The Dupire Local Volatility Model

the one model that fits every vanilla, and the one promise it breaks

Martin Burke (with Claude)

Draft

Abstract

Black–Scholes assigns one volatility to an asset; the option market quotes a different implied volatility for every strike and every maturity. That gap—the volatility *smile*—is not noise, and no single number can close it. Bruno Dupire's 1994 insight was that we do not need a single number: if we let instantaneous volatility be a deterministic function $\sigma_{\text{loc}}(t, S)$ of time and spot, then there is exactly *one* such function that reprices the entire surface of traded vanilla options, and it can be read straight off market prices by a single formula. This note builds that result from the ground up. We start from the smile itself, recover the risk-neutral density from option prices (Breeden–Litzenberger), derive Dupire's forward equation and the formula it yields, and then re-read that formula two ways—in terms of quoted implied volatilities, and as a conditional average of the "true" instantaneous volatility. We then show how the formula must be amended once interest rates are stochastic—where a correlation-dependent correction term appears and the model-free "read it off the surface" property is partly lost—and how the model is used to hedge: the behaviour of its Greeks, the backbone that biases its delta, and the vega matrix that turns an exotic's volatility risk into a basket of vanillas. We close with what local volatility is excellent at (a perfect, unique, arbitrage-free fit to today's surface) and the one thing it gets wrong—the dynamics of the *future* smile—which is exactly where stochastic volatility takes over. Appendices collect the supporting detail kept out of the main line of argument: the Black–Scholes formula and backward PDE, the Fokker–Planck equation and the full forward-equation derivation, the implied-variance form with its short-maturity "rule of two," the SVI/SSVI surface parametrisation used to generate every figure, the change the formula needs once interest rates are stochastic, and the risk decompositions behind the Greeks. The local-volatility model is extracted from a synthetic market surface and then Monte-Carlo simulated to confirm it reprices that surface.

Contents

1 The problem Black–Scholes leaves behind

1.1 One number, and why it is not enough

The Black–Scholes model represents a stock price as geometric Brownian motion (GBM). Under the risk-neutral measure $\mathbb{Q}$,[1] with a constant risk-free rate r and dividend yield q,

$$dS_t \;=\; (r-q)\,S_t\,dt \;+\; \sigma\,S_t\,dW_t, \tag{1}$$

where W_t is a Brownian motion and σ, the volatility, is a single constant. This is the workhorse every quant meets first, and its appeal is that one parameter prices every European option through one closed-form expression, the Black–Scholes formula $C_{\mathrm{BS}}(S_0, K, T; \sigma)$ for a call of strike K and maturity T (written out in Appendix A).

The trouble is empirical. Take the market prices of all the calls and puts on a single underlying, and ask: *what constant σ would Black–Scholes need to reproduce each one?* That number—the value of σ that makes the formula match the observed price—is the *implied volatility* $\sigma_{\mathrm{BS}}(K,T)$. If GBM were literally true, every option on the same underlying would imply the *same* σ, and a plot of implied volatility against strike would be a flat horizontal line. It never is.

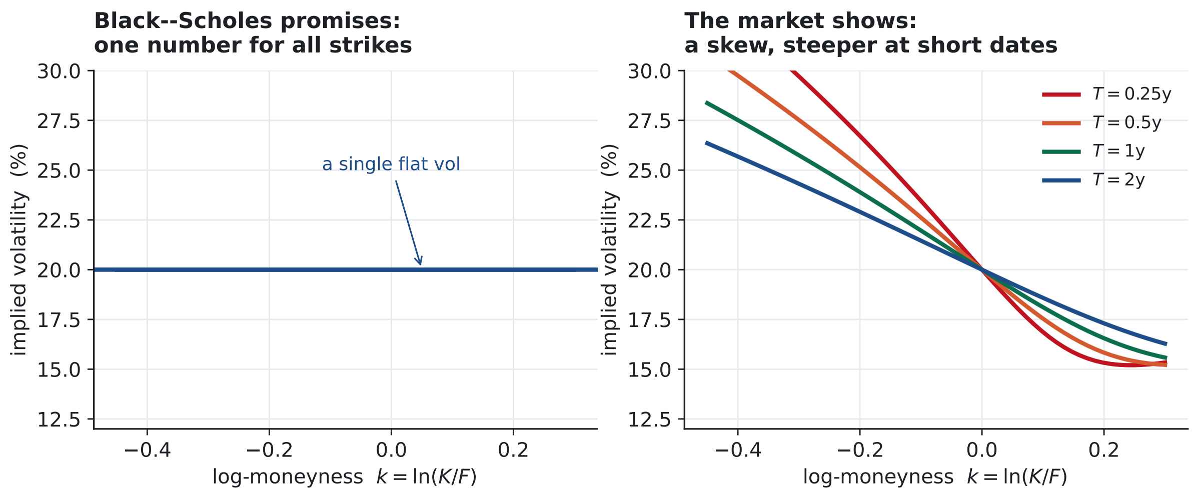

Figure 1: The defect in one picture. **Left:** Black–Scholes assumes a single volatility, so it predicts one flat line of implied vol across all strikes. **Right:** the market instead shows a downward-sloping curve—the *skew*—that is steeper at short maturities and flatter at long ones. Strike is shown as log-moneyness $k = \ln(K/F)$ (motivated in the text), so $k = 0$ is at-the-money. One number cannot reproduce these shapes.

Figure 1 shows what the market actually looks like. At a fixed maturity the curve $K \mapsto \sigma_{\mathrm{BS}}(K,T)$ is called the *smile* (for equity indices it is really a one-sided *skew*: low-strike puts are dearer, in implied-vol terms, than high-strike calls). Stack these curves across maturities and you get the *implied-volatility surface* of Figure 2—a two-dimensional sheet $\sigma_{\mathrm{BS}}(K,T)$ that the desk re-marks every day from traded option prices. This surface is the raw material for everything that follows; it is what the model must fit.

[1] The *risk-neutral measure* is the probability measure under which every traded asset, discounted at the risk-free rate, is a martingale—a fair game with no drift edge. Pricing by discounted expectation under $\mathbb{Q}$ is equivalent to pricing by replication, and it is why the stock's real-world expected return is replaced here by the financing cost $r-q$: under $\mathbb{Q}$ the drift of any tradable is pinned by no-arbitrage, not by risk appetite.

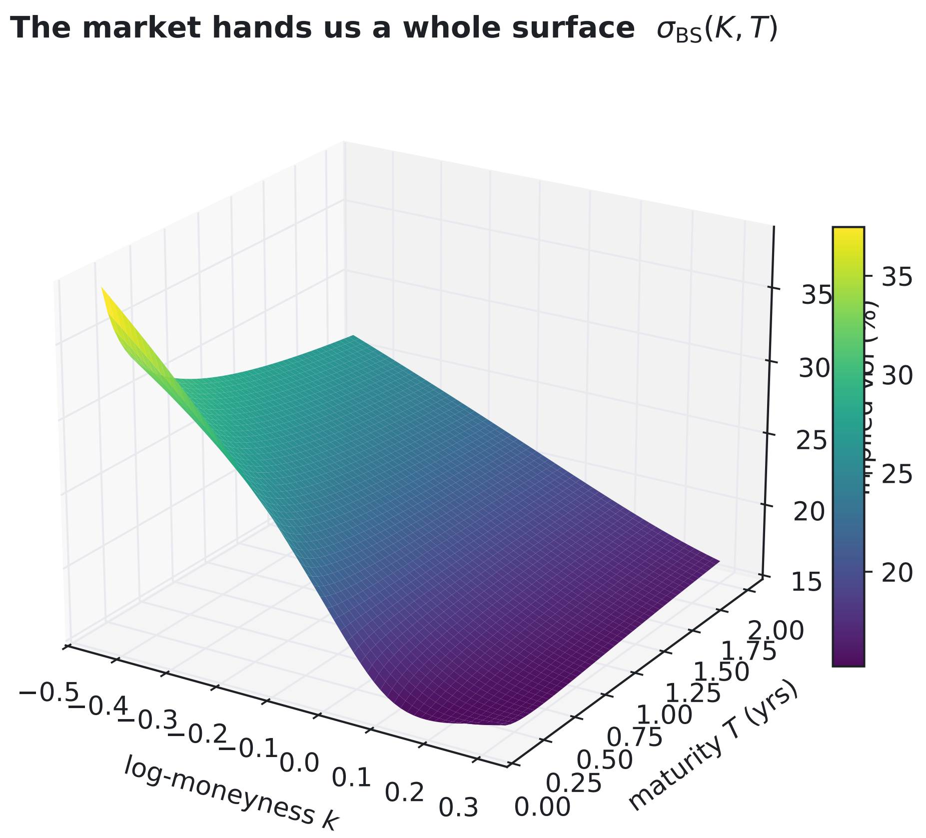

Figure 2: The implied-volatility surface $\sigma_{\mathrm{BS}}(K, T)$: one implied volatility for every strike *and* every maturity. Black–Scholes would render this as a flat horizontal plane. The market's version slopes and curves—high in the low-strike, short-maturity corner, flattening out as maturity grows. (This surface is generated from an arbitrage-free SSVI parametrisation; see Section 7.)

Pinning down the units: log-moneyness. Because the whole note is about the shape of this surface, it is worth fixing the horizontal axis. We do not plot implied volatility against the raw strike K but against *log-moneyness*,

$$k \;=\; \ln(K/F), \tag{2}$$

where F is the *forward price* of the underlying for delivery at T.[2] Three reasons make this the natural unit rather than a cosmetic one. First, dividing by the forward *re-centres* the picture: $k = 0$ is exactly at-the-money-forward, so smiles at different maturities and different spot levels become directly comparable, all pinned to a common origin. Second, taking the logarithm matches the model: GBM makes *log*-returns the natural coordinate—a 10% rise and a 10% fall are symmetric in k but not in K—so a lognormal world would be flat in k, and any curvature is a genuine departure. Third, k is dimensionless, independent of the quote currency or the absolute price level. With the axis fixed, $k < 0$ is the low-strike (put) wing, $k > 0$ is the high-strike (call) wing, and the *skew* at maturity T is the slope of the smile at the money,

$$\mathcal{S}(T) \;=\; \left.\frac{\partial\, \sigma_{\mathrm{BS}}(K, T)}{\partial k}\right|_{k=0}. \tag{3}$$

For equity indices $\mathcal{S}(T) < 0$: the smile slopes *down* to the right, as in Figure 1. A flat smile is zero skew; a steeply down-sloping one is a large (negative) skew. That single quantity, and how

[2] The forward price is the price agreed today for delivery of the stock at T; under the risk-neutral measure it is the expected future spot, $F = \mathbb{E}[S_T] = S_0\, e^{(r-q)T}$. With zero rates and dividends, as we assume from Section 3 on, $F = S_0$.

it behaves today versus on a future date, is the thread running through the note.

1.2 Why the smile exists

Before fixing the model it helps to know why the line bends in the first place. The constant-σ GBM of (1) produces a *lognormal* distribution for S_T, which is thin-tailed and symmetric in log-space. Real return distributions are neither: equity returns have a fat left tail (crashes are larger and more sudden than rallies) and are negatively skewed. The market prices that reality in. A fat left tail makes low-strike puts more valuable than lognormal maths says they should be, and the only way to express "more valuable" in the single dial Black–Scholes offers is to quote a *higher implied volatility* for those strikes. The skew is the market's way of overwriting the lognormal assumption, one strike at a time. Any model that hopes to match the surface must therefore generate a non-lognormal, skewed distribution for S_T—and must do so consistently across every maturity at once.

2 The local-volatility idea

Dupire's move is disarmingly simple. Keep the one-factor diffusion of Black–Scholes, but stop insisting that volatility be constant. Let it be a deterministic *function* of the current time and the current spot:

$$\mathrm{d}S_t \;=\; (r-q)\,S_t\,\mathrm{d}t \;+\; \sigma_{\mathrm{loc}}(t,S_t)\,S_t\,\mathrm{d}W_t. \tag{4}$$

Here $\sigma_{\mathrm{loc}}(t,S)$ is the *local volatility*: the instantaneous volatility the stock will experience *if* it finds itself at level S at time t. It is still one Brownian motion, still a Markov diffusion, still a complete market.[3] All the comforts of Black–Scholes survive. What changes is that a *surface* of volatility values, rather than a single one, is now available to shape the distribution of S_T.

The word "local" is the key. In (1) volatility is global: the same σ applies everywhere. In (4) volatility is attached to a *place* in the (t,S) plane. A path that drifts down into a high-volatility region will diffuse faster there; one that climbs into a calm region will slow down. Figure 3 shows such a field, with simulated spot paths threading through it: the colour at each point is the volatility a path would feel were it to pass through. The downward-sloping equity skew corresponds to a field that is hot at low spot and cool at high spot, exactly as drawn.

[3] *Markov* means the future depends on the present only through the current spot S_t—there is no hidden second state (such as a random variance) carrying memory, so (t,S) alone determines the dynamics. *Complete* means every payoff can be replicated, and hence priced and hedged, by trading the stock and cash alone; with a single Brownian driver and one tradable, the market stays complete exactly as in Black–Scholes. This is what a stochastic-volatility model gives up (Section 6).

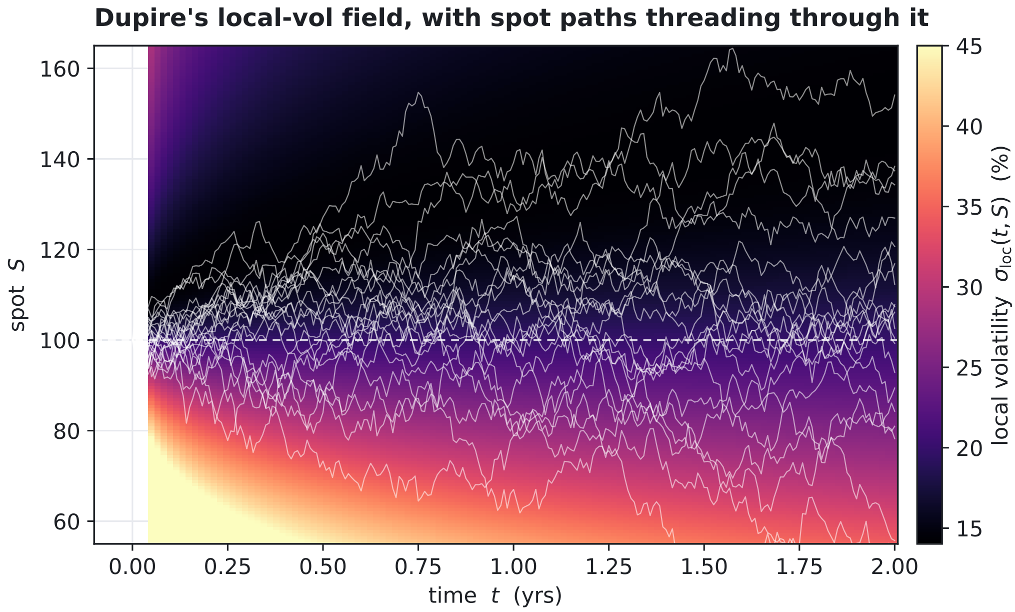

Figure 3: A local-volatility field $\sigma_{\rm loc}(t,S)$ (colour), with Monte-Carlo spot paths drawn in white. Volatility is no longer one number but a value attached to each point of the time–spot plane: low at high spot, high at low spot, which is what produces the equity skew. Each path feels the volatility of wherever it happens to be. This particular field was *extracted* from the surface of Figure 2 using the Dupire formula derived in Section 4.

Two questions immediately arise, and the rest of the note answers them. *First*—is there a function $\sigma_{\rm loc}(t,S)$ that reproduces the *entire* market surface, and if so, is it unique? *Second*—if it exists, how do we find it, ideally without guessing and re-simulating? Dupire's theorem answers both at once: yes, such a function exists and is unique, and it is given by an explicit formula in the traded option prices. To get there we first need to see what those prices are really telling us.

3 What the option market is really quoting: a density

A European call of strike K and maturity T pays $(S_T - K)^+$. Under the risk-neutral measure its price today is the discounted expected payoff. To keep the algebra clean we set $r = q = 0$ for the rest of the derivation (so the forward equals spot, $F = S_0$, and discount factors are 1); the general case is recorded in Appendix F. Then

$$C(K,T) \;=\; \mathbb{E}\big[(S_T - K)^+\big] \;=\; \int_K^\infty (S-K)\,p_T(S)\,\mathrm{d}S, \tag{5}$$

where $p_T(S)$ is the *risk-neutral density* of S_T—the probability density the market assigns to where the stock lands at time T. Equation (5) says the call price is just an average of its payoff over that density.

The remarkable thing is that the relationship inverts. The option prices, as a function of strike, *contain* the density. Differentiate (5) once with respect to the strike K (the lower limit of the integral),

$$\frac{\partial C}{\partial K} \;=\; -\int_K^\infty p_T(S)\,\mathrm{d}S \;=\; -\,\mathbb{Q}\big(S_T > K\big), \tag{6}$$

so the slope of the call curve in strike is (minus) the probability of finishing in the money. Differentiate once more and the density itself appears:

$$\frac{\partial^2 C}{\partial K^2} \;=\; p_T(K). \tag{7}$$

This is the *Breeden–Litzenberger* identity. Its meaning is worth saying in plain terms: the curvature of the call-price curve, strike by strike, *is* the risk-neutral probability density of the underlying.

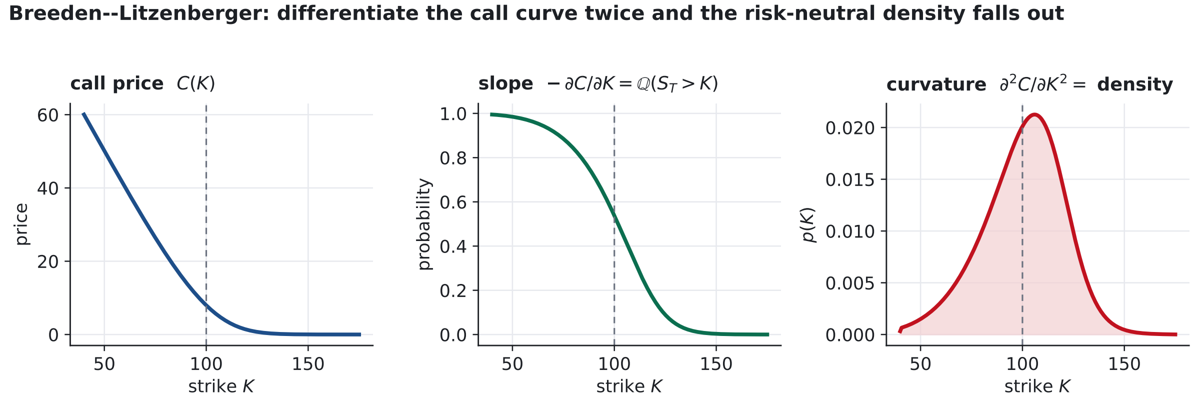

Figure 4: Breeden–Litzenberger, computed from the $T = 1$ slice of our surface. **Left:** the call price $C(K)$ falls from near-spot value to zero as the strike rises. **Middle:** its slope, negated, is the probability the stock ends above K. **Right:** its curvature is the risk-neutral density $p_T(K)$—note the fat left tail and the peak shifted below $S_0 = 100$, the non-lognormal shape the skew encodes. Differentiating the price curve twice recovers the market's full view of where the stock might land.

There is an intuition behind (7) that needs no calculus. A *butterfly* spread—long a call at $K - \Delta$, short two at K, long one at $K + \Delta$—pays off only if the stock lands in a narrow band around K, and its cost is precisely the discretised second difference of C in strike. So the price of a thin butterfly at K is the market's price for "S_T lands near K," which is exactly the probability mass there. Take the limit $\Delta \to 0$ and the butterfly price becomes $\partial^2 C/\partial K^2$, the density. Figure 4 carries this out numerically: differentiate the call curve once for the exceedance probability, twice for the density.

The upshot is that each maturity slice of the surface is, after a change of coordinates, a probability distribution for S_T. The market is not quoting volatilities so much as quoting an entire *family* of distributions, one per maturity. Dupire's formula is the bridge from that family to the dynamics that generate it.

4 Dupire's forward equation

4.1 Two ways to look at an option-pricing PDE

Every quant knows the Black–Scholes PDE: fix a *single* option—one strike, one maturity—and its value $V(t, S)$ as a function of calendar time and spot satisfies a partial differential equation that is solved *backward* in time from the known payoff at maturity. The unknowns are t and S; the strike and maturity are fixed parameters of the contract. This is the natural view for pricing one derivative.

Dupire turned the picture inside out (Figure 5). Hold the *starting point* $(t = 0, S = S_0)$ fixed—it is just today's market—and instead let the call price $C(K, T)$ vary over the grid of *strikes and maturities*. That object is observable directly: it *is* the price surface. Dupire showed it obeys its own PDE, run *forward* in maturity, in which the local volatility appears as a coefficient. Because everything in that equation except σ_{loc} is observable, it can be solved *for* σ_{loc}. The backward equation prices one option given the volatility; the forward equation extracts the volatility given the option prices.

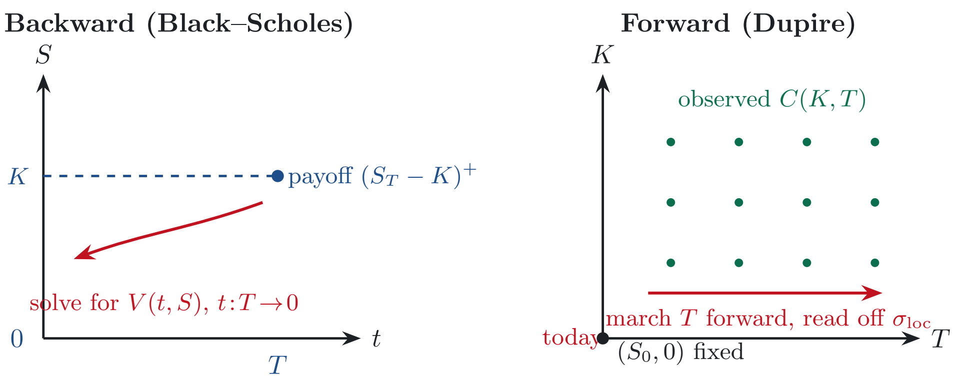

Figure 5: Two views of the same market. **Left:** the Black–Scholes backward PDE fixes one contract (one strike K, one maturity T) and solves for its value over all (t, S), running back from the payoff. **Right:** Dupire fixes today's spot and lets the call price range over all strikes and maturities; the resulting forward PDE marches outward in maturity and exposes the local volatility as the one unknown coefficient.

4.2 The equation, and the formula

The derivation has two ingredients, both already in hand: the Breeden–Litzenberger identity (7), and the fact that the density $p_T(S)$ of a diffusion evolves forward in time according to the *Fokker–Planck* (forward Kolmogorov) equation. For the local-vol SDE (4) with $r = q = 0$ this reads

$$\frac{\partial p}{\partial T} = \frac{1}{2}\frac{\partial^2}{\partial S^2}\Big[\sigma^2_{\text{loc}}(T,S)\, S^2\, p(T,S)\Big]. \tag{8}$$

Fokker–Planck is the conservation law for probability: it says how the cloud of possible spot values spreads and drifts as maturity advances (a one-paragraph refresher is in Appendix B). Figure 6 shows that cloud explicitly for our surface—a near-spike of certainty at the shortest maturity, spreading and developing the fat left tail of the equity skew as maturity grows. This is exactly the family of densities of Figure 4, now stacked across maturities; the Fokker–Planck equation is the rule that carries one into the next.

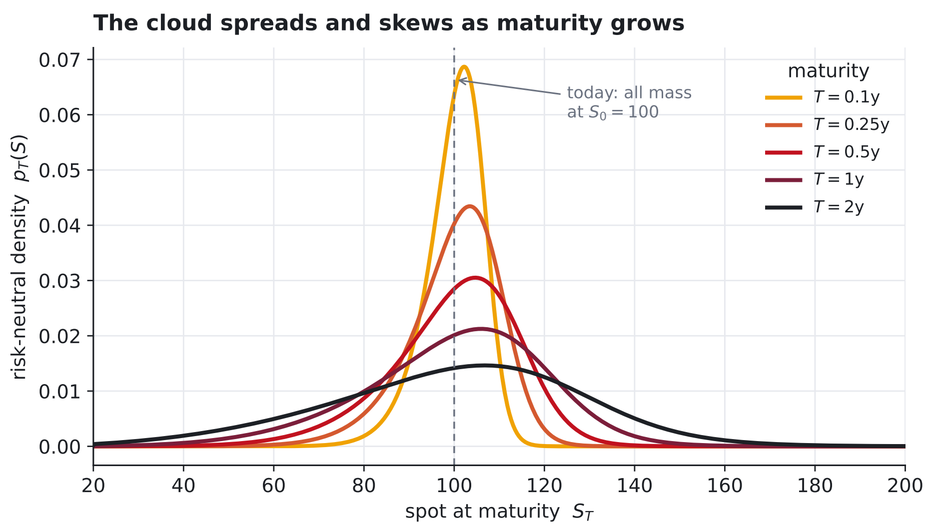

Figure 6: What the Fokker–Planck equation governs. Each curve is the risk-neutral density $p_T(S)$ of the spot at a different maturity, recovered from the surface by Breeden–Litzenberger (7). Today the stock is known to sit at $S_0 = 100$; as maturity lengthens the distribution spreads out and leans to the left (the fat-tailed, skewed shape the smile encodes). Dupire's local volatility is the diffusion coefficient that reproduces precisely this forward march of densities.

Differentiating the call price (5) in maturity, substituting (8), and integrating by parts twice (the full steps are in Appendix C) collapses everything to a strikingly simple statement,

$$\boxed{\frac{\partial C}{\partial T} \;=\; \frac{1}{2}\,\sigma_{\mathrm{loc}}^2(K,T)\,K^2\,\frac{\partial^2 C}{\partial K^2}}. \tag{9}$$

This is *Dupire's forward equation.* Read left to right it is a forward PDE for the price surface. But every term except σ_{loc} is observable from the market, so we simply solve for the local volatility:

$$\boxed{\sigma_{\mathrm{loc}}^2(K,T) \;=\; \frac{\dfrac{\partial C}{\partial T}}{\dfrac{1}{2}\,K^2\,\dfrac{\partial^2 C}{\partial K^2}}}. \tag{10}$$

This is the *Dupire formula*—the centrepiece of the model. Given a sufficiently smooth surface of call prices, it returns the unique local volatility at every (K,T) directly, with no iteration and no guessing.

Reading the formula. Both pieces have a plain meaning. The numerator, $\partial C/\partial T$, is the rate at which a call gets more valuable as you lengthen its maturity: hold the strike fixed and ask one more day to expiry, and this is what the option is worth. More time means more room for the stock to move, so it is always positive. The denominator, $\frac{1}{2}K^2\,\partial^2 C/\partial K^2$, is $\frac{1}{2}K^2$ times the risk-neutral density at the strike (by Breeden–Litzenberger, (7))—in words, how much probability mass the market parks right at K. So Dupire's formula says:

$$\sigma_{\mathrm{loc}}^2(K,T) \;=\; \frac{\text{rate the call gains value with maturity}}{\frac{1}{2}K^2 \times (\text{probability mass sitting at the strike})}. \tag{11}$$

Where the market puts a lot of probability mass and the call is barely gaining value with time, local volatility must be low; where mass is thin but the call still gains value quickly, local volatility must be high. The local-volatility surface is exactly the field that makes those two market-observed quantities consistent with a diffusion.

Existence and uniqueness. Equation (10) answers both questions from Section 2. If the surface is free of arbitrage the right-hand side is well defined and positive (the numerator and the curvature are both positive), so a local-volatility function *exists*; and because the formula returns a single value at each (K,T), it is *unique*. There is one and only one local-volatility diffusion consistent with a given arbitrage-free surface of European prices. That is a strong statement, and the source of both the model's appeal and its limits.

4.3 Proof of concept: extract, simulate, recover

A formula is only convincing if it works end to end. Figure 7 closes the loop. We start from the implied-volatility surface of Figure 2, apply Dupire's formula (10) to get the local-volatility field of Figure 3, then run a Monte-Carlo simulation of the local-vol SDE (4) forward through that field and re-imply the volatilities of the options it produces. The simulated points land back on the input curves to a fraction of a volatility point. The model really does reprice the surface it was built from—Dupire's theorem, demonstrated rather than asserted.

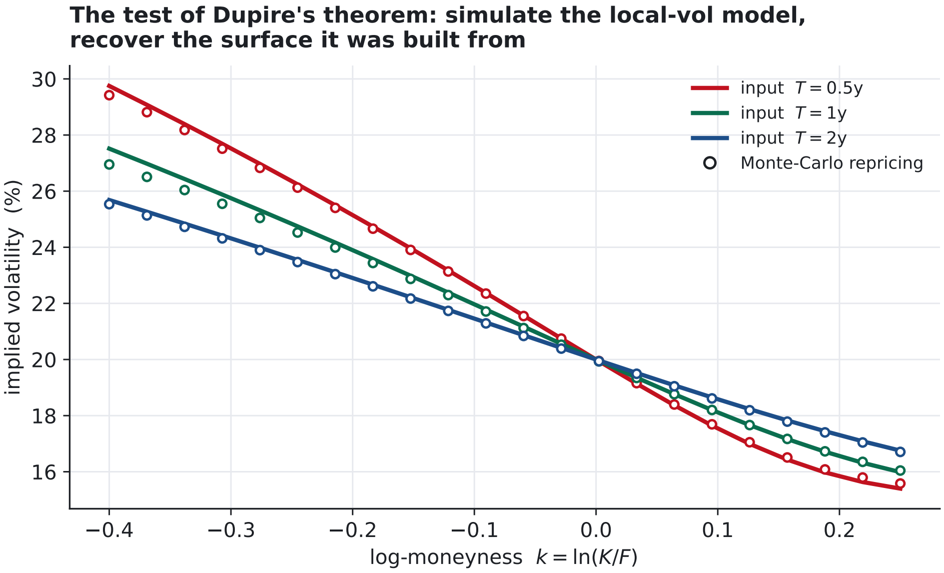

Figure 7: The round trip. Solid lines are the input market smiles at three maturities. Open circles are the implied volatilities recovered by Monte-Carlo-simulating the local-vol model whose field was extracted from those very smiles. They coincide to within $\sim$0.5 volatility points (the residual is discretisation and Monte-Carlo noise). The local-volatility model fits the surface exactly, by construction.

5 The same formula in the language of implied volatility

Equation (10) is written in call prices, but desks store and interpolate the surface in *implied volatilities*, not prices. It is therefore standard to rewrite Dupire in terms of $\sigma_{\mathrm{BS}}(K,T)$. The cleanest version uses not implied volatility itself but *total implied variance*,

$$w(k,T) \;=\; \sigma_{\mathrm{BS}}^2(k,T)\,T, \qquad k = \ln(K/F). \tag{12}$$

This is the implied variance accumulated to maturity, and it is the right variable for a concrete reason: variance, not volatility, is what scales linearly with time under Brownian motion. In a flat Black–Scholes world the variance of the log-spot grows as $\sigma^2 T$, so w would be exactly linear in T; real surfaces bend away from that line, but w remains the quantity in which the no-arbitrage conditions take their simplest form (Section 7). Substituting the Black–Scholes price written in terms of w into (10) and working through the chain rule gives Gatheral's form,

$$\sigma_{\mathrm{loc}}^2(k,T) \;=\; \frac{\partial_T w}{1 - \dfrac{k}{w}\,\partial_k w + \dfrac{1}{4}\left(-\dfrac{1}{4} - \dfrac{1}{w} + \dfrac{k^2}{w^2}\right)(\partial_k w)^2 + \dfrac{1}{2}\,\partial_k^2 w}. \tag{13}$$

It looks forbidding, but it is the same content as (10) expressed in the coordinates a trading system actually holds; the change of variables is sketched in Appendix D. Every figure in this note that needed a local volatility was produced with (13).

The "rule of two." The implied-vol form exposes a fact worth carrying around as intuition. Near the money, and to leading order, the *local* volatility skew is about *twice* the *implied* volatility skew. The reason is that implied volatility at strike K is a kind of *average* of local volatility over all the paths and spot levels between today's spot and K, whereas local volatility is the value at the endpoint K itself. Averaging a sloping line halves its slope—the running average of a line that climbs at rate a from the money out to k has itself only climbed by $ak/2$ at that point, i.e. at half the rate—so to produce a given implied slope, the underlying local slope has to be roughly double. The short-maturity expansion that turns this hand-waving into a precise statement,

$$\sigma_{\mathrm{loc}}(k,0) \;=\; \sigma_{\mathrm{BS}}(k,0) + k\,\partial_k\sigma_{\mathrm{BS}}(k,0), \tag{14}$$

is worked out in Appendix D. Figure 8 shows the two curves together: the local-vol smile is visibly steeper through the money, by close to a factor of two. The practical lesson is that local volatility is a *more extreme* object than the implied surface that generates it—its wings and its skew are exaggerated, which is one reason the extraction is numerically delicate (Section 7).

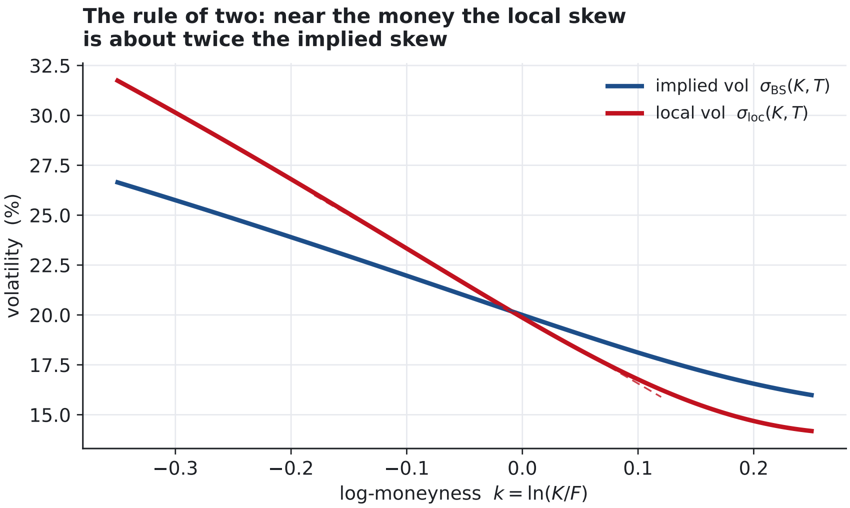

Figure 8: The rule of two. The implied-volatility smile (blue) and the local-volatility smile (red) at the same maturity, with their at-the-money tangents (dashed). Local volatility passes through the same at-the-money point but slopes down about twice as steeply, because implied vol is an average of local vol along the way while local vol is the value at the destination.

6 Local volatility as an average

So far local volatility is a curve-fitting device: the field that happens to reprice the surface. There is a deeper reading that explains both why it works and why, ultimately, it disappoints. Suppose the "true" instantaneous variance of the stock is some genuinely random process v_t (this is what a stochastic volatility model assumes—variance with a life of its own). One can ask: what is the *deterministic* function of (t, S) that, plugged into a one-factor diffusion, would reproduce the same European option prices as the true random process? The answer, due to Gyöngy, is a conditional expectation:

$$\boxed{\sigma^2_{\mathrm{loc}}(K,T) \;=\; \mathbb{E}\big[\, v_T \;\big|\; S_T = K \,\big]}. \tag{15}$$

Local variance at a given strike and maturity is the *average of the true instantaneous variance over exactly those scenarios in which the stock is at that strike at that time.* Dupire's formula and Gyöngy's identity describe the same object from two directions: the former extracts it from prices, the latter explains what it *is*.

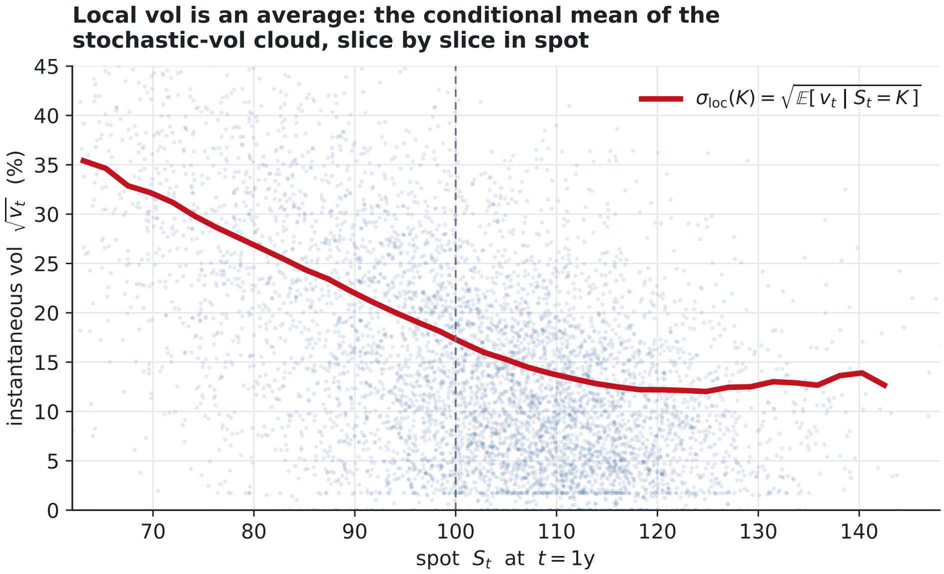

Figure 9: Local volatility is a conditional average. Each faint blue dot is one Monte-Carlo scenario of a stochastic-volatility model at $T = 1$: a pair (spot, instantaneous vol). The cloud is wide—at any given spot the true volatility takes a whole range of values. The red curve is the conditional mean $\sqrt{\mathbb{E}[v_T \mid S_T = K]}$, slice by slice in spot. *That curve is the local volatility.* It is the best deterministic stand-in for a random process, one spot-column at a time.

Figure 9 makes (15) concrete. The blue cloud is the joint distribution of spot and instantaneous volatility under a stochastic-vol model; the red curve is its column-by-column average. Local volatility keeps the *mean* volatility at each spot level and discards the *spread* around it. That is precisely the right thing to do if all you care about is repricing today's European options—a European payoff at T depends only on the distribution of S_T, and (15) guarantees that distribution is matched. But it is the wrong thing to do if you care about how volatility and spot *co-move through time*, because the discarded spread is exactly that information. This is the seed of the model's one real weakness, which we turn to next.

7 Calibration in practice

The Dupire formula is exact, but it is also a recipe for amplifying noise. It asks for the *second* derivative of the price surface in strike and the *first* in maturity. Differentiating twice turns small wiggles in the input into large swings in the output, and market quotes are discrete, noisy, and exist only at a handful of strikes and maturities. Applied naively to raw quotes, (10) returns a jagged, sometimes negative, "variance"—which is meaningless. Practitioners therefore never differentiate raw quotes. The work is in building a smooth, arbitrage-free surface first, and only then applying the formula (or, equivalently, solving the forward PDE (9) on a grid). Three requirements shape that step.

- **No butterfly arbitrage.** The density (7) must be non-negative everywhere, i.e. the call price must be *convex* in strike. Violate this and the Dupire denominator goes negative and the local variance is imaginary.
- **No calendar arbitrage.** Total implied variance (12) must be non-decreasing in maturity: a longer-dated option cannot be worth less in variance terms than a shorter-dated one at the same moneyness. Violate this and the Dupire numerator $\partial C/\partial T$ goes negative.
- **Sensible wings.** Far out-of-the-money there are few or no quotes, yet diffusing paths still reach those strikes, so the surface must be extrapolated in a controlled, arbitrage-free way.

Meeting all three at once by hand is hopeless, so the market-standard approach is to fit the surface with a *parametric form* that has smoothness and arbitrage-freedom built into its functional shape. The dominant such form is *SVI*—"stochastic-volatility-inspired," so named because its shape coincides with the large-maturity smile of the Heston stochastic-volatility model, not because it is itself stochastic.[4] For a single maturity, SVI models the total variance $w(k)$ as a hyperbola in log-moneyness: linear in each wing (far in- and out-of-the-money the smile is roughly straight in k) joined by a smooth, convex bottom around the money. Five parameters set the overall level, the slopes of the left and right wings, the position of the minimum, and how rounded the bottom is—enough freedom to fit any one observed smile, while convexity (no butterfly arbitrage) holds automatically for sensible parameter ranges. Its surface extension, *SSVI*, ties all the maturities together through a single at-the-money variance term-structure θ_T and a one-parameter shape function, which makes the whole sheet free of *calendar* arbitrage by construction as well. The formulas, the meaning of each parameter, and the exact values used to generate every figure in this note are collected in Appendix E. (Because our surface is SSVI, it is arbitrage-free to begin with, which is why the round-trip of Figure 7 is so clean: the extracted local variance is smooth and positive everywhere.) With a good surface in hand, the local-vol field follows immediately from (13), and the model is ready to price exotics by simulation or by solving its backward PDE. The pipeline is summarised in Figure 10.

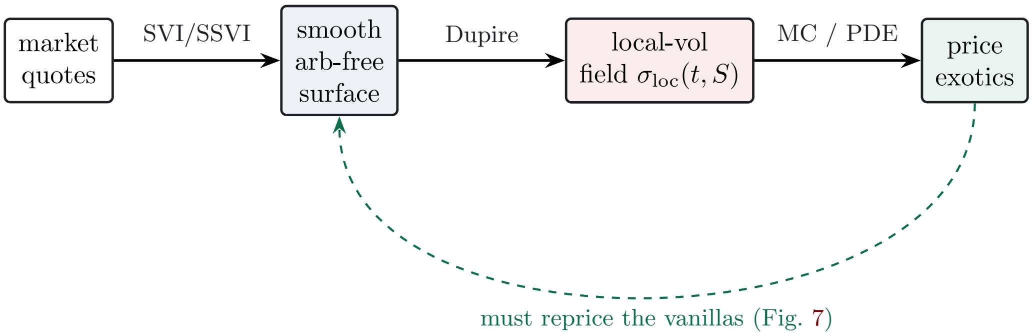

Figure 10: The calibration pipeline. Raw quotes are first fit to a smooth, arbitrage-free surface (SVI/SSVI); Dupire's formula then turns that surface into the local-vol field; the field prices exotics by Monte-Carlo or PDE. The dashed arrow is the consistency check: simulating the field must return the vanilla prices it came from.

8 Extension: stochastic interest rates

Everything so far assumed the discount rate was known in advance—constant, or at worst a deterministic curve. For short-dated equity options that is harmless: over a few months the randomness of rates barely registers next to the randomness of the stock. It stops being harmless for *long-dated* options and, especially, for *hybrid* products whose payoff depends on both equity and rates (a convertible bond, an equity note with a funding leg, anything with a multi-year horizon). There the assumption that rates are frozen is exactly the one that fails, and the Dupire formula needs amending. This section sets out how, and why the clean "read it off the surface" property is partly lost.

[4] SVI was introduced inside Merrill Lynch by Jim Gatheral and popularised in his book [6]; the arbitrage-free surface version, SSVI, is due to Gatheral and Jacquier [7].

8.1 What actually changes

Make the short rate its own stochastic process—take the Hull–White model as the standard workhorse—correlated with the spot:

$$\begin{aligned} \mathrm{d}S_t &= (r_t - q)\, S_t\, \mathrm{d}t + \sigma_{\mathrm{loc}}(t, S_t)\, S_t\, \mathrm{d}W_t^S, \\ \mathrm{d}r_t &= a\,(b_t - r_t)\, \mathrm{d}t + \sigma_r\, \mathrm{d}W_t^r, \qquad \mathrm{d}W_t^S\, \mathrm{d}W_t^r = \rho_{S,r}\, \mathrm{d}t. \end{aligned} \tag{16}$$

The whole of the original derivation rested on discounting being a known constant that could be pulled outside the expectation. It no longer can. The discount factor is now random,

$$D_T \;=\; \exp\Big(- \int_0^T r_s\, \mathrm{d}s \Big), \tag{17}$$

and a call price is a genuine *joint* expectation of the discount factor and the payoff together,

$$C(K, T) \;=\; \mathbb{E}^{\mathbb{Q}}\big[\, D_T\, (S_T - K)^+ \,\big], \tag{18}$$

the two factors free to move together. Three quantities that were trivial before now carry information. Today's zero-coupon bond price and the forward price are

$$P(0, T) \;=\; \mathbb{E}^{\mathbb{Q}}[D_T], \qquad F(0, T) \;=\; \frac{\mathbb{E}^{\mathbb{Q}}[D_T\, S_T]}{P(0, T)}, \tag{19}$$

so the forward is no longer S_0 grown at a known rate but is read from the bond market. And the Breeden–Litzenberger curvature now returns the *discounted* density,

$$\frac{\partial^2 C}{\partial K^2} \;=\; \mathbb{E}^{\mathbb{Q}}\big[\, D_T\, \delta(S_T - K) \,\big], \tag{20}$$

which entangles the law of S_T with the discounting along each path. The clean separation of "payoff" from "discount" that made the original derivation collapse so neatly is gone, and that is what the new term will measure.

8.2 The generalized formula, before the covariance

Repeating the forward-equation derivation with D_T kept inside the expectation (the Itô–Tanaka steps are in Appendix G) gives, before any tidying,

$$\tfrac{1}{2}\, \sigma_{\mathrm{loc}}^2(K, T)\, K^2\, \frac{\partial^2 C}{\partial K^2} \;=\; \frac{\partial C}{\partial T} + qC - qK\, \frac{\partial C}{\partial K} \;-\; K\, \mathbb{E}^{\mathbb{Q}}\big[\, D_T\, r_T\, \mathbf{1}_{S_T > K} \,\big]. \tag{21}$$

Every term but the last was present (in some guise) with deterministic rates. The last term is new, and it has a concrete meaning: it is the strike K times

$$\mathbb{E}^{\mathbb{Q}}\big[\, D_T\, r_T\, \mathbf{1}_{S_T > K} \,\big], \tag{22}$$

the discounted average of the short rate taken *only over the scenarios where the call pays*—the indicator $\mathbf{1}_{S_T > K}$ restricts the average to those paths. It appears because differentiating the discount factor (17) in maturity brings down a factor of $-r_T$: each extra instant of maturity is an extra instant of discounting at whatever the short rate then happens to be. So the option's sensitivity to maturity now carries the short rate, weighted by the payoff. The question is why this should reduce to a *covariance*.

8.3 Why a covariance appears

The new term (22) is the expectation of a *product*—the short rate r_T times the in-the-money indicator $\mathbf{1}_{S_T>K}$. Any expectation of a product of two random quantities splits into the product of their means plus their covariance,

$$\mathbb{E}[XY] \;=\; \mathbb{E}[X]\,\mathbb{E}[Y] \;+\; \mathrm{Cov}(X,Y). \tag{23}$$

This is the whole story. Apply it to $X = r_T$ and $Y = \mathbf{1}_{S_T>K}$. Moving to the T-forward measure $\mathbb{Q}_T$ to keep the discounting tidy[5] and using two facts about that measure—that the expected short rate is today's forward rate, and that the probability of finishing in the money is the (normalised) call-price slope,

$$\mathbb{E}^{\mathbb{Q}_T}[r_T] \;=\; f(0,T), \qquad \mathbb{E}^{\mathbb{Q}_T}[\mathbf{1}_{S_T>K}] \;=\; \mathbb{Q}_T(S_T>K) \;=\; -\frac{1}{P(0,T)}\frac{\partial C}{\partial K}, \tag{24}$$

the split (23) becomes

$$K\,\mathbb{E}^{\mathbb{Q}}\big[D_T\,r_T\,\mathbf{1}_{S_T>K}\big] \;=\; \underbrace{-\,K\,f(0,T)\,\partial_K C}_{\text{means}} \;+\; \underbrace{K\,P(0,T)\,\mathrm{Cov}^{\mathbb{Q}_T}\big(r_T,\mathbf{1}_{S_T>K}\big)}_{\text{covariance}}. \tag{25}$$

The first piece—the product of the means, what you would get if the rate and the option's outcome were unrelated—recombines with the existing drift term and simply upgrades the constant rate of the old formula to today's forward curve $f(0,T)$. The second piece is irreducible. Substituting (25) into (21) and solving gives the generalized Dupire formula,

$$\boxed{\sigma^2_{\text{loc}}(K,T) \;=\; \frac{\partial_T C + qC + \big(f(0,T)-q\big)\,K\,\partial_K C \;-\; K\,P(0,T)\,\mathrm{Cov}^{\mathbb{Q}_T}\big(r_T,\mathbf{1}_{S_T>K}\big)}{\tfrac{1}{2}\,K^2\,\partial_K^2 C}}. \tag{26}$$

The first three numerator terms are exactly the deterministic-rate formula (Appendix F) with the forward curve in place of a constant rate. Everything stochastic rates add is the one covariance term.

What the covariance is asking. Strip away the constants and the covariance

$$\mathrm{Cov}^{\mathbb{Q}_T}\big(r_T,\mathbf{1}_{S_T>K}\big) \tag{27}$$

poses a single question: *is the short rate systematically high or low in exactly those scenarios where the option finishes in the money?* If rates tend to rise with the stock ($\rho_{S,r}>0$), then on the paths that push the stock above K the rate is also high, the covariance is positive, and the discounting applied to the winning payoffs is heavier than the surface alone would suggest. If rates fall with the stock ($\rho_{S,r}<0$), the reverse. Figure 11 makes this visible: it plots the average short rate *conditional on the terminal spot* against the spot, and contrasts it with the unconditional average. Over the in-the-money region the two differ—upward for positive correlation, downward for negative—and that gap, integrated against the density, is the covariance.

[5] Working under $\mathbb{Q}_T$—the measure whose numeraire is the T-maturity bond—is the standard device for stochastic rates: it turns the awkward discounted expectation into an ordinary one, $\mathbb{E}^{\mathbb{Q}}[D_T(\cdot)] = P(0,T)\,\mathbb{E}^{\mathbb{Q}_T}[\cdot]$. Its defining property is that the forward rate is the expected future short rate, $f(0,T) = \mathbb{E}^{\mathbb{Q}_T}[r_T]$.

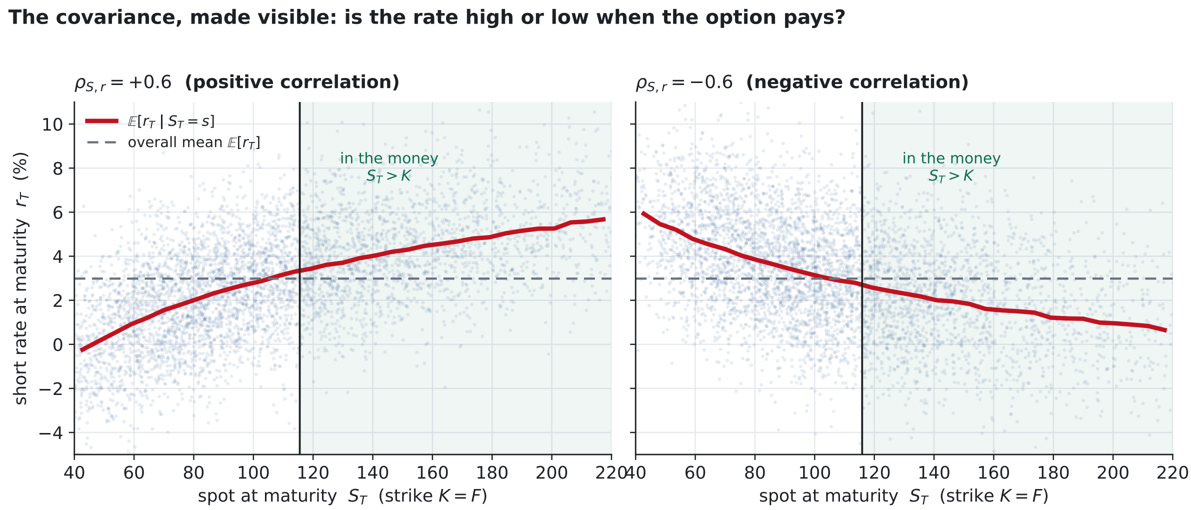

Figure 11: The covariance term (27), made visible, from the hybrid model at $T = 5$. Each blue dot is one scenario: its terminal spot S_T and short rate r_T. The red curve is the average rate *given* the spot, $\mathbb{E}[r_T \mid S_T = s]$; the dashed line is the overall average $\mathbb{E}[r_T]$; the shaded band is the in-the-money region $S_T > K$. **Left ($\rho_{S,r} = +0.6$):** where the call pays, the rate averages well above the overall mean—a positive covariance. **Right ($\rho_{S,r} = -0.6$):** it averages below—a negative covariance. With deterministic or spot-independent rates the red curve would be flat and the covariance zero.

8.4 Reading the correction term

The decomposition makes the term's behaviour easy to read off:

- it *vanishes* when rates are deterministic (r_T is then a constant, so the covariance is zero, the red curve in Figure 11 is flat) and (26) reduces to the familiar formula—a reassuring consistency check;
- it *also vanishes* when rates are stochastic but *independent* of the spot ($\rho_{S,r} = 0$ again flattens the red curve);
- otherwise it is driven by the *spot–rate correlation* $\rho_{S,r}$, and because it depends on the strike K it reshapes the local-volatility surface rather than merely shifting its level.

The decisive consequence is conceptual. With deterministic rates, the Dupire formula reads the local volatility *straight off the vanilla surface*, with no model assumptions—that was its great virtue. The covariance (27) cannot be observed from the vanilla surface alone: it depends on the dynamics of rates and on the correlation $\rho_{S,r}$. So under stochastic rates the local volatility is no longer model-free. The surface pins it down only *once a rate model and a correlation have been chosen*; change the assumed $\rho_{S,r}$ and the local volatility consistent with the same option prices changes with it.

8.5 Seeing the effect on prices

Figure 12 makes the size of the correction concrete. We hold the local volatility *flat* at 20% and simulate the hybrid model (16) for three correlations, so that every departure of the resulting implied volatility from 20% is the pure footprint of stochastic rates. Two effects appear. Even at zero correlation the implied volatility drifts above the local volatility as maturity grows—bond volatility adds variance to the forward, a convexity effect. Correlation then fans the curves out: positive $\rho_{S,r}$ lifts implied volatility further, negative $\rho_{S,r}$ pulls it down, several volatility points either way by ten years. The right panel turns the warning around: if one ignored rates and applied the ordinary Dupire formula to that implied curve, the local volatility "recovered" would be wrong by the same margin—the gap that the covariance term (27) is there to absorb.

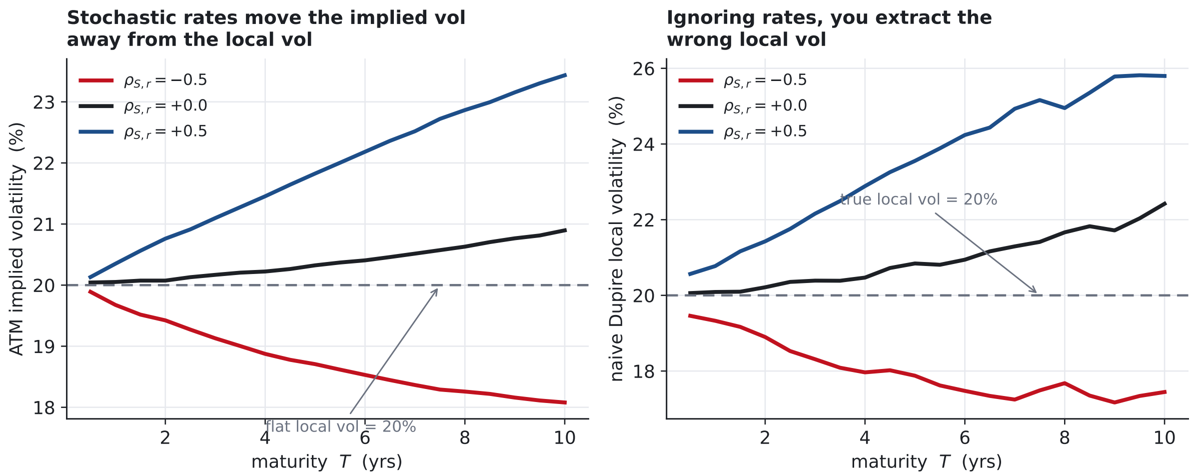

Figure 12: Stochastic rates and the Dupire formula, from the hybrid model (16) with a *flat* 20% local volatility and a Hull–White short rate ($a = 0.10$, $\sigma_r = 1.5\%$). **Left:** the at-the-money implied volatility the model actually produces, by maturity and spot–rate correlation. With the local volatility flat, every move away from the dashed 20% line is caused by rates—a convexity drift present even at $\rho_{S,r} = 0$, plus a correlation tilt. **Right:** the local volatility one would (wrongly) extract by feeding those prices into the deterministic-rate formula; it misses the true 20% by up to several volatility points. The discrepancy is what the covariance term (27) accounts for.

8.6 What this means in practice

The fix is not to abandon local volatility but to calibrate it *jointly* with the rate model. One fixes the Hull–White parameters and the correlation $\rho_{S,r}$ from the rates market and a view on the equity–rate link, then solves for the local-volatility surface that reprices the equity vanillas *under those dynamics*—in effect restoring the covariance (27). This is the *local-volatility Hull–White* hybrid, and it is calibrated numerically rather than by a closed form: either a two-factor forward (Fokker–Planck) PDE in spot and rate, or a particle Monte-Carlo method that estimates the conditional covariance (27) on the fly and feeds it back into the local volatility. The headline is simply stated: stochastic rates do not break the local-volatility *model*, but they do break the model-free Dupire *formula*—calibration becomes a computation that depends on the rate dynamics, not a closed-form read-off.

9 Risk management: the Greeks in a local-volatility model

A calibrated model earns its keep at the hedging desk, not the pricing screen. Once the local-volatility surface is fixed, the model is used to compute the *Greeks*—the sensitivities of a position to spot, to volatility, and to time—and those sensitivities tell the trader what to buy and sell to stay hedged. Local volatility changes the character of every Greek, and one of them, the delta, carries an assumption subtle enough, and consequential enough, to be the main reason the model is handled with care.

9.1 Delta: what is held fixed when spot moves?

In Black–Scholes the delta is unambiguous because there is only one volatility. In a local-volatility model there is a whole surface, so to differentiate the price in spot one must first say what happens to volatility when spot moves. The model's own answer is clean: the local-volatility *field* $\sigma_{\text{loc}}(t,S)$ is a fixed function of time and spot, so it is held fixed, and

$$\Delta_{\text{LV}} \;=\; \left.\frac{\partial V}{\partial S}\right|_{\sigma_{\text{loc}}(\cdot,\cdot)\ \text{fixed}} \tag{28}$$

is the *model delta*, or "sticky local vol" delta. But what a trader feels is not the unobservable field; it is the quoted implied surface, which also moves when spot moves. Splitting the total spot sensitivity into the move at fixed implied volatility plus the move the implied volatility itself makes,

$$\Delta \;=\; \underbrace{\frac{\partial V}{\partial S}}_{\text{Black–Scholes delta}} \;+\; \underbrace{\frac{\partial V}{\partial \sigma_{\text{BS}}}}_{\text{vega}} \cdot \frac{\mathrm{d}\sigma_{\text{BS}}}{\mathrm{d}S}, \tag{29}$$

exposes the crux: the delta depends on $\mathrm{d}\sigma_{\text{BS}}/\mathrm{d}S$, the assumed *dynamics of the smile*. Three conventions are standard—*sticky strike* (the smile is fixed in absolute strike, $\mathrm{d}\sigma_{\text{BS}}/\mathrm{d}S = 0$ at fixed K), *sticky moneyness* (the smile rides with spot), and the local-volatility model's own prediction. They give materially different hedges, so it matters which one the model implies.

9.2 The backbone

The cleanest way to compare the conventions is to ask how the at-the-money implied volatility moves as spot moves to a new level—a curve the desk calls the *backbone*. Local volatility makes a definite prediction. For a short-dated option the at-the-money implied volatility is approximately the local volatility at the current spot,

$$\sigma_{\text{BS}}(\text{ATM},T) \;\approx\; \sigma_{\text{loc}}(0,S) \qquad (\text{short } T), \tag{30}$$

so as spot moves the at-the-money volatility slides along the local-volatility curve $\sigma_{\text{loc}}(0,\cdot)$. By the rule of two of Section 5, that curve is about *twice* as steep as the implied skew. The local-volatility backbone is therefore roughly twice the skew, against the skew itself for sticky strike and zero for sticky moneyness (Figure 13).

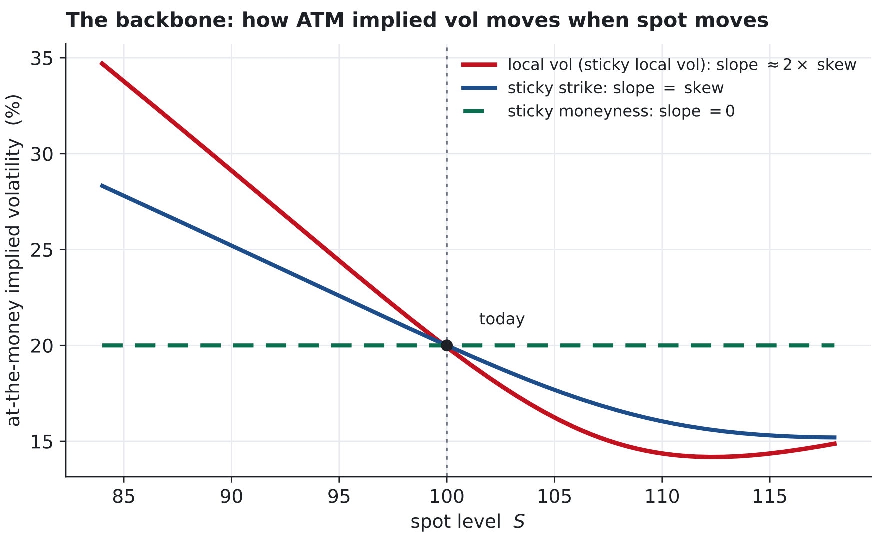

Figure 13: The backbone: how the at-the-money implied volatility moves as spot moves away from today (black dot). *Sticky moneyness* (green) keeps it constant; *sticky strike* (blue) slides it along the existing skew; the *local-volatility* model (red) makes it ride the local-vol curve, about twice as steep (the rule of two, Section 5). The three conventions disagree most exactly where hedging matters—when spot has moved.

This is where the trouble is. Empirically, the at-the-money volatility of an equity index moves with spot at roughly the skew slope, or less—broadly sticky-strike behaviour. The local-volatility backbone, at twice the skew, is too steep; in some regimes it even implies the smile sliding the *wrong way* relative to spot [9]. Because the backbone enters the delta through (29), an overstated backbone means a biased delta.

9.3 The delta in practice

Figure 14 shows the bias directly. The local-volatility model delta, computed by bumping spot and re-pricing with the field held fixed, sits well below the naïve sticky-strike Black–Scholes delta—by as much as fourteen delta points around the money for the downward equity skew used throughout. The sign follows from (29): vega is positive, the backbone $d\sigma_{BS}/dS$ is negative for a downward skew, so the correction lowers the call delta.

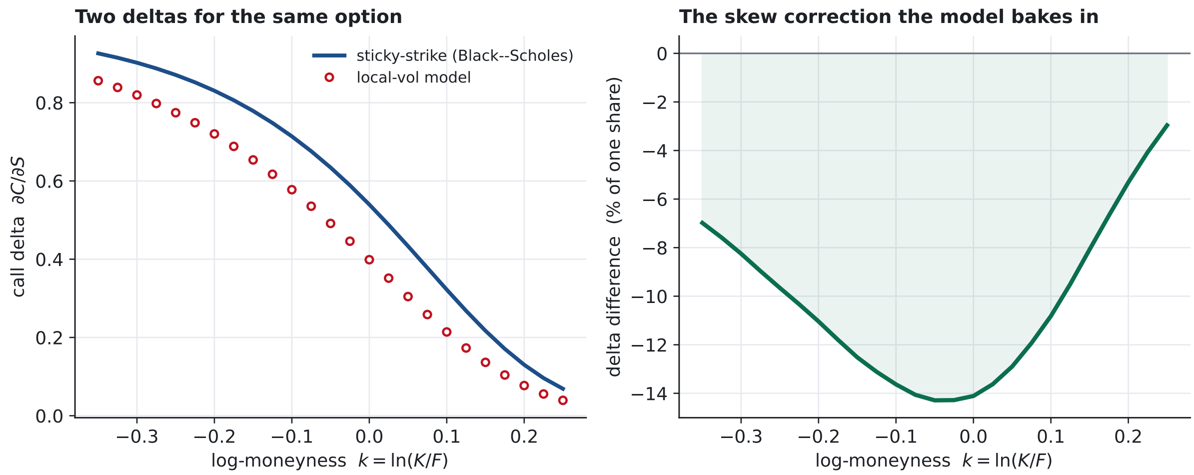

Figure 14: Left: the call delta against moneyness under two conventions—the local-volatility model (red, from a same-random-numbers spot bump with the field held fixed) and sticky-strike Black–Scholes (blue). **Right:** their difference, the skew correction the model bakes into its delta, reaching about -14% of a share near the money. Which is the better hedge is an empirical question, and for index options the answer is usually *not* the raw local-vol delta.

The practical response is to stop taking the delta the model offers at face value. Two routes are common. One is the *minimum-variance delta* [10]: keep the decomposition (29) but replace the model's backbone with the one actually observed in the market (estimated by regressing implied-volatility moves on spot moves), so the hedge matches realised smile dynamics. The other is to draw the delta from a richer model—a stochastic or local-stochastic volatility model (Section 10)—whose smile dynamics are more realistic, while still using local volatility for the exact vanilla fit and for pricing. Either way, the lesson is to trust local volatility for what it does well and to correct the delta.

9.4 Vega is a surface, not a number

In Black–Scholes there is one volatility, so there is one vega. In a local-volatility model there is no single volatility to bump—the input is the entire implied surface. The meaningful risk is therefore the sensitivity to each point of that surface: bump the implied volatility of one pillar $\sigma_{BS}(K_i, T_j)$, recalibrate the local-volatility field, reprice, and record the change. The result is not one number but a *vega matrix* across strikes and maturities,

$$\mathcal{V}_{ij} = \frac{\partial V}{\partial \sigma_{BS}(K_i, T_j)}. \tag{31}$$

This is more feature than nuisance. Because the model is calibrated *to* the traded vanillas, each entry of (31) is the sensitivity to a genuinely traded instrument, so the vol risk of an exotic comes pre-decomposed into a basket of vanilla hedges—buy or sell the indicated amount of each pillar option and the first-order volatility risk is neutralised. The shape of the matrix tells the desk *where* on the surface the exposure lives; its sum is close to the overall vega. Reporting and hedging vol risk bucket by bucket, in vanillas, is the day-to-day reality of running a local-volatility book.

9.5 Gamma, theta, and the break-even

Gamma and theta keep their usual meanings—$\Gamma = \partial^2 V/\partial S^2$ and the bleed of value with calendar time—but local volatility ties them together in a particularly useful way. For a delta-hedged position the pricing equation (34) turns the profit and loss over a short interval (taking $r = q = 0$ for clarity) into

$$\mathrm{d}\Pi \;=\; \tfrac{1}{2}\,\Gamma\, S^2 \left(\sigma_{\mathrm{R}}^2 - \sigma_{\mathrm{loc}}^2(t,S)\right) \mathrm{d}t, \tag{32}$$

where σ_{R} is the volatility actually realised over the interval (derived in Appendix H). The reading is exact and practical: a delta-hedged local-volatility book earns the gap between realised variance and the *local* variance at the spot and time it currently occupies. The break-even is not some average volatility but $\sigma_{\mathrm{loc}}(t,S)$—the model tells the trader, point by point along the path, the volatility they need to realise to come out flat. Where gamma is large the book is most exposed to that gap.

9.6 The workflow

Putting it together, running a local-volatility book settles into a routine. Calibrate the field to the vanilla surface (an exact, direct fit). Price the exotic and read its risk: a delta—smile-corrected or taken from a richer model—a vega *matrix* hedged with a basket of vanillas, and a gamma/theta position whose break-even is the local volatility underfoot. As the market moves, recalibrate the field and attribute the profit and loss across delta, the vega buckets, gamma/theta, and a recalibration residual. The recurring discipline is the one the Greeks make precise: lean on local volatility for the static fit and for expressing vol risk as vanillas, and be sceptical of the delta it hands you, because its backbone is too steep.

10 What local volatility gets right, and the one thing it gets wrong

What it gets right. Within the world of European options, local volatility is close to perfect. It fits the entire vanilla surface *exactly*, not approximately; the fit is *unique*; the model stays a complete, one-factor, arbitrage-free market in which Greeks and hedges are well defined; and the calibration is a direct formula, not a search. For any product whose value is fixed by the distribution of the spot at each single date S_T (as a European option's is)—and for fast, consistent repricing and hedging of the vanilla book—it is hard to beat and remains a desk staple three decades on.

The one thing it gets wrong. The weakness is dynamic, and Section 6 already named its cause: local volatility keeps the average volatility at each spot and throws away the spread. The cleanest symptom is the *forward smile*—the smile the model predicts will prevail not today but starting on some future date. Empirically, equity smiles are persistent: a year from now the market will still show a pronounced skew. Local volatility predicts the opposite. Because its only mechanism for skew is the slope of the $\sigma_{\mathrm{loc}}(t,S)$ field, and because diffusion spreads paths across a *range* of spot levels before the forward-start date arrives, the level-dependence gets averaged out and the forward skew decays. Figure 15 shows this flattening, simulated from the very model we calibrated: the one-year smile starting one year from now is markedly flatter than today's one-year smile.

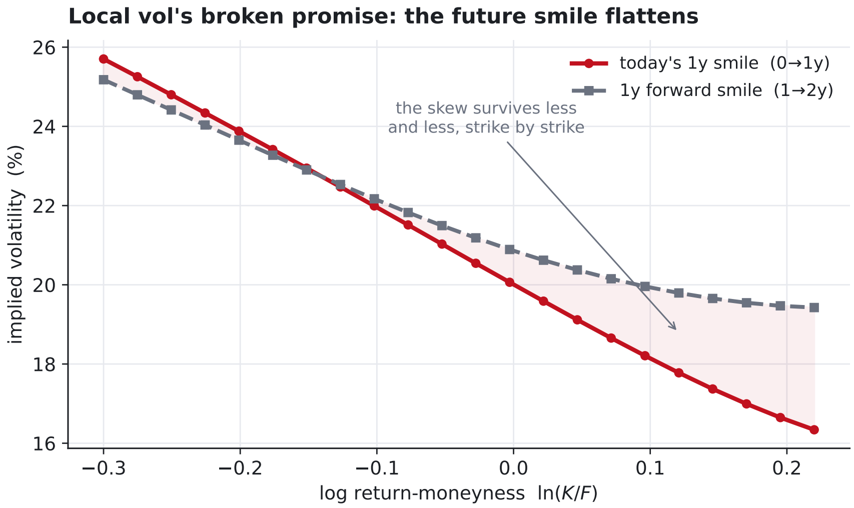

Figure 15: Local volatility's broken promise. The red curve is today's one-year smile; the grey curve is the one-year smile the *same* calibrated model predicts for a window starting one year out. The forward smile is conspicuously flatter—diffusion has averaged the spot-dependence away. Real markets keep their skew; local volatility does not. This is the defect that motivates stochastic and local-stochastic volatility models.

Table 1 puts numbers on the picture. Measured as the 90–110 skew—the gap in implied volatility between the 90% and 110% strikes[6]—the skew falls by roughly 40% over a one-year forward start, from 3.8 to 2.3 volatility points, even though the model was calibrated to reproduce today's skew exactly.

	ATM vol (%)	90–110 skew (vol pts)
today's 1y smile $(0 \to 1\text{y})$	20.0	3.8
1y forward smile $(1 \to 2\text{y})$	20.9	2.3

Table 1: Monte-Carlo skew term structure of the calibrated local-volatility model (the smiles of Figure 15). The model fits today's skew by construction, yet its forward skew has decayed by about 40%—the symptom of the averaged-away spot-dependence. A real equity market would keep the skew roughly intact.

Where this leads. The cure is to give volatility back the randomness local vol averaged out. A *stochastic-volatility* model (Heston, SABR) lets variance be its own mean-reverting random process correlated with spot; the spread that Figure 9 discarded is restored, the forward smile persists, but a handful of parameters can no longer bend themselves into the shape of an entire surface, so the exact fit is lost. *Local-stochastic volatility* (LSV) marries the two: a stochastic variance for realistic dynamics, multiplied by a deterministic *leverage function* chosen—via exactly the conditional-expectation identity (15)—to restore the perfect vanilla fit. Local volatility is thus not a dead end but the foundation: the object every richer equity model is built to contain as a special case. Whoever understands the Dupire formula and its conditional-expectation reading has the key to all of them.

[6] A standard desk shorthand for the slope of the smile: a single number, in volatility points, that is larger for a steeper down-sloping smile and zero for a flat one. It reads off the same information as the at-the-money slope $\mathcal{S}(T)$ of (3), just discretely.

A The Black–Scholes formula and the backward PDE

The closed-form price. For constant volatility σ and a call of strike K and maturity T (rate r, dividend yield q), the Black–Scholes price referenced in Section 1 is

$$\begin{aligned} C_{\mathrm{BS}}(S_0, K, T; \sigma) &= S_0\, e^{-qT}\, N(d_1) \;-\; K\, e^{-rT}\, N(d_2), \\ d_{1,2} &= \frac{\ln(S_0/K) + (r - q \pm \frac{1}{2}\sigma^2)\, T}{\sigma\sqrt{T}}, \end{aligned} \tag{33}$$

with N the standard normal distribution function. The implied volatility $\sigma_{\mathrm{BS}}(K, T)$ of Section 1 is the value of σ that makes (33) equal the market price; because C_{BS} is strictly increasing in σ, that value is unique and well defined.

The backward PDE. For a local-vol diffusion (4), a self-financing replicating portfolio plus Itô's lemma gives the value $V(t, S)$ of any European claim as the solution of the backward equation

$$\frac{\partial V}{\partial t} + (r - q)\, S\, \frac{\partial V}{\partial S} + \frac{1}{2}\, \sigma^2_{\mathrm{loc}}(t, S)\, S^2\, \frac{\partial^2 V}{\partial S^2} - rV \;=\; 0, \tag{34}$$

solved backward from the payoff $V(T, S) = (S - K)^+$ with K, T fixed. Contrast the unknowns: (34) varies over (t, S) for a fixed contract; Dupire's (9) varies over (K, T) for a fixed start. The constant coefficient σ recovers the textbook Black–Scholes PDE.

B The Fokker–Planck equation in one paragraph

For a diffusion $\mathrm{d}S_t = \mu(t, S_t)\, \mathrm{d}t + \Sigma(t, S_t)\, \mathrm{d}W_t$, the density $p(t, S)$ of S_t evolves *forward* in time by the Fokker–Planck (forward Kolmogorov) equation

$$\frac{\partial p}{\partial t} \;=\; -\frac{\partial}{\partial S}\big[\mu(t, S)\, p\big] + \frac{1}{2}\, \frac{\partial^2}{\partial S^2}\big[\Sigma^2(t, S)\, p\big]. \tag{35}$$

It is a conservation law: the first term on the right transports probability mass along the drift, the second diffuses it. For the local-vol SDE with $r = q = 0$ we have $\mu = 0$ and $\Sigma = \sigma_{\mathrm{loc}}(t, S)\, S$, which is exactly (8). Where the backward equation (34) propagates a payoff back to today, the forward equation (35) propagates today's certainty (the stock is at S_0) forward into a spreading cloud of possibilities—the densities of Figure 4.

C Derivation of the Dupire formula

Start from the undiscounted call (5) with $r = q = 0$ and differentiate in maturity:

$$\frac{\partial C}{\partial T} \;=\; \int_K^\infty (S - K)\, \frac{\partial p}{\partial T}(T, S)\, \mathrm{d}S. \tag{36}$$

Insert Fokker–Planck (8) and write $g(S) = \frac{1}{2}\, \sigma^2_{\mathrm{loc}}(T, S)\, S^2\, p(T, S)$, so the integrand becomes $(S - K)\, g''(S)$. Integrate by parts twice. The boundary terms vanish—at $S = K$ because the factor $(S - K)$ is zero, and at $S \to \infty$ because a well-behaved density and its derivatives decay faster than $(S - K)$ grows:

$$\begin{aligned} \int_K^\infty (S - K)\, g''(S)\, \mathrm{d}S &= \Big[(S - K)\, g'(S)\Big]_K^\infty - \int_K^\infty g'(S)\, \mathrm{d}S \\ &= -\big[g(S)\big]_K^\infty \\ &= g(K). \end{aligned} \tag{37}$$

Therefore
$$\frac{\partial C}{\partial T} \;=\; g(K) \;=\; \frac{1}{2}\,\sigma^2_{\rm loc}(T,K)\,K^2\,p(T,K). \tag{38}$$

Finally substitute Breeden–Litzenberger (7), $p(T,K) = \partial^2 C/\partial K^2$, to reach Dupire's forward equation (9), and solve for $\sigma^2_{\rm loc}$ to obtain the formula (10).

D The implied-variance form and the rule of two

This appendix supports Section 5: how the price-form Dupire formula (10) becomes Gatheral's implied-variance form (13), and where the factor of two in the skew comes from.

The change of variables. The Black–Scholes price of a call depends on its strike and maturity only through the total implied variance $w = w(k,T)$ and the log-moneyness k: writing $C = C_{\rm BS}\big(k, w(k,T)\big)$, the $d_{1,2}$ of (33) become $d_{1,2} = -k/\sqrt{w} \pm \frac{1}{2}\sqrt{w}$. The three derivatives that appear in Dupire's formula (10) can therefore be expanded by the chain rule in terms of $\partial_T w$, $\partial_k w$ and $\partial_k^2 w$. Substituting $\partial_T C = \partial_w C\,\partial_T w$ for the numerator, and the analogous two-derivative expansion for the denominator $\partial_K^2 C$, and cancelling the common Black–Scholes vega factors, collapses (10) to (13). The manipulation is purely mechanical—no new finance enters—which is why the body states the result and defers the bookkeeping here.

The short-maturity limit (rule of two). In the limit $T \to 0$ the relationship between the two smiles becomes especially clean. A classical result[7] states that the zero-maturity implied volatility $\sigma_{\rm BS}(k,0)$ is the harmonic mean of local volatility between 0 and k,

$$\frac{k}{\sigma_{\rm BS}(k,0)} \;=\; \int_0^k \frac{\mathrm{d}y}{\sigma_{\rm loc}(y,0)}. \tag{39}$$

Differentiating (39) in k and solving for $\sigma_{\rm loc}$ gives the exact short-maturity inversion, and its first-order (small-skew) expansion,

$$\sigma_{\rm loc}(k,0) \;=\; \frac{\sigma_{\rm BS}(k,0)^2}{\sigma_{\rm BS}(k,0) - k\,\partial_k\sigma_{\rm BS}(k,0)} \;\approx\; \sigma_{\rm BS}(k,0) + k\,\partial_k\sigma_{\rm BS}(k,0). \tag{40}$$

At the money ($k=0$) the two volatilities coincide. Differentiating the right-hand expansion once more and evaluating at $k=0$,

$$\partial_k\sigma_{\rm loc}(0,0) \;=\; 2\,\partial_k\sigma_{\rm BS}(0,0), \tag{41}$$

so the local skew is twice the implied skew at the money—the rule of two. It is exact only in the short-maturity limit and near the money; at longer maturities the ratio drifts below two (in the $T=1$ figure of Section 5 it is about 1.7), but the qualitative message—local volatility is the more extreme object—always holds.

E The SVI and SSVI parametrisations

This appendix supports Section 7 and records the surface used for every figure. SVI fits one maturity slice; SSVI ties the slices into an arbitrage-free sheet.

[7] Berestycki, Busca and Florent [5]; see also [6]. The statement is that as $T \to 0$ the implied volatility is the *harmonic mean* of the local volatility along the path in log-moneyness from the money out to k.

Raw SVI (one maturity). The total variance is modelled as

$$w(k) \;=\; a + b\Big[\rho\,(k-m) + \sqrt{(k-m)^2 + s^2}\Big]. \tag{42}$$

The five parameters carry distinct geometric jobs: a shifts the whole curve *up or down* (the overall variance level); b sets how sharply the two wings rise (their asymptotic slopes are $b(1-\rho)$ on the left and $b(1+\rho)$ on the right); $\rho \in (-1,1)$ tilts the curve, making the left wing steeper than the right for $\rho < 0$ (the equity skew); m slides the minimum left or right; and s controls how *rounded* the bottom is (large s a gentle valley, small s a sharp kink). The far wings are linear in k and the bottom is smooth and convex, which is exactly the observed smile shape and which keeps each slice free of butterfly arbitrage for sensible parameter ranges.

SSVI (the whole surface). To couple maturities, SSVI writes the slice at maturity T using a single at-the-money total-variance term structure $\theta_T = \sigma^2_{\mathrm{BS}}(0,T)\,T$ and a shape function $\varphi(\theta)$,

$$w(k,T) \;=\; \frac{\theta_T}{2}\Big[\,1 + \rho\,\varphi(\theta_T)\,k + \sqrt{\big(\varphi(\theta_T)\,k + \rho\big)^2 + 1 - \rho^2}\,\Big], \qquad \varphi(\theta) = \frac{\eta}{\theta^{\gamma}\,(1+\theta)^{1-\gamma}}. \tag{43}$$

A single ρ now sets the skew across all maturities, while $\varphi(\theta)$ makes the skew decay with maturity as observed. The surface is free of butterfly arbitrage when $\theta_T\,\varphi(\theta_T)\,(1+|\rho|) < 4$ and free of calendar arbitrage when θ_T is non-decreasing in T—both easy to enforce, which is the whole point of the parametrisation [7].

Values used in this note. Every figure uses $S_0 = 100$, $r = q = 0$ (so $F = S_0$), a flat at-the-money volatility of 20% (hence $\theta_T = 0.04\,T$, increasing in T, so calendar-arbitrage-free), and shape parameters $\eta = 0.85$, $\gamma = 0.40$, $\rho = -0.65$. These give a realistic downward equity skew that steepens at short maturities, and they satisfy the no-arbitrage bounds above throughout the plotted region.

F The general formula with rates and dividends

Restoring a constant rate r and dividend yield q, the same argument (now with the drift term of (35) present) yields the full Dupire formula,

$$\sigma^2_{\mathrm{loc}}(K,T) \;=\; \frac{\dfrac{\partial C}{\partial T} + (r-q)\,K\,\dfrac{\partial C}{\partial K} + q\,C}{\dfrac{1}{2}\,K^2\,\dfrac{\partial^2 C}{\partial K^2}}. \tag{44}$$

The extra numerator terms account for the deterministic drift of the forward; when $r = q = 0$ they vanish and (44) reduces to (10). In practice it is cleaner to quote everything in forward moneyness and use the total-variance form (13), which absorbs rates and dividends into the definition of the forward F.

G Derivation with stochastic rates

This appendix derives the generalized formula (26) of Section 8. The method is the same forward-equation argument as Appendix C, but with the random discount factor $D_T = \exp(-\int_0^T r_s\,\mathrm{d}s)$ carried through; the tool that handles the non-smooth payoff is the Itô–Tanaka formula.

Differentiate the discounted call. Write $C(K,T) = \mathbb{E}^{\mathbb{Q}}[D_T(S_T - K)^+]$. By the Itô–Tanaka formula, the kink in the payoff contributes a local-time term,

$$\mathrm{d}(S_t - K)^+ = \mathbf{1}_{S_t>K}\,\mathrm{d}S_t + \tfrac{1}{2}\,\delta(S_t - K)\,\sigma_{\mathrm{loc}}^2(t,S_t)\,S_t^2\,\mathrm{d}t, \tag{45}$$

while the discount factor evolves as $\mathrm{d}D_t = -r_t\,D_t\,\mathrm{d}t$. Applying the product rule to $D_t(S_t - K)^+$, inserting the spot dynamics $\mathrm{d}S_t = (r_t - q)S_t\,\mathrm{d}t + \sigma_{\mathrm{loc}}S_t\,\mathrm{d}W_t^S$, taking $\mathbb{E}^{\mathbb{Q}}$ (the $\mathrm{d}W^S$ term is a martingale and drops), and reading off the maturity derivative,

$$\frac{\partial C}{\partial T} = \mathbb{E}^{\mathbb{Q}}\Big[D_T\,\mathbf{1}_{S_T>K}\big(r_TK - q\,S_T\big)\Big] + \tfrac{1}{2}\,\sigma_{\mathrm{loc}}^2(K,T)\,K^2\,\mathbb{E}^{\mathbb{Q}}\big[D_T\,\delta(S_T - K)\big], \tag{46}$$

where the first bracket used $-r_T(S_T - K)^+ + (r_T - q)S_T\mathbf{1}_{S_T>K} = \mathbf{1}_{S_T>K}(r_TK - qS_T)$.

Identify the strike derivatives. Differentiating $C = \mathbb{E}^{\mathbb{Q}}[D_T(S_T - K)^+]$ in K gives the discounted analogues of the Breeden–Litzenberger relations,

$$\frac{\partial C}{\partial K} = -\,\mathbb{E}^{\mathbb{Q}}\big[D_T\,\mathbf{1}_{S_T>K}\big], \qquad \frac{\partial^2 C}{\partial K^2} = \mathbb{E}^{\mathbb{Q}}\big[D_T\,\delta(S_T - K)\big], \tag{47}$$

and hence $\mathbb{E}^{\mathbb{Q}}[D_TS_T\mathbf{1}_{S_T>K}] = C - K\,\partial_K C$. Substituting these into (46) and rearranging,

$$\tfrac{1}{2}\,\sigma_{\mathrm{loc}}^2(K,T)\,K^2\,\frac{\partial^2 C}{\partial K^2} = \frac{\partial C}{\partial T} + qC - qK\,\frac{\partial C}{\partial K} - K\,\mathbb{E}^{\mathbb{Q}}\big[D_T\,r_T\,\mathbf{1}_{S_T>K}\big]. \tag{48}$$

Move to the forward measure. Finally, $\mathbb{E}^{\mathbb{Q}}[D_T\,X] = P(0,T)\,\mathbb{E}^{\mathbb{Q}_T}[X]$ for any X, and under $\mathbb{Q}_T$ the forward rate is the expected short rate, $f(0,T) = \mathbb{E}^{\mathbb{Q}_T}[r_T]$, while (47) gives $\mathbb{E}^{\mathbb{Q}_T}[\mathbf{1}_{S_T>K}] = -\partial_K C/P(0,T)$. Splitting the expectation of a product into mean-times-mean plus covariance,

$$K\,\mathbb{E}^{\mathbb{Q}}\big[D_Tr_T\mathbf{1}_{S_T>K}\big] = -K\,f(0,T)\,\frac{\partial C}{\partial K} + K\,P(0,T)\,\mathrm{Cov}^{\mathbb{Q}_T}\big(r_T, \mathbf{1}_{S_T>K}\big). \tag{49}$$

Inserting (49) into (48) and solving for σ_{loc}^2 yields (26). Setting r deterministic kills the covariance and recovers Appendix F; setting $r \equiv 0$, $q = 0$ recovers the basic formula (10).

H Risk decompositions for the Greeks

This appendix supports Section 9: the delta decomposition (29) and the delta-hedged profit-and-loss identity (32).

The delta decomposition. The market value of a position is a function of spot both directly and through the implied volatility the desk reads off the surface, $V = V\big(S, \sigma_{\mathrm{BS}}(S)\big)$. The total derivative in spot is the chain rule,

$$\frac{\mathrm{d}V}{\mathrm{d}S} = \frac{\partial V}{\partial S} + \frac{\partial V}{\partial \sigma_{\mathrm{BS}}}\,\frac{\mathrm{d}\sigma_{\mathrm{BS}}}{\mathrm{d}S}, \tag{50}$$

which is (29): the Black–Scholes delta at frozen volatility, plus vega times the rate at which the implied volatility moves with spot. Every hedging convention is a choice of that last factor.

The delta-hedged P&L. Hold the claim and short $\Delta = \partial_S V$ units of spot, $\Pi = V - \Delta\, S$, and take $r = q = 0$. Over a short interval Itô's lemma gives

$$\mathrm{d}\Pi \;=\; \mathrm{d}V - \Delta\,\mathrm{d}S \;=\; \frac{\partial V}{\partial t}\,\mathrm{d}t + \tfrac{1}{2}\,\frac{\partial^2 V}{\partial S^2}\,(\mathrm{d}S)^2, \tag{51}$$

the first-order spot terms cancelling by the hedge. The realised move supplies $(\mathrm{d}S)^2 = \sigma_{\mathrm{R}}^2\, S^2\,\mathrm{d}t$, while the backward pricing equation (34) with $r = q = 0$ supplies $\partial_t V = -\tfrac{1}{2}\,\sigma_{\mathrm{loc}}^2(t,S)\,S^2\,\partial_S^2 V$. Substituting both into (51),

$$\mathrm{d}\Pi \;=\; \tfrac{1}{2}\,S^2\,\frac{\partial^2 V}{\partial S^2}\big(\sigma_{\mathrm{R}}^2 - \sigma_{\mathrm{loc}}^2(t,S)\big)\,\mathrm{d}t, \tag{52}$$

which is (32) with $\Gamma = \partial^2 V/\partial S^2$. The delta-hedged book earns the gap between realised variance and the local variance at its current spot and time, scaled by the dollar gamma.

References

[1] B. Dupire, *Pricing with a Smile*, Risk 7(1), 18–20, 1994.

[2] E. Derman and I. Kani, *Riding on a Smile*, Risk 7(2), 32–39, 1994.

[3] D. Breeden and R. Litzenberger, *Prices of State-Contingent Claims Implicit in Option Prices*, Journal of Business 51(4), 621–651, 1978.

[4] I. Gyöngy, *Mimicking the One-Dimensional Marginal Distributions of Processes Having an Itô Differential*, Probability Theory and Related Fields 71, 501–516, 1986.

[5] H. Berestycki, J. Busca and I. Florent, *Asymptotics and Calibration of Local Volatility Models*, Quantitative Finance 2(1), 61–69, 2002.

[6] J. Gatheral, *The Volatility Surface: A Practitioner's Guide*, Wiley, 2006.

[7] J. Gatheral and A. Jacquier, *Arbitrage-Free SVI Volatility Surfaces*, Quantitative Finance 14(1), 59–71, 2014.

[8] L. B. G. Andersen and V. V. Piterbarg, *Interest Rate Modeling*, Atlantic Financial Press, 2010 (hybrid local-volatility/short-rate calibration).

[9] P. S. Hagan, D. Kumar, A. S. Lesniewski and D. E. Woodward, *Managing Smile Risk*, Wilmott Magazine, 84–108, 2002.

[10] L. Bergomi, *Stochastic Volatility Modeling*, Chapman & Hall/CRC, 2016.