Chapter 5

Black Scholes

In these notes we price a European option, and pay attention to which parts of the argument are doing work. The asset's expected return disappears, for a reason the lemma below makes precise. The formula's two terms turn out to be the same probability measured under two different numeraires, which is visible in the derivation as a completed square before chapter 6 gives it a name. We then read the equation for what it says about a hedged position — that theta and gamma are one quantity seen twice — and collect the approximations a trader carries in place of the formula.

5.1 Black Scholes Option Pricing

Definition 5.1 (Black-Scholes market)**.** The Black-Scholes market has a constant risk-free rate r, so that the discount factor is $D_t = e^{-rt}$, and one traded asset following a geometric Brownian motion,

$$\frac{dS_t}{S_t} = \mu\,dt + \sigma\,dW_t, \tag{5.1}$$

with σ constant.

Lemma 5.2 (The drift is not a parameter)**.** *Under the risk-neutral measure of chapter 4, $\mu = r$.*

Proof. That measure is characterised by D_tS_t being a martingale, so its drift vanishes. By Itô's product rule, using $dD_t = -rD_t\,dt$ and that D has no diffusion term,

$$d(D_tS_t) = -rD_tS_t\,dt + D_t\big(\mu S_t\,dt + \sigma S_t\,dW_t\big) = (\mu - r)D_tS_t\,dt + \sigma D_tS_t\,dW_t,$$

and the drift is zero for all t only if $\mu = r$. □

So the asset's expected return, the one quantity a reader might expect to matter most, does not appear in any price. Chapter 6 explains what has happened: changing measure moves drifts and cannot touch volatilities, so μ was never an observable of the option in the first place. Under the risk-neutral measure (5.1) integrates to

$$S_T = S_t \exp\left(\left(r - \tfrac{1}{2}\sigma^2\right)\tau + \sigma\sqrt{\tau}\,Z\right), \qquad \tau = T - t, \quad Z \sim N(0,1). \tag{5.2}$$

Theorem 5.3 (Black and Scholes)**.** *In the market of definition 5.1, a European call struck at K and expiring at T is worth*

$$V_t = S_t\,N(d_1) \;-\; Ke^{-r\tau}\,N(d_2), \tag{5.3}$$

where N is the standard normal distribution function and

$$d_{1,2} = \frac{1}{\sigma\sqrt{\tau}}\left[\ln\frac{S_t}{K} + \Big(r \pm \tfrac{1}{2}\sigma^2\Big)\tau\right], \qquad \textit{so that} \qquad d_1 = d_2 + \sigma\sqrt{\tau}. \tag{5.4}$$

Proof. Since D_tV_t is a martingale, $V_t = e^{-r\tau}\,\mathbb{E}\big[(S_T - K)^+ \,\big|\, \mathcal{F}_t\big]$. Substituting (5.2), the option finishes in the money exactly when

$$Z > \frac{1}{\sigma\sqrt{\tau}}\left[\ln\frac{K}{S_t} - \Big(r - \tfrac{1}{2}\sigma^2\Big)\tau\right] = -d_2,$$

so with φ the standard normal density the expectation splits into two integrals over the same region:

$$V_t = \underbrace{S_t\int_{-d_2}^{\infty} e^{-\frac{1}{2}\sigma^2\tau+\sigma\sqrt{\tau}\,z}\,\varphi(z)\,dz}_{\text{the asset leg}} \;-\; \underbrace{Ke^{-r\tau}\int_{-d_2}^{\infty}\varphi(z)\,dz}_{\text{the cash leg}}. \tag{5.5}$$

The cash leg is immediate: $\int_{-d_2}^{\infty}\varphi = N(d_2)$ by the symmetry of φ.

The asset leg needs one manipulation. Collect the exponentials and complete the square:

$$e^{-\frac{1}{2}\sigma^2\tau+\sigma\sqrt{\tau}z}\,\varphi(z) = \frac{1}{\sqrt{2\pi}}\exp\big(-\tfrac{1}{2}\sigma^2\tau + \sigma\sqrt{\tau}z - \tfrac{1}{2}z^2\big) = \frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac{1}{2}\big(z-\sigma\sqrt{\tau}\big)^2\Big) = \varphi\big(z-\sigma\sqrt{\tau}\big).$$

The integrand is a normal density again, *shifted* by $\sigma\sqrt{\tau}$. Substituting $y = z - \sigma\sqrt{\tau}$ moves the limit with it,

$$\int_{-d_2}^{\infty}\varphi\big(z-\sigma\sqrt{\tau}\big)\,dz = \int_{-d_2-\sigma\sqrt{\tau}}^{\infty}\varphi(y)\,dy = N\big(d_2+\sigma\sqrt{\tau}\big) = N(d_1)$$

by symmetry and (5.4), which is (5.3). □

Structure (Completing the square is a change of measure)**.** What happened is that a density multiplied by an exponential in its own variable came back as the *same* density about a different centre. That is exactly a Girsanov shift: multiplying by $e^{\sigma\sqrt{\tau}z-\sigma^2\tau/2}$ — which is a mean-one positive random variable, so a legitimate Radon-Nikodym derivative — is changing measure, and under the new measure Z has mean $\sigma\sqrt{\tau}$ instead of zero.

So (5.5) is not one expectation but two, taken under two *different* measures. The cash leg is priced in units of the money market account and $N(d_2)$ is the probability of exercise there. The asset leg is priced in units of the asset itself, and $N(d_1)$ is the probability of exercise *under that measure*. The two normal distribution functions are the same event — the option finishing in the money — measured by two different numeraires, which is why they differ by precisely the volatility term $\sigma\sqrt{\tau}$ that separates d_1 from d_2. Chapter 6 derives the formula that way from the start and draws the two probabilities against each other; here the shift arrives unbidden, out of completing a square.

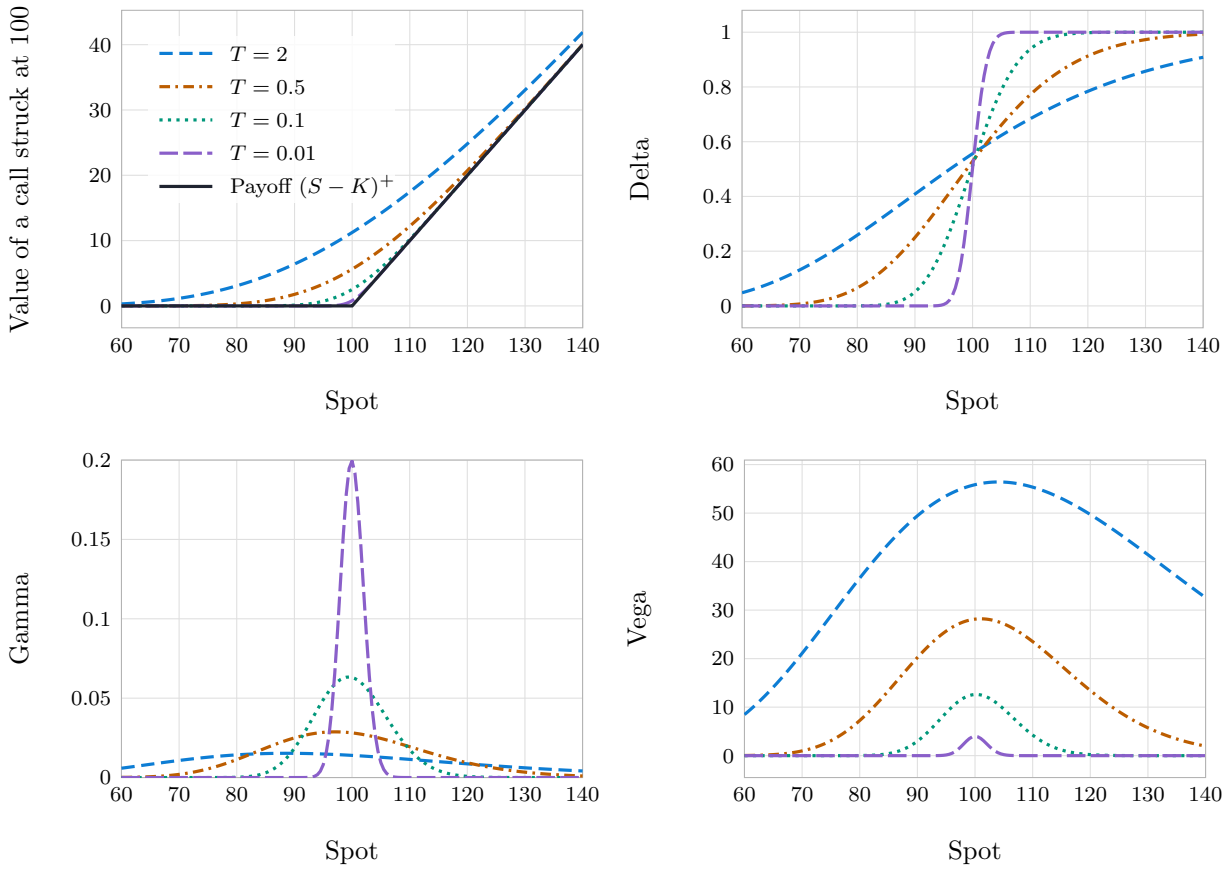

Figure 5.1: A call struck at 100 at twenty percent volatility and zero rates, with its first three sensitivities, all drawn at one maturity at a time. As expiry approaches the value falls onto $(S-K)^+$ and the curve above the payoff — the whole of the option's time value — is squeezed out of it. The derivatives say the same thing more sharply. Delta stiffens from a slope into the step that a forward would have, so a long dated option hedges like a fraction of the stock and a short dated one like all or none of it. Gamma concentrates over the strike and grows without bound, which is why the axis has to be pinned to see anything at all. And vega goes to zero everywhere at once: with no time left there is nothing for volatility to act on, so nothing to be uncertain about and nothing to pay for.

Drawn by `greeks`.

Formula (5.3) is a smooth function of the spot at every positive maturity, and the payoff it is a price for is not smooth at all. The two have to meet, since an option a moment before expiry is worth what it is about to pay, and watching them meet is the most direct account of what the extra value is.

Three things in that picture stand out, each a fact about options rather than about this formula. The kink in the payoff is the same non-smooth point that made the generator of chapter 3 awkward, and the diffusion is what smooths it — time to expiry and volatility enter only through $\sigma\sqrt{\tau}$, so shortening the maturity and lowering the volatility do the identical thing to the picture. Gamma growing without bound as vega vanishes is not a paradox but the same statement twice, and *Theta Is Gamma* below makes the relation exact. And a delta that approaches a step is a warning: it is the derivative of a function converging to one that has no derivative at the strike, which is where hedging a short dated option near its strike becomes genuinely hard rather than merely expensive.

5.2 What the Model Was Not Needed For

Theorem 5.4 (Put-call parity)**.** *Let C and P be the prices of a European call and put on the same underlying, with the same strike K and the same expiry T. Then in any market without arbitrage,*

$$C - P = P(0,T)\,\big(F - K\big), \tag{5.6}$$

with F the forward price to T and $P(0,T)$ the discount factor. Under the assumptions of this chapter, where the underlying pays nothing and rates are constant, $F = S_0 e^{rT}$ and (5.6) reads $C - P = S_0 - Ke^{-rT}$.

Proof. At expiry, for every value of S_T,

$$(S_T - K)^+ - (K - S_T)^+ = S_T - K,$$

since exactly one of the two options finishes in the money and the other is worth nothing. So a long call and a short put have, in every state, the payoff of a forward struck at K. Two portfolios with identical payoffs in every state have identical prices, by the no-arbitrage argument of chapter 4, and the forward struck at K is worth $P(0,T)(F-K)$. □

Remark (Read what is absent)**.** The proof used no model. It did not ask how S gets to S_T, whether the path is continuous, what the volatility is, or whether one exists. It is an identity between payoffs, and it therefore survives every criticism made of the Black-Scholes model in the chapters that follow — jumps, stochastic volatility, a smile, all of it. (5.6) holds in a market with any of those, because it never had an opinion about them.

This is the same distinction chapter 15 will build a whole chapter on. Some prices are fixed by static replication, and for those a model can only add an assumption to information that was already complete. Others are not, and those are what a model is for. Knowing which is which is most of knowing when to trust one.

Remark (Three things it is used for)**.** - *A call and a put at the same strike have the same implied volatility.* Since (5.6) holds for market prices and, being a model, also for Black-Scholes prices at any volatility, subtracting the two shows the volatility that reprices the call reprices the put. This is why chapter 9 can speak of *the* smile as a function of strike alone,

rather than of strike and option type, and why a desk may quote whichever of the pair is out of the money and let parity supply the other.

- *The forward can be read out of option prices.* Rearranging, $F = K + (C - P)/P(0,T)$. The option market therefore carries its own opinion about the forward, and it need not agree with the one computed from a curve. Where they differ the difference is a borrow cost, a dividend assumption, or a stale quote.

- *It is the first thing to check in an implementation.* A pricer that violates parity is wrong in a way that needs no market data to detect, and the check costs one line. Both `black76` and `Model::price` are tested against it.

Remark (When it appears to fail)**.** Quoted prices routinely violate (5.6) on a screen, and the violation is almost never an arbitrage. It is the forward: an equity with an uncertain dividend, a stock that is expensive to borrow, or a pair of quotes that were not observed at the same instant.

The right reading is that (5.6) has three market inputs and the least reliable of them is F. So rather than concluding the market is mispriced, one solves for the F that makes parity hold and asks whether that number is plausible. Used this way the relation is not a test the market can fail; it is a measuring instrument.

Exercise. Verify that the formula derived above satisfies (5.6), using $N(-x) = 1 - N(x)$ on both terms. Then deduce, without any further calculation, that the put's delta is $N(d_1) - 1$ and that a call and a put at the same strike have the same gamma and the same vega. (Hint: differentiate (5.6); the right hand side has no volatility in it, and its derivative in S_0 is one.)

5.3 The Option's Own Diffusion

The derivation above computed an expectation. There is a second route to the same answer that says more about what the option *is*, and it starts by asking what process the option price itself follows.

The option is worth $V(t, S_t)$, a function of time and of a diffusion, and chapter 3 already says what the drift of such a thing is. Writing $\mathcal{L}$ for the generator of the stock,

$$dV = \left(\frac{\partial V}{\partial t} + \mathcal{L}V \right) dt + \sigma S \frac{\partial V}{\partial S}\, dW_t,$$

the extra $\partial V/\partial t$ appearing because V depends on time directly as well as through S. So the option is itself an Itô diffusion, driven by the same Brownian motion as the stock — which is chapter 4's replication argument appearing as a statement about coefficients rather than about portfolios.

Now impose no arbitrage. The option is a tradable, so by chapter 4 its discounted price is a martingale in the risk neutral measure, which is to say its drift is rV:

$$\frac{\partial V}{\partial t} + \mathcal{L}V = rV. \tag{5.7}$$

That is the Black-Scholes equation, and how little was needed to get here deserves a pause: the drift of a function of a diffusion, and the statement that a traded thing drifts at the riskless rate. Written out, with $\mathcal{L}$ the generator under the risk neutral dynamics,

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV. \tag{5.8}$$

Structure (The same equation for every model in these notes)**.** Equation (5.7) contains nothing specific to Black-Scholes. Substituting a different generator gives the pricing equation of a different model, and that is the only step: local volatility in chapter 9, the short rate models of chapter 8, the quasi-Gaussian models of chapter 12 all satisfy (5.7) with their own $\mathcal{L}$.

So none of those chapters derives a pricing equation again. They specify a generator, which is what specifying a model means.

This is the Black-Scholes partial differential equation, and the correspondence just used — that the expectation of a payoff under a diffusion solves a PDE of this shape, and conversely — is the Feynman-Kac theorem. It is the bridge the whole subject walks back and forth across: an expectation is something to simulate, a PDE is something to solve on a grid, and Feynman-Kac says they are the same object seen from two sides. Chapter 9's Dupire formula is derived from the forward version of the same correspondence.

5.4 Theta Is Gamma

Equation (5.8) looks like a technical detour until it is read as a statement about a trading position, at which point it becomes the most useful thing in this chapter.

Take $r = 0$, which costs nothing and removes the financing terms. Writing $\Theta = \partial V/\partial t$ and $\Gamma = \partial^2 V/\partial S^2$ in the trader's notation, (5.8) reduces to

$$\Theta = -\frac{1}{2}\sigma^2 S^2 \Gamma. \tag{5.9}$$

These are not two independent quantities that happen to be related. They are the same number, with a sign. A position with positive gamma has negative theta, in exact proportion, and the constant of proportionality is the variance.

To see why that matters, hold the option, hedge it, and account for a single day.

Calculation 5.5 (The daily profit and loss of a hedged option)**.** Hold one option and short Δ shares. Over a short interval, expanding the option's value to second order,

$$\begin{aligned} d\Pi &= dV - \Delta\, dS \\ &= \left(\Theta\, dt + \Delta\, dS + \frac{1}{2}\Gamma (dS)^2\right) - \Delta\, dS \\ &= \Theta\, dt + \frac{1}{2}\Gamma (dS)^2, \end{aligned}$$

the first-order terms cancelling, which is what the hedge was for. Substituting (5.9),

$$d\Pi = \frac{1}{2}\Gamma S^2 \left[\left(\frac{dS}{S}\right)^2 - \sigma^2\, dt\right]. \tag{5.10}$$

Read the bracket. The first term is the move the market actually made, squared. The second is the move the option was priced for. A hedged option earns the difference between them, scaled by gamma.

So an option is not a bet on direction — the hedge removed that — and it is not really a bet on volatility either. It is a bet on *realised variance against implied variance*, settled a day at a time. Buy an option, hedge it, and you make money on every day the market moves more than the price assumed and lose on every day it moves less.

Remark (What the premium is)**.** Equation (5.10) also closes a question the replication argument of chapter 4 left open. Rebalancing a hedge buys after a rise and sells after a fall, which loses money on every round trip; that loss is the $\frac{1}{2}\Gamma(dS)^2$ term, and its expected size over the life of the option is

$$\mathbb{E}\left[\int_0^T \frac{1}{2}\Gamma S^2\sigma^2\,dt\right],$$

which by (5.9) is exactly the accumulated theta — that is, the premium. The option premium is not a fee for optionality in the abstract. It is the amount the seller expects to lose rehedging, paid up front.

Which is also why the seller cares about the frequency of large moves rather than only their variance. The premium compensates for the average of $(dS)^2$ and for nothing else, so a distribution with the same average and fatter tails hands over the same premium against a worse experience — and since that is a claim about two distributions rather than about one model, the next calculation shows both.

Calculation 5.6 (The same premium, two worlds)**.** Sell the same one year at-the-money call in two markets whose total variance is identical, so the Black-Scholes price is identical and the seller takes in the same 7.97 either way. In the first all of that variance is diffusive. In the second half of it arrives as compensated jumps at a rate of two a year. Hedge both two hundred and fifty times at the same Black-Scholes delta.[1]

	All diffusion	**Half in jumps**
Mean profit and loss	−0.00	+0.12
Standard deviation	0.44	2.70
Worst one per cent, averaged	−1.50	−12.54
Worst outcome seen	−4.31	−32.86
Fraction of paths that made money	49.9%	63.8%

The premium was fair in both — the mean is zero to within the simulation error, which is what (5.10) guarantees once the variance is matched. Everything else differs. The tail is eight times worse, and the position *wins more often*: a seller in the jump world is right on nearly two thirds of paths and loses thirty-three on the one that goes wrong.

Remark (What that pattern is)**.** The last row is the one to remember, because the shape it describes recurs throughout these notes. A position that makes a little money most of the time and loses a great deal rarely is not a good position that occasionally has bad luck; it is a short option, whether or not anybody wrote one down. Chapter 22 finds the same profile in six trades that have nothing else in common, and chapter 24 is about measuring it, since a mean and a standard deviation describe the left column adequately and the right column not at all.

5.5 Numbers Carried in the Head

The formula of §5.1 is not something anyone evaluates mentally, and traders do not try. What they carry instead is a handful of approximations, all of which come from the same place: at the money, the normal density is flat near zero, so $N(x) \approx \frac{1}{2} + n(0)x$ with $n(0) = 1/\sqrt{2\pi} \approx 0.4$.

[1] `pathwise::hedge_under_jumps`.

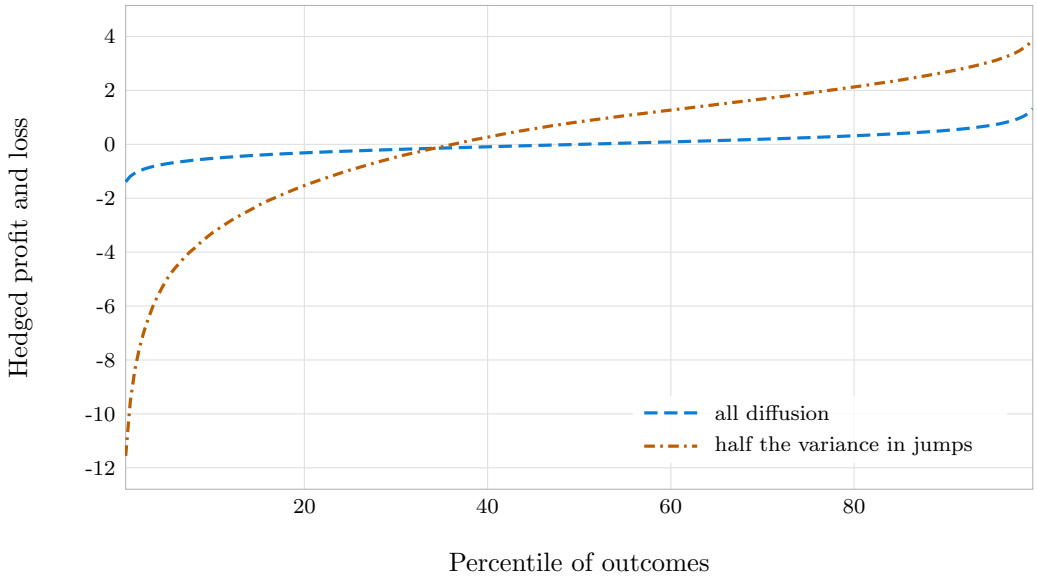

Figure 5.2: Hedged profit and loss at each percentile, selling the same option for the same premium into two markets with the same total variance. The diffusive curve is nearly flat — almost every outcome is close to zero, which is what a working hedge looks like. The jump curve is above it across most of the range and falls off a cliff at the left. Same average of $(dS)^2$, same price, and an entirely different thing to live through, which is why a seller cares about the shape of the distribution and not only its second moment.

Drawn by `hedge_under_jumps`.

Calculation 5.7 (The at-the-money option price)**.** With $r = q = 0$ and $K = F$, we have $d_{1,2} = \pm\frac{1}{2}\sigma\sqrt{T}$, so

$$\begin{aligned} C &= F\big[N(d_1) - N(d_2)\big] \\ &\approx F\Big[\big(\tfrac{1}{2} + n(0)\tfrac{\sigma\sqrt{T}}{2}\big) - \big(\tfrac{1}{2} - n(0)\tfrac{\sigma\sqrt{T}}{2}\big)\Big] \\ &= F\, n(0)\, \sigma\sqrt{T} \;\approx\; 0.4\, F\, \sigma\sqrt{T}. \end{aligned}$$

The error is under half a percent for ordinary parameters and only reaches one percent at volatilities above fifty. Differentiating the same expression in σ gives the vega, and the rest follow:

at-the-money price	$0.4\, F\sigma\sqrt{T}$
vega	$0.4\, F\sqrt{T}$ per unit of volatility
delta	about $\frac{1}{2}$
gamma	$0.4\,/\,(F\sigma\sqrt{T})$
daily break-even move	$\sigma/\sqrt{252} \approx \sigma/16$

The last line is the one used most. Since $\sqrt{252} = 15.87$, a volatility quoted annually divides by roughly sixteen to give the daily move it implies. Twenty vol is a one and a quarter percent day; thirty-two vol is a two percent day. By (5.10) that is exactly the break-even: a day that moves less than $\sigma/16$ loses money for the holder of a hedged option, and a day that moves more makes it.

Example 5.1 (A trade, decided in one's head)**.** An index is at 4,000 and its three month at-the-money options are quoted at 16 volatility. Then

- the option costs about $0.4 \times 4{,}000 \times 0.16 \times \sqrt{0.25} = 128$ index points, a little over three percent of spot;
- its vega is about $0.4 \times 4{,}000 \times 0.5 = 800$ points per unit of volatility, so eight points per volatility point;
- and it breaks even on a daily move of $16/16 = 1\%$, which is 40 points.

So the decision to buy it is a view that the index will move more than forty points a day, on average, over the next three months — which is a question about the world rather than about the model, and can be answered from a chart. The whole apparatus of this chapter has been reduced to a comparison a trader can make between two sentences of a phone call.

5.6 Why the Formula Is a Price

We wrote down a payoff, took an expectation under a particular measure, and called the answer a price. Why should anyone pay it?

The answer is the argument of chapter 4, carried into continuous time: because the option can be *manufactured*. If holding Δ_t shares and adjusting continuously reproduces the payoff exactly, then the option and that strategy are the same object, and two things that pay identically must cost identically or there is an arbitrage. The formula is not a forecast of what the option will be worth. It is the cost of building one.

Two features of the result follow from this, and both surprise people meeting them for the first time.

Remark (The drift is absent, and that is the point)**.** The real world drift μ appears nowhere in the formula. Two investors who disagree completely about whether the stock will rise must still agree on the option's price — because neither of them is being asked to forecast anything. They are being asked what it costs to build a payoff, and the hedge removes the drift from the answer just as it removes it from the portfolio.

This is the continuous time version of the observation in chapter 4 that the historic probabilities did not appear in the binomial price. It is the same fact, and it has the same cause.

Remark (Continuously is not a word about the world)**.** Replication is exact only if the position is adjusted continuously, and nobody does that. Real hedging happens at discrete times, so real replication is imperfect, and the difference is a genuine risk carried by anyone who sells an option.

How imperfect, and in what way, is a question a simulation can answer.

Read the two panels together, because they say different things.

The upper one is reassuring. The error does go to zero, and at the classical rate: quadruple the number of rebalances and the spread halves. Daily hedging of a one year option leaves a standard deviation of a few tenths of a percent of spot, against a premium of about eight percent. Replication is not a fiction.

The lower one is the warning, and it has to be read carefully, because the asymmetry is a property of the tail and not of the body. The mean is zero — the premium is right — and within one and a half standard deviations the two sides carry very nearly equal mass. The difference sits at the thousandth quantile. Hedging eight times, the worst outcomes reach four and a half standard deviations below the mean while the best stop short of three, against a symmetric reference that would go the same distance either way. That is why the panel is drawn on a log scale: on a linear one the region where the two sides differ is indistinguishable from zero.

So a desk that sells options and hedges them discretely collects a little most of the time and occasionally takes a large loss. But notice what else the panel shows, because it qualifies the warning: the asymmetry *shrinks as the rebalancing gets finer*, from a skewness of -0.42 at eight rebalances to -0.18 at a hundred and twenty-eight. This left tail is the price of trading discretely, and trading less discretely reduces it.

What does not shrink is the left tail of Calculation 5.6, because a jump is a jump however often the hedge is adjusted. The two are different warnings, and should be kept apart: this one is about how often you trade, that one about what the world does between your trades. Only the second is a reason to care about the frequency of large moves more than a formula containing only σ would suggest.

References

- Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637–654.

- Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1), 141–183.

- Karatzas, I., & Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer.

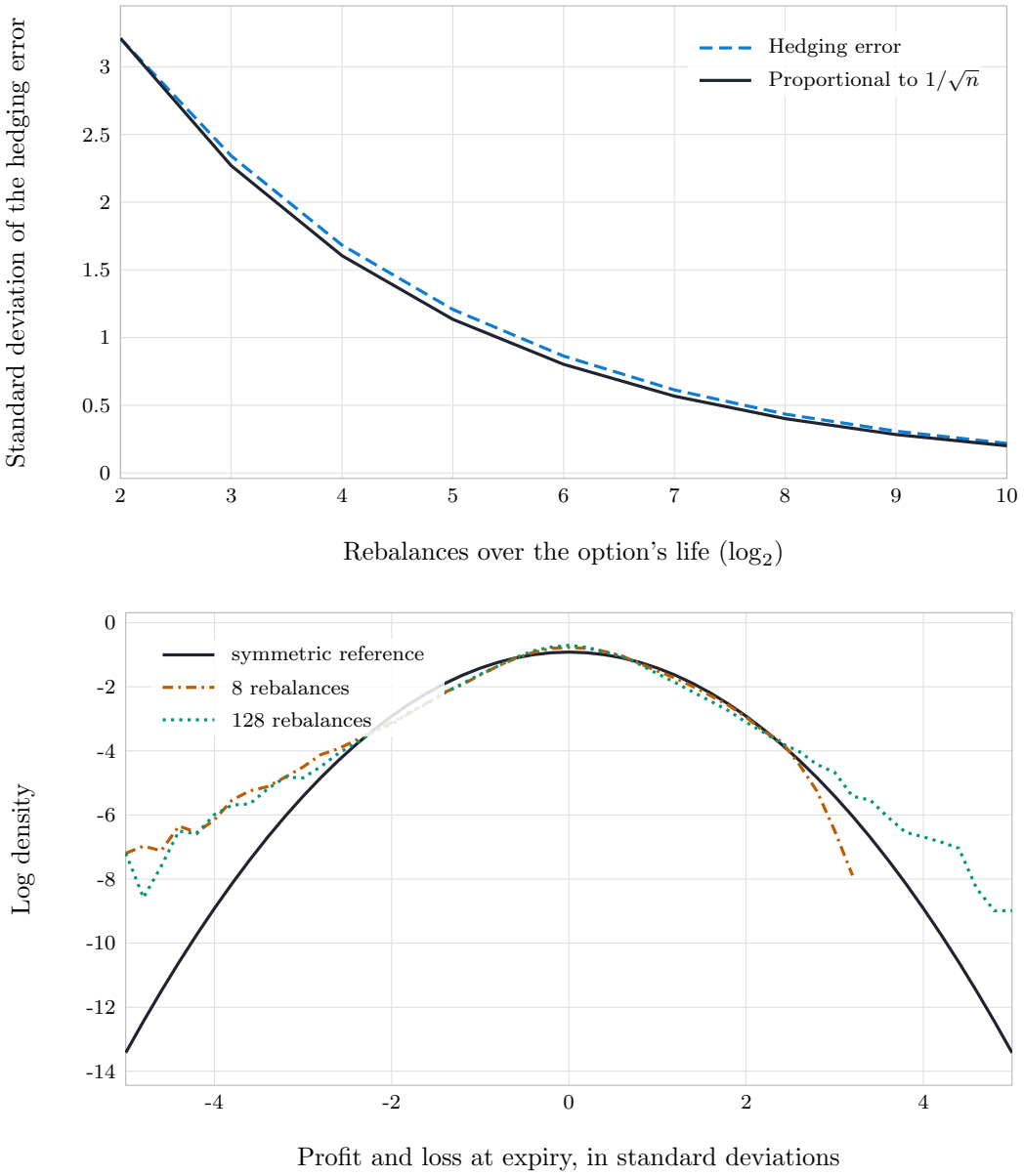

Figure 5.3: Selling a one year at-the-money call for its Black-Scholes price and hedging it n times before expiry. Above: the spread of the outcomes falls in proportion to $1/\sqrt{n}$ — the reference curve is anchored at the coarsest point rather than fitted, so its agreement further along is a prediction. Below: the shape of the profit and loss, standardised so that only the shape is being compared, and drawn against a symmetric reference on a log scale — the asymmetry lives in the last tenth of a per cent, where a linear density is visually zero. At eight rebalances the left arm reaches five standard deviations while the right stops at three; by a hundred and twenty-eight the two are close to level again, which says the skew is the cost of trading discretely rather than a permanent feature of the position.

Drawn by `delta_hedge` and `hedge_statistics`.