This section begins the third segment of the course: a discussion of the volatility skew/smile – its sources and consequences. We start with a general discussion of the phenomenon, and its relation to “fat tails,” following Hull’s Chapter 15. There are three main approaches to modeling the volatility skew/smile quantitatively: (a) local vol models, (b) jump-diffusion models, and (c) stochastic vol models. We have time for just a very brief introduction to each; you’ll learn much more in the course Case Studies. Today we focus on local vol models, explaining why it is “in principle” easy but in practice quite difficult to extract a local volatility function from market data on calls. The heart of the matter is “Dupire’s equation.”

The implied volatility skew/smile. Consider call options on an underlying which earns no dividends. We assume the interest rate $r$ is constant. We write

$$C(S_0, K, T) = \text{market price of a call option with strike } K \text{ and maturity } T$$

where $S_0$ is the spot price and the current time is $t = 0$. Now define

$$C_{BS}(S_0, K, \sigma, T) = \text{Black-Scholes value of the call, using constant volatility } \sigma.$$

Then the implied volatility $\sigma_I(S_0, K, T)$ is defined by the equation

$$C(S_0, K, T) = C_{BS}(S_0, K, \sigma_I(S_0, K, T), T).$$

Since the Black-Scholes value of a call is a monotone function of $\sigma$, the implied volatility is well-defined. If the constant-vol Black-Scholes model were “correct,” i.e. if it gave the actual market values of call options, then $\sigma_I$ would be constant, independent of $S_0$, $K$, and $T$.

In fact however $\sigma_I$ is not constant. The “volatility skew/smile” refers to its dependence on $K$. Typically, for equities, $\sigma_I$ decreases as $K$ increases. For foreign exchange the typical behavior is different: $\sigma_I$ is smallest when $K \approx S_0$ so its graph looks like a “smile.”

The definition of implied vol depends on the choice of payoff. But if we used puts rather than calls we would get the same implied vols, by put-call parity. (We use here the fact that put-call parity is model-independent!).

Hull discusses at length how the skew/smile reflects “fat tails” in the risk-neutral probability distribution. No need to repeat his discussion here. But note what underlies it. Since prices are linear, the value of any option with maturity $T$ is a linear function of its payoff at time $T$. We recognize from the relation

$$\text{option value} = e^{-rT}E_{RN}[\text{payoff}]$$
that (for a complete market model) this linear relation is expressed by integration against
the risk-neutral probability density times $e^{-rT}$. In other words, if the payoff is $f(S_T)$ and
the risk-neutral probability density of $S_T$ at time $T$ (given price $S_0$ at time 0) is $p(\xi, T; S_0)$ then

$$\text{option value} = e^{-rT} \int_{-\infty}^{\infty} f(\xi) p(\xi, T; S_0) \, d\xi.$$ 

The volatility smile associated with foreign currency rates reflects the fact that the market-
place uses a $p$ with “fatter tails” than a lognormal distribution. This is reasonable, since
empirical studies suggest that in the real world, large moves are more likely than a log-
normal distribution would suggest. (Of course the empirical studies are of subjective, not
risk-neutral, probabilities; but change of measure changes the drift, not the volatility.) The
skew associated with a stock can be attributed to a different effect: the volatility tends to
increase when the stock price decreases, since a large decrease in the stock price is a sign
that the company is in trouble.

Implied volatilities depend on the time-to-maturity too. Thus we can speak of the “term
structure of implied volatility,” i.e. the dependence of the skew/smile on maturity.

What’s the goal here? Well, it’s natural to ask how the basic Black-Scholes model can be
changed to make it consistent with the observed prices of calls and puts of all strikes and
maturities. More practically: it’s natural to ask whether we can use information about the
implied vol skew/smile for improved pricing and/or hedging. Concerning pricing: don’t
expect too much – we use European option prices to deduce the implied volatility, so we
can’t expect to get anything new about the prices of European options. But we can hope
that a model calibrated using European options will do better than basic Black-Scholes for
valuing other, more exotic options (Americans, barriers, etc). And we can hope that an
improved model will provide improved hedges, even for Europeans.

OK, the goal is clear. But what changes of Black-Scholes should we consider? There are
basically three different approaches:

(a) **Local volatility models.** This approach relaxes the hypothesis that $\sigma$ be constant,
permitting it instead to be a (deterministic) function of $S$ and $t$. It has the advantage
of staying very close to the original Black-Scholes framework: the market is still
complete, options can still be hedged, etc. One disadvantage: the function $\sigma(S, t)$ is
difficult to extract from market data in a stable way.

(b) **Jump-diffusion models.** Why must the underlying execute a diffusion process? It
is obvious that unexpected news can make the market react abruptly. This is best
modelled by permitting the underlying to jump. This approach has advantage of being
intuitively plausible. One disadvantage: it involves two sources of randomness (the
jumps and the diffusion). So hedging cannot be achieved by trading the underlying
alone.

(c) **Stochastic volatility models.** Returning briefly to (a), let’s suppose the underlying
follows a diffusion process, i.e. there are no jumps. Why must the volatility be
deterministic? Market data are at least consistent with the idea that volatility is
itself random. Typically the process used to model volatility is mean-reverting. We
are thus led to study a system of two coupled SDE’s, for the price of the underlying and
the value of the volatility. This approach has the advantage of capturing relatively well
the observed time-series behavior of asset dynamics. Again, it has the disadvantage
of using two sources of randomness, so hedging cannot be achieved by trading the
underlying alone.

Each of these approaches has its proponents; most people believe (b) and (c) are closer
to the truth than (a). This is not a purely academic question: the different approaches
give, in many cases, different prices (for exotics) and different hedging strategies (even for
Europeans). A recent review of jump-diffusion and stochastic vol models, focusing on how
well they fit the observed skew/smile and term structure of implied volatility, is: S.R. Das
Anal. 34 (1999), 211-239 (available through JSTOR).

*****************

Local volatility models. As noted above, the local vol framework pursues the idea that the
only flaw in the Black-Scholes framework was the assumption that \( \sigma \) was constant. Put
differently: the local vol framework assumes the underlying solves an SDE of the form

\[
\frac{dS}{S} = r dt + \sigma(S, t) dw
\]

under the risk-neutral measure, where \( \sigma(S, t) \) is some deterministic function of \( S \) and \( t \). The
idea is to deduce the form of \( \sigma(S, t) \) from market data on European options (e.g. the implied
volatilities of calls, of various strikes and maturities). Once we know \( \sigma(S, t) \) we can price
and hedge options by using familiar means – e.g. by solving the associated Black-Scholes
PDE numerically. (For Europeans this would reproduce the prices used to calibrate the
model, but it would give us also the hedge; for American or barrier options we would get
predictions of prices as well as information on hedging.)

The good news is: the idea is internally consistent. Indeed: if the underlying really does
follow a diffusion of the form (1), then the prices of European call options (of all strikes and
maturities) provide, in principle, sufficient information to determine \( \sigma(S, t) \).

The bad news is: this program is rather difficult to pursue in practice. To see the difficulty
in its simplest form, suppose \( \sigma = \sigma(t) \) depends only on time. Then \( S(T) \) is lognormal, and
its drift and volatility depend only on \( \int_0^T \sigma^2(s) ds \). So from the prices of options (of all
maturities) we would get the value of \( \int_0^T \sigma^2(s) ds \) as a function of the maturity \( T \). But to
get \( \sigma(t) \) itself we would have to differentiate the results of our observations with respect
to \( T \). Alas, differentiation is very unstable. Real market data is noisy; therefore are our
“observations” of \( \int_0^T \sigma^2(s) ds \) are imperfect and/or noisy. The process of differentiation
amplifies the noise a lot – leading, if we’re not careful, to large errors in the estimation of
\( \sigma \).

I said “if we’re not careful.” What does it mean to be careful? Well, suppose you have deter-
mined, by noisy measurements, a function \( f(T) \) which is your best estimate of \( \int_0^T \sigma^2(s) ds \).
A good way to find \( \sigma(s) \) is to minimize, over some reasonable class of candidates, the error
between your measurement and the prediction of that candidate. In practice our “mea-
measurements” would not be continuous functions of time at all; rather they would be given at
selected maturities \(\{T_j\}_{j=1}^N\). So one could minimize (numerically)

\[
\sum_{j=1}^N \left| f(T_j) - \int_0^{T_j} \sigma^2(s) \, ds \right|^2
\]

over an appropriate (not-too-high-dimensional) family of candidate functions \(\sigma(s)\), defined
e.g. by splines.

Let’s turn now to the more general setting where \(\sigma = \sigma(S, t)\). Clearly the task of finding
\(\sigma(S, t)\) from market data will be at least as difficult as the one discussed just a moment
ago. Let’s nevertheless ask: can we understand the mathematical form of the problem in
a direct, transparent way – as we achieved above for the case when \(\sigma\) depends only on \(t\)?
Remarkably, the answer is yes, in the following sense. Suppose the underlying solves (1),
and suppose the stock price today is \(S_0\). Let \(C(K, T)\) be resulting value of a call with strike
\(K\) and maturity \(T\). Then \(C\) solves Dupire’s equation

\[
C_T - \frac{1}{2} \sigma^2(K, T) K^2 C_{KK} + rKC_K = 0,
\]

for all \(T > 0\) and \(K > 0\), with initial condition

\[
C(K, 0) = (S_0 - K)_+
\]

and boundary condition

\[
C(0, T) = S_0.
\]

Thus: market data on calls gives us the solution of the Dupire PDE, a PDE in “strike
and maturity space”, whose “diffusion coefficient” is \(\frac{1}{2} K^2\) times the unknown function \(\sigma^2\)
evaluated at “stock price” \(K\) and “time” \(T\). (Dupire’s equation was first derived in B.
Dupire, Pricing with a smile, Risk 7(1) 18-20, 1994. Analogous binomial-tree discussions
were given in E. Derman and I. Kani, Riding on a smile, Risk 7(2) 32-39, 1994, and in M.
is presented by Hull in Section 20.4.)

Formally, we can give a formula for \(\sigma\) by reorganizing Dupire’s equation:

\[
\sigma^2(K, T) = \frac{C_T + rKC_K}{\frac{1}{2} K^2 C_{KK}}.
\]

This formula is useless in practice, since our estimates of \(C_T, C_K,\) and \(C_{KK}\) will be hopelessly
inaccurate. But Dupire’s equation can still be used to estimate local vol from market data,
by using a more robust scheme analogous to (2): see e.g. L. Jiang, Q. Chen, L. Wang, and
J. Zhang, A new well-posed algorithm to recover implied local volatility, Quant. Finance 3
(2003) 451-457; also Y. Achdou and O. Pironneau, Volatility smile by multilevel least square,

Let’s explain why Dupire’s equation is valid. Easy things first: the boundary and initial
conditions are obvious. Indeed, a call with strike \(K = 0\) is equivalent to the stock itself,
so its value is $S_0$; and a call with maturity $T = 0$ has value equal to its payoff, namely $(S_0 - K)_+$.

There are two key ingredients to the derivation of Dupire’s equation:

(i) The forward Kolmogorov equation for the probability density of the underlying. Let $p(\xi, \tau)$ be the probability that the underlying has value $\xi$ at time $\tau$, given that it has value $S_0$ at time 0. Then $p$ solves the forward Kolmogorov PDE

$$p_t - \left( \frac{1}{2} \sigma^2(\xi, t) \xi^2 p \right)_{\xi\xi} + r(\xi p)_{\xi} = 0$$

for $t > 0$, with a “delta-function” at $\xi = S_0$ as initial data. (I suppose you learned about the forward Kolmogorov PDE in Stochastic Calculus; for a review of this topic, see Section 1 of my PDE for Finance lecture notes.)

(ii) The fact that the second derivative of the call payoff is a delta function. Thus:

differentiating the formula

$$C(K, T) = e^{-rT} E[(S - K)_+] = e^{-rT} \int (\xi - K)_+ p(\xi, T) \, d\xi$$

twice with respect to $K$ gives $C_{KK}(K, T) = e^{-rT}p(K, T)$.

Starting from these ingredients, the argument is easy. Differentiating with respect to $T$ the equation $e^{rT}C_{KK}(K, T) = p(K, T)$, we get

$$C_{KKT} + rC_{KK} = e^{-rT} p_T.$$

Combining this with the forward Kolmogorov equation and remembering that $e^{-rT} p(K, T) = C_{KK}$ we get

$$C_{KKT} + rC_{KK} = (\frac{1}{2} \sigma^2(K, T) K^2 C_{KK})_{KK} - r(KC_{KK})_K.$$

Integrate once in $K$ to get

$$C_{KT} + rC_K - (\frac{1}{2} \sigma^2(K, T) K^2 C_{KK})_K + rKC_{KK} = a(T)$$

where the right hand side is an unknown function of $T$ alone. Now, $KC_{KK} = (KC_K)_K - C_K$, so the preceding equation can be rewritten as

$$C_{KT} - (\frac{1}{2} \sigma^2(K, T) K^2 C_{KK})_K + r(KC_K)_K = a(T).$$

Therefore we can integrate again, obtaining

$$C_T - \frac{1}{2} \sigma^2(K, T) K^2 C_{KK} + rKC_K = a(T)K + b(T)$$

where $b(T)$ is another unknown function of $T$ alone. Finally, observe that as $K \to \infty$ the value of a call tends to zero – and so do its derivatives $C_T$, $C_K$, and $C_{KK}$ – since the probability density $p(\xi, \tau)$ decays as $\xi \to 0$. So the left hand side of (4) tends to zero as $K \to \infty$. Therefore the right hand side must do the same. This implies $a(T) = b(T) = 0$ for all $T$. Thus finally

$$C_T - \frac{1}{2} \sigma^2(K, T) K^2 C_{KK} + rKC_K = 0,$$

which is Dupire’s equation.
A different viewpoint, which I won’t pursue at any length, is the following. Rather than considering only constant $\sigma$ (corresponding to geometric Brownian motion) or considering arbitrary $\sigma(S,t)$ (as we have done above), one can consider simple functional forms for $\sigma(S,t)$ with just a few free parameters. Exact option pricing formulas can be given for some classes of such $\sigma$. If we’re lucky, there will be a choice of the parameters for which the associated smile/skew resembles what is seen in the marketplace. The problem with this approach is that its output depends strongly on the specific class of $\sigma$’s considered. So one can question whether its predictive value is much greater than the original constant-volatility framework. A good survey of work in this direction can be found in A. Lipton’s book *Mathematical Methods for Foreign Exchange: a Financial Engineer’s Approach*, World Scientific, 2001.