

Using the results from the Fundamental Transform for a covered call option and put-call parity, we have for $0 < k_i < 1$

$$call(S, v, \tau) = Se^{-q\tau} - Ke^{-r\tau} \frac{1}{2\pi} \int_{ik_i-\infty}^{ik_i+\infty} e^{-ikX} \frac{\hat{H}(k, v, \tau)}{k^2 - ik} dk \tag{1.52}$$

where $\tau = T - t$, and $X = \ln(\frac{Se^{-q\tau}}{Ke^{-r\tau}})$ represents the adjusted moneyness of the option. For the special "At the Money"[35] case where $X = 0$ we have

$$call(S, v, \tau) = Ke^{-r\tau} \left[1 - \frac{1}{2\pi} \int_{ik_i-\infty}^{ik_i+\infty} \frac{\hat{H}(k, v, \tau)}{k^2 - ik} dk \right]. \tag{1.53}$$

As we previously said for a deterministic volatility case, we know the Fundamental Transform

$$\hat{H}(k, v, \tau) = \exp[-c(k)U(v, \tau)]$$

with $U(v, \tau) = \int_0^\tau v(t)dt$ and as before $c(k) = \frac{1}{2}(k^2 - ik)$, which in the special case of the Square-Root model (1.51), will provide us with

$$U(v, \tau) = \frac{\omega}{\theta}\tau + \left(v - \frac{\omega}{\theta}\right)\left(\frac{1 - e^{-\theta\tau}}{\theta}\right).$$

This shows once again that $\hat{H}(k)$ is analytic in k over the entire complex plane.

Now if we let $\tau \to \infty$ we can write the approximation

$$\frac{call(S, v, \tau)}{Ke^{-r\tau}} \approx 1 - \frac{1}{2\pi} \int_{ik_i-\infty}^{ik_i+\infty} \exp\left[-c(k)\frac{\omega}{\theta}\tau - c(k)\frac{1}{\theta}\left(v - \frac{\omega}{\theta}\right)\right] \frac{dk}{k^2 - ik}. \tag{1.54}$$

We can either calculate this integral exactly using the Black-Scholes theory, or take the minimum where $c'(k_0) = 0$, meaning $k_0 = \frac{i}{2}$, and perform a Taylor approximation parallel to the real axis around the point $k = k_r + \frac{i}{2}$, which will give us

$$\frac{call(S, v, \tau)}{Ke^{-r\tau}} \approx 1 - \frac{2}{\pi} \exp\left(-\frac{\omega}{8\theta}\tau\right) \exp\left[-\frac{1}{8\theta}\left(v - \frac{\omega}{\theta}\right)\right] \times \int_{-\infty}^{\infty} \exp\left(-k_r^2 \frac{\omega}{2\theta}\tau\right) dk_r$$

[35] This is different from the usual definition of At the Money calls where $S = K$. This vocabulary is borrowed from Alan Lewis.

The integral being a Gaussian, we will get the result

$$\frac{call(S,v,\tau)}{Ke^{-r\tau}} \approx 1 - \sqrt{\frac{8\theta}{\pi\omega\tau}} \exp\left[-\frac{1}{8\theta}\left(v - \frac{\omega}{\theta}\right)\right] \exp\left(-\frac{\omega}{8\theta}\tau\right) \tag{1.55}$$

which finishes our deterministic approximation case.

The Stochastic Case

For the stochastic volatility (SV) case, Lewis uses the same Taylor expansion. He notices that for the deterministic case we had

$$\hat{H}(k,v,\tau) = \exp[-c(k)U(v,\tau)] \approx \exp[-\lambda(k)\tau]u(k,v)$$

for $\tau \to \infty$ where

$$\lambda(k) = c(k)\frac{\omega}{\theta}$$

and

$$u(k,v) = \exp\left[-c(k)\frac{1}{\theta}\left(v - \frac{\omega}{\theta}\right)\right].$$

If we *suppose* that this identity holds for the SV case as well, we can use the PDE (1.44) and interpret the result as an *eigenvalue-eigenfunction* identity with the eigenvalue $\lambda(k)$ and the eigenfunction $u(k,v)$. This assumption is reasonable since the first Taylor approximation term for the stochastic process *is* deterministic.

Indeed, introducing the operator

$$\Lambda(u) = -\frac{1}{2}a^2(v)\frac{d^2u}{dv^2} - [\tilde{b}(v) - ik\rho(v)a(v)\sqrt{v}]\frac{du}{dv} + c(k)vu$$

we have

$$\Lambda(u) = \lambda(k)u. \tag{1.56}$$

Now the idea would be to perform a Taylor expansion around the minimum k_0 where $\lambda'(k_0) = 0$. Lewis shows that such k_0 is always situated on the imaginary axis. This property is referred to as the "ridge" property.

The Taylor expansion along the real axis could be written as

$$\lambda(k) = \lambda(k_0 + k_r) \approx \lambda(k_0) + \frac{1}{2}k_r^2\lambda''(k_0).$$

Note that we are dealing with a *minimum* and therefore $\lambda''(k_0) > 0$. Using the previous second-order approximation for $\lambda(k)$ we get

$$\frac{call(S,v,\tau)}{Ke^{-r\tau}} \approx 1 - \frac{u(k_0,v)}{k_0^2 - ik_0}\frac{1}{\sqrt{2\pi\lambda''(k_0)\tau}}\exp[-\lambda(k_0)\tau].$$

We can then move from the special "At the Money" case to the general case by reintroducing $X = \ln(\frac{Se^{-q\tau}}{Ke^{-r\tau}})$ and we will finally obtain

$$\frac{call(S,v,\tau)}{Ke^{-r\tau}} \approx e^X - \frac{u(k_0,v)}{k_0^2 - ik_0}\frac{1}{\sqrt{2\pi\lambda''(k_0)\tau}}\exp[-\lambda(k_0)\tau - ik_0X] \quad (1.57)$$

which completes our determination of the asymptotic closed-form in the general case.

For the special case of the Square-Root model, taking the risk-neutral case $\gamma = 1$, we have[36]

$$\lambda(k) = -\omega g^*(k) = \frac{\omega}{\xi^2}[\sqrt{(\theta + ik\rho\xi)^2 + (k^2 - ik)\xi^2} - (\theta + ik\rho\xi)]$$

which also allows us to calculate $\lambda''(k)$. Also

$$u(k,v) = \exp[g^*(k)v]$$

where we use the notations from (1.45) and we pose

$$g^* = g - d.$$

The k_0 such that $\lambda'(k_0) = 0$ is

$$k_0 = \frac{i}{1-\rho^2}\left(\frac{1}{2} - \frac{\rho}{\xi}\left[\theta - \frac{1}{2}\sqrt{4\theta^2 + \xi^2 - 4\rho\theta\xi}\right]\right)$$

which together with (1.57) provides us with the result for $call(S,v,\tau)$ in the asymptotic case under the Square-Root stochastic volatility model.

Note that for $\xi \to 0$ and $\rho \to 0$ we find again the deterministic result $k_0 \to \frac{i}{2}$.

[36]We can go back to the general case $\gamma \leq 1$ by replacing θ with $\sqrt{\theta^2 - \gamma(1-\gamma)\xi^2} + (1-\gamma)\rho\xi$ since this transformation is independent from k altogether.

A Series Expansion on Volatility-of-Volatility

Another asymptotic approach for the stochastic volatility model suggested by Lewis [185] is a Taylor expansion on the volatility-of-volatility. There are two possibilities for this; we can perform the expansion *either* for the option price *or* for the implied-volatility directly.

In what follows, we consider the former approach. Once again, we use the fundamental transform $H(k, V, \tau)$ with $H(k, V, 0) = 1$ and

$$\frac{\partial H}{\partial \tau} = \frac{1}{2}a^2(v)\frac{\partial^2 H}{\partial v^2} + (\tilde{b}(v) - ik\rho(v)a(v)\sqrt{v})\frac{\partial H}{\partial v} - c(k)vH$$

and $c(k) = \frac{1}{2}(k^2 - ik)$.

We then pose $a(v) = \xi\eta(v)$ and expand $H(k, V, \tau)$ on powers of ξ and finally apply the inverse Fourier Transform to obtain an expansion on the call price.

With our usual notations $\tau = T - t$, $X = \ln\left(\frac{S}{K}\right) + (r - q)\tau$ and $Z(V) = V\tau$, the series will be

$$C(S, V, \tau) = c_{BS}(S, v, \tau) + \xi\tau^{-1}J_1\tilde{R}_{11}\frac{\partial c_{BS}(S, v, \tau)}{\partial V} + \xi^2\left[\tau^{-2}J_3\tilde{R}_{20} + \tau^{-1}J_4\tilde{R}_{12} + \frac{1}{2}\tau^{-2}J_1^2\tilde{R}_{22}\right]\frac{\partial c_{BS}(S, v, \tau)}{\partial V} + O(\xi^3)$$

where $v(V, \tau)$ is the deterministic variance

$$v(V, \tau) = \frac{\omega}{\theta} + \left(V - \frac{\omega}{\theta}\right)\left(\frac{1 - e^{-\theta\tau}}{\theta\tau}\right),$$

and $\tilde{R}_{pq} = R_{pq}(X, v(V, \tau), \tau)$ with R_{pq} given polynomials of (X, Z) of degree four at most, and J_n's known functions of (V, τ).

The explicit expressions for all these functions are given in the third chapter of the Lewis book [185].

The obvious advantage of this approach is its speed and stability. The issue of lack of time-homogeneity of the parameters $\Psi = (\omega, \theta, \xi, \rho)$ could be addressed by performing one calibration per time-interval. In this case, for each time-interval $[t_n, t_{n+1}]$ we will have one set of parameters $\Psi_n = (\omega_n, \theta_n, \xi_n, \rho_n)$ and depending on what maturity T we are dealing with, we will use one or the other parameter-set.

We compare the values obtained from this series-based approach with those from a mixing Monte-Carlo method in Figure 1.1. We are taking the example that Heston studied in [141]. The graph shows the difference

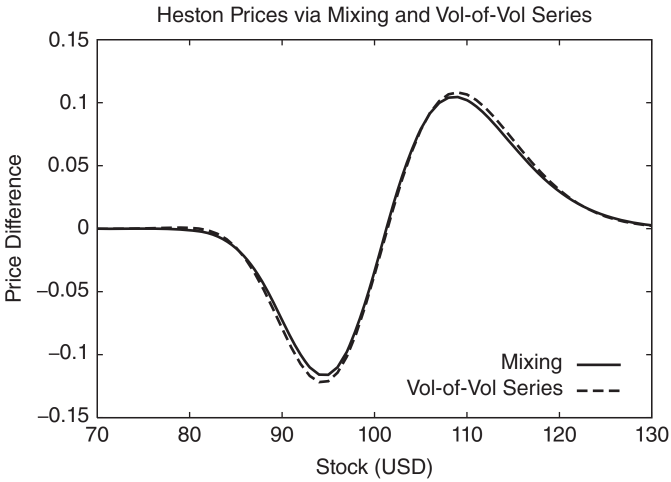


FIGURE 1.1 Comparing the volatility-of-volatility series expansion with the Monte-Carlo mixing model. The graph shows the price difference $C(S, V, \tau) - c_{BS}(S, V, \tau)$. We are taking $\xi = 0.10$ and $\rho = -0.50$. This example was used in the original Heston paper.

$C(S, V, \tau) - c_{BS}(S, V, \tau)$ for a fixed $K = 100$ USD and $\tau = 0.50$ years. The other inputs are $\omega = 0.02$, $\theta = 2.00$, $\xi = 0.10$, $\rho = -0.50$, $V = 0.01$, and $r = q = 0$.

As we can see, the true value of the call is *lower* than the Black-Scholes value for the OTM region. The higher ξ and $|\rho|$ are, the larger this difference will be.

In Figures 1.2 and 1.3 we reset the correlation ρ to zero to have a symmetric distribution, but we use a volatility-of-volatility of $\xi = 0.10$ and $\xi = 0.20$ respectively. As discussed, the parameter ξ is the one creating the leptokurticity phenomenon. A higher volatility-of-volatility causes higher valuation for Far-from-the-Money options.[37]

Unfortunately, the series approximation in Figure 1.1 becomes poor as soon as the volatility of volatility becomes larger than 0.40 and the maturity becomes of the order of one year. This case is not unusual at all and therefore makes the use of this method limited. This is why the method of choice remains the Inversion of the Fourier Transform, as previously described.

[37] Also note that the gap between the closed-form series and the Monte-Carlo model increases with ξ. Indeed the accuracy of the expansion decreases as ξ becomes larger.

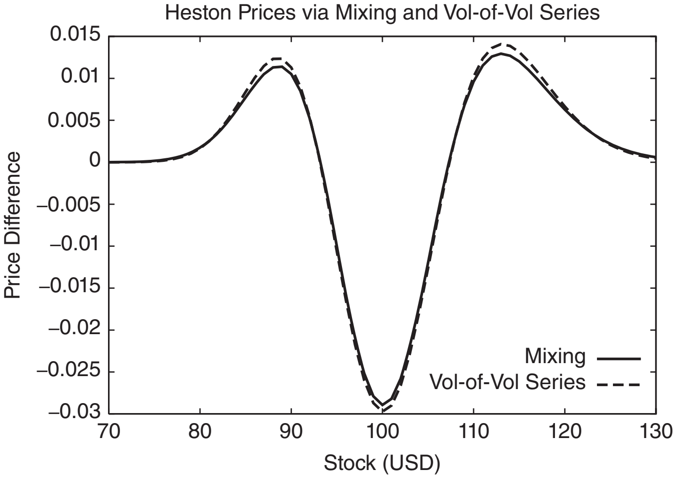


FIGURE 1.2 Comparing the volatility-of-volatility series expansion with the Monte-Carlo mixing model. The graph shows the price difference $C(S, V, \tau) - c_{BS}(S, V, \tau)$. We are taking $\xi = 0.10$ and $\rho = 0$. This example was used in the original Heston paper.

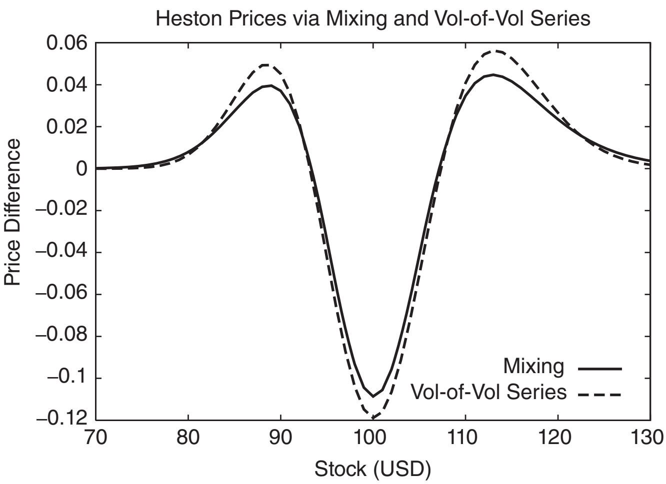


FIGURE 1.3 Comparing the volatility-of-volatility series expansion with the Monte-Carlo mixing model. The graph shows the price difference $C(S, V, \tau) - c_{BS}(S, V, \tau)$. We are taking $\xi = 0.20$ and $\rho = 0$. This example was used in the original Heston paper.

LOCAL VOLATILITY STOCHASTIC VOLATILITY MODELS

As previously mentioned, one known issue with SV models such as Heston (or other) is that given their parametric form they typically cannot match options markets exactly, which would create residual PNL. LV, on the other hand, matches options market prices exactly by construction, however does not have the proper dynamics for many cases such as forward skew or…

One possible compromise as suggested by [216] is to have a stochastic volatility model with a superimposed local volatility component. We will use the same notations as [216] here.

In general, having a stock process

$$dS = (r - q)dt + \sigma(S, t)Z(t)SdW_S(t)$$

with, for example

$$d\ln Z = \kappa(\theta(t) - \ln Z)dt + \lambda dW_Z(t),$$

we can view the $Z(t)$ term as the stochastic volatility piece giving the desired dynamics albeit partially, the $\sigma(S, t)Z(t)$ term, the local volatility (without the expectations as per below), and the $\sigma(S, t)$ term the residual.

More precisely, as we had already seen, the local variance is

$$\sigma^2_{LV}(K, T) = \sigma^2(K, T)E[Z^2(T)|S(T) = K]$$

and separately we know that

$$\sigma^2_{LV}(K, T) = \frac{\frac{\partial C}{\partial T} + (r - q)\frac{\partial C}{\partial K} + qC}{\frac{1}{2}\frac{\partial^2 C}{\partial K^2}};$$

therefore, having the local volatility we can get our residual piece via

$$\sigma^2(K, T) = \frac{\sigma^2_{LV}(K, T)}{E[Z^2(T)|S(T) = K]}.$$

And writing this $E[Z^2(T)|S(T) = K] = \Psi(K, T)$ we can calculate it as

$$\Psi(K, T) = \frac{\int_0^\infty Z^2 p(K, Z, T)dZ}{\int_0^\infty p(K, Z, T)dZ}$$

with $p(S, V, T)$, the forward joint transition density of S and Z.

This density can then be determined via the forward Kolmogorov equation sequentially as the authors suggest or alternatively via a particle filter algorithm in a Monte-Carlo simulation.

The calibration of the model therefore includes the usual off-line SV part to estimate $\kappa, \lambda, \theta(t)$ via least-squares and the on-line part via the aforementioned sequence for the local volatility piece.

STOCHASTIC IMPLIED VOLATILITY

As we have already mentioned, a few authors [45, 59, 224] try to model the Black-Scholes implied volatility (instead of the instantaneous variance) as a stochastic variable. In what follows in this section, we will use the notations of [45].

Assuming zero rates, borrow, and dividends, we can write for the spot process

$$dS_t = S_t \theta_t dW_t$$

where θ_t is the instantaneous volatility[38] and is stochastic.

Calling σ_t the stochastic implied volatility for strike K and maturity T we have the price of a call option at time t as

$$C_t = C(t, T, K) = \phi(S_t, \sigma_t(T, K), T - t, K)$$

where ϕ is the usual Black Scholes pricing function

$$\phi(S, \sigma, \tau, K) = S\mathcal{N}(h_1) - K\mathcal{N}(h_2)$$

with the usual

$$h_1 = \frac{\log \frac{S}{K} + \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}}$$

$$h_2 = \frac{\log \frac{S}{K} - \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}},$$

and we assume the implied volatility satisfies the SDE

$$d\sigma_t = m_t(T, K, S_t, \theta_t, \sigma_t)dt + v_t(T, K, S_t, \theta_t, \sigma_t)dZ_t$$

where m_t and v_t are the drift and volatility process for the implied volatility.

[38]The *square-root* of our usual instantaneous variance.

Other than the usual boundedness and positivity constraints, we have the key constraint reducing the degrees of freedom

$$\sigma_t(t, S_t) = \theta_t.$$

Indeed, for a zero time to maturity, the implied volatility process should converge toward the instantaneous volatility.

As shown in [45] one can see that the call option C_t follows the SDE

$$dC_t = \mathcal{N}(h_1)S_t\theta_t dW_t + \sqrt{T-t}\mathcal{N}'(h_1)v_t dZ_t$$
$$+ \sqrt{T-t}\mathcal{N}'(h_1)\left[m_t + \frac{\theta_t^2}{2\sigma_t(T-t)}\right.$$
$$\left.-\frac{\sigma_t}{2(T-t)} + \frac{h_1h_2|v_t|^2}{2\sigma_t} - \frac{h_2\theta_t v_t}{\sigma_t\sqrt{T-t}}\right].$$

The traded option price (under zero rates) should be a martingale, giving the condition

$$m_t = \frac{1}{2\sigma_t(T-t)}[\sigma_t^2 - \theta_t^2 - (T-t)h_1h_2|v_t|^2 + 2\sqrt{T-t}h_2\theta_t v_t]$$

which will lead to

$$dC_t = \mathcal{N}(h_1)S_t\theta_t dW_t + \sqrt{T-t}S_t\mathcal{N}'(h_1)v_t dZ_t.$$

Now assuming

$$v_t = \sigma_t u_t,$$

we effectively have

$$d\sigma_t = \frac{1}{2\sigma_t(T-t)}\left(\sigma_t^2 + 1/4\sigma_t^4(T-t)^2|u_t|^2 - \sigma_t^2(T-t)\theta_t u_t\right.$$
$$\left.-\left|\theta_t + u_t\log\frac{K}{S_t}\right|^2\right)dt + \sigma_t u_t dZ_t$$

Based on an alternate formulation, [45] shows there is existence and uniqueness of the solution to this system of equations.

The SDE for the implied volatility is

$$d\sigma_t = h(\sigma_t, v_t, u_t)dt + \sigma_t u_t dZ_t$$

with the constraint:

$$\sqrt{v_t} = \sigma(t; t, S_t)$$

and dynamics

$$d\sigma_t = \frac{1}{2(T-t)\sigma_t}\left[\sigma_t^2 - \left(\sqrt{v_t} + u_t \ln\frac{K}{S_t}\right)^2\right]dt + \frac{1}{8}(T-t)\sigma_t^3 u_t^2 dt + \sigma_t u_t dZ_t$$

and for example the vol of implied vol u_t follows some SDE such as

$$du_t = f(u_t)dt + \Upsilon dZ_t^{\perp}$$

where Υ is a model parameter.

We will use this model later to estimate the dynamics of the instantaneous volatility from historic option prices. As we will see, the quality of observations for historic option prices (or equivalently implied volatilities) are superior to the ones for the stock prices.

JOINT SPX AND VIX DYNAMICS

In order to have an SV model representing not only SPX but also the VIX dynamics, we need more than one factor of stochasticity. VIX index was created by CBOE and futures and options on VIX have increased in trading volume over the past ten years tremendously. The actual definition of VIX is based on SPX puts and calls as explained in

`http://www.cboe.com/micro/VIX/vixintro.aspx`

but one can represent the (nontraded) VIX spot in an abstract way by

$$V_t = \sqrt{E_t^Q\left[\int_t^{t+\tau} v_u du\right]}$$

where τ always corresponds to 30 calendar days, and v_t is the instantaneous variance of SPX.

Following the same notations, we can write VIX futures (with maturity T) as

$$F_t^T = E_t^Q[V_T],$$

and in the same manner they define options on futures with same maturity T.

There have been several models suggested in literature. A popular one is the Bergomi [35] model. We are using the same notations as the author in this section.

The idea is to have a two-factor SV model with abstract OU processes

$$dX_t = k_1 X_t dt + dW_t^X$$
$$dY_t = k_1 Y_t dt + dW_t^Y$$

with $\langle dW^X, dW^Y \rangle = \rho dt$, $\langle dW^X, dW \rangle = \rho_{SX} dt$, $\langle dW^Y, dW \rangle = \rho_{SY} dt$ where dW is the Wiener process for the equity (SPX).

We then take the linear combination

$$x_t^T = \alpha_\theta[(1-\theta)e^{-k_1(T-t)}X_t + \theta e^{-k_2(T-t)}Y_t]$$

with α_θ a normalization factor equal to $1/\sqrt{(1-\theta)^2 + \theta^2 + 2\rho\theta(1-\theta)}$.

Defining the forward variance as the stochastic variable

$$\xi_t^T = E\left[\int_t^T v_u du | t\right],$$

we set

$$\xi_t^T = \xi_0^T f^T(x_t^T, t)$$

where ξ_0^T is the variance swap level to maturity T, and f^T is a function to be determined parametrized for each maturity T.

Note that X_t, Y_t, and therefore x_t^T are by construction drift-less.[39]

The local martingale condition (zero drift) on ξ_t^T provides the Feynman-Kac PDE

$$\frac{\partial f^T}{\partial t} + \frac{\sigma^2(T-t)}{2}\frac{\partial^2 f^T}{\partial x^2} = 0$$

[39]It is important (as the author mentions) to note that unlike in interest rates markets, in equities one cannot take one function for $S_t = f(W_t, t)$ since we typically have multiple expirations T_i for contracts on the same traded asset S.

with

$$\sigma^2(\tau) = \alpha_\theta^2[(1-\theta)^2 e^{-2k_1\tau} + \theta^2 e^{-2k_2\tau} + 2\rho\theta(1-\theta)e^{-(k_1+k_2)\tau}].$$

In practice, we would choose one f^T per VIX future maturity interval $[T_i, T_{i+1}]$.

The author then suggests a weighted exponential form for f^T such as

$$f^T(x,T) = (1-\gamma_T)e^{\omega_T x}e^{-\frac{\omega_T^2 h(t,T)}{2}} + \gamma_T e^{\beta_T\omega_T x}e^{-\frac{\beta_T^2\omega_T^2 h(t,T)}{2}}$$

with $h(t,T) = \int_{T-t}^{T}\sigma^2(\tau)d\tau$.

Note that the instantaneous variance of the equity spot is simply ξ_t^t, so the risk-neutral dynamics are

$$dS = (r-q)Sdt + S\sqrt{\xi_t^t}dW_t.$$

Note that in practice, working with discrete maturities T_i for VIX futures, we can define a function $\sigma_i(S,t)$ with

$$dS = (r-q)Sdt + S\sigma_i\left(\frac{S}{S_{T_i}}, \frac{1}{T_{i+1}-T_i}\int_{T_i}^{T_{i+1}}\xi_{T_i}^T dT\right)dW_t$$

which will make the calculations less cumbersome.

PURE-JUMP MODELS

Variance Gamma

An alternative point of view is to drop the diffusion assumption altogether and replace it with a pure-jump process. Note that this is different from the jump-diffusion process previously discussed. Madan et al. suggested the following framework called Variance-Gamma (VG) in [192].

We would have the log-normal-like stock process

$$d\ln S_t = (\mu_S + \omega)dt + X(dt;\sigma,\nu,\theta)$$

where, as before, μ_S is the real-world statistical drift of the stock log-return and $\omega = \frac{1}{\nu}\ln(1-\theta\nu-\sigma^2\nu/2)$.

As for $X(dt;\sigma,\nu,\theta)$ it has the following meaning:

$$X(dt;\sigma,\nu,\theta) = B(\gamma(dt,1,\nu);\theta,\sigma)$$

where $B(dt;\theta,\sigma)$ would be a Brownian Motion with drift θ and volatility σ. In other words

$$B(dt;\theta,\sigma) = \theta dt + \sigma\sqrt{dt}N(0,1),$$

and $N(0,1)$ is a standard Gaussian realization.

The time-interval at which the Brownian Motion is considered is not dt but $\gamma(dt,1,\nu)$, which is a random realization following a Gamma distribution with a mean 1 and variance-rate ν.

The corresponding probability density function is

$$f_\nu(dt,\tau) = \frac{\tau^{\frac{dt}{\nu}-1}e^{-\frac{\tau}{\nu}}}{\nu^{\frac{dt}{\nu}}\Gamma\left(\frac{dt}{\nu}\right)}$$

where $\Gamma(x)$ is the usual Gamma function.

Note that the stock log-return density could actually be *integrated* for the VG model, and the density of $\ln(S_t/S_0)$ is known and could be implemented via $K_\alpha(x)$ the modified Bessel function of the second kind.

Indeed, calling $z = \ln(S_k/S_{k-1})$ and $h = t_k - t_{k-1}$ and posing $x_h = z - \mu_S h - \frac{h}{\nu}\ln(1-\theta\nu-\sigma^2\nu/2)$ we have

$$p(z|h) = \frac{2\exp(\theta x_h/\sigma^2)}{\nu^{\frac{h}{\nu}}\sqrt{2\pi}\sigma\Gamma\left(\frac{h}{\nu}\right)}\left(\frac{x_h^2}{2\sigma^2/\nu+\theta^2}\right)^{\frac{h}{2\nu}-\frac{1}{4}}$$
$$\times K_{\frac{h}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{x_h^2(2\sigma^2/\nu+\theta^2)}\right).$$

What is more, as Madan et al. show, the option valuation under VG is fairly straightforward and admits an analytically tractable closed-form which can be implemented via the above modified Bessel function of second kind and a degenerate hypergeometric function. All details are available in [192].

Remark on the Gamma Distribution The Gamma Cumulative Distribution Function (CDF) could be defined as

$$P(a,x) = \frac{1}{\Gamma(a)}\int_0^x e^{-t}t^{a-1}dt.$$

Note that with our notations

$$F_\nu(h, x) = F(h, x, \mu = 1, \nu)$$

with

$$F(h, x, \mu, \nu) = \frac{1}{\Gamma\left(\frac{\mu^2 h}{\nu}\right)} \left(\frac{\mu}{\nu}\right)^{\frac{\mu^2 h}{\nu}} \int_0^x e^{-\frac{\mu t}{\nu}} t^{\frac{\mu^2 h}{\nu}-1} dt.$$

In other words

$$F(h, x, \mu, \nu) = P\left(\frac{\mu^2 h}{\nu}, \frac{\mu x}{\nu}\right).$$

The behavior if this CDF is displayed in Figure 1.4 for different values of the parameter $a > 0$ and for $0 < x < +\infty$.

Using the inverse of this CDF we can have a simulated data-set for the Gamma law:

$$x^{(i)} = F_\nu^{-1}(h, \mathcal{U}^{(i)}[0, 1])$$

with $1 \leq i \leq N_{sims}$ and $\mathcal{U}^{(i)}[0, 1]$ a uniform random realization between zero and one.

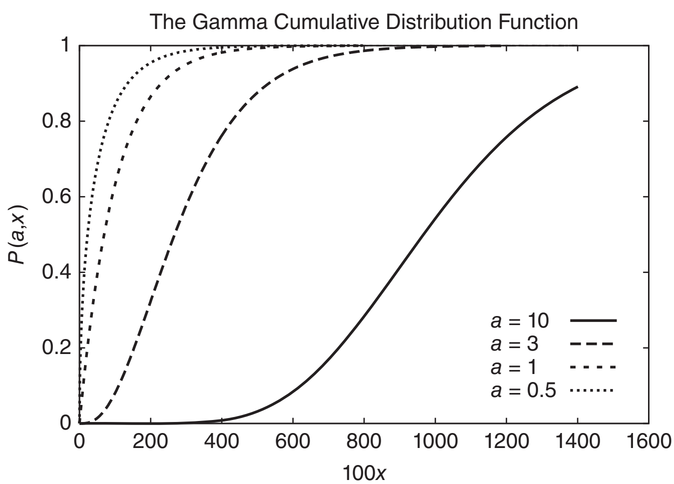


FIGURE 1.4 The Gamma Cumulative Distribution Function $P(a, x)$ for various values of the parameter a. The implementation is based on code available in "Numerical Recipes in C."

Stochastic Volatility vs. Time-Changed Processes As mentioned in [24], this alternative formulation leading to time-changed processes, is closely related to the previously discussed stochastic volatility approach in the following way:

Taking the above VG stochastic differential equation

$$d \ln S_t = (\mu_S + \omega)dt + \theta\gamma(dt, 1, \nu) + \sigma\sqrt{\gamma(dt, 1, \nu)}N(0, 1),$$

one could consider $\sigma^2\gamma(t, 1, \nu)$ as the integrated-variance and define $v_t(\nu)$ the instantaneous-variance as

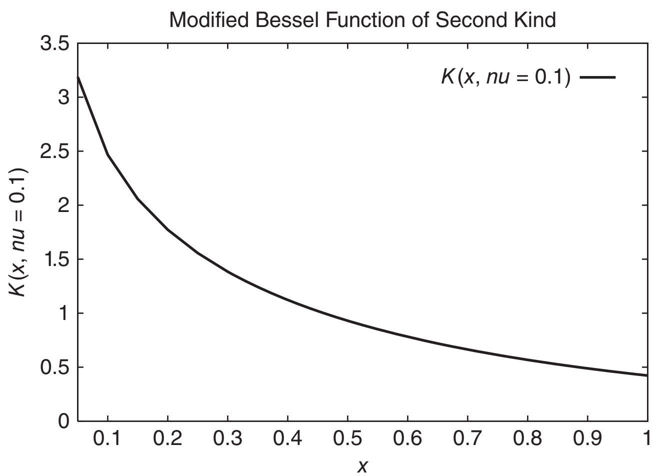
$$\sigma^2\gamma(dt, 1, \nu) = v_t(\nu)dt$$

in which case, we would have

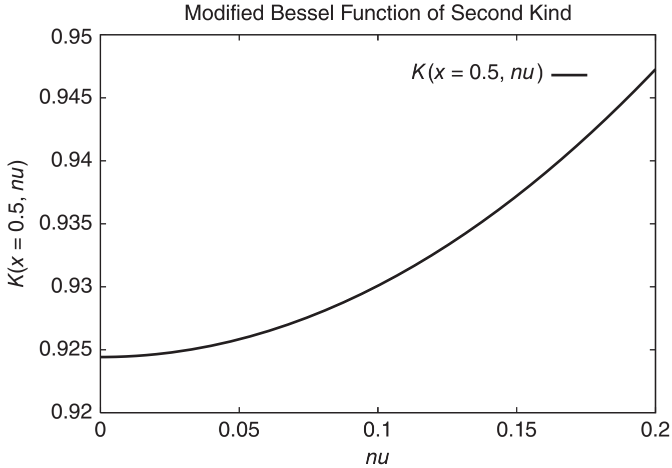
$$\begin{aligned} d \ln S_t &= (\mu_S + \omega)dt + (\theta/\sigma^2)v_t(\nu)dt + \sqrt{v_t(\nu)dt}N(0, 1) \\ &= (\mu_S + \omega + (\theta/\sigma^2)v_t(\nu))dt + \sqrt{v_t(\nu)}dZ_t \end{aligned}$$

where dZ_t is a Brownian Motion.

This last equation is a traditional stochastic volatility one.



The Modified Bessel Function of Second Kind for a given parameter. The implementation is based on code available in "Numerical Recipes in C."



The Modified Bessel Function of Second Kind as a function of the parameter. The implementation is based on code available in "Numerical Recipes in C."

Variance Gamma with Stochastic Arrival

An extension to the VG model would be a Variance Gamma model with Stochastic Arrival (VGSA), which would include the volatility *clustering* effect. This phenomenon (also represented by GARCH) means that a high (low) volatility will be followed by a series of high (low) volatilities.

In this approach, we replace the dt in the previously defined $f_\nu(dt, \tau)$ with $y_t dt$ where y_t follows a Square-Root (CIR) process

$$dy_t = \kappa(\eta - y_t)dt + \lambda\sqrt{y_t}dW_t$$

where the Brownian Motion dW_t is independent from other processes in the model.

This is therefore a VG process where the arrival time itself is stochastic. The mean-reversion of the Square-Root process will cause the volatility persistence effect that is empirically observed. Note that (not counting μ_S) the new model parameter-set is $\Psi = (\kappa, \eta, \lambda, \nu, \theta, \sigma)$.

Option Pricing under VGSA The option pricing could be carried out via a Monte-Carlo simulation algorithm under the risk-neutral measure, where, as before, μ_S is replaced with $r - q$. We first would simulate the path of y_t by writing

$$y_k = y_{k-1} + \kappa(\eta - y_{k-1})\Delta t + \lambda\sqrt{y_{k-1}}\sqrt{\Delta t}Z_k$$

then calculate

$$Y_T = \sum_{k=0}^{N-1} y_k \Delta t$$

and finally apply *one-step* simulations

$$T^* = F_\nu^{-1}(Y_T, \mathcal{U}[0,1])$$

and[40]

$$\ln S_T = \ln S_0 + (r - q + \omega)T + \theta T^* + \sigma\sqrt{T^*}B_k.$$

Note that we have two normal random-variables B_k, Z_k as well as a Gamma-distributed random variable T^*, and that they are all uncorrelated.

Once the stock price S_T is properly simulated, we can calculate the option price as usual.

The Characteristic Function As previously discussed, another way to tackle the option-pricing issue would be to use the characteristic functions.

For VG, the characteristic function is

$$\Psi(u,t) = \mathbf{E}[e^{iuX(t)}] = \left(\frac{1}{1 - i\frac{\nu}{\mu}u}\right)^{\frac{\mu 2}{\nu}t}.$$

Therefore, the log-characteristic function could be written as

$$\psi(u,t) = \ln(\Psi(u,t)) = t\psi(u,1).$$

In other words,

$$\mathbf{E}[e^{iuX(t)}] = \Psi(u,t) = \exp(t\psi(u,1)).$$

Using which, the VGSA characteristic function becomes

$$\mathbf{E}[e^{iuX(Y(t))}] = \mathbf{E}[\exp(Y(t)\psi(u,1))] = \phi(-i\psi(u,1))$$

with $\phi()$ the CIR characteristic function, namely

$$\phi(u_t) = \mathbf{E}[\exp(iuY_t)] = A(t,u)\exp(B(t,u)y_0)$$

[40]This means that T in VG, is replaced with Y_T. The rest remains identical.

where

$$A(t,u) = \frac{\exp(\kappa^2\eta t/\lambda^2)}{[\cosh(\gamma t/2) + \kappa/\gamma \sinh(\gamma t/2)]^{\frac{2\kappa\eta}{\lambda^2}}}$$

$$B(t,u) = \frac{2iu}{\kappa + \gamma \coth(\gamma t/2)}$$

and

$$\gamma = \sqrt{\kappa^2 - 2\lambda^2 iu}.$$

This allows us to determine the VGSA characteristic function, which we can use to calculate options prices via numeric Fourier inversion as described in [52] and [55].

Variance Gamma with Gamma Arrival Rate

For the Variance Gamma with Gamma Arrival Rate (VGG), as before, the stock process under the risk-neutral framework is

$$d\ln S_t = (r - q + \omega)dt + X(h(dt); \sigma, \nu, \theta)$$

with $\omega = \frac{1}{\nu}\ln(1 - \theta\nu - \sigma^2\nu/2)$ and

$$X(h(dt); \sigma, \nu, \theta) = B(\gamma(h(dt), 1, \nu); \theta, \sigma)$$

and the general Gamma Cumulative Distribution Function for $\gamma(h, \mu, \nu)$ is

$$F(\mu, \nu; h, x) = \frac{1}{\Gamma\left(\frac{\mu^2 h}{\nu}\right)}\left(\frac{\mu}{\nu}\right)^{\frac{\mu^2 h}{\nu}} \int_0^x e^{-\frac{\mu t}{\nu}} t^{\frac{\mu^2 h}{\nu} - 1} dt$$

and here $h(dt) = dY_t$ with Y_t is also Gamma-distributed

$$dY_t = \gamma(dt, \mu_a, \nu_a).$$

The parameter-set is therefore $\Psi = (\mu_a, \nu_a, \nu, \theta, \sigma)$.