CHAPTER 15

Volatility as an Asset Class and the Smile

1. Introduction to Volatility as an Asset Class

Acceptance of implied volatility as an asset class is growing. The main players are (1) institutional investors, (2) hedge funds, and (3) banks. This increased liquidity facilitates the engineering of structured products with embedded volatility. Also, standardized trading in volatility of volatility and skew becomes possible.

EXAMPLE:

It appears that institutional investors are migrating to four types of strategies for going long exposure to implied volatility. Among the largest institutional investors, variance swap based strategies are the most popular.

Variance swaps offer the easiest and most liquid way to get exposure to volatility. Institutional investors and hedge funds are the target audience for the new services offered in this field. These services enable customers to trade variance swaps through Bloomberg terminals. Volumes in the inter-dealer market and with clients prompted the move. Variance swaps are used to go long or short volatility on an index or equity with a selected maturity.

Pure volatility instruments, such as volatility swaps and variance swaps, make sense for institutional investors, because volatility is both a diversifier on the downside and a global hedge on an equity portfolio.

Institutional investors such as pension funds and insurance companies clearly need to diversify. While they are moving to other asset classes, such as hedge funds, they also do not want to reduce their exposure to equity markets, particularly if there is a good chance of equity markets performing well. With this in mind, they are increasingly turning to long-term volatility strangles.

The main external driver of the current ongoing rise in volatility is M & A activity.

> *Requests for forward volatility strategies to hedge structured products are also on the rise, particularly among private banks. These strategies fit their needs, as dealers sold a lot of forward volatility certificates and warrants to them last year.*
>
> *The launch of newly listed volatility products, such as the Chicago Board Options Exchange's soon-to-launch options on the CBOE S&P500 Volatility Index (VIX), was a key driver of investor demand for volatility products simply because it made it easier to trade volatility. The many investors who cannot trade OTC markets and the demand for similarly structured OTC products both point to a healthy take-up of the CBOE's VIX option contract. This is significant because trading volatility in its pure form as an asset class is established. This may well be a catalyst for encouraging trading in volatility of volatility and skew. (IFR, 2004)*

2. Volatility as Funding

For market professionals and hedge funds, the issue of how to *fund* an investment is as important as the investment itself. After all, a hedge fund would look for the "best way" to borrow funds to carry a position. The best way may sometimes carry a negative interest. In other words, the hedge fund would make money from the investment *and* from the funding itself.

The normal floating Libor funding one is accustomed to think about is "risk-free,"[1] but at the same time may not always carry the lowest funding cost. Suppose a practitioner starts with the standard floating Libor-referenced loan that is rolled over at intervals of length δ in order to fund a long position and then show how volatility can be used as an alternative funding strategy. Also, suppose a long position involves buying a straight (default-free) Eurobond with coupon r_{t_0}. The market professional borrows N and buys the bond. The outcome will be similar to an interest rate swap.

Now suppose the bond under consideration is the liquid emerging market benchmark Brazil-40. In Figure 15-1 this is represented as if it has annual coupon payments over four settlement dates. In general, hedge funds use strategies other than using straightforward Libor funding to buy the bond. One common strategy is called *relative value* trade. Suppose the hedge fund has calculated that the Venezuelan benchmark Eurobond may lose value during the investment period.[2] Then the hedge fund will search for the Venezuelan bond in the repo market, "borrow" the bond (instead of borrowing USD) and then sell it to generate the needed cash of N. Using this cash the hedge fund buys the Brazilian bond. The Venezuelan bond has a coupon of R_{t_0} as the Brazilian bond assumed to be trading *at par* value $N = 100$.

The value of the Venezuelan bond may decline during the investment period and the hedge fund can cover the short bond position at a lower price than the original N.[3]

Now consider the alternative shown in Figure 15-2. If the purpose is funding a position, then why not select an appropriate *volatility*, sell options of value N, and then *delta*-hedge these option positions? In fact, this would fund the bond position with volatility. We analyze it below.

First we know from Chapter 8 that *delta*-hedged short option positions are convex exposures that will pay the *gamma*. These payouts are unknown initially. As market volatility is observed, the hedge is dynamically adjusted, and depending on the market volatility the hedge fund will face a cash outflow equal to *gamma*. To the hedge fund this is similar to paying floating money market interest rates.

[1] See the section on the zero in finance in Chapter 5.

[2] Both bonds are assumed to be in the same currency, say USD, and have similar maturities.

[3] The difference $r_{t_0} - R_{t_0}$ is known as the carry of the position. It could be positive or negative. Obviously positions with positive carry can be continued longer.

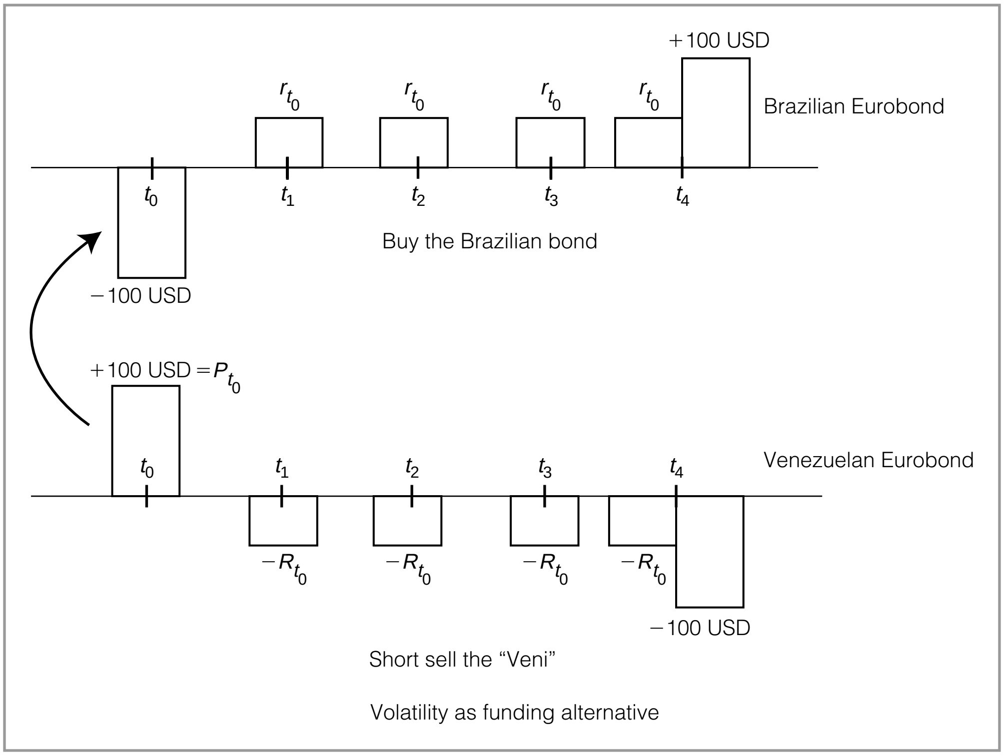

FIGURE 15-1

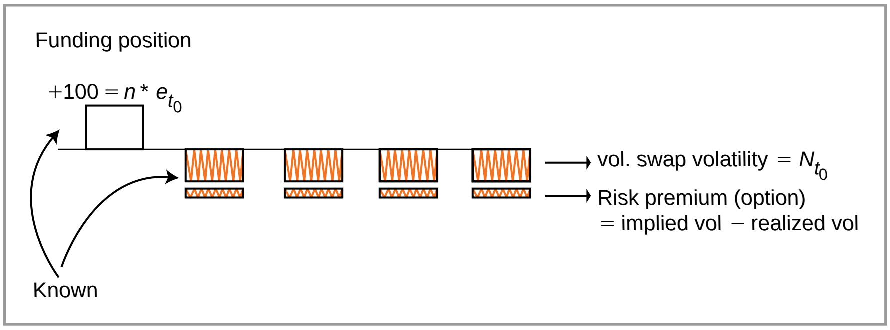

FIGURE 15-2

Note one difference between loan cash flows and volatility cash flows: In volatility funding there is no payback of the principal N at the end of the contract. In this sense the N is borrowed and then paid back gradually over time as *gamma* gains. One example is provided below from the year 2005.

EXAMPLE:

Merrill notes "one of the most overcrowded trades in the market has been to take advantage of the long term trading range," by selling volatility and "earning carry via mortgage-backed securities."

Market professionals use options as funding vehicles for their positions. The main problem with this is that in many cases option markets may not have the depth needed in order to sell large chunks of options. If such selling depresses prices (i.e. volatility), then this idea may be hard to implement no matter how attractive it looks at the outset.

3. Smile

Options were introduced as volatility instruments in Chapter 8. This is very much in line with the way traders think about options. We showed that when we deal with options as volatility instruments mathematically we arrived at the same formula, in this case the same partial differential equation (PDE) as the Black-Scholes PDE. *Mathematically* the approach was identical to the standard textbook treatment that considers options as directional instruments.[4] Yet, although the interpretation in Chapter 8 is more in line with the way traders and option markets think, in that discussion there was still a major missing component.

It turns out that everything else being the same, an out-of-the-money put or call has a higher implied volatility than an ATM call or put. This effect, alluded to several times up to this point, is called the *volatility smile* and is discussed in this chapter. However, in order to do this in this chapter we adopt still another interpretation of options as instruments.

The discussion in Chapter 8 showed that the option price (after some adjustments for interest receipts and payments) is actually related to the *expected gamma* gains due to volatility in the underlying. The interpretation we use in this chapter will show that these expected gains will depend on the option's *strike*. One cost to pay for this interesting result is the need for a different mathematical approach. The advantage is that the smile will be the *natural* outcome. A side advantage is that we will discuss a dynamic hedging strategy other than the well-known *delta*-hedging. In fact, we start the chapter with a discussion of options from a more "recent" point of view which uses the so-called *dirac delta* functions. It is perhaps the best way of bringing the smile explicitly in option pricing.

4. Dirac *Delta* Functions

Consider the integral of the Gaussian density with mean K given below

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{1}{2}\frac{(x-K)^2}{\beta^2}} dx = 1 \tag{1}$$

where β^2 is the "variance" parameter. Let $f(x)$ denote the density:

$$f(x) = \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{1}{2}\frac{(x-K)^2}{\beta^2}} \tag{2}$$

We will use the $f(x)$ as a mathematical *tool* instead of representing a probability density associated with a financial variable. To see how this is done, suppose we consider the values of β that sequentially go from one toward zero. The densities will be as shown in Figure 15-3. Clearly, if β is very small, the "density" will essentially be a spike at K, but still will have an area under it that adds up to one.

[4] On one hand, in this textbook approach, calls are regarded as a bet in increasing prices, and put a bet on decreasing prices. This, however, would be true under the risk-adjusted probability and leaves the wrong impression that calls and puts are different in some sense. On the other hand, the volatility interpretation shows that the calls and puts are in fact the same from the point of view of volatility.

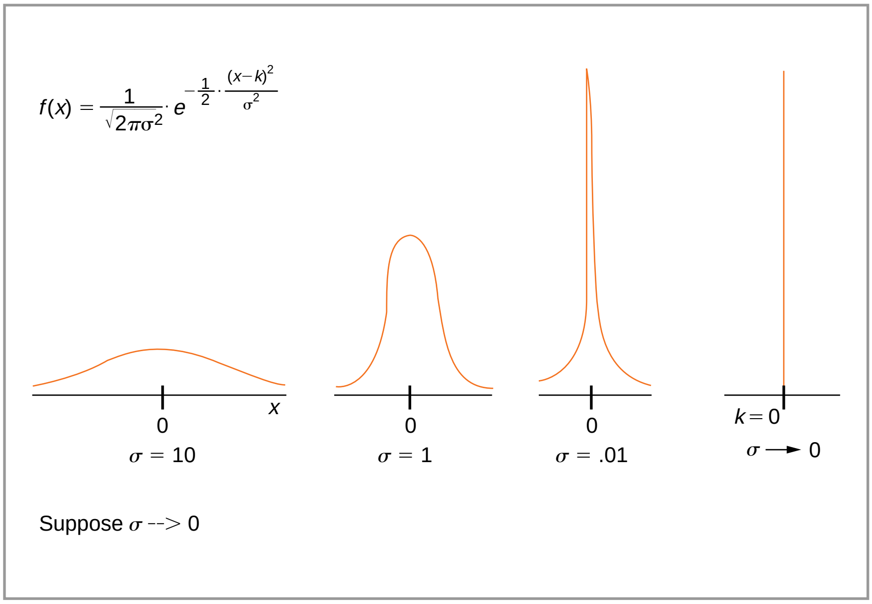

FIGURE 15-3

Now consider the "expectations" calculated with such an $f(x)$. Let $C(x_t)$ be a random value that depends on the random variable x_t, indexed by the time t. Then we can write

$$E[C(x_t)] = \int_{-\infty}^{\infty} C(x_t) f(x_t) dx_t \tag{3}$$

Now we push the β toward zero. The density $f(x_t)$ will become a spike at K. This means that all values of $C(x_t)$ will be multiplied by a probability of almost zero, except the ones around $x_t = K$. After all, at the limit the $f(.)$ is nonzero only around $x_t = K$. Thus at the limit we obtain

$$\lim_{\beta \to 0} \int_{-\infty}^{\infty} C(x_t) f(x_t) dx_t = C(K) \tag{4}$$

The integral of the product of a function $C(x_t)$ and of the $f(x_t)$ as β goes to zero *picks up* the value of the function at $x_t = K$.

Hence we define the *Dirac delta function* as

$$\delta_K(x) = \lim_{\beta \to 0} f(x, K, \beta) \tag{5}$$

Remember that the β determines how close the $f(x)$ is to a spike at K. The integral can then be rewritten as

$$\int_{-\infty}^{\infty} C(K) \delta_K(x) dx_t = C(K) \tag{6}$$

This integral shows the most useful property of dirac *delta* function for our purposes. Essentially, the dirac *delta* picks up the value of $C(x_t)$ at the point $x_t = K$. We now apply this property to option payoffs at expiration.

5. Application to Option Payoffs

The major advantage of the dirac *delta* functions, interpreted as the limits of distributions, is in differentiating functions that have points that cannot be differentiated in the usual sense. There are many such points in option trading. The payoff at the strike K is one example. Knock-in, knock-out barriers is another example. Dirac *delta* will be useful for discussing derivatives at those points.

Before we proceed, for simplicity we will assume in this section that interest rates are equal to zero:

$$r_t = 0 \tag{7}$$

We also assume that the underlying S_T follows the risk-neutral SDE, which in this case will be given by

$$dS_T = \sigma\left(S_t\right) S_t dW_t \tag{8}$$

Note that with interest rates being zero, the drift is eliminated and that the volatility is *not* of the Black-Scholes form. It depends on the random variable S_t. Let

$$\begin{aligned} f(S_T) &= \max[S_T - K, 0] \\ &= (S_T - K)^+ \end{aligned} \tag{9}$$

be the vanilla call option payoff shown in Figure 15-4. The function is not differentiable at $S_T = K$, yet its first order derivative is like a step function. More interestingly, the *second* order derivative can be interpreted as a dirac *delta* function. These derivatives are shown in Figures 15-4 and 15-5.

Now write the equivalent of Ito's Lemma in a setting where functions have kinks as in the option payoff case. This is called *Tanaka's formula* and essentially extends Ito's Lemma to functions that cannot be differentiated at all points. We can write

$$d(S_t - K)^+ = \frac{\partial (S_t - K)^+}{\partial S_t} dS_t + \frac{1}{2}\frac{\partial^2 (S_t - K)^+}{\partial S_t^2}\sigma(S_t)^2 dt \tag{10}$$

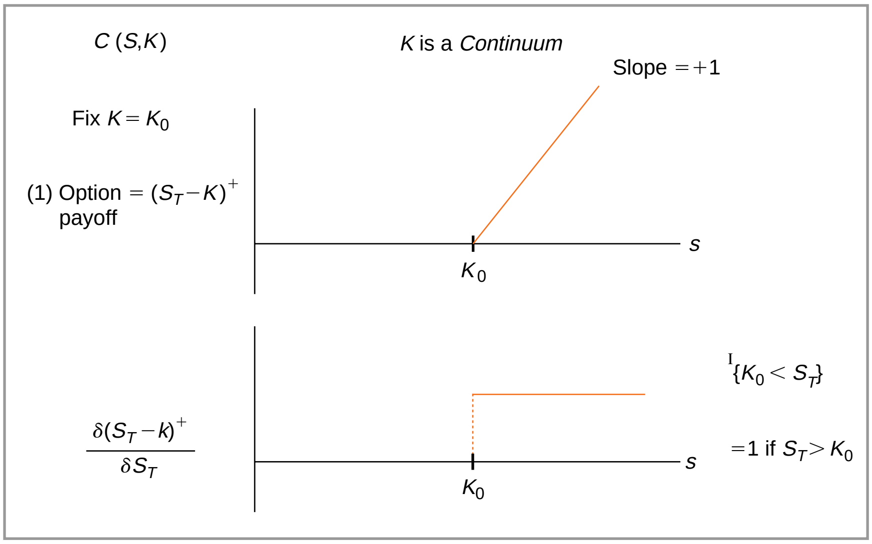

FIGURE 15-4

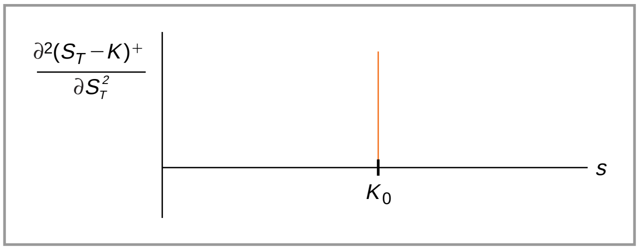

FIGURE 15-5

where we define

$$\frac{\partial (S_t - K)^+}{\partial S_t} = 1_{S_t > K} \tag{11}$$

$$\frac{\partial^2 (S_t - K)^+}{\partial S_t^2} = \delta_K(S_t) \tag{12}$$

Taking integrals from t_0 to T we get:

$$(S_T - K)^+ = (S_{t_0} - K)^+ + \int_{t_0}^{T} 1_{S_t > K} dS_t + \frac{1}{2} \int_{t_0}^{T} \frac{\partial^2 (S_t - K)^+}{\partial S_t^2} \sigma(S_t)^2 dt \tag{13}$$

where the first term on the right-hand side is the time value of the option at time t_0, and is known with certainty. We also know that with zero interest rates, the option price $C(S_{t_0})$ will be given by

$$C(S_{t_0}) = E_{t_0}^{\tilde{P}} (S_T - K)^+ \tag{14}$$

Now, using the risk-adjusted probability $\tilde{P}$, (1) apply the expectation operator to both sides of equation (13), (2) change the order of integration and expectation, and (3) use the property of dirac *delta* functions in eliminating the terms valued at points other than $S_t = K$. We obtain the characterization of the option price as:

$$\begin{aligned} E_t^P\Big[(S_T - K)^+\Big] &= (S_{t_0} - K)^+ + \int_{t_0}^{T} \sigma\left(K\right)^2 \phi_t\left(K\right) dt \\ &= C\left(S_{t_0}\right) \end{aligned} \tag{15}$$

where $\phi_t(.)$ is the continuous density function that corresponds to the risk-adjusted probability of S_t.[5] This means that the time value of the option depends (1) on the intrinsic value of the option, (2) on the *time spent* around K during the life of the option, and (3) on the *volatility at that strike*, $\sigma(K)$.

The main point for us is that this expression shows that the option price depends *not* on the overall volatility, but on the volatility of S_t *around* K. This is exactly what the notion of volatility smile is.

5.1. An Interpretation of Dynamic Hedging

There are many dynamic strategies that replicate an option's final payoff. The best known is *delta* hedging. In *delta* hedging the financial engineer will buy or sell the $delta = D_t$ units

[5] We assume that a density exists.

of the underlying, borrow any necessary funds, and adjust the D_t as the underlying S_t moves over time. As $t \to T$, the expiration date, this will duplicate the option's payoff. This is the case because, as the time value goes to zero the option price merges with $(S_T - K)^+$.

However, there is an alternative dynamic hedging procedure that is similar to the approach adopted in the previous section. The dynamic hedging technique, called stop-loss strategy, is as follows.

In order to *replicate* the payoff of the long call, hold *one* unit of S_t if $K < S_t$. Otherwise hold *no* S_t. This strategy requires that as S_t crosses level K, we keep adjusting the position as soon as possible. Either buy one unit of S_t, or sell the S_t immediately as S_t crosses the K *from left to right* or *from right to left* respectively. The P/L of this position is given by the term

$$\frac{1}{2}\int_{t_0}^{T} \frac{\partial^2 (S_{t_0} - K)^+}{\partial S_t^2} \sigma (S_t)^2 \, dt \tag{16}$$

Clearly the switches at $S_t = K$ cannot be done instantaneously at zero cost. The trader is moving with time Δ while the underlying Wiener process is moving at a *faster* rate $\sqrt{\Delta}$. These adjustments are shown in Figures 15-6 and 15-7. The resulting hedging cost is the options value.

6. Breeden-Litzenberger Simplified

The so-called Breeden-Litzenberger Theorem is an important result that shows how one can back out risk-adjusted probabilities from liquid arbitrage-free option prices. In this section we discuss a trader's approach to Breeden-Litzenberger. This approach will show the theoretical relevance of some popular option strategies used in practice. Below, we use a simplified framework which could be generalized in a straightforward way. However, we will not generalize these results, but instead in the following section use the dirac *delta* approach to prove the Breeden-Litzenberger Theorem.

Consider a simple setting where we observe prices of four liquid European call options, denoted by $\{C_t^1, \ldots, C_t^4\}$. The options all expire at time T with $t < T$. The options have

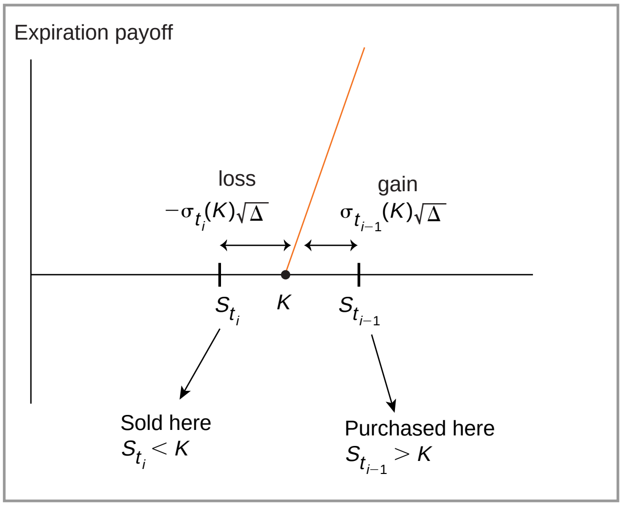

FIGURE 15-6

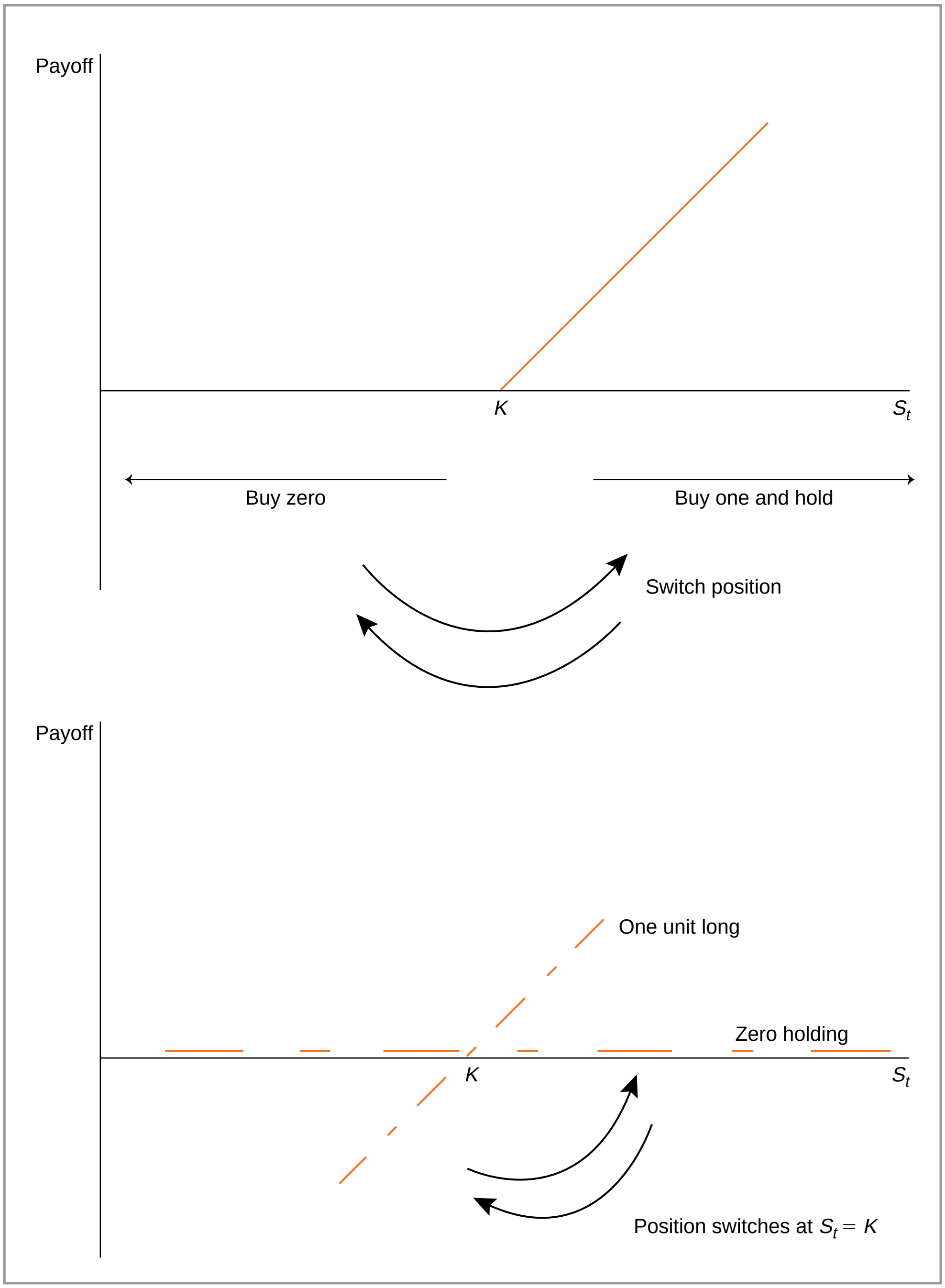

FIGURE 15-7

strike prices denoted by $\{K^1 < \cdots < K^4\}$ with

$$K^i - K^{i-1} = \Delta K \tag{17}$$

Hence, the strike prices are equally spaced. Apart from the assumption that these options are written on the same underlying S_t which does not pay dividends, we make no distributional assumption about S_t. In fact the volatility of S_t can be stochastic and the distribution is not necessarily log normal.

Finally we use the Libor rate L_t to discount cash flows to be received at time T. The discount factor will then be given by

$$\frac{1}{(1 + L_t\delta)} \tag{18}$$

Next we define a simple probability space. We assume that the strike prices define the four *states of the world* where S_T can end up. Hence the state space is discrete and is assumed to be made of only four states, $\{\omega^1, . . . , \omega^4\}$.[6]

$$\omega^i = K^i \tag{19}$$

We then have four risk-adjusted probabilities associated with these states defined as follows:

$$p^1 = P\left(S_T = K^1\right) \tag{20}$$

$$p^2 = P\left(S_T = K^2\right) \tag{21}$$

$$p^3 = P\left(S_T = K^3\right) \tag{22}$$

$$p^4 = P\left(S_T = K^4\right) \tag{23}$$

The arbitrage-free pricing of Chapter 11 can be applied to these vanilla options:

$$C_t^i = \frac{1}{(1 + L_t\delta)} E_t^{\tilde{P}} \left[(S_T - K, 0)^+\right] \tag{24}$$

The straightforward application of this formula using the probabilities p^i gives the following pricing equations, where possible payoffs are weighed by the corresponding probabilities.

$$C_t^1 = \frac{1}{(1 + L_t\delta)} \left[p^2 \Delta K + p^3 (2\Delta K) + p^4 (3\Delta K)\right] \tag{25}$$

$$C_t^2 = \frac{1}{(1 + L_t\delta)} \left[p^3 (\Delta K) + p^4 (2\Delta K)\right] \tag{26}$$

$$C_t^3 = \frac{1}{(1 + L_t\delta)} \left[p^4 (\Delta K)\right] \tag{27}$$

Next we calculate the first differences of these option prices.

$$C_t^1 - C_t^2 = \frac{1}{(1 + L_t\delta)} \left[p^2 \Delta K + p^3 (\Delta K) + p^4 (\Delta K)\right] \tag{28}$$

$$C_t^2 - C_t^3 = \frac{1}{(1 + L_t\delta)} \left[p^3 (\Delta K) + p^4 (\Delta K)\right] \tag{29}$$

Finally, we calculate the *second* difference and obtain the following interesting result:

$$\left(C_t^1 - C_t^2\right) - \left(C_t^2 - C_t^3\right) = \frac{1}{(1 + L_t\delta)} p^2 \Delta K \tag{30}$$

Divide by ΔK twice to obtain

$$\frac{\Delta^2 C}{\Delta K^2} = \frac{1}{(1 + L_t\delta)\,\Delta K} p^2 \tag{31}$$

[6] The following discussion can continue unchanged by assuming n discrete states.

where

$$\Delta^2 C = \left(C_t^1 - C_t^2\right) - \left(C_t^2 - C_t^3\right) \tag{32}$$

This is the well-known Breeden-Litzenberger result in this very simple environment. It has interesting implications for the options trader.

Note that

$$\left(C_t^1 - C_t^2\right) - \left(C_t^2 - C_t^3\right) = \left(C_t^1 + C_t^3\right) - 2C_t^2 \tag{33}$$

In other words, this is an option position that is long two wings and short the center twice. In fact this is a butterfly centered at K_2. It turns out that the arbitrage-free market value of this butterfly multiplied by the $(1 + L_t\delta)\Delta K$ gives the risk-adjusted probability that the underlying S_t will end up at state K_2. Letting $\Delta K \to 0$ we get

$$\frac{\partial^2 C}{\partial K^2} = \frac{1}{(1 + L_t\delta)} \phi\left(S_T = K\right) \tag{34}$$

where $\phi(S_T = K)$ is the (conditional) risk adjusted *density* of the underlying at time T.[7]

This discussion illustrates one reason why butterflies are traded as vanilla instruments in option markets. They yield the probability associated with their center. Below we prove the Breeden-Litzenberger result using the dirac *delta* function.

6.1. The Proof

The idea behind the Breeden-Litzenberger result has been discussed before. It rests on the fact that by using liquid and arbitrage-free options prices we can back out the risk-adjusted probabilities associated with various states of the world in the future. The probabilities will relate to the future values of the underlying price, the S_T.

The theorem asserts that (a) if a continuum of European vanilla option prices exist for all $0 \leq K$, and (b) if the function giving the $C(S_t, K)$ is twice differentiable with respect to K, then we have

$$\frac{\partial^2 C}{\partial K^2} = \frac{1}{(1 + L_t\delta)} \phi\left(S_T = K\right) \tag{35}$$

Where $\phi\left(S_T = K|S_{t_0}\right)$ is the conditional risk-adjusted density of the S_T. In other words, if we had a continuum of vanilla option prices, we could obtain the risk-adjusted density with a straightforward differentiation. We now prove this using the dirac *delta* function $\delta_K(S_T)$.

Apply the twice differential operator to the definition of both sides of the arbitrage-free price $C(S_t, K)$. By definition, this means

$$\frac{\partial^2}{\partial K^2} C(S_t, K) = \frac{1}{(1 + L_t\delta)} \frac{\partial^2}{\partial K^2} \int_0^\infty (S_T - K)^+ \phi(S_T) dS_T \tag{36}$$

Assuming that we can interchange the operators and realizing that $\phi\left(S_T\right)$ does not depend on the K we obtain

$$\frac{\partial^2}{\partial K^2} C\left(S_t, K\right) = \frac{1}{(1 + L_t\delta)} \int_0^\infty \frac{\partial^2}{\partial K^2} \left(S_T - K\right)^+ \phi\left(S_T\right) dS_T \tag{37}$$

[7] Remember that if the density at x_0 is $f(x_0)$, then $f(x_0)dx$ is the probability of ending around x_0. In other words we have

$$p^2 \sim \phi\left(S_T = K^2\right) \Delta K$$

But

$$\frac{\partial^2}{\partial K^2}(S_T - K)^+ = \delta_K(S_T) \tag{38}$$

is a dirac *delta*, which means that

$$\frac{\partial^2}{\partial K^2}C(S_t, K) = \frac{1}{(1+L_t\delta)}\int_0^{\infty} \delta_K(S_T)\phi(S_T)\,dS_T \tag{39}$$

The previous discussion and equation (4) tells us that in this integral the $\phi(S_T)$ is being multiplied by zero everywhere except for $S_T = K$. Thus,

$$\frac{\partial^2}{\partial K^2}C(S_t, K) = \frac{1}{(1+L_t\delta)}\phi(S_T = K) \tag{40}$$

To recover the risk-adjusted density just take the second partial of the European vanilla option prices with respect to K. This is the Breeden-Litzenberger result.

7. A Characterization of Option Prices as *Gamma* Gains

The question then is, how does a trader "characterize" an option using these hedging gains? First of all, in liquid option markets the order flow determines the price and the trader does not have to go through a pricing exercise. But still, can we use these trading gains to represent the frame of mind of an options trader?

The discussion in the previous section provides a hint about this issue. The trader buys or sells an option with strike price K. The cash needed for this transaction is either borrowed or lent. Then the trader *delta* hedges the option. Finally, this hedge is adjusted as the underlying price fluctuates *around the initial* S_{t_0}.

According to this, the trader could add the (discounted) future gains (payouts) from these hedge adjustments and this would be the true *time-value* of the option, besides interest or other expenses. The critical point is that these future gains need to be calculated at the initial *gamma*, evaluated at the initial S_{t_0}, and adjusted for passing time.

We can explain this statement. First, for simplicity assume interest rates are equal to zero. We then let the price of the vanilla call be denoted by $C(S_t, t)$. Then by definition we have

$$C(S_{t_0}, T) = \text{Max}\left[S_{t_0} - K, 0\right] \tag{41}$$

This will be the future value of the option if the underlying ended up at the S_{t_0} at time T. Now, this value is equal to the initial price plus how much the time value has changed between t_0 and T,

$$C(S_{t_0}, T) = C(S_{t_0}, t_0) + \int_{t_0}^{T} \left.\frac{\partial C}{\partial t}\right|_{S_t = S_{t_0}} dt \tag{42}$$

Now, we know from the Black-Scholes partial differential equation that

$$\frac{\partial C}{\partial t} = \frac{1}{2}\frac{\partial^2 C}{\partial S_t^2}\sigma_t^2(S_t, t) \tag{43}$$

Substituting and reorganizing equation (42) above becomes

$$C(S_{t_0}, t_0) = \text{Max}\left[S_{t_0} - K, 0\right] + \int_{t_0}^{T} \frac{1}{2}\frac{\partial^2 C(S_{t_0}, t)}{\partial S_t^2}\sigma_t^2(S_{t_0}, t)\,dt \tag{44}$$

Note that on the right-hand side the integral is evaluated at the constant S_{t_0} so we don't need to take expectations.

According to this, the trader is valuing the option at time t_0 by adding the intrinsic value and the *gamma* gains evaluated with a *gamma* at S_{t_0} and a volatility centered at S_{t_0}. Still, the time t changes and this will change the *gamma* over time. Thus, it is important to realize that the trader is *not* valuing the option by looking at the expected value of the future *gamma* gains evaluated at *random* future S_t. The *gamma* is evaluated at the initial S_{t_0}.

7.1. Relationship to Tanaka's Formula

The discussion above is also consistent with the option interpretation obtained using Tanaka's formula. Consider the value of the option as shown above, again

$$C(S_{t_0}, t_0) = \text{Max}\left[S_{t_0} - K, 0\right] + \int_{t_0}^{T} \frac{1}{2}\frac{\partial^2 C\left(S_{t_0}, t\right)}{\partial S_t^2}\sigma_t^2\left(S_{t_0}, t\right)dt \tag{45}$$

Now we know from Breeden-Litzenberger that

$$\frac{\partial^2 C\left(S_{t_0}, t\right)}{\partial S_t^2}\sigma_t^2\left(S_{t_0}, t\right) = \Phi(S_{t_0}, t|S_{t_0}) \tag{46}$$

where the $\Phi(S_{t_0}, t|S_{t_0})$ is the risk-adjusted density of S_t at time t. Substituting this in the option value

$$C(S_{t_0}, t_0) = \text{Max}\left[S_{t_0} - K, 0\right] + \int_{t_0}^{T} \frac{1}{2}\sigma_t^2\left(S_{t_0}, t\right)\Phi(S_{t_0}, t|S_{t_0})dt \tag{47}$$

This is the same equation we obtained by using dirac *delta* functions along with the Tanaka formula. The second term on the right-hand side was called *local time*. In this case the local time is evaluated for the ATM option with strike $K = S_{t_0}$.

8. Introduction to the Smile

Markets trade *many* options with the same underlying, but different strike prices and different expirations. Does the difference in strike price between options that are identical in every other aspect have any important implications?

At first, the answer to this question seems to be no. After all, vanilla options are written on an underlying, with say, price S_t, and this price will have only *one* volatility at any time t, regardless of the strike price K_i. Hence, it appears that, regardless of the differences in the strike price, the implied volatility of options written on the same underlying, with the same expiration, should be the same.

Yet, this first impression is wrong. In reality, options that are identical in every respect, except for their strike, in general, have *different* implied volatilities. Overall, the more out-of-the-money a call (put) option is, the higher is the corresponding implied volatility. This well-established empirical fact is known as the volatility *smile*, or volatility *skew*, and has major implications for hedging, pricing, and marking-to-market of many important instruments. In the remainder of this chapter, we discuss the volatility smile and skews using caps and floors as vehicles for conveying the main ideas. This will indirectly give us an opportunity to discuss the engineering of this special class of convex instruments.

From this point and on, in this chapter we will use the term *smile* only. This will be the case even when the smile is, in fact, a one-sided *skew*. However, whenever relevant, we will point out the differences.

9. Preliminaries

The volatility smile has important implications for trading, hedging, and pricing financial instruments. To illustrate how far things have come in this area, we look at a position taken with the objective of benefiting from abnormal conditions regarding the volatility smile.

We can trade stocks, bonds, or, as we have seen before, the slope of the yield curve. We may, for example, expect that the long-term yields will decline *relative to* the short-term yields. This is called a flattening of the yield curve and it invites curve-flattening strategies that (short) sell short maturities, and buy long maturities. This can be done using cash instruments (i.e., bonds) or swaps.

In any case, such trades have become routine in financial markets. A more recent relative value trade relates to the volatility smile. Consider the following episode.

EXAMPLE:

Over the last month, European equity options traders have seen interest by contrarian investors, namely hedge funds, in buying at-the-money volatility and selling out-of-the-money volatility to take advantage of a skew in volatility levels in certain markets.

. . . the skew trade involves an investor buying at-the-money vol and selling out-of-the-money vol. Due to supply/demand pressures, the level of out-of-the-money vol sometimes rises higher than normal. In other words, the spread between out, and at-the-money-volatility increases, causing a so-called skew. Investors put on the trade in anticipation of the skew dissipating.

[A trader] explained that along with the bull run of US and European equity markets has come a sense of unease among some investors regarding a downturn. Many have thus sought protection via over-the-counter put contracts. Because out-of-the-money puts are usually cheaper than at-the-money puts, many investors have opted for the former. The heavy volume has caused out-of-the-money vol levels to rise. Many investors want crash protection but today puts are too expensive. So, instead of buying today at 100 they buy puts at 80. (Based on an article in Derivatives Week).

According to this example, equity investors that had heavily invested during the "stock market bubble" of the 1990s were looking for *crash protection*. They were long equities and would have suffered significant losses if markets crashed. Instead of selling the stocks that they owned, they bought put options. With a put an investor has the right to *sell* the underlying stocks at a predetermined price, say, K. If market price declines below K, the investor would have some protection.

According to the reading, the large number of investors who were willing to buy puts increased, first, the at-the-money (ATM) volatility.[8] The ATM options became expensive. To lower the cost of insurance sought, investors instead bought options that were, according to the reading, 20% out-of-the-money. These options were cheaper. But as more and more investors bought them, the out-of-the-money volatility started to increase relative to at-the-money volatility of the same option series. This led to an abnormally steep skew.[9]

[8] This is, of course, somewhat circular. If there is a fear of a crash, one would normally expect such an increase in the volatility anyway.

[9] These out-of-the-money puts would still be cheaper in monetary terms when compared with ATM puts. Only, the volatility that they imply would be higher. Another way of saying this is that if we plugged the ATM volatility into the Black-Scholes formula for these out-of-the-money options, they would end up being even cheaper.

The reading suggests that this "abnormal" skew may have attracted some hedge funds who expected the abnormality to disappear in the longer run. According to one theory, these funds sold out-of-the-money volatility and bought ATM volatility. This position will make money if the skew *flattens* and out-of-the-money volatility decreases *relative to* the ATM volatility.[10] As this example shows, the skew, or smile, should be considered as an integral part of financial markets activity. However, as we see in this section, its existence causes several complications and difficulties in financial modeling and in risk management.

10. A First Look at the Smile

The volatility smile can be a confusing notion, and we need to discuss some preliminary ideas before getting into the mechanics of pricing and market applications. It is well known that the Black-Scholes assumptions are not very realistic. And yet, the Black-Scholes formula is routinely used by options traders, although these traders know better than anybody else that the assumptions behind the model are problematic. One of the major Black-Scholes assumptions, for example, is that volatility is *constant* during the life of an option. How can a trader still use the Black-Scholes formula if the realized volatility is known to fluctuate significantly during the life of the option?

If this Black-Scholes assumption is violated, wouldn't the price given by the Black-Scholes formula be "wrong," and, hence, the volatility implied by the formula be erroneous? This question needs to be carefully considered. In the end, we will see that there really are no inconsistencies in traders' behavior. We can explain this as follows.

1. First, note that the Black-Scholes formula is *simple* and depends on a small number of parameters. In fact, the only major parameter that it depends on is the volatility, σ. A simple formula has some advantages. It is easy to understand and remember. But, more importantly, it is also easy to realize *where* or *when* it may go wrong. A simple formula permits developing ways to correct for any inaccuracies *informally* by making subjective adjustments during trading. The Black-Scholes formula has one parameter, and it may be easier to remember how to "adjust" this parameter to cover for the imperfections of the formula.[11]
2. An important aspect of the Black-Scholes formula is that it has become a *convention*. In other words, it has become a *standard* among professionals and also in computer platforms. The formula provides a way to connect a volatility quote to a dollar value attached to this quote. This way traders use the *same* formula to put a dollar value on a volatility number quoted by the market. This helps in developing common platforms for hedging, risk managing, and trading volatility.
3. Thus, once we accept that the use of the Black-Scholes formula amounts to a convention, and that traders differ in their selection of the value of the parameter σ, then the critical process is no longer the option price, but the volatility. This is one reason why in many markets, such as caps, floors, and swaptions markets, the volatility is quoted *directly*.

[10] However, a similar effect would be observed if investors were unwinding their previous insurance bought when markets were at, say, 120 and buying new insurance at $K = 80$. This is equivalent to rolling over their protection. Hence, it is difficult to tell what the real driving force behind this observation is, namely, whether it is due to speculative relative value plays or simply rolling over the positions.

[11] In the theory of prediction, there is the notion of *parsimony*. During a prediction exercise it is costly to have too many parameters because errors are more likely to occur. The notion applies to the numerical calculation of complex options prices. If a model has fewer parameters to be calibrated, the likelihood of making mistakes decreases.

One way to account for the imperfections of the Black-Scholes assumptions would be for traders to adjust the volatility parameter.

4. However, the convention creates new risks. Once the underlying is the volatility process, another issue emerges. For example, traders could add a *risk premium* to quoted volatilities. Just like the risk premium contained in asset prices, the quotes on volatility may incorporate a risk premium.

The volatility smile and its generalization, the *volatility surface*, could then contain a great deal of information concerning the implied volatilities and any arbitrage relations between them. Trading, pricing, hedging, and arbitraging of the smile thus become important.

11. What Is the Volatility Smile?

Consider the Black-Scholes world with vanilla European call and put options written on the equity price (index), S_t, that expire on the same date T. Let K_i denote the ith strike of the option series; and σ_i the *constant* Black-Scholes instantaneous (implied) volatility coefficient for the strike K_i. Finally, let r be the constant risk-free rate.

The Black-Scholes setting makes many assumptions beyond that of constant volatility. In particular, the underlying equity does not make any dividend payments, and there are no transaction costs, tax issues, or regulatory costs. Finally, S_t is assumed to follow the *geometric* stochastic differential equation (SDE)

$$dS_t = \mu S_t dt + \sigma S_t dW_t \qquad t \in [0, \infty) \tag{48}$$

where W_t is a Wiener process defined under the probability P. Here, the parameter μ may also depend on S_t. The crucial assumption is that the *diffusion* component is given by σS_t. This is the assumption that we are concerned with in this chapter. The Black-Scholes setting assumes that the absolute volatility during an infinitesimally small interval dt is given (heuristically) by[12]

$$\sqrt{E_t^P[(dS_t - \mu S_t dt)^2]} = \sigma S_t \sqrt{dt} \tag{49}$$

Thus, for a small interval, Δ, we can write the *percentage* volatility approximately as

$$\frac{\sqrt{E_t^P[(\Delta S_t - \mu S_t \Delta)^2]}}{S_t} \cong \sigma\sqrt{\Delta} \tag{50}$$

According to this, as S_t changes, the percentage volatility during intervals of length Δ remains approximately constant.

In this environment, a typical put option's price is given by the Black-Scholes formula

$$P(S_t, K, \sigma, r, T) = -S_t N(-d_1) + K e^{-r(T-t)} N(-d_2) \tag{51}$$

with

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(\frac{1}{2}\sigma^2 + r\right)(T-t)}{\sigma\sqrt{T-t}} \tag{52}$$

$$d_2 = d_1 - \sigma\sqrt{(T-t)} \tag{53}$$

[12] Here, dS_t is an infinitesimally small change in the price. It is only a symbolic way of writing small changes, and the expectation of such infinitesimal increments can only be heuristic.

Suppose the markets quote implied volatility, σ. To get the monetary value of an option with strike K_i, the trader will put the current values of S_t, t, and r and the quoted value of the implied volatility σ_i at which the trade went through, in this formula. According to this interpretation, the Black-Scholes formula is used to assign a dollar value to a quoted volatility. Conversely, given the correct option price $P(.)$, the implied volatility, σ_i, for the K_i-put could be extracted.

We can now define the volatility smile within this context. Consider a series of *T*-expiration, *liquid*, and arbitrage-free out-of-the-money put option prices indexed by the strike prices K_i, denoted respectively by P_{K_i}:

$$P_{K_1}, \ . \ . \ . \ , P_{K_n} \tag{54}$$

for

$$K_n < \cdots < K_1 < K_0 = S_t \tag{55}$$

According to this, the K_0-put is at-the-money (ATM) and, as K_i decreases, the puts go deeper out-of-the-money. See Figure 15-8, for an example.

Then, given the (bid or ask) option prices, we can use the Black-Scholes formula *backward* and extract the σ_i that the trader used to conclude the deal on the P_{K_i}. If the assumptions of the Black-Scholes world are correct, all the implied volatilities would turn out to be the same

$$\sigma_{K_0} = \sigma_{K_1} = \cdots = \sigma_{K_n} = \sigma \tag{56}$$

since the put options would be identical except for their strike price. Thus, in a market that conforms to the Black-Scholes world, the traders would use the *same* σ in the Black-Scholes put formula to obtain each $P_{K_i}, i = 0, . \ . \ . , n$. Going backward, we would then recover the same constant σ from the prices.[13]

Yet, if we conducted this exercise in reality with observed option prices, we would find that the implied volatilities would satisfy

$$\sigma_{K_0} < \sigma_{K_1} < \cdots < \sigma_{K_n} \tag{57}$$

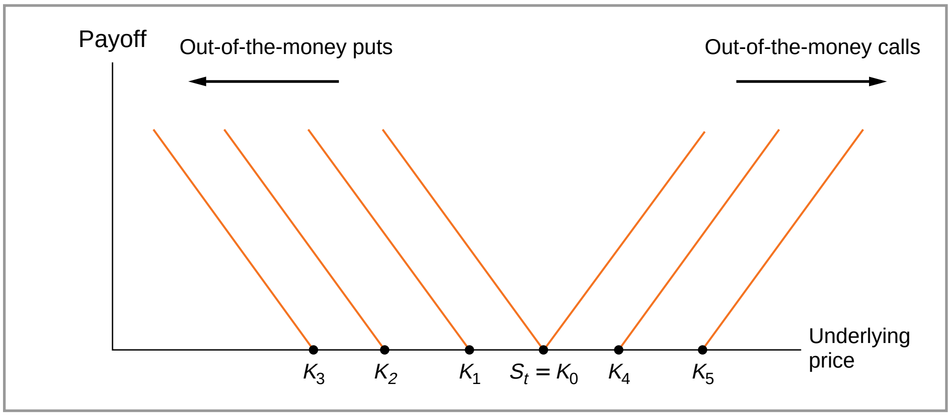

FIGURE 15-8

[13] This exercise requires that the put values were indeed obtained at the same time *t* and were identical in all other aspects except for the K_i.

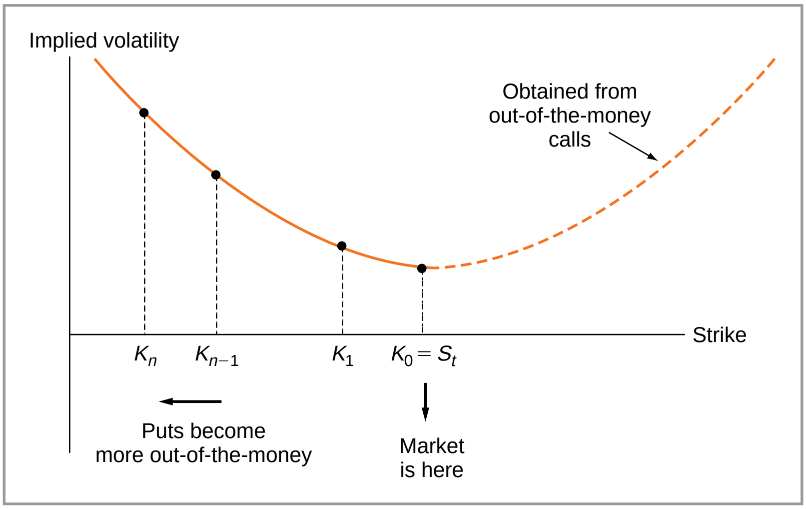

FIGURE 15-9

In other words, the more out-of-the-money the put option is, the higher the corresponding implied volatility. As a result, we would obtain a "smiley" curve.

We can also use the implied vols from progressively out-of-the-money call options and obtain, depending on the underlying instrument, the second half of the smile, as shown in Figure 15-9.

11.1. *Some Stylized Facts*

Volatility smiles observed in reality seem to have the following characteristics,

1. Options written on *equity* indices yield, in general, a nonsymmetric *one-sided* "smile" as shown in Figure 15-10a. For this reason, they are often called *skews*.
2. The *FX markets* are quite different in this respect. They yield a more or less *symmetric* smile, as in Figure 15-10b. However, the smile will rarely be exactly symmetric and it is routine practice in foreign exchange markets to trade this asymmetry using risk reversals.
3. Options on various *interest rates* yield a more *monotonous* one-sided smile than the equity indices. The fact that "smile" patterns vary from market to market would suggest, on the surface, that there are different explanations involved.

It is also natural to think that the *dynamics* of the smile vary depending on the sector. This point is relevant for risk management, running swaption, cap/floor books, and volatility trading. But, before we discuss it, we consider an example.

EXAMPLE:

Table 15-1 displays all the options written on the S&P100 index with a very short expiration. These data were obtained from live quotes early in the morning, so few trades had passed. However, the option bid-ask quotes were live, in the sense that reasonably sized trades could be conducted on them.

When the data were gathered, the underlying was trading at 589.14. We use 12 out-of-the-money puts and 9 out-of-the-money calls to obtain the Black-Scholes implied volatilities.

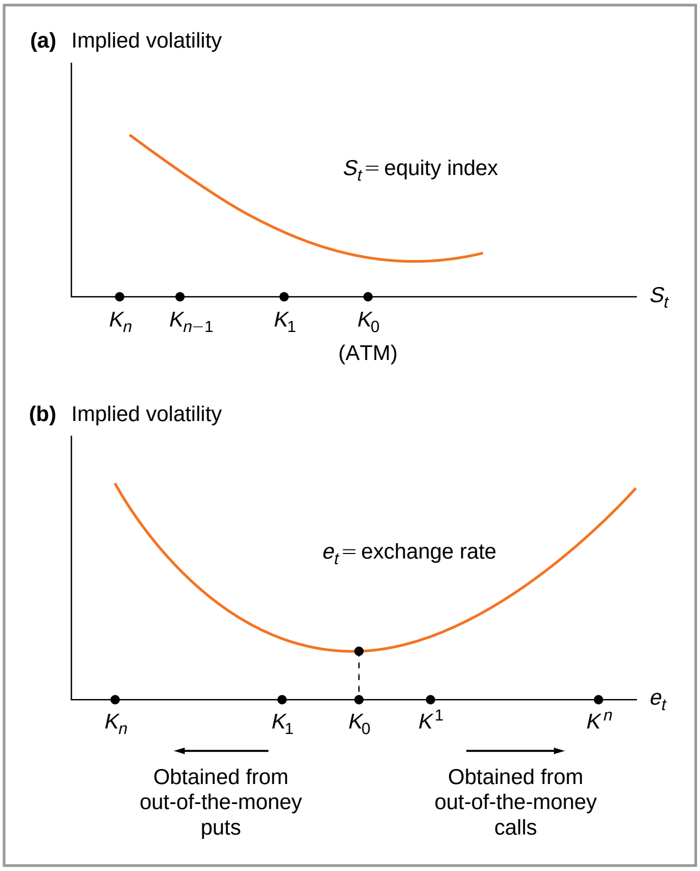

FIGURE 15-10

The interest rate is taken to be 1.98%, and the time to expiration was 8/365. Using these values and the bid prices for the options given in the table, the equations

$$C(S_t, K_i, r, T, \sigma_i) = C_i \tag{58}$$

$$P(S_t, K_j, r, T, \sigma_j) = P_j \tag{59}$$

were solved for the implied vols of calls $\{\sigma_i\}$ *and the implied vols of puts* $\{\sigma_j\}$, *the* C_i *and the* P_j *being observed option prices.*

The resulting vols were plotted against $\frac{K_i}{S_t}$ *in Figure 15-11. We see a pronounced smile. For example, the January 400 put, which traded at about 32% out-of-the-money, had a volatility of about 26%, while the ATM option traded at an implied volatility of 18.5%.*

OEX options are of American style and this issue was ignored in the example above. This would introduce an upward bias in the calculated volatilities. This bias is secondary for our purpose, but in real trading should be corrected. One correction has been suggested by Barone-Adesi and Whaley (1987).

TABLE 15-1. OEX Options with January 18, 2002, Expiration

Calls	Bid	Ask	Vol	Puts	Bid	Ask	Vol
Jan 550	39.5	41.5	0	Jan 550	0.45	0.75	0
Jan 555	34.8	36.3	0	Jan 555	0.65	0.95	0
Jan 560	30	31.5	0	Jan 560	0.9	1.2	0
Jan 565	25.2	26.7	0	Jan 565	1.25	1.55	0
Jan 570	20.6	22.1	0	Jan 570	1.8	2.1	0
Jan 575	16.3	17.8	0	Jan 575	2.3	3	0
Jan 580	13	13.5	0	Jan 580	3.4	4.1	2
Jan 585	9.1	9.8	0	Jan 585	5	5.7	5
Jan 590	6.1	6.8	50	Jan 590	7.6	7.9	5
Jan 595	4.1	4.5	12	Jan 595	10.1	10.8	25
Jan 600	2.5	2.8	3	Jan 600	13.1	14.5	0
Jan 605	1.2	1.5	0	Jan 605	17.2	18.7	0
Jan 610	0.55	0.85	1	Jan 610	21.7	23.2	0
Jan 615	0.25	0.55	0	Jan 615	26.6	28.1	0
Jan 620	0.2	0.35	1	Jan 620	31.4	32.9	0
Jan 625	0.05	0.2	0	Jan 625	36.3	37.8	0
Jan 630	0	0.15	0	Jan 630	41	43	0
Jan 635	0	0.1	0	Jan 635	46	48	0
Jan 640	0	0.1	0	Jan 640	51	53	0
Jan 645	0	0.1	0	Jan 645	56.5	57.5	0
Jan 650	0	0.1	0	Jan 650	60.5	63.5	0
Jan 660	0	0.05	0	Jan 660	70.5	73.5	0
Jan 680	0	0.05	0	Jan 680	90.5	93.5	0

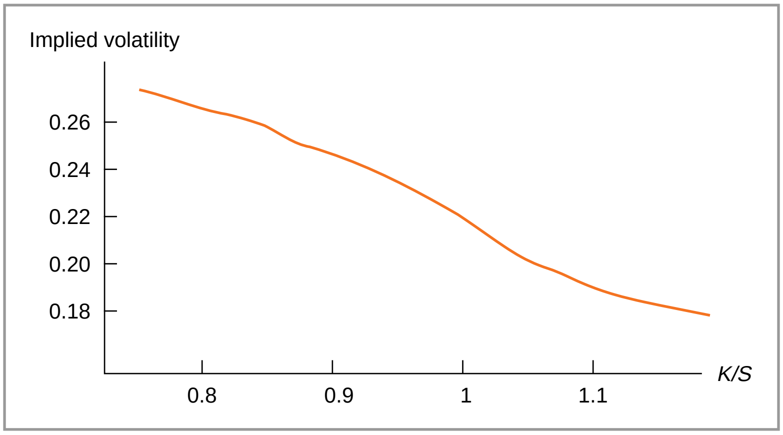

FIGURE 15-11

11.2. *How Can We Define Moneyness?*

The way a smile is plotted varies from one market to another. The implied volatility, denoted by σ_i, always appears on the y-axis. Unless stated otherwise, we extract this volatility from the Black-Scholes formula in equity or FX markets, and from the Black formula in the case of interest rates. The implied volatilities are then treated as if they were *random* and *time varying*.

What to put on the horizontal axis is a more delicate question and eventually depends on how we define "moneyness" of an option. Sometimes the smile is plotted against *moneyness* measured by the ratio of the strike price to the current market price, K_i/S_t. If the smile is a function of how much the option is out-of-the-money only, then this normalization will stabilize the smile in the sense that as S_t changes, the smile for that particular option series may be more or less invariant. But there are almost always factors other than moneyness that affect the smile, and some practitioners define moneyness differently.

For example, sometimes the smile is plotted against $K_i e^{-r(T-t)}/S_t$. For short-dated options, this makes little difference, since $r(T-t)$ will be a small number. For longer-dated options, the difference is more relevant. By including this discount factor, market practitioners hope to eliminate the effect of the changes in the remaining life of the option.

Sometimes the horizontal axis represents the option's *delta*. FX traders take the size of *delta* as a measure of moneyness. This practice can be challenged on the grounds that an option's *delta* depends on more variables than just moneyness. It also depends, for example, on the instantaneous implied volatility. Yet, as we will see later, there are some *deltas* at which volatility trading is particularly liquid in FX markets.

The reader should note that some of the smiles in Figures 15-11 and 15-12 are plots of the implied volatility against the *strike* only. Also, these curves relate to a particular time t and

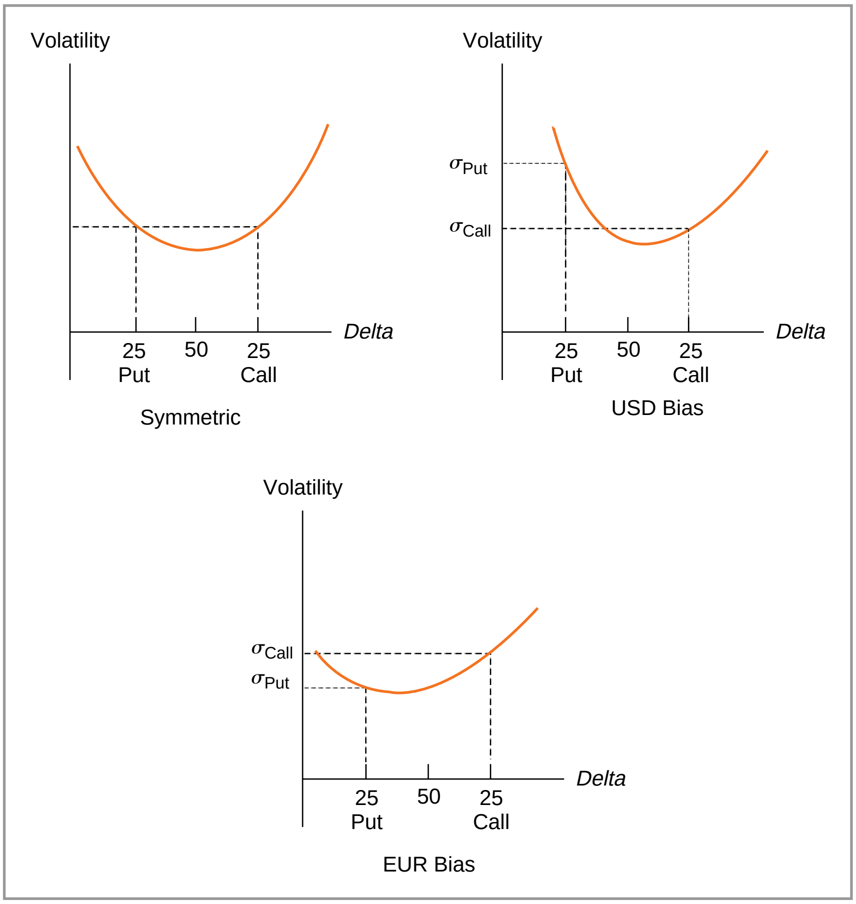

FIGURE 15-12

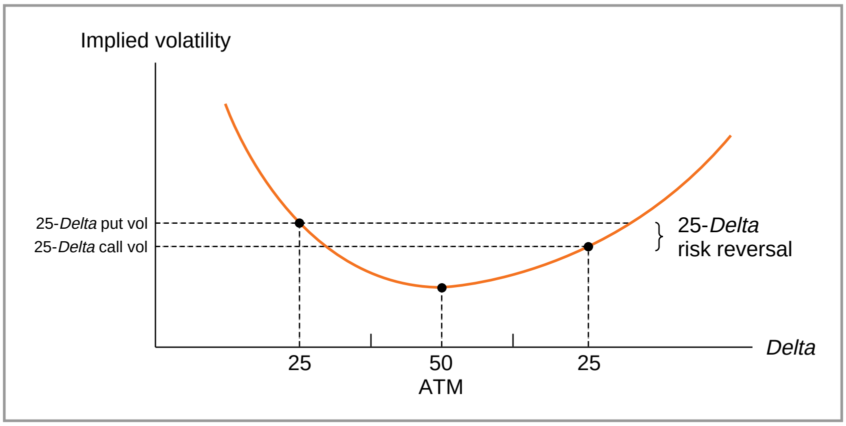

FIGURE 15-13

expiration date T. As the latter change, the smile will, in general, shift. It is quite important to know how changes in time t and expiration date T affect the smile.

11.3. *Replicating the Smile*

The volatility smile is a plot of the implied volatility of options that are alike in all respects except for their moneyness. The basic shape of the volatility smile has two major characteristics. The first relates to the extent of symmetry in the smile. The second is about how "pronounced" the curvature is. There are good approximations for measuring both characteristics.

First, consider the issue of symmetry. Figure 15-12 shows three smiles for FX markets. One is symmetric, the other two are asymmetric with different *biases*. If the smile is symmetric, the volatilities across similarly out-of-the-money puts and calls will be the same. This means that if a trader buys a call and sells a put with the same moneyness, the structure will have zero value. Such positions were called risk reversals in Chapter 10. A symmetric smile implies that a zero cost risk reversal can be achieved by buying and selling similarly out-of-the-money options. In the case of asymmetric smiles, puts and calls with similar moneyness are sold at *different* implied volatilities, and this is labeled a *bias*. Thus, a risk reversal is one way of measuring the bias in a volatility smile.

The way risk reversals measure the symmetry of the volatility smile is shown in Figure 15-13. We use the *delta* of the option as a measure of its moneyness on the x-axis. ATM options would have a *delta* of around 50 and would be in the "middle" of the x-axis. The volatility of the 25-*delta* risk reversal gives the difference between the volatilities of a 25-*delta* put and a 25-*delta* call as indicated in the graph. We can write this as

$$\sigma(25\text{-}delta\ RR\ \text{spread}) = \sigma(25\text{-}delta\ \text{put}) - \sigma(25\text{-}delta\ \text{call}) \tag{60}$$

where, $\sigma(25\text{-}delta\ RR\ \text{spread})$, $\sigma(25\text{-}delta\ \text{put})$, and $\sigma(25\text{-}delta\ \text{call})$ indicate, respectively, the implied volatilities of a risk reversal, a 25-*delta* put, and a 25-*delta* call.

The curvature of the smile can be measured using a *butterfly*. Consider the sale of an ATM put and an ATM call along with the purchase of one 25-*delta* out-of-the-money put and a 25-*delta* out-of-the-money call. This butterfly has a payoff diagram that should be familiar from Chapter 10. The position consists of buying two symmetric out-of-the-money volatilities and selling two ATM volatilities. If there were no smile effects, these volatilities would all be the same and the net volatility position would be zero. On the other hand, the more pronounced the smile becomes, the higher the out-of-the-money volatilities would be relative to the ATM volatilities, and the net volatility position would become more and more positive. Figure 15-14

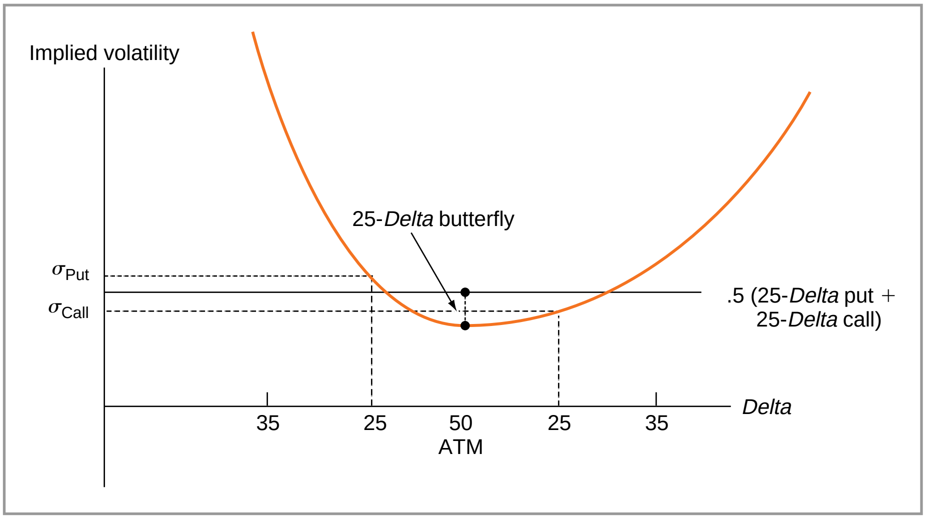

FIGURE 15-14

shows how a butterfly measures the magnitude of the curvature in a smile. The following equality holds:

$$\sigma(25\text{-}delta \text{ butterfly spread}) = \sigma(25\text{-}delta \text{ put}) + \sigma(25\text{-}delta \text{ call}) - 2\sigma(\text{ATM}) \quad (61)$$

where the $\sigma(25\text{-}delta \text{ butterfly spread})$ and $\sigma(\text{ATM})$ are the butterfly and the ATM implied volatilities, respectively.

11.3.1. Contractual Equations

Chapters 3 and 4 of this book dealt with contractual equations for simple assets. The equalities discussed in the preceding paragraphs now permit considering quite different types of contractual equations. In fact, we can rearrange equalities shown in equations (60) and (61) to generate some contractual equations for out-of-the-money implied volatilities:

$$\boxed{\begin{array}{l}25\text{-}delta \text{ put}\\ \text{volatility}\end{array}} = \boxed{\text{ATM volatility}} + \frac{1}{2}\boxed{\begin{array}{l}25\text{-}delta\ RR\\ \text{volatility spread}\end{array}} + \frac{1}{2}\boxed{\begin{array}{l}25\text{-}delta \text{ butterfly}\\ \text{volatility spread}\end{array}} \quad (62)$$

$$\boxed{\begin{array}{l}25\text{-}delta \text{ call}\\ \text{volatility}\end{array}} = \boxed{\text{ATM volatility}} - \frac{1}{2}\boxed{\begin{array}{l}25\text{-}delta\ RR\\ \text{volatility spread}\end{array}} + \frac{1}{2}\boxed{\begin{array}{l}25\text{-}delta \text{ butterfly}\\ \text{volatility spread}\end{array}} \quad (63)$$

These equalities can be used to determine out-of-the-money volatilities in the case of vanilla options. For example, if ATM, *RR* and butterfly volatilities are liquid, we can use these equations to "calculate" 25-*delta* call and put volatilities. However, it has to be noted that for exotic options, adjusting the volatility parameter this way will not work. This issue will be discussed at the end of the chapter.

12. Smile Dynamics

There are at least two types of smile "dynamics." In the first, we would fix the time parameter t and consider options with longer and longer expirations, T. In the second case, we would keep T constant, but let time t pass and study how changes in various factors affect the volatility smile. In particular, we can observe *if* changes in S_t affect the smile when the moneyness K_i/S_t is kept constant.

We first keep t fixed and increase T. We consider *two* series of options that trade at the same time t. Both series have comparable strikes, but one series has a relatively longer maturity. How would the smiles implied by the two series of options with expirations, say, T_1, T_2, compare with each other?

The second question of interest is how the smile of the *same* option series moves over time as S_t changes. In particular, would the smile be a function of the ratio K_i/S_t only, or would it also depend on the level of S_t over and above the moneyness?

The answers to these questions change depending on which underlying asset is considered. This is because there is more than one explanation for the existence of the smile, and for different sectors, different explanations seem to prevail. Thus, before we analyze the smile dynamics and its properties any further, we need to discuss the major explanations advanced for the existence of the volatility smile.

13. How to Explain the Smile

The volatility smile is an empirical phenomenon that violates the assumptions of the Black-Scholes world. At the same time, the volatility smile is related to the implied volatilities obtained *from* the Black-Scholes formula. This may give rise to confusion. The smile suggests that the Black-Scholes formula is not valid, while at the same time, the trader obtains the smile using the very same Black-Scholes formula. Is there an internal inconsistency?

The answer is no. To clarify the point, we use an analogy that is unrelated to the present discussion, but illustrates what market conventions are. Consider the 3-month Libor rate L_t. What is the present value of, say, \$100 that will be received in 3 months' time? We saw in Chapter 3 that all we need to do is calculate the ratio:

$$\frac{100}{(1 + L_t \frac{1}{4})} \tag{64}$$

An economist who is used to a different de-compounding may disagree and use the following present value formula:

$$\frac{100}{(1 + L_t)^{\frac{1}{4}}} \tag{65}$$

Who is right? The answer depends on the market convention. If L_t is quoted under the condition that formula (64) be used, then formula (65) would be wrong if used with the *same* L_t. However, we can always calculate a new L_t^* using the equivalence:

$$\frac{100}{(1 + L_t \frac{1}{4})} = \frac{100}{(1 + L_t^*)^{\frac{1}{4}}} \tag{66}$$

Then, the formula

$$\frac{100}{(1+L_t^*)^{\frac{1}{4}}} \tag{67}$$

used with L_t^* would *also* yield the correct present value. The point is, the market is quoting an interest rate L_t with the condition that it is used with formula (64). If for some odd reason a client wants to use formula (65), then the market would quote L_t^* instead of L_t. The result would be the same since, whether we use formula (64) with L_t, as the market does, or formula (65) with L_t^*, we would obtain the same present value. In other words, the question of which formula is correct depends on *how* the market quotes the variable of interest.

This goes for options also. The Black-Scholes formula may be the wrong formula if we substitute one particular volatility, but may give the right answer if we use another volatility. And the latter may be different than the real-world volatility at that instant. But traders can still use a particular volatility to obtain the right option price from this "wrong" formula, just as in the earlier present value example. This particular volatility, when associated with the Black-Scholes formula, may give the correct value for the option even though the assumptions leading to the formula are not satisfied.

Thus, suppose the arbitrage-free option price obtained under the "correct" assumptions is given by

$$C(S_t, t, T, K, \sigma_t^*, \theta_t) \tag{68}$$

where K is the strike price, T is the expiration date, and S_t is the underlying asset price. The (vector) variable θ_t represents all the other parameters that enter the "correct" formula and that may not be taken into account by the Black-Scholes world. For example, the volatility may be stochastic, and some parameters that influence the volatility dynamics may indirectly enter the formula and be part of θ_t.[14] The critical point here is the meaning that is attached to σ_t^*. We assume for now that it is the correct instantaneous volatility as of time t.

The (correct) pricing function in equation (68) may be more complex and may not even have a closed form solution in contrast to the Black-Scholes formula, $F(S_t, t, \sigma)$. Suppose traders ignore equation (68) but prefer to use the formula $F(S_t, t, \sigma)$, even though the latter is "wrong." Does this mean traders will calculate the wrong price?

Not necessarily. The "wrong" formula $F(S_t, t, \sigma)$ can very well yield the same option price as $C(S_t, t, K, \sigma_t^*, \theta_t)$ if the trader uses in $F(S_t, t, \sigma)$, another volatility, σ, such that the two formulas give the same correct price:

$$C(S_t, t, K, \sigma_t^*, \theta_t) = F(S_t, t, \sigma) \tag{69}$$

Thus, we may be able to get the correct option price from the "unrealistic" Black-Scholes formula if we proceed as follows:

1. We quote the K_i-strike option volatilities σ_i directly at every instant t, *under the condition that the Black-Scholes formula be used* to obtain the option value. Then, liquid and arbitrage-free markets will supply "correct" observations of the ATM volatility σ_0.[15]

[14] In the case of a mean-reverting stochastic volatility model, we will have

$$d\sigma_t = \lambda(\sigma_0 - \sigma_t)dt + \kappa\sigma_t dW_t$$

where σ_0, κ, and λ are, respectively, the long-run average volatility, the volatility of the volatility, and the speed of mean reversion. The θ_t in formula (68) will include λ, σ_0, and the, possibly time varying, γ_t.

[15] Especially FX markets quote such implied volatilities and active trading is done on them.

2. For out-of-the-money options, we use the Black-Scholes formula with a new volatility denoted by $\sigma(\frac{S}{K_i}, S)$, and

$$\sigma\left(\frac{S}{K_i}, S\right) = \sigma_0 + f\left(\frac{S}{K_i}, S\right) \tag{70}$$

where $f(.)$ is, in general, positive and implies a smile effect. The adjustment made to the ATM volatility, σ_0, is such that when $\sigma(\frac{S}{K_i}, S)$ is used in the Black-Scholes formula, it gives the correct value for the K_i strike option:

$$F\left(S_t, t, K_i, \sigma_0 + f\left(\frac{S}{K_i}, S\right)\right) = C\left(S_t, t, K_i, \sigma_t^*, \theta_t\right) \tag{71}$$

The adjustment factor $f(\frac{S}{K_i}, S)$ is determined by the trader's experience, knowledge, and the trading environment at that instant. The relationships between risk reversal, butterfly, and ATM volatilities discussed in the previous section can also be used here.[16]

The trader, thus, *adjusts* the volatility of the non-ATM options so that the wrong formula gives the correct answer, even though what is used in the Black-Scholes formula may not be the "correct" instantaneous realized volatility of the S_t process.

The $f(\frac{S}{K_i}, S)$ is, therefore, an adjustment required by the imperfections of the Black-Scholes formula in adequately representing the real-world environment. The upshot is that when we plot $\sigma(\frac{S}{K_i}, S)$ against K_i/S or K_i we get a smile, or a skew curve, depending on the time and the sector we are working with.

For what types of situations should the volatilities be adjusted? At least three inconsistencies of the Black-Scholes assumptions with the real world can be corrected for by adjusting the volatilities across the strike K_i. The first is the lognormal process assumption. The second is the fact that if asset prices fall dramatically during a relatively short period of time, this could increase the "fear factor" and volatility would increase. The third involves the organizational and regulatory assumptions concerning financial markets. We discuss these in more detail next.

13.1. Case 1: Nongeometric Price Processes

Suppose the underlying obeys the true risk-neutral dynamics described by the SDE:

$$dS_t = rS_t dt + \sigma S_t^\alpha dW_t \qquad t \in [0, \infty) \tag{72}$$

With $\alpha = 1$, S_t would be lognormal. Everything else being conformable to the Black-Scholes world, there would be no smile in the implied volatilities.

The case of $\alpha < 1$ would require an adjustment to the volatility coefficient used in the Black-Scholes formula as the strike changes. This is true, since, unlike in the case of $\alpha = 1$, now the percentage volatility is dependent on the level of S_t. We divide by S_t to obtain

$$\frac{dS_t}{S_t} = rdt + \sigma S_t^{\alpha-1} dW_t \qquad t \in [0, \infty) \tag{73}$$

The percentage volatility is given by the term $\sigma S_t^{\alpha-1}$. This percentage volatility will be a decreasing function of S_t if $\alpha < 1$. As S_t declines, the *percentage* volatility increases. Thus, the trader needs to use higher implied volatility parameters in the Black-Scholes formula

[16] For convenience, the t subscripts are ignored in these formulas.

for put options with lower and lower strike prices. This means that the more out-of-the-money the put option is, the higher the volatility used in the Black-Scholes formula must be.

This illustrates the idea that although the trader knows that the Black-Scholes world is far from reality, the volatility is adjusted so that the original Black-Scholes framework is preserved and that a "wrong" formula can still give the correct option value.

13.2. Case 2: Possibility of Crash

Suppose a put option series has an expiration of two months. All options are identical except for their strike. They run from ATM to deep out-of-the-money. Suppose also that the current level of S_t is 100. The liquid put options have strikes 90, 80, 70, and 60.

Here is what the 90-strike option implies. If the option expires in-the-money, then the market would have fallen by at least 10% in two months. This is a big fall, perhaps, but not a disaster. In contrast, if the 60-strike put expires in-the-money, this would imply a 40% drop in two months. This is clearly an unusual event, and unusual events lead to sudden spikes in volatility. Thus, the 60-strike option is relatively more associated with events that are labeled as crises and, everything else the same, this option would, in all likelihood, be in-the-money when the volatility is very high. But when this option becomes in-the-money, its *gamma*, which originally is close to zero, will also be higher. Thus, the trader who sells this option would have higher cash payouts due to *delta* hedge adjustments. So, to compensate for these potentially higher cash payments, the trader would use higher and higher vol parameters in put options that are more and more out-of-the-money, and, hence, are more and more likely to be associated with a crisis situation.

This explanation is consistent with the smiley shapes observed in reality. Note that in FX markets, sudden drops *and* sudden increases would mean higher volatility because in each case *one* of the observed currencies could be falling dramatically. So the smile will be more or less symmetric. But in the case of equity markets, a sudden increase in equity prices may be an important event, but not a crisis at all. For traders (excluding the shorts) this is a "happy" outcome, and the volatility may not increase much. In contrast, when asset prices suddenly crash, this increases the fear factor and the volatilities may spike. Thus, in equity markets the smile is expected to be mostly one-sided if this explanation is correct. It turns out that empirical data support this contention. Out-of-the-money equity puts have a smile; but out-of-the-money equity calls exhibit almost no smile.

EXAMPLE:

Consider Table 15-2 which displays the prices of options with June 2002 expiry, on January 10, 2002, and ignore issues related to Americanness or any possible payouts. These data are collected at the same time as those discussed in the earlier example. In this case, the options are longer dated and expire in about 6 months. First, we obtain the volatility smile for these data.

The data are collected at the same instant, and since the current value of the underlying index is the same in each case, the division by S_{t_0} is not a major issue, but we still prefer to graph the volatility smile against the $\frac{K}{S}$.

We extract ask prices for the eight out-of-the-money puts and consider the 600-put as being in-the-money. This way we can calculate nine implied vols. The price data that we use are shown in Table 15-2. We consider first the out-of-the-money put asking prices listed in the sixth column of this table. This will give nine prices.

TABLE 15-2. OEX Options with June 21, 2002, Expiry (collected 9:46 CBOT on January 10, 2002)

Calls	Bid	Ask	Puts	Bid	Ask
Jun 440	153.4	156.4	Jun 440	4.2	4.8
Jun 460	134.8	137.8	Jun 460	5.6	6.3
Jun 480	116.7	119.7	Jun 480	7.4	8.1
Jun 500	99.2	102.2	Jun 500	9.9	10.6
Jun 520	82.6	85.6	Jun 520	12.9	14.4
Jun 540	67.2	69.7	Jun 540	17.2	18.7
Jun 560	52.7	55.2	Jun 560	22.7	24.2
Jun 580	39.8	41.8	Jun 580	29.3	31.3
Jun 600	28.6	30.6	Jun 600	38.3	40.3
Jun 620	19.9	21.4	Jun 620	49.5	51.5
Jun 640	12.8	14.3	Jun 640	62.2	64.7
Jun 660	8	8.7	Jun 660	76.9	79.9
Jun 680	4.7	5.4	Jun 680	93.7	96.7
Jun 700	2.55	3.2	Jun 700	111.6	114.6

Ignoring other complications that may exist in reality, we use the Black-Scholes formula straightforwardly with

$$S_{t_0} = 589.15, r = 1.90\%, t = \frac{152}{365} = 0.416 \tag{74}$$

We solve the equations

$$P(589.15, K_i, 1.90, \sigma_{K_i}, 0.416) = P_{K_i} \qquad i = 1, \ldots, 9 \tag{75}$$

and obtain the nine implied volatilities σ_{K_i}. Using Mathematica, we obtain the following result, which shows the value of K_i/S and the corresponding implied vols for out-of-the-money puts:

$\frac{K}{S}$	Vol
0.74	0.26
0.78	0.26
0.81	0.26
0.84	0.25
0.88	0.25
0.91	0.24
0.95	0.23
0.98	0.22
1.01	0.21

This is shown in Figure 15-15. Clearly, as the moneyness of the puts decreases, the volatility increases. Option market makers will conclude that, if in 6 months, U.S. equity markets were to drop by 25%, then the fear factor would increase volatility from 21% to 26%. By selecting the seven out-of-the-money call prices, we get the implied vols for out-of-the-money calls.

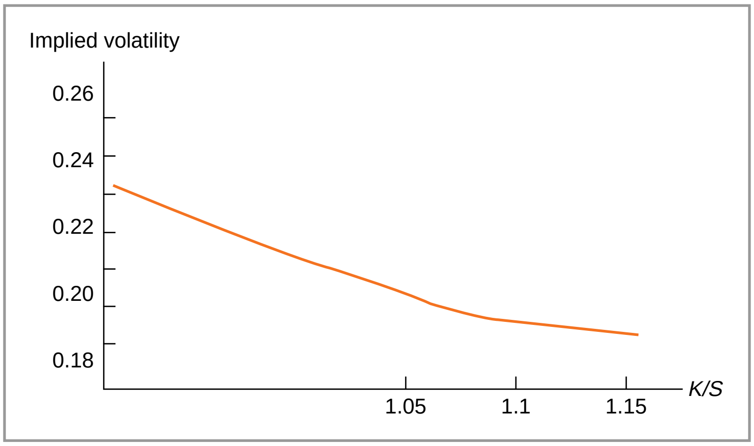

FIGURE 15-15

$\frac{K}{S}$	**Vol**
0.98	*0.23*
1.01	*0.22*
1.05	*0.21*
1.08	*0.20*
1.12	*0.19*
1.15	*0.19*
1.18	*0.18*

Here, the situation is different. We see that as moneyness of the calls decreases, the volatility also decreases.

Option market makers may now think that if, in 6 months, U.S. equity markets were to increase by 20%, then the fear factor would decrease and so would volatility.

The fear of a crash that leads to a smile phenomenon can, under some conditions, be represented analytically using the so-called jump processes. We discuss this modeling approach next.

13.2.1. Modeling Crashes

Consider again the standard geometric Brownian motion case:

$$dS_t = rS_t dt + \sigma S_t dW_t \qquad t \in [0, \infty) \tag{76}$$

W_t is a Wiener process under the risk-neutral probability $\tilde{P}$. Now, keep the volatility parameterization the same, but instead, add a jump component as discussed in Lipton (2002). For example, let

$$dS_t = rS_t dt + \sigma S_t dW_t + S_t \left[(e^j - 1)dJ_t - \lambda m dt\right] \qquad t \in [0, \infty) \tag{77}$$

Some definitions are needed regarding the term $(e^j - 1)dJ_t - \lambda m dt$. The j is the size of a *random* logarithmic jump. The size of the jump is not related to the occurrence of the jump, which is represented by the term dJ_t. If the jump is of size zero, then $(e^j - 1) = 0$ and the jump term does not matter.

The term dJ_t is a Poisson-type process. In general, at time *t*, it equals zero. But, with "small" probability, it can equal one. The probability of this happening depends on the length of the interval we are looking at, and on the size of the *intensity coefficient* λ. The jump can heuristically be modeled as follows

$$dJ_t = \begin{cases} 0 & \text{with probability } (1-\lambda dt) \\ 1 & \text{with probability } \lambda dt \end{cases} \tag{78}$$

where $0 < dt$ is an infinitesimally short interval. Finally, m is the expected value of $(e^j - 1)$:

$$E_t^{\tilde{P}}[(e^j - 1)] = m \tag{79}$$

Thus, we see that, for an infinitesimal interval we can heuristically write

$$E_t^{\tilde{P}}[(e^j - 1)dJ_t] = E_t^{\tilde{P}}[(e^j - 1)]E_t^{\tilde{P}}[dJ_t] \tag{80}$$

$$= m[0.(1-\lambda dt) + 1.\lambda dt] \tag{81}$$

$$= m\lambda dt \tag{82}$$

According to this, the expected value of the term $(e^j - 1)dJ_t - \lambda m dt$ is zero.

This jump-diffusion model captures some crash phenomena. Stock market crashes, major defaults, 9/11–type events, and currency devaluations can be modeled as rare but discrete events that lead to jumps in prices.

The way these types of jumps create a smile can be heuristically explained as follows: In a world where the Black-Scholes assumptions hold, with a geometric S_t process, a constant volatility parameter $\tilde{\sigma}$, and *no* jumps, the volatility trade yields the arbitrage relation:

$$\frac{1}{2}C_{ss}\tilde{\sigma}^2 S^2 + C_t + rC_s S - rC = 0 \tag{83}$$

With a jump term added to the geometric process as in equation (77), the corresponding arbitrage relation becomes

$$\frac{1}{2}C^*_{ss}\sigma^2 S^2 + C^*_t + (r - \lambda m)C^*_s S - rC^* + \lambda E_t^{\tilde{P}}[C^*(Se^j, t) - C^*(S,t)] = 0 \tag{84}$$

where $\tilde{P}$ is the risk-neutral probability. Suppose we decide to use, as a convention, the Black-Scholes formula, but believe that the true PDE is the one in equation (83). Then, we would select $\tilde{\sigma}$ such that the Black-Scholes formula yields the same option value as the other PDE would yield.

For example, out-of-the-money options will have much smaller *gammas*, C_{ss}. If the expected jump is negative, then $\tilde{\sigma}$ will be bigger, and the more out-of-the-money the options are. As the expiration date *T* increases, C_{ss} will increase and the smile will be less pronounced.

13.3. *Other Explanations*

Many other effects can cause a volatility smile. One is *stochastic volatility*. Consider a local volatility specification using

$$dS_t = \mu S_t dt + \sigma S_t^\alpha dW_t \qquad t \in [0, \infty) \tag{85}$$

with, say, $\alpha < 1$. In this specification, percentage volatility *will* be stochastic since it depends on the random variable S_t. But often this specification does not express what is meant by models of stochastic volatility. What is captured by stochastic volatility is a situation where an additional Wiener process dB_t, possibly correlated with dW_t, affects the dynamics of percentage volatility. For example, we can write

$$dS_t = \mu S_t dt + \sigma_t S_t dW_t \qquad t \in [0, \infty) \tag{86}$$

$$d\sigma_t = a(\sigma_t, S_t)dt + \kappa\sigma_t dB_t \qquad t \in [0, \infty) \tag{87}$$

where κ is the parameter representing the (constant) percentage volatility of the volatility of S_t. In this model, the volatility itself is driven by some random increments that originate in the volatility market *only*. These shocks are only partially correlated with the innovation terms dW_t affecting the price data.

It can be shown that stochastic volatility generates a volatility smile. In fact, with stochastic volatility, we can perform an analysis similar to the PDE with a jump process (see Lipton (2002)). The result will essentially be similar. However, it is important to emphasize that, everything else being the same, this model may be incomplete in the sense that there may not be enough instruments to hedge the risks associated with dW_t and dB_t completely, and form a risk-free, self-financing portfolio. The jump-diffusion model discussed in the previous section may entail the same problem. To the extent that the jump part and the diffusion part are affected by different processes, the model may not be complete.

13.3.1. Structural and Regulatory Explanations

Tax effects (Merton, 1976) and the capital requirements associated with carrying out-of-the-money options in options books may also lead to a smile in implied volatility. We briefly touch on the second point.

The argument involves the concept of *gamma*. A negative *gamma* position is considered to be more risky, the more out-of-the-money the option is. Essentially, negative *gamma* means that the market maker has sold options and *delta* hedged them, and that he or she is paying the *gamma* through the rebalancing of this hedge. If the option is deep out-of-the-money, *gamma* would be close to zero. Yet, if the option suddenly becomes in-the-money, the *gamma* could spike, especially if the option is about to expire. This may cause significant losses. Out-of-the-money options, therefore, involve substantial risk and require more capital. Due to such costs, the market maker may want to sell the out-of-the-money option at a higher price than warranted by the ATM volatility.

14. The Relevance of the Smile

The volatility smile is important in financial engineering for at least *three* reasons.

First, if we associate a volatility smile with all the risk factors, and if this smile shifts randomly over time, then we may be able to *trade* it, take spread positions, and arbitrage it. The smile dynamics, thus, imply new opportunities for a market professional.

The second reason for the relevance of the volatility smile is that it may contain important information about the dynamics of the underlying realized volatility processes. With a volatility smile, pricing and hedging may become much more complicated, especially if the instrument has characteristics of an exotic option. Is volatility constant or a stochastic process? If the latter is the case, then what type of stochastic process is it? Are there jumps or is a process with Wiener-type increments a sufficiently good approximation? These questions are important for risk management and pricing.

Third, the creation of new products and synthetics must pay attention to the causes of the smile. If the smile is the result of conventions and practices adopted by market professionals rather than resulting from the underlying volatility processes, we must take these conventions into account. We now discuss these issues in more detail and provide some examples to the uses of the volatility smile.

15. Trading the Smile

The volatility smile is actively traded to a different extent in different sectors. The smile is an integral part of daily trading in the FX sector. Here, market practitioners routinely quote risk reversals, which relate to the symmetry in exchange rate volatility and butterflies related to the curvature of the smile. Traders trade and arbitrage these effects. The volatility smile is also traded in the equity sector. Traders arbitrage volatility across stock market indices, and in doing this, sometimes trade the smile indirectly. At other times, this trade is direct. The smile relating to a risk may be too steep and is expected to flatten. The trader then sells the deep out-of-the-money options and buys those that are closer to being at-the-money. In the interest rate sector, volatility smile is mainly traded due to its risk management and hedging implications for cap/floor positions and swaption books.

Smiles can be of interest to investors who may want to take positions on the slope and the curvature of the volatility smile, thinking that the market has under- or overemphasized one of the underlying parameters. In the following example, traders are putting together *skew swaps* that will trade *realized* skews against *implied* skews.

EXAMPLE:

As the skew in volatility between out-of-the-money puts and calls on Standard & Poor's 500 index has grown, street traders are looking to capture discrepancies between the realized and implied skews of the options. One trader in New York noted interest in a skew swap on the S&P500 from hedge funds trading volatility. The swaps—which traders believe would be a first—would offer the realized skew of puts and calls in return for the implied skew.

Currently, the S&P skew is above 30—if strikes on puts and calls are moved by 10%, the volatility would [increase by] 3%, explained one structurer. This compares with a level of 15 at the beginning of October, which is in line with the historical levels of 15–20.

One structurer who had tried to put together a skew swap noted that there is no mathematical formula that can capture implied skews for any period of time. He also admitted to being stumped by hedging the product. "To hedge this, we would have to close every night with a vol swap on the deal and that can't always be done," he said. A rival noted that one popular trade to capture flattening skews is selling out-of-the-money puts and calls and buying at-the-money puts and calls. (Derivatives Week, November 1998)

One interesting point in this reading is that, at least in this particular case, the observed smile (skew) is characterized by multiplying a *linear* relationship with a slope of .3. According to the traders, if moneyness decreases by 10%, volatility increases by 3%. Traders expect this relationship to be around .15 during normal times. Hence, the smile is expected to flatten.

16. Pricing with a Smile

Pricing and hedging are fairly closely related activities, at least in abstract settings. Once an asset is replicated with liquid securities, the price of the asset is the cost of the replicating portfolio plus some profit margin. At several points in the previous chapters, we saw that assets can be replicated using a series of options with different strike prices. This was the method applied for finding a hedge for a volatility swap in Chapter 14, for example. The replicating portfolio was made of a weighted sum of relevant options with the same characteristics except for their strikes. In Chapter 11 we saw that option portfolios could replicate statically almost any future payoff function. Again, these options were similar except for their strike prices.

The potential use of options with different strikes makes the volatility smile a crucial parameter in forming hedging portfolios and in pricing complex instruments. In fact, if implied volatilities depend on the moneyness of the option, then the volatility parameters used in formulas for replicating portfolios would automatically change as the markets move and the moneyness of the replicating options changes. The critical point is that this will be true even if the underlying *realized* volatility remains the same. This section presents two examples of this phenomenon.

The first involves the class of interest rate derivatives known as caps and floors. These are among the most widely traded instruments. Their hedging and pricing are influenced in a crucial way if there is a volatility smile. The second involves the pricing and hedging of exotic options. Here also, the methodology and market practices crucially depend on the existence of the volatility smile.

We start this discussion with a simple framework for caps/floors. We use a limited number of settlement dates to motivate the main points and the importance of the smile. A general treatment of these instruments can be found in a number of excellent texts.[17]

17. Exotic Options and the Smile

The second major category of instruments where the existence of the volatility smile can change pricing and hedging practices significantly is exotic options. In this section, we consider a simple knock-out call that is representative of the main ideas we want to convey. Due to the contractual equation between vanilla options, knock-out calls, and knock-in calls, our discussion immediately extends to knock-in calls as well. At the end of this section, we briefly discuss digital options and how the existence of the volatility smile affects them.

17.1. *A Hedge for a Barrier*

Knock-out calls were discussed in Chapters 8 and 10. As a reminder, a simple knock-out call is similar to a European vanilla call with strike K and expiration T, written on the underlying S_t, except that the option will cease to exist if, during the life of the option, S_t falls below barrier H:

$$S_t < H \qquad t \in [t_0, T] \tag{88}$$

The price of a knock-out barrier approaches the price of a vanilla call as the option becomes more in-the-money. However, as the underlying approaches the barrier, the value of the knock-out will approach zero.

There are several ways of hedging knock-out options used by practitioners. Here, we explain a hedge that (1) has nice financial engineering implications, and (2) shows clearly the important role played by the smile.

Suppose we bought the corresponding vanilla call *and* sold a carefully chosen out-of-the money vanilla put with strike $K^*, K^* < K$, with a very precise objective. We want the put

[17] For example, see Hull (2002), Musiela and Rutkowski (1998), and Rebonato (2002).

and the call be such that, as S_t approaches the barrier H, the portfolio's value becomes zero. This portfolio, which is, in fact, a type of risk reversal, approximately replicates the knock-out option. If S_t moves away from the H, the put becomes more out-of-the-money, and its value will decline. Then the portfolio looks more and more like a vanilla call. This is what the knock-out option accomplishes anyway. On the other hand, as S_t approaches H, the put becomes more valuable. If it is carefully chosen, at H the value of the put position can equal the value of the vanilla call and the portfolio would have zero value. This is, again, what the knock-out option accomplishes near H. As S_t falls below H, the portfolio has to be liquidated.

Thus, the portfolio of

$$\{\text{Short } x \text{ units of } K^*\text{-Put, Long one } K\text{-Call}\} \tag{89}$$

replicates the knock-out call if x and K^* are appropriately selected. One way to do this is to use the "symmetry" principle.

Suppose there is no smile effect and that all options with different strikes that belong to a series have the *same* volatility. Then we can choose x and K^* as follows: We want the value of x units of the K^*-put to equal the value of the vanilla call when $H \leftarrow S_t$. This can be achieved by choosing K^* such that

$$K^* K = H^2 \tag{90}$$

The prices of these options are assumed to satisfy

$$\frac{K^*\text{-put}}{K\text{-call}} = \sqrt{\frac{K^*}{K}} \tag{91}$$

This means that if *x* is chosen so that

$$x = \frac{K}{K^*} \tag{92}$$

then the value of x units of these K^*-puts would equal the value of the K-call as S_t approaches H. As a result, the portfolio would replicate the knock-out call, except that once the barrier is hit, the portfolio needs to be liquidated.

17.2. *Effects of the Smile*

Consider first the effect of a stable volatility smile on this procedure. If the smile does not shift over time, then it is easy to incorporate the effect into the previous replicating portfolio. Suppose the smile is downward sloping over $[K^*, K]$. Then, one could plug different volatility parameters in Black-Scholes formulas for each vanillla option. The same K^*-put selected earlier would be relatively more valuable than in the case of a flat smile. This means that the knock-out option could be sold at a cheaper price. If the smile was upward sloping over the range, then the reverse would be true and the knock-out would be more expensive. Hence, the smile has a direct effect that needs to be taken into account in the pricing and hedging of the knock-out.

There is a second effect of the smile as well. Suppose the smile is *not* stable during the life of the option, and that it shifts as time passes. Then the logic of replication would fall apart since a smile that shifts over time would make the relative values of the call and the put differ from the originally intended ratio as S_t approaches *H*. Given that in most markets the smile *is* unstable over time, the hedging technique by this replication is questionable. The pricing of the knock-out would also be unreliable.

17.2.1. An Example of Technical Difficulties

We can look at the complications introduced by the volatility smile using knock-out pricing formulas in case all standard assumptions are satisfied. The pricing formula for knock-outs was given in Chapter 8. In particular, the price of a down-and-out call written on a stock S_t, satisfying all standard assumptions, and paying no dividends, was given by

$$C^b(t) = C(t) - J(t) \tag{93}$$

where $C(t)$ is the value of a vanilla call, given by the standard Black-Scholes formula, and where $J(t)$ is the "discount" that needs to be applied because the option may die if S_t falls below H during the life of the contract. The formula for $J(t)$ was

$$J(t) = S_t \left(\frac{H}{S_t}\right)^{\frac{2(r-\frac{1}{2}\sigma^2)}{\sigma^2}+2} N(c_1) - Ke^{-r(T-t)} \left(\frac{H}{S_t}\right)^{\frac{2(r-\frac{1}{2}\sigma^2)}{\sigma^2}} N(c_2) \tag{94}$$

where

$$c_{1,2} = \frac{\log \frac{H^2}{S_t K} + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \tag{95}$$

Note that just like the Black-Scholes formula, this pricing function contains a *single* volatility parameter σ. In the plain vanilla case, this parameter could be manipulated to make the formula yield the correct answer, even when the underlying assumptions do not hold. Thus, in a smile environment with nonconstant volatilities, the trader could use *one* value for volatility and make the formula yield the correct answer. In fact, this was used in pricing caps and floors even though the volatilities of the individual forward rates that are relevant to these instruments were different.

Unfortunately, this approach is not guaranteed to work for the down-and-out call pricing formula given here if a volatility smile exists and if this smile is continuously (and stochastically) shifting over time. In fact, there may not be a single *well-behaved* volatility value to replace in equation (94) to obtain the correct price of the down-and-out call. Even when the realized and ATM implied volatilities remain the same, if the slope of the smile changes, the price of the down-and-out call may change as in the previous argument. In fact, if the smile becomes less negatively sloped, the short put component of the replicating portfolio will become relatively less expensive and the value of the down-and-out call may increase. Hence, in the case of exotic options, the relationship between volatility adjustments and the correct option price may become much more complex as smile effects become significant.

17.2.2. Pricing Exotics

Actual trading takes place in the presence of a volatility smile, and the prices of exotic options need to be set so as to take into account the future costs and benefits of adjusting the hedge to the exotic option. With the presence of smile, as time passes, the *vega* hedge of an exotic option book needs to be adjusted for the reasons explained earlier. As this happens, depending on the net position of the option book, the trader may realize some net cash gains or losses. The present value of these "expected" cash gains and losses needs to be incorporated into the market price. At the end, the market price of the exotic may be higher or lower than the theoretical price indicated by equation (94).

17.3. The Case of Digital Options

Chapter 10 showed that a theoretical replication of a European digital call with strike K and expiration T would be to buy $1/h$ units of the vanilla call with strike K and to sell $1/h$ units of the vanilla call with strike $K + h$. In this case h would be the minimum tick in a futures market.

If there is a volatility smile, then the prices of these vanilla calls would need to be adjusted since they have different strikes and, therefore, different volatilities. Of course, if h is small, this volatility difference would be small as well, but then $1/h$ would be large and the position would involve buying and selling a large number of vanilla calls. With such numbers, small variations in volatilities can make a difference on the end result.[18]

17.4. Another Application: Risk Reversals

One of the most liquid ways of trading the smile is using risk reversals from FX markets. Consider options written on an exchange rate e_t. Fix the expiration at T, and arrange the puts and calls by their strike price K_i. Then calculate these options' *deltas* and consider a grid of reasonable *deltas*. We use the options' *deltas* to represent the moneyness.

A typical smile for these exchange rate options will then look like the one shown in Figure 15-13. It is a "symmetric" curve, and is plotted in a two-dimensional graph having the percentage volatility on the vertical axis and the option's *delta* on the *x*-axis. In particular, consider the 25, 50, and 75-*delta* options.[19]

We look at the following example.

EXAMPLE:

Activity in the dollar/yen foreign exchange markets over the past fortnight has emphasized the severe complexities associated with pricing exotic options. More importantly, it has provided sophisticated option houses with an opportunity to test their pricing theories against each other and against their less-advanced competitors. . . .

However, it was not the decline in the spot rate itself that provided the interest for options dealers but the resulting risk-reversal position. Risk-reversal is an expression of the directional preference in the market. If spot is expected to fall, as in the case of dollar/yen, then there will be greater demand for puts relative to calls, and so the volatility trader will pay a higher price for the puts than the calls. The upshot of this, said one commentator, is that volatility is not constant, as assumed by the standard Black-Scholes option-pricing model, but instead changes according to the option delta.

One-month dollar/yen risk reversal shot up to nearly 3 last week and has continued to hover at levels not seen since the summer of 1995. This extraordinary situation enabled sophisticated traders to find value in the pricing of their so-called naive counterparts.

"A lot of banks must have learned a lot about risk-reversal over the past few days," said one trader. According to market insiders, the less advanced houses failed to adequately account for the effects of risk-reversal in pricing and hedging exotic structures such as knock-out and path-dependent options.

[18] In addition, note that this hedge requires selling and buying volatilities and, hence, is subject to variations in the volatility bid-ask spreads.

[19] Practitioners multiply the Black-Scholes *delta* by 100 in their daily usage of this concept.

> *They also asserted that simple off-the-shelf option pricing software was unable to cope with pricing exotic derivatives in a risk-reversal scenario. These packages did not allow the user to enter different volatilities for different deltas and so failed to capture the nuances of exotics pricing.*
>
> *However, other commentators argued that this too was an over-simplification and that other "third-order" effects came into play when hedging certain types of options in very high risk-reversal scenarios. They added that the third-order effects meant the barrier described might not be overpriced by the Black-Scholes user, merely mispriced. (IFR, Issue 1188 June 1997)*

As this reading illustrates, risk reversals are creations of the volatility smile in FX markets and are heavily traded. But, as indicated in the reading, market practitioners involved in risk reversal trading are clearly dealing with the underlying smile dynamics. The dynamics can become very complex. Pricing and hedging such positions on exotic options may become much more difficult.

18. Conclusions

The volatility smile is a fascinating topic in finance. Yet, it is also a complex phenomenon and more research needs to be done on its causes and on the ways to model it. This chapter offered a simple introduction. However, it has illustrated some of the essential points associated with this topic.

Suggested Reading

The text by ***Brigo and Mercurio*** *(2001) deals with the volatility smile in the interest rate sector. For equity smiles, the reader should consult the papers by* ***Derman*** *et al. (1994) and (1996), at minimum. A comprehensive treatment of the volatility smile is provided in* ***Lipton*** *(2002).* ***Taleb*** *(1996) is the source that we used to discuss most market practices. There is also a flurry of papers dealing with empirical and theoretical issues involving the volatility smile. One recent source is* ***Johnson and Lee*** *(2003). The cited sources contain further references on the previous research.*

Exercises

1. Consider the following table displaying the bid-ask prices for all options on the OEX index passed on January 10, 2002, at 9:46. These options have February 22, 2002, expiry and at the time of data collection, the underlying was at 589.14.

Calls	Bid	Ask	Vol	Puts	Bid	Ask	Vol
Feb 400	188.9	191.9	0	Feb 400	0.05	0.2	0
Feb 420	169	172	0	Feb 420	0.1	0.4	0
Feb 440	149.2	152.2	0	Feb 440	0.25	0.55	0
Feb 460	129.4	132.4	0	Feb 460	0.45	0.75	0
Feb 480	109.6	112.6	0	Feb 480	0.8	1.1	0
Feb 500	90.2	92.7	0	Feb 500	1.4	1.7	0
Feb 520	71	73.5	0	Feb 520	2.5	2.8	0
Feb 530	61.6	64.1	0	Feb 530	2.8	3.5	0
Feb 540	52.4	54.9	0	Feb 540	3.7	4.4	0
Feb 550	43.8	45.8	0	Feb 550	4.9	5.6	1
Feb 560	35.4	37.4	0	Feb 560	6.6	7.3	0
Feb 570	27.9	29.4	0	Feb 570	8.9	9.6	0
Feb 580	20.8	22	0	Feb 580	11.8	12.8	0
Feb 590	14.8	15.8	0	Feb 590	15.8	16.8	1
Feb 600	10	10.7	1	Feb 600	20.8	22	0
Feb 610	6.1	6.8	0	Feb 610	27.1	28.6	0
Feb 615	4.6	5.3	0	Feb 615	31	32	0
Feb 620	3.4	4.1	0	Feb 620	34.3	36.3	0
Feb 630	1.9	2.2	0	Feb 630	42.8	44.8	0
Feb 640	0.9	1.2	0	Feb 640	52	54	0
Feb 650	0.4	0.7	0	Feb 650	61.4	63.9	0
Feb 660	0.15	0.45	0	Feb 660	71.2	73.7	0
Feb 680	0	0.25	0	Feb 680	90.8	93.8	0
Feb 700	0	0.2	100	Feb 700	110.8	113.8	0

   (a) Using the out-of-the-money ask prices for the puts, calculate the implied volatility for the relevant strikes. Plot the volatility smile against K/S.

   (b) Using the out-of-the-money bid prices for the puts, calculate the implied volatility for the relevant strikes. Plot the volatility smile against K/S. Are the bid-ask spreads for these vols constant?

   (c) Using the out-of-the-money ask prices for the calls, calculate the implied vol for the relevant strikes. Plot the volatility smile against K/S. When you put this figure together with the out-of-the-money put volatilities, do you obtain a smile or a skew?

2. Consider the following statement:

   One prop trader noted that cap/floor volatility should be slightly higher than swaptions. Corporates buy caps and investors sell swaptions through callable bonds, said one London-based prop trader. The market is structurally short caps and long swaptions.

(a) Swaptions are options to get into swaps in a predetermined data, at a predetermined rate. Explain why, according to this reading, cap/floor volatility should be higher than swaption volatility.
(b) What are some plausible reasons for the market to be structurally short of caps and long on swaptions?
(c) What would this statement mean in terms of hedging and risk management?