CHAPTER 1

Stochastic Volatility and Local Volatility

In this chapter, we begin our exploration of the volatility surface by introducing stochastic volatility—the notion that volatility varies in a random fashion. Local variance is then shown to be a conditional expectation of the instantaneous variance so that various quantities of interest (such as option prices) may sometimes be computed as though future volatility were deterministic rather than stochastic.

STOCHASTIC VOLATILITY

That it might make sense to model volatility as a random variable should be clear to the most casual observer of equity markets. To be convinced, one need only recall the stock market crash of October 1987. Nevertheless, given the success of the Black-Scholes model in parsimoniously describing market options prices, it's not immediately obvious what the benefits of making such a modeling choice might be.

Stochastic volatility (SV) models are useful because they explain in a self-consistent way why options with different strikes and expirations have different Black-Scholes implied volatilities—that is, the "volatility smile." Moreover, unlike alternative models that can fit the smile (such as local volatility models, for example), SV models assume realistic dynamics for the underlying. Although SV price processes are sometimes accused of being *ad hoc*, on the contrary, they can be viewed as arising from Brownian motion subordinated to a random clock. This clock time, often referred to as *trading time*, may be identified with the volume of trades or the frequency of trading (Clark 1973); the idea is that as trading activity fluctuates, so does volatility.

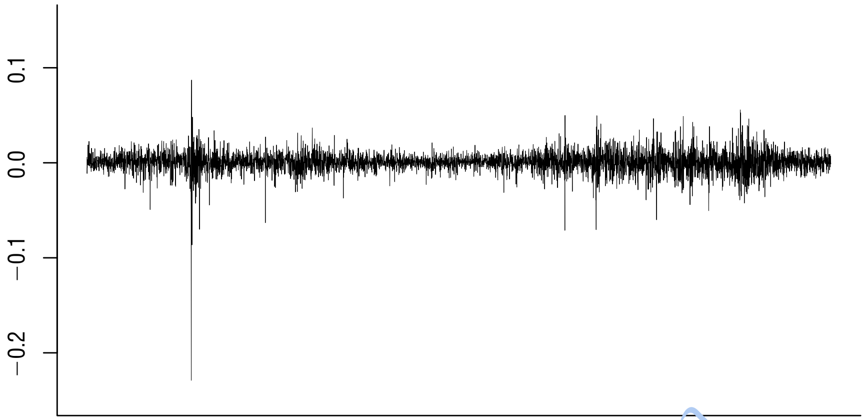

FIGURE 1.1 SPX daily log returns from December 31, 1984, to December 31, 2004. Note the -22.9% return on October 19, 1987!

From a hedging perspective, traders who use the Black-Scholes model must continuously change the volatility assumption in order to match market prices. Their hedge ratios change accordingly in an uncontrolled way: SV models bring some order into this chaos.

A practical point that is more pertinent to a recurring theme of this book is that the prices of exotic options given by models based on Black-Scholes assumptions can be wildly wrong and dealers in such options are motivated to find models that can take the volatility smile into account when pricing these.

In Figure 1.1, we plot the log returns of SPX over a 15-year period; we see that large moves follow large moves and small moves follow small moves (so-called “volatility clustering”). In Figure 1.2, we plot the frequency distribution of SPX log returns over the 77-year period from 1928 to 2005. We see that this distribution is highly peaked and fat-tailed relative to the normal distribution. The Q-Q plot in Figure 1.3 shows just how extreme the tails of the empirical distribution of returns are relative to the normal distribution. (This plot would be a straight line if the empirical distribution were normal.)

Fat tails and the high central peak are characteristics of mixtures of distributions with different variances. This motivates us to model variance as a random variable. The volatility clustering feature implies that volatility (or variance) is auto-correlated. In the model, this is a consequence of the mean reversion of volatility.*

*Note that simple jump-diffusion models do not have this property. After a jump, the stock price volatility does not change.

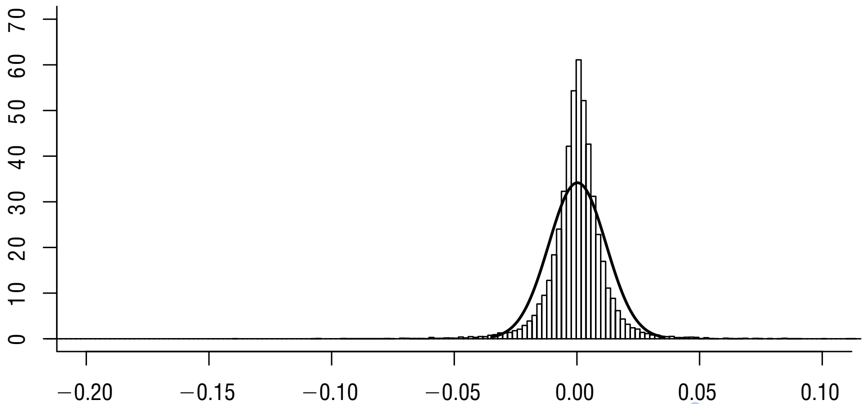

FIGURE 1.2 Frequency distribution of (77 years of) SPX daily log returns compared with the normal distribution. Although the −22.9% return on October 19, 1987, is not directly visible, the x-axis has been extended to the left to accommodate it!

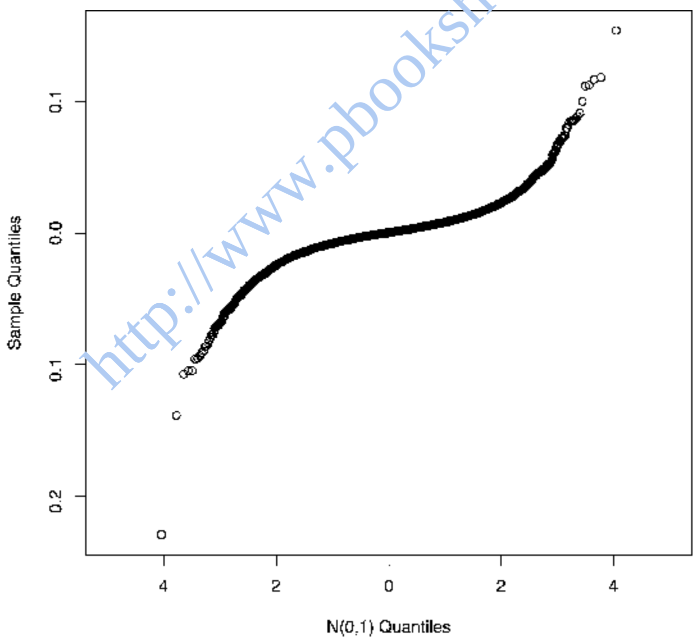

FIGURE 1.3 Q-Q plot of SPX daily log returns compared with the normal distribution. Note the extreme tails.

There is a simple economic argument that justifies the mean reversion of volatility. (The same argument is used to justify the mean reversion of interest rates.) Consider the distribution of the volatility of IBM in 100 years time. If volatility were not mean reverting (i.e., if the distribution of volatility were not stable), the probability of the volatility of IBM being between 1% and 100% would be rather low. Since we believe that it is overwhelmingly likely that the volatility of IBM would in fact lie in that range, we deduce that volatility must be mean reverting.

Having motivated the description of variance as a mean reverting random variable, we are now ready to derive the valuation equation.

Derivation of the Valuation Equation

In this section, we follow Wilmott (2000) closely. Suppose that the stock price S and its variance v satisfy the following SDEs:

$$dS_t = \mu_t \, S_t \, dt + \sqrt{v_t} \, S_t \, dZ_1 \tag{1.1}$$

$$dv_t = \alpha(S_t, v_t, t) \, dt + \eta \, \beta(S_t, v_t, t) \, \sqrt{v_t} dZ_2 \tag{1.2}$$

with

$$\langle dZ_1 \, dZ_2 \rangle = \rho \, dt$$

where μ_t is the (deterministic) instantaneous drift of stock price returns, η is the volatility of volatility and ρ is the correlation between random stock price returns and changes in v_t. dZ_1 and dZ_2 are Wiener processes.

The stochastic process (1.1) followed by the stock price is equivalent to the one assumed in the derivation of Black and Scholes (1973). This ensures that the standard time-dependent volatility version of the Black-Scholes formula (as derived in Section 8.6 of Wilmott (2000) for example) may be retrieved in the limit $\eta \to 0$. In practical applications, this is a key requirement of a stochastic volatility option pricing model as practitioners' intuition for the behavior of option prices is invariably expressed within the framework of the Black-Scholes formula.

In contrast, the stochastic process (1.2) followed by the variance is very general. We don't assume anything about the functional forms of $\alpha(\cdot)$ and $\beta(\cdot)$. In particular, we don't assume a square root process for variance.

In the Black-Scholes case, there is only one source of randomness, the stock price, which can be hedged with stock. In the present case, random changes in volatility also need to be hedged in order to form a riskless portfolio. So we set up a portfolio Π containing the option being priced, whose value we denote by $V(S, v, t)$, a quantity $-\Delta$ of the stock and a quantity $-\Delta_1$ of another asset whose value V_1 depends on volatility. We have

$$\Pi = V - \Delta\, S - \Delta_1\, V_1$$

The change in this portfolio in a time dt is given by

$$\begin{aligned}
d\Pi = &\left\{ \frac{\partial V}{\partial t} + \frac{1}{2}\, v\, S^2\, \frac{\partial^2 V}{\partial S^2} + \rho\, \eta\, v\, \beta\, S\, \frac{\partial^2 V}{\partial v\, \partial S} + \frac{1}{2} \eta^2 v \beta^2 \frac{\partial^2 V}{\partial v^2} \right\} dt \\
&- \Delta_1 \left\{ \frac{\partial V_1}{\partial t} + \frac{1}{2}\, v\, S^2\, \frac{\partial^2 V_1}{\partial S^2} + \rho\, \eta\, v\, \beta\, S\, \frac{\partial^2 V_1}{\partial v\, \partial S} + \frac{1}{2}\, \eta^2\, v\, \beta^2\, \frac{\partial^2 V_1}{\partial v^2} \right\} dt \\
&+ \left\{ \frac{\partial V}{\partial S} - \Delta_1\, \frac{\partial V_1}{\partial S} - \Delta \right\} dS \\
&+ \left\{ \frac{\partial V}{\partial v} - \Delta_1\, \frac{\partial V_1}{\partial v} \right\} dv
\end{aligned}$$

where, for clarity, we have eliminated the explicit dependence on t of the state variables S_t and v_t and the dependence of α and β on the state variables. To make the portfolio instantaneously risk-free, we must choose

$$\frac{\partial V}{\partial S} - \Delta_1 \frac{\partial V_1}{\partial S} - \Delta = 0$$

to eliminate dS terms, and

$$\frac{\partial V}{\partial v} - \Delta_1 \frac{\partial V_1}{\partial v} = 0$$

to eliminate dv terms. This leaves us with

$$\begin{aligned}
d\Pi &= \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} v\, S^2 \frac{\partial^2 V}{\partial S^2} + \rho\, \eta\, v\, \beta\, S \frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2} \eta^2\, v \beta^2 \frac{\partial^2 V}{\partial v^2} \right\} dt \\
&\quad - \Delta_1 \left\{ \frac{\partial V_1}{\partial t} + \frac{1}{2} v\, S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho \eta v\, \beta\, S \frac{\partial^2 V_1}{\partial v \partial S} + \frac{1}{2} \eta^2\, v\, \beta^2 \frac{\partial^2 V_1}{\partial v^2} \right\} dt \\
&= r\, \Pi\, dt \\
&= r (V - \Delta S - \Delta_1 V_1)\, dt
\end{aligned}$$

where we have used the fact that the return on a risk-free portfolio must equal the risk-free rate r, which we will assume to be deterministic for our purposes. Collecting all V terms on the left-hand side and all V_1 terms on the right-hand side, we get

$$\frac{\frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho \eta v \beta S \frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2} \eta^2 v \beta^2 \frac{\partial^2 V}{\partial v^2} + r S \frac{\partial V}{\partial S} - r V}{\frac{\partial V}{\partial v}}$$

$$= \frac{\frac{\partial V_1}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho \eta v \beta S \frac{\partial^2 V_1}{\partial v \partial S} + \frac{1}{2} \eta^2 v \beta^2 \frac{\partial^2 V_1}{\partial v^2} + r S \frac{\partial V_1}{\partial S} - r V_1}{\frac{\partial V_1}{\partial v}}$$

The left-hand side is a function of V only and the right-hand side is a function of V_1 only. The only way that this can be is for both sides to be equal to some function f of the *independent* variables S, v and t. We deduce that

$$\frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho \eta v \beta S \frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2} \eta^2 v \beta^2 \frac{\partial^2 V}{\partial v^2} + r S \frac{\partial V}{\partial S} - r V$$
$$= - \left(\alpha - \phi \beta \sqrt{v} \right) \frac{\partial V}{\partial v} \tag{1.3}$$

where, without loss of generality, we have written the arbitrary function f of S, v and t as $\left(\alpha - \phi \beta \sqrt{v} \right)$, where α and β are the drift and volatility functions from the SDE (1.2) for instantaneous variance.

The Market Price of Volatility Risk $\phi(S, v, t)$ is called the market price of volatility risk. To see why, we again follow Wilmott's argument.

Consider the portfolio Π_1 consisting of a delta-hedged (but not vega-hedged) option V. Then

$$\Pi_1 = V - \frac{\partial V}{\partial S} S$$

and again applying Itô's lemma,

$$d\Pi_1 = \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho \eta v \beta S \frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2} \eta^2 v \beta^2 \frac{\partial^2 V}{\partial v^2} \right\} dt$$
$$+ \left\{ \frac{\partial V}{\partial S} - \Delta \right\} dS + \left\{ \frac{\partial V}{\partial v} \right\} dv$$

Because the option is delta-hedged, the coefficient of dS is zero and we are left with

$$
\begin{aligned}
& d\Pi_1 - r\,\Pi_1\,dt \\
&\quad = \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho \eta v \beta S \frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2} \eta^2 v \beta^2 \frac{\partial^2 V}{\partial v^2} + rS \frac{\partial V}{\partial S} - r\,V \right\} dt \\
&\qquad + \frac{\partial V}{\partial v}\,dv \\
&\quad = \beta\,\sqrt{v}\,\frac{\partial V}{\partial v}\,\left\{ \phi(S, v, t)\,dt + \eta dZ_2 \right\}
\end{aligned}
$$

where we have used both the valuation equation (1.3) and the SDE (1.2) for v. We see that the extra return per unit of volatility risk dZ_2 is given by $\phi(S, v, t)\,dt$ and so in analogy with the Capital Asset Pricing Model, ϕ is known as the *market price of volatility risk*.

Now, defining the *risk-neutral drift* as

$$\alpha' = \alpha - \beta\,\sqrt{v}\,\phi$$

we see that, as far as pricing of options is concerned, we could have started with the risk-neutral SDE for v,

$$dv = \alpha'\,dt + \beta\,\sqrt{v}\,dZ_2$$

and got identical results with no explicit price of risk term because we are in the risk-neutral world.

In what follows, we always assume that the SDEs for S and v are in risk-neutral terms because we are invariably interested in fitting models to option prices. Effectively, we assume that we are imputing the risk-neutral measure directly by fitting the parameters of the process that we are imposing.

Were we interested in the connection between the pricing of options and the behavior of the time series of historical returns of the underlying, we would need to understand the connection between the statistical measure under which the drift of the variance process v is α and the risk-neutral process under which the drift of the variance process is α'. From now on, we act as if we are risk-neutral and drop the prime.

LOCAL VOLATILITY

History

Given the computational complexity of stochastic volatility models and the difficulty of fitting parameters to the current prices of vanilla options, practitioners sought a simpler way of pricing exotic options consistently with the volatility skew. Since before Breeden and Litzenberger (1978), it was understood (at least by floor traders) that the risk-neutral density could be derived from the market prices of European options. The breakthrough came when Dupire (1994) and Derman and Kani (1994)* noted that under risk neutrality, there was a unique diffusion process consistent with these distributions. The corresponding unique state-dependent diffusion coefficient $\sigma_L(S,t)$, consistent with current European option prices, is known as the *local volatility function*.

It is unlikely that Dupire, Derman, and Kani ever thought of local volatility as representing a model of how volatilities actually evolve. Rather, it is likely that they thought of local volatilities as representing some kind of average over all possible instantaneous volatilities in a stochastic volatility world (an "effective theory"). Local volatility models do not therefore really represent a separate class of models; the idea is more to make a simplifying assumption that allows practitioners to price exotic options consistently with the known prices of vanilla options.

As if any proof were needed, Dumas, Fleming, and Whaley (1998) performed an empirical analysis that confirmed that the dynamics of the implied volatility surface were not consistent with the assumption of constant local volatilities.

Later on, we show that local volatility is indeed an average over instantaneous volatilities, formalizing the intuition of those practitioners who first introduced the concept.

A Brief Review of Dupire's Work

For a given expiration T and current stock price S_0, the collection $\{C(S_0, K, T)\}$ of undiscounted option prices of different strikes yields the risk-neutral density function φ of the final spot S_T through the relationship

$$C(S_0, K, T) = \int_K^\infty dS_T\, \varphi(S_T, T; S_0)\, (S_T - K)$$

Differentiate this twice with respect to K to obtain

$$\varphi(K, T; S_0) = \frac{\partial^2 C}{\partial K^2}$$

*Dupire published the continuous time theory and Derman and Kani, a discrete time binomial tree version.

so the Arrow-Debreu prices for each expiration may be recovered by twice differentiating the undiscounted option price with respect to K. This process is familiar to any option trader as the construction of an (infinite size) infinitesimally tight butterfly around the strike whose maximum payoff is one.

Given the distribution of final spot prices S_T for each time T conditional on some starting spot price S_0, Dupire shows that there is a unique risk neutral diffusion process which generates these distributions. That is, given the set of all European option prices, we may determine the functional form of the diffusion parameter (local volatility) of the unique risk neutral diffusion process which generates these prices. Noting that the local volatility will in general be a function of the current stock price S_0, we write this process as

$$\frac{dS}{S} = \mu_t \, dt + \sigma \left(S_t, t; S_0\right) \, dZ$$

Application of Itô's lemma together with risk neutrality, gives rise to a partial differential equation for functions of the stock price, which is a straightforward generalization of Black-Scholes. In particular, the pseudo-probability densities $\varphi \left(K, T; S_0\right) = \frac{\partial^2 C}{\partial K^2}$ must satisfy the Fokker-Planck equation. This leads to the following equation for the undiscounted option price C in terms of the strike price K:

$$\frac{\partial C}{\partial T} = \frac{\sigma^2 \, K^2}{2} \frac{\partial^2 C}{\partial K^2} + \left(r_t - D_t\right) \left(C - K \frac{\partial C}{\partial K}\right) \tag{1.4}$$

where r_t is the risk-free rate, D_t is the dividend yield and C is short for $C\left(S_0, K, T\right)$.

Derivation of the Dupire Equation

Suppose the stock price diffuses with risk-neutral drift $\mu_t \left(= r_t - D_t\right)$ and local volatility $\sigma \left(S, t\right)$ according to the equation:

$$\frac{dS}{S} = \mu_t \, dt + \sigma \left(S_t, t\right) \, dZ$$

The undiscounted risk-neutral value $C\left(S_0, K, T\right)$ of a European option with strike K and expiration T is given by

$$C\left(S_0, K, T\right) = \int_K^\infty dS_T \, \varphi \left(S_T, T; S_0\right) \, \left(S_T - K\right) \tag{1.5}$$

Here $\varphi(S_T, T; S_0)$ is the pseudo-probability density of the final spot at time T. It evolves according to the Fokker-Planck equation:

$$\frac{1}{2}\frac{\partial^2}{\partial S_T^2}\left(\sigma^2 S_T^2\,\varphi\right) - \frac{\partial}{\partial S_T}\left(\mu\,S_T\,\varphi\right) = \frac{\partial \varphi}{\partial T}$$

Differentiating with respect to K gives

$$\frac{\partial C}{\partial K} = -\int_K^\infty dS_T\,\varphi(S_T, T; S_0)$$

$$\frac{\partial^2 C}{\partial K^2} = \varphi(K, T; S_0)$$

Now, differentiating (1.5) with respect to time gives

$$\begin{aligned}\frac{\partial C}{\partial T} &= \int_K^\infty dS_T\left\{\frac{\partial}{\partial T}\varphi\ (S_T, T; S_0)\right\}(S_T - K)\\ &= \int_K^\infty dS_T\left\{\frac{1}{2}\frac{\partial^2}{\partial S_T^2}\left(\sigma^2 S_T^2\varphi\right) - \frac{\partial}{\partial S_T}(\mu\,S_T\,\varphi)\right\}(S_T - K)\end{aligned}$$

Integrating by parts twice gives:

$$\begin{aligned}\frac{\partial C}{\partial T} &= \frac{\sigma^2 K^2}{2}\varphi + \int_K^\infty dS_T\,\mu\,S_T\,\varphi\\ &= \frac{\sigma^2 K^2}{2}\frac{\partial^2 C}{\partial K^2} + \mu\ (T)\left(C - K\frac{\partial C}{\partial K}\right)\end{aligned}$$

which is the Dupire equation when the underlying stock has risk-neutral drift μ. That is, the forward price of the stock at time T is given by

$$F_T = S_0\,\exp\left\{\int_0^T dt\,\mu_t\right\}$$

Were we to express the option price as a function of the forward price $F_T = S_0 \exp\left\{\int_0^T \mu(t)dt\right\}^*$, we would get the same expression minus the drift term. That is,

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2 K^2\frac{\partial^2 C}{\partial K^2}$$

*From now on, $\mu(T)$ represents the risk-neutral drift of the stock price process, which is the risk-free rate $r(T)$ minus the dividend yield $D(T)$.

where C now represents $C\left(F_T, K, T\right)$. Inverting this gives

$$\sigma^2(K, T, S_0) = \frac{\frac{\partial C}{\partial T}}{\frac{1}{2}\, K^2\, \frac{\partial^2 C}{\partial K^2}} \tag{1.6}$$

The right-hand side of equation (1.6) can be computed from known European option prices. So, given a complete set of European option prices for all strikes and expirations, local volatilities are given uniquely by equation (1.6).

We can view equation (1.6) as a *definition* of the local volatility function regardless of what kind of process (stochastic volatility for example) actually governs the evolution of volatility.

Local Volatility in Terms of Implied Volatility

Market prices of options are quoted in terms of Black-Scholes implied volatility $\sigma_{BS}\left(K, T; S_0\right)$. In other words, we may write

$$C\left(S_0, K, T\right) = C_{BS}\left(S_0, K, \sigma_{BS}\left(S_0, K, T\right), T\right)$$

It will be more convenient for us to work in terms of two dimensionless variables: the Black-Scholes implied total variance w defined by

$$w\left(S_0, K, T\right) := \sigma_{BS}^2\left(S_0, K, T\right)\, T$$

and the log-strike y defined by

$$y = \log\left(\frac{K}{F_T}\right)$$

where $F_T = S_0 \exp\left\{\int_0^T dt\, \mu(t)\right\}$ gives the forward price of the stock at time 0. In terms of these variables, the Black-Scholes formula for the future value of the option price becomes

$$\begin{aligned} C_{BS}\left(F_T, y, w\right) &= F_T\left\{N\left(d_1\right) - e^y N\left(d_2\right)\right\} \\ &= F_T\left\{N\left(-\frac{y}{\sqrt{w}} + \frac{\sqrt{w}}{2}\right) - e^y N\left(-\frac{y}{\sqrt{w}} - \frac{\sqrt{w}}{2}\right)\right\} \end{aligned} \tag{1.7}$$

and the Dupire equation (1.4) becomes

$$\frac{\partial C}{\partial T} = \frac{v_L}{2}\left\{\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y}\right\} + \mu\left(T\right)\, C \tag{1.8}$$

with $v_L = \sigma^2(S_0, K, T)$ representing the local variance. Now, by taking derivatives of the Black-Scholes formula, we obtain

$$\begin{aligned}
\frac{\partial^2 C_{BS}}{\partial w^2} &= \left(-\frac{1}{8} - \frac{1}{2\,w} + \frac{y^2}{2\,w^2}\right)\frac{\partial C_{BS}}{\partial w} \\
\frac{\partial^2 C_{BS}}{\partial y \partial w} &= \left(\frac{1}{2} - \frac{y}{w}\right)\frac{\partial C_{BS}}{\partial w} \\
\frac{\partial^2 C_{BS}}{\partial y^2} - \frac{\partial C_{BS}}{\partial y} &= 2\,\frac{\partial C_{BS}}{\partial w}
\end{aligned} \tag{1.9}$$

We may transform equation (1.8) into an equation in terms of implied variance by making the substitutions

$$\begin{aligned}
\frac{\partial C}{\partial y} &= \frac{\partial C_{BS}}{\partial y} + \frac{\partial C_{BS}}{\partial w}\frac{\partial w}{\partial y} \\
\frac{\partial^2 C}{\partial y^2} &= \frac{\partial^2 C_{BS}}{\partial y^2} + 2\frac{\partial^2 C_{BS}}{\partial y \partial w}\frac{\partial w}{\partial y} + \frac{\partial^2 C_{BS}}{\partial w^2}\left(\frac{\partial w}{\partial y}\right)^2 + \frac{\partial C_{BS}}{\partial w}\frac{\partial^2 w}{\partial y^2} \\
\frac{\partial C}{\partial T} &= \frac{\partial C_{BS}}{\partial T} + \frac{\partial C_{BS}}{\partial w}\frac{\partial w}{\partial T} = \frac{\partial C_{BS}}{\partial w}\frac{\partial w}{\partial T} + \mu(T)\; C_{BS}
\end{aligned}$$

where the last equality follows from the fact that the only explicit dependence of the option price on T in equation (1.7) is through the forward price $F_T = S_0 \exp\left\{\int_0^T dt\, \mu(t)\right\}$. Equation (1.4) now becomes (cancelling $\mu(T)\; C$ terms on each side)

$$\begin{aligned}
&\frac{\partial C_{BS}}{\partial w}\frac{\partial w}{\partial T} \\
&= \frac{v_L}{2}\left\{-\frac{\partial C_{BS}}{\partial y} + \frac{\partial^2 C_{BS}}{\partial y^2} - \frac{\partial C_{BS}}{\partial w}\frac{\partial w}{\partial y} + 2\frac{\partial^2 C_{BS}}{\partial y \partial w}\frac{\partial w}{\partial y}\right. \\
&\qquad \left. + \frac{\partial^2 C_{BS}}{\partial w^2}\left(\frac{\partial w}{\partial y}\right)^2 + \frac{\partial C_{BS}}{\partial w}\frac{\partial^2 w}{\partial y^2}\right\} \\
&= \frac{v_L}{2}\frac{\partial C_{BS}}{\partial w}\left\{2 - \frac{\partial w}{\partial y} + 2\left(\frac{1}{2} - \frac{y}{w}\right)\frac{\partial w}{\partial y}\right. \\
&\qquad \left. + \left(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2}\right)\left(\frac{\partial w}{\partial y}\right)^2 + \frac{\partial^2 w}{\partial y^2}\right\}
\end{aligned}$$

Then, taking out a factor of $\frac{\partial C_{BS}}{\partial w}$ and simplifying, we get

$$\frac{\partial w}{\partial T} = v_L \left\{ 1 - \frac{y}{w}\frac{\partial w}{\partial y} + \frac{1}{4}\left(-\frac{1}{4} - \frac{1}{w} + \frac{y^2}{w^2}\right)\left(\frac{\partial w}{\partial y}\right)^2 + \frac{1}{2}\frac{\partial^2 w}{\partial y^2} \right\}$$

Inverting this gives our final result:

$$v_L = \frac{\frac{\partial w}{\partial T}}{1 - \frac{y}{w}\frac{\partial w}{\partial y} + \frac{1}{4}\left(-\frac{1}{4} - \frac{1}{w} + \frac{y^2}{w^2}\right)\left(\frac{\partial w}{\partial y}\right)^2 + \frac{1}{2}\frac{\partial^2 w}{\partial y^2}} \tag{1.10}$$

Special Case: No Skew*

If the skew $\frac{\partial w}{\partial y}$ is zero, we must have

$$v_L = \frac{\partial w}{\partial T}$$

So the local variance in this case reduces to the forward Black-Scholes implied variance. The solution to this is, of course,

$$w(T) = \int_0^T v_L(t)\, dt$$

Local Variance as a Conditional Expectation of Instantaneous Variance

This result was originally independently derived by Dupire (1996) and Derman and Kani (1998). Following now the elegant derivation by Derman and Kani, assume the same stochastic process for the stock price as in equation (1.1) but write it in terms of the forward price $F_{t,T} = S_t \exp\left\{\int_t^T ds\, \mu_s\right\}$:

$$dF_{t,T} = \sqrt{v_t} F_{t,T} dZ \tag{1.11}$$

Note that $dF_{T,T} = dS_T$. The undiscounted value of a European option with strike K expiring at time T is given by

$$C(S_0, K, T) = \mathbb{E}\left[(S_T - K)^+\right]$$

*Note that this implies that $\frac{\partial}{\partial K}\sigma_{BS}(S_0, K, T)$ is zero.

Differentiating once with respect to K gives

$$\frac{\partial C}{\partial K} = -\mathbb{E}\left[\theta\left(S_T - K\right)\right]$$

where $\theta(\cdot)$ is the Heaviside function. Differentiating again with respect to K gives

$$\frac{\partial^2 C}{\partial K^2} = \mathbb{E}\left[\delta\left(S_T - K\right)\right]$$

where $\delta(\cdot)$ is the Dirac δ function.

Now a formal application of Itô's lemma to the terminal payoff of the option (and using $dF_{T,T} = dS_T$) gives the identity

$$d\left(S_T - K\right)^+ = \theta\left(S_T - K\right)\, dS_T + \frac{1}{2}\, v_T\, S_T^2\, \delta\left(S_T - K\right)\, dT$$

Taking conditional expectations of each side, and using the fact that $F_{t,T}$ is a martingale, we get

$$dC = d\mathbb{E}\left[\left(S_T - K\right)^+\right] = \frac{1}{2}\mathbb{E}\left[v_T\, S_T^2\, \delta\left(S_T - K\right)\right] dT$$

Also, we can write

$$\mathbb{E}\left[v_T S_T^2\, \delta\left(S_T - K\right)\right] = \mathbb{E}\left[v_T \,|S_T = K\right]\, \frac{1}{2} K^2\, \mathbb{E}\left[\delta\left(S_T - K\right)\right]$$
$$= \mathbb{E}\left[v_T \,|S_T = K\right]\, \frac{1}{2} K^2\, \frac{\partial^2 C}{\partial K^2}$$

Putting this together, we get

$$\frac{\partial C}{\partial T} = \mathbb{E}\left[v_T \,|S_T = K\right]\, \frac{1}{2} K^2\, \frac{\partial^2 C}{\partial K^2}$$

Comparing this with the definition of local volatility (equation (1.6)), we see that

$$\sigma^2(K, T, S_0) = \mathbb{E}\left[v_T \,|S_T = K\right] \tag{1.12}$$

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