

Derivatives: Contracts, Models, and Markets

Part VII: Exotics and Structured Products

Wulin Suo

Draft — October 2, 2026

Contents

† indicates an elective chapter or section; the preface explains what that means.

VII	Exotics and Structured Products	1
34	Exotic Options	2
34.1	What Makes an Option Exotic	2
34.2	Options That Are Barely Exotic	3
34.3	Barrier Options	6
34.4	Why Barriers Break Trees	8
34.5	Static Replication	9
34.6	Discrete Monitoring	10
34.7	Barriers and the Volatility Surface	11
34.8	Digitals and the Overhedge	12
34.9	Lookbacks and Asians	16
34.10	Baskets, Rainbows, and Exchange Options	19
34.11	Worst-of Options	20
34.12	Cliquets, and the Cap That Reverses the Sign	21
34.13	What They Have in Common	23
34.14	Who Trades Exotics, and Why	25
	Summary	25
	Further Reading	26
	Practice Questions	26
	Further Questions	27
35	Structured Products and Autocallables	29
35.1	Taking a Note Apart	29
35.2	Reverse Convertibles and Autocallables	33
35.3	What the Return Distribution Looks Like	35
35.4	The Secondary Market	36
35.5	Other Families	37
35.6	Issuer Hedging, and the Discontinuity	39
35.7	When Many Notes Share a Barrier	40
35.8	Worst-of Underlyings	41
35.9	The Issuer's Credit	42
35.10	Where the Fees Are	42
35.11	How to Evaluate a Note	43
35.12	Suitability, and What Disclosure Cannot Fix	44

35.13 Who Buys These, and Why	44
Summary	45
Further Reading	45
Practice Questions	46
Further Questions	46
Subject Index	49

Part VII

Exotics and Structured Products

Two chapters on payoffs that are designed rather than listed. The first covers the exotic families and the replication arguments that price them; the second follows those payoffs into the retail market, where the same structure is sold as a note with a headline coupon. Read together they make one point: the exotic is priced by the dealer and bought by someone reading a different document.

34 Exotic Options

barriers, lookbacks, Asians, baskets, and the arguments that price them

35 Structured Products and Autocallables

the same payoffs wrapped as notes, and what the wrapper hides

Chapter 34

Exotic Options

Six chapters have deferred material to this one. Everything in it trades over the counter, in the market Section 10.10 described; an exotic option is a dealer's product, and that fact shapes how it is priced, hedged, and sold.

Section 15.2 said every payoff whose value depends on S and t alone satisfies the Black–Scholes–Merton equation without modification, and that a path-dependent one adds a state variable. Section 18.6 said vanna and volga matter here. Section 18.13 said that where gamma diverges, dynamic hedging becomes an impossibility problem and a static hedge is the only answer. Section 19.12 said barrier prices depend on how the surface *evolves*, which no static surface specifies. Section 20.4 promised that a barrier priced on a tree whose nodes straddle the barrier converges appallingly. Section 30.7 said the replication argument for exotic payoffs is developed here.

The unifying question is which exotics are genuinely harder than vanillas and which merely look it. The answer is not about payoff complexity. It is about whether the payoff can be replicated by instruments the market quotes, and Sections 34.5 and 34.13 give the test.

34.1 WHAT MAKES AN OPTION EXOTIC

Three features separate an exotic from a vanilla, with very different consequences. Two turn on the path rather than the endpoint, so the notation comes first. Throughout, S_t is the underlying and K the strike, and the path statistics are

$$M_T = \max_{0 \leq t \leq T} S_t, \quad m_T = \min_{0 \leq t \leq T} S_t, \quad A_T = \frac{1}{T} \int_0^T S_t dt$$

the running maximum, the running minimum, and the average.

Path dependence. The payoff depends on the route, not only the destination. A **barrier option** is a vanilla multiplied by an indicator:

$$\text{down-and-out call} = (S_T - K)^+ \mathbf{1}_{\{m_T > H\}}, \quad \text{down-and-in call} = (S_T - K)^+ \mathbf{1}_{\{m_T \leq H\}} \quad (34.1)$$

with the up-variants replacing m_T by M_T . Written this way, in–out parity—knock-in plus knock-out equals the vanilla, equation (34.10) in Section 34.3—stops being a result and becomes an observation: the two indicators sum to one, so the two options sum to the vanilla.

A **lookback option** replaces the strike or the terminal price by an extreme of the path:

$$\begin{aligned} \text{floating call} &= S_T - m_T, & \text{floating put} &= M_T - S_T, \\ \text{fixed call} &= (M_T - K)^+, & \text{fixed put} &= (K - m_T)^+ \end{aligned} \quad (34.2)$$

The floating pair can never finish out of the money, which is most of the reason they cost what they cost. An **Asian option** replaces the terminal price by the average A_T in one of two places:

$$\text{average price} = (A_T - K)^+, \quad \text{average strike} = (S_T - A_T)^+ \quad (34.3)$$

and a **cliquet** sums period returns $r_i = S_{t_i}/S_{t_{i-1}} - 1$ over reset dates $0 = t_0 < t_1 < \dots < t_n = T$, each floored at F and capped at C :

$$\text{cliquet} = \sum_{i=1}^n \min(\max(r_i, F), C) \quad (34.4)$$

The cap is inside the sum rather than outside it, and Section 34.12 is about what that does to the sign of the volatility exposure.

This is where the model risk lives, because a model calibrated to terminal distributions does not pin down paths. Five sections of this chapter go to the barrier alone for that reason—Sections 34.3 through 34.7—and Section 34.9 takes the lookback and the Asian together.

Multiple underlyings. **Basket options, rainbow options, and exchange options**, of which four matter:

$$\begin{aligned} \text{basket} &= \left(\sum_i w_i S_i(T) - K \right)^+, & \text{exchange} &= (S_1(T) - S_2(T))^+, \\ \text{worst-of put} &= (K - \min_i S_i(T))^+, & \text{best-of call} &= \left(\max_i S_i(T) - K \right)^+ \end{aligned} \quad (34.5)$$

Here $S_i(T)$ is the price of the i th asset at expiration and w_i the number of units of it the basket holds; the basket of Example 34.8 is $(S_1 + S_2)/2$, so $w_1 = w_2 = 1/2$. The new input is correlation, which Section 19.7 established is not observable in the way volatility is. It enters every one of these payoffs and is quoted in none of them, which is Section 34.10's subject and the reason the worst-of put in Section 34.11 behaves as it does.

Non-standard exercise or payoff. Bermudans, compounds, choosers, and the digital pair, the last of these carrying the identity that ties it to Black–Scholes:

$$\begin{aligned} \text{cash-or-nothing} &= Q \mathbf{1}_{\{S_T > K\}}, & \text{asset-or-nothing} &= S_T \mathbf{1}_{\{S_T > K\}}, \\ (S_T - K)^+ &= S_T \mathbf{1}_{\{S_T > K\}} - K \mathbf{1}_{\{S_T > K\}} \end{aligned} \quad (34.6)$$

A vanilla call is one asset-or-nothing less K cash-or-nothings, which is Section 28.10's reading of $N(d_1)$ and $N(d_2)$ written as a payoff rather than as a price. These are usually the easy ones, because a single-asset payoff at a single date is a function of the terminal distribution, which the vanilla market already reveals.

The taxonomy sorts the chapter: the third group is priced almost model-independently, the first two are not. A desk that treats “exotic” as one category prices some trades far too carefully and others not carefully enough.

34.2 OPTIONS THAT ARE BARELY EXOTIC

Several instruments carry exotic names and almost no model risk.

Packages are combinations of vanillas: Chapter 12's spreads, straddles, and collars, sold as one contract. They are priced by adding up the pieces, and the only thing exotic about them is the ticket.

Non-standard American options—exercisable only in windows, or with a strike that steps—need Chapter 20’s lattice but no new theory.

Compound options, an option on an option, have closed forms under Black–Scholes and are genuinely sensitive to volatility in a way vanillas are not: a compound option is a bet on volatility twice over, since both the outer and inner values depend on it. Their vega is correspondingly large relative to their price. Geske’s formula for a call on a call—the right, on payment of K_1 at t_1 , to acquire a call struck at K_2 expiring at T —is

$$c_{\text{call}} = S_0 e^{-qT} M(a_1, b_1; \rho) - K_2 e^{-rT} M(a_2, b_2; \rho) - K_1 e^{-rt_1} N(a_2), \quad \rho = \sqrt{t_1/T} \quad (34.7)$$

where $M(\cdot, \cdot; \rho)$ is the bivariate standard normal distribution function,

$$a_1 = \frac{\ln(S_0/S^*) + (r - q + \sigma^2/2)t_1}{\sigma\sqrt{t_1}}, \quad b_1 = \frac{\ln(S_0/K_2) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}},$$

$a_2 = a_1 - \sigma\sqrt{t_1}$, $b_2 = b_1 - \sigma\sqrt{T}$, and S^* is the asset price at t_1 at which the inner call is worth exactly K_1 . That critical price is the only part of the formula requiring numerical work, and it is a one-dimensional root search on Black–Scholes. The correlation entering M is $\sqrt{t_1/T}$ because one Brownian motion carries the price through both dates, which is the whole reason a compound option is a two-dimensional integral rather than a one-dimensional one. The other three differ only in signs. Writing p_{call} for a put on a call, and c_{put} and p_{put} for a call and a put on a put,

$$\begin{aligned} p_{\text{call}} &= K_2 e^{-rT} M(-a_2, b_2; -\rho) - S_0 e^{-qT} M(-a_1, b_1; -\rho) + K_1 e^{-rt_1} N(-a_2) \\ c_{\text{put}} &= K_2 e^{-rT} M(-a_2, -b_2; \rho) - S_0 e^{-qT} M(-a_1, -b_1; \rho) - K_1 e^{-rt_1} N(-a_2) \\ p_{\text{put}} &= S_0 e^{-qT} M(a_1, -b_1; -\rho) - K_2 e^{-rT} M(a_2, -b_2; -\rho) + K_1 e^{-rt_1} N(a_2) \end{aligned}$$

Two things change with the inner option rather than the outer one. In the last two, S^* is the price at which the inner put is worth K_1 , not the inner call, so it is a different root search; and a put loses value as the asset rises, so the outer option is exercised below S^* for a call on a put and above it for a put on a put.

Only two of the four are independent. Holding a call on a call and writing a put on the same call delivers that call at t_1 against K_1 paid, whatever the asset does in between, so

$$c_{\text{call}} - p_{\text{call}} = c(S_0, K_2, T) - K_1 e^{-rt_1}, \quad c_{\text{put}} - p_{\text{put}} = p(S_0, K_2, T) - K_1 e^{-rt_1}$$

and those two identities are the cheapest check available on an implementation of any of the four.

Figure 34.1 puts the volatility claim in numbers. At a 10% volatility the compound is worth 3.02 against the vanilla’s 6.81, and its value moves **2.2 times** as much for a given relative move in volatility; by 50% that ratio is down to 1.2. The compound’s elasticity barely moves—it sits near 1.1 throughout, so its value is close to proportional to volatility—while the vanilla’s climbs from 0.50 to 0.86. Paying twice for volatility does not buy a larger vega; it buys a vega that stays large relative to a much smaller price.

Adding the second premium at its present value, $K_1 e^{-rt_1} = 4.88$, gives what ending up with the same call costs in total: 7.90 against the vanilla’s 6.81 at a 10% volatility, 22.59 against 21.79 at 50%. The total exceeds the vanilla at every volatility—committing to exercise buys the identical instrument, so it cannot be cheaper—and the gap is what the right to walk away at t_1 is worth: **1.10** at 10%, narrowing to **0.80** at 50%, because the higher the volatility the less likely the holder is to decline.

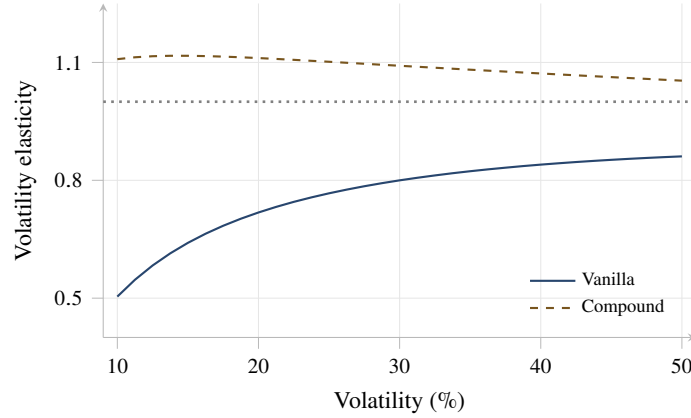


Figure 34.1: $\sigma (\partial V / \partial \sigma) / V$ at $S_0 = K_2 = 100$, outer premium $K_1 = 5$ paid at $t_1 = 0.5$, $T = 1$, $r = 5\%$, no dividends.

Chooser options let the holder decide at t_1 whether the option is a call or a put, both struck at K and both expiring at T . A simple chooser decomposes exactly into vanillas. Write $c(S, K, T)$ and $p(S, K, T)$ for Black–Scholes–Merton values. At the choice date the holder takes whichever is worth more, so the chooser is then worth

$$\max\{c(S_{t_1}, K, T - t_1), p(S_{t_1}, K, T - t_1)\}$$

Put–call parity at t_1 gives $p - c = Ke^{-r(T-t_1)} - S_{t_1}e^{-q(T-t_1)}$, the value of a short forward, which carries no optionality at all. Substituting it,

$$\begin{aligned} \max\{c, p\} &= c + \max\{0, p - c\} \\ &= c(S_{t_1}, K, T - t_1) + e^{-q(T-t_1)} \max\left(Ke^{-(r-q)(T-t_1)} - S_{t_1}, 0\right) \end{aligned}$$

where $e^{-q(T-t_1)}$ has been taken outside the maximum. Both terms are now instruments the vanilla market prices. The first is a call struck at K expiring at T , carried through the choice date untouched, and so worth $c(S_0, K, T)$ today. The second is $e^{-q(T-t_1)}$ puts struck at $Ke^{-(r-q)(T-t_1)}$, and they expire at t_1 rather than at T , because the maximum is taken on the price at the choice date. Hence

$$V_0 = c(S_0, K, T) + e^{-q(T-t_1)} p\left(S_0, Ke^{-(r-q)(T-t_1)}, t_1\right) \quad (34.8)$$

Nothing is left over, which makes it a good exercise and a poor example of an exotic.

Forward-start options are struck at the money on a future date. Under Black–Scholes their value scales with the spot, so an option struck at αS_{t_1} and expiring at T is worth

$$V_0 = e^{-qt_1} c(S_0, \alpha S_0, T - t_1) \quad (34.9)$$

a vanilla's price times e^{-qt_1} , with no dependence on anything observed between now and t_1 . Under any realistic model it is not, because it depends on the *forward* smile—and Section 21.3 measured how badly local volatility gets that wrong. A cliquet is a strip of forward-starts, so it is the purest instrument in this chapter for trading forward volatility, and the hardest to price for exactly the same reason.

34.3 BARRIER OPTIONS

Barriers are the largest exotic market, and they take up most of this chapter.

A barrier option is a vanilla that is extinguished (**knock-out**) or created (**knock-in**) if the underlying touches a level H . With up and down, in and out, and calls and puts, there are eight basic types, and the in-out parity

$$\text{knock-in} + \text{knock-out} = \text{vanilla} \quad (34.10)$$

means only four need to be priced.

They exist because they are cheaper. A down-and-out call costs less than the vanilla because it can be extinguished, and a buyer who believes the barrier will not be touched is paying for the part of the distribution they think matters.

All eight have closed forms under Black–Scholes with continuous monitoring, and the whole family is assembled from four expressions. Write $\phi = +1$ for a call and -1 for a put, $\eta = +1$ for a down-barrier and -1 for an up-barrier, and set

$$\lambda = \frac{r - q + \sigma^2/2}{\sigma^2}, \quad \begin{aligned} x_1 &= \frac{\ln(S_0/K)}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}, & x_2 &= \frac{\ln(S_0/H)}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}, \\ y_1 &= \frac{\ln(H^2/(S_0K))}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}, & y_2 &= \frac{\ln(H/S_0)}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}. \end{aligned}$$

Then

$$\begin{aligned} A &= \phi S_0 e^{-qT} N(\phi x_1) - \phi K e^{-rT} N(\phi(x_1 - \sigma\sqrt{T})), \\ B &= \phi S_0 e^{-qT} N(\phi x_2) - \phi K e^{-rT} N(\phi(x_2 - \sigma\sqrt{T})), \\ C &= \phi S_0 e^{-qT} \left(\frac{H}{S_0}\right)^{2\lambda} N(\eta y_1) - \phi K e^{-rT} \left(\frac{H}{S_0}\right)^{2\lambda-2} N(\eta(y_1 - \sigma\sqrt{T})), \\ D &= \phi S_0 e^{-qT} \left(\frac{H}{S_0}\right)^{2\lambda} N(\eta y_2) - \phi K e^{-rT} \left(\frac{H}{S_0}\right)^{2\lambda-2} N(\eta(y_2 - \sigma\sqrt{T})), \end{aligned} \quad (34.11)$$

and the four knock-in values are read off the table below. Knock-outs follow from the parity (34.10).

Knock-in option	Value, by where the barrier sits relative to the strike	
Down-and-in call	$H \leq K: C$	$H > K: A - B + D$
Up-and-in put	$H \geq K: C$	$H < K: A - B + D$
Up-and-in call	$H \geq K: B - C + D$	$H < K: A$
Down-and-in put	$H \leq K: B - C + D$	$H > K: A$

Three things can be read off (34.11) directly.

A is the vanilla, since $x_1 = d_1$; the barrier enters only through B , C , and D . B is A with the barrier replacing the strike inside the normal arguments—the strike still sets the payoff, but the barrier now sets the boundary. C and D are the reflections of A and B in the barrier, weighted by the powers of H/S_0 that the reflection principle produces—and those powers are where the drift asymmetry of Section 34.5 comes from.

The two entries equal to A are the vanilla itself, because a barrier lying between the spot and the strike is one the option cannot finish in the money without having crossed: an up-and-in call with $H < K$ knocks in on every path that pays it anything.

And the formula contains the static hedge. When $r = q$ the exponent λ collapses to $1/2$, the two weights become H/S_0 and S_0/H , and the down-and-in call's C reads term for term as

K/H puts struck at H^2/K . Section 34.5's reflection result is visible in the formula, which is a good reason to write the formula out rather than call a library.

The formulas assume no rebate. A rebate R paid on knock-out is valued separately, and for a down barrier both versions have closed forms in the same λ as (34.11). If it is paid at expiration it is a digital on the running minimum, paying if $m_T \leq H$, and is worth Re^{-rT} times the risk-neutral probability that the barrier is touched,

$$P(m_T \leq H) = N\left(\frac{\ln(H/S_0)}{\sigma\sqrt{T}} - (\lambda - 1)\sigma\sqrt{T}\right) + \left(\frac{H}{S_0}\right)^{2\lambda-2} N\left(\frac{\ln(H/S_0)}{\sigma\sqrt{T}} + (\lambda - 1)\sigma\sqrt{T}\right)$$

whose weight $(H/S_0)^{2\lambda-2}$ is the one that already appears in C and D . If it is paid on the touch it is a first-passage claim, discounted from a random date, and is worth

$$R \left[\left(\frac{H}{S_0}\right)^{\lambda-1+\psi} N(z) + \left(\frac{H}{S_0}\right)^{\lambda-1-\psi} N(z - 2\psi\sigma\sqrt{T}) \right]$$

where

$$\psi = \sqrt{(\lambda - 1)^2 + \frac{2r}{\sigma^2}}, \quad z = \frac{\ln(H/S_0)}{\sigma\sqrt{T}} + \psi\sigma\sqrt{T}$$

An up barrier reverses the signs of the arguments of N in both. For the down-and-out call of Example 34.1 below, $\lambda = 1.3$, and with no dividends $\psi = \lambda$ exactly; the barrier is touched with probability 0.652, so a rebate of 1 paid at expiration is worth $e^{-0.05} \times 0.652 = 0.620$, and the same rebate paid on the touch, which arrives earlier on average, is worth 0.643. Either is added to the no-rebate value to price the contract with it.

Example 34.1 What the barrier removes

$S_0 = 100$, $K = 100$, $H = 90$, $r = 5\%$, $\sigma = 25\%$, $T = 1$, continuously monitored.

Vanilla call	12.335999
Down-and-out call	9.111221
Down-and-in call (the difference)	3.224778

The down-and-in value is the first row of the table above, the single expression C with $\lambda = 1.3$: $(H/S_0)^{2\lambda} = 0.760380$ and $(H/S_0)^{2\lambda-2} = 0.938740$, $y_1 = -0.517884$, and

$$C = 100 \times 0.760380N(-0.517884) - 100e^{-0.05} \times 0.938740N(-0.767884) = 3.224778$$

The barrier removes 26.1% of the premium—a discount large enough to be the entire reason the instrument is bought, and large enough that the pricing had better be right.

Two observations shape everything after this.

The knock-out is not simply a cheaper call. Approaching the barrier makes the option less valuable while approaching the strike makes it more valuable, and between those two pulls a knock-out's delta can exceed one in magnitude and its gamma can change sign—Figure 34.4 measures both. *A reverse knock-out*—barrier above the strike, so the option is knocked out when it is deep in the money—has a discontinuous payoff at the barrier and is the most difficult vanilla-like instrument in the market.

And the price depends on the whole path distribution. The vanilla depends only on the terminal distribution, which Section 19.2 showed the option chain reveals. The barrier depends on the probability of touching, which it does not.

34.4 WHY BARRIERS BREAK TREES

Section 20.4 promised this, and the failure is instructive rather than merely inconvenient.

Example 34.2 More steps, worse answer

Price Example 34.1's down-and-out call— $S_0 = K = 100$, $H = 90$, $r = 5\%$, $\sigma = 25\%$, $T = 1$, no dividends, European exercise—on a Cox–Ross–Rubinstein tree, knocking out at any node at or below H . The tree can knock out only at a time step, while the analytic answer is the continuously monitored one: **9.111221**.

Steps	Tree	Error	Effective barrier
50	9.142046	+0.031	89.94
51	10.494157	+1.383	86.93
100	9.944477	+0.833	88.25
200	9.144117	+0.033	89.94
400	9.438414	+0.327	89.36
1600	9.153506	+0.042	89.92

Fifty-one steps is forty-five times worse than fifty. And 1600 steps is worse than 200.

The mechanism is entirely in the last column. The tree can only knock out at a node, so the barrier it actually enforces is the highest node level at or below H . As n changes, that level jumps around: at $n = 50$ it is 89.94 and at $n = 51$ it is 86.93. The tree is pricing a different contract at every step count, and the error is dominated by which contract rather than by discretization.

The error is also one-signed. The effective barrier is always at or below the true one, so the option is always over-priced—which is worse than a random error, because averaging two grids does not help and Richardson extrapolation, which Section 20.3 already warned about, makes it worse.

Now fix it. Choose n so that a layer of nodes falls exactly on the barrier. In a CRR tree the node levels are $S_0 u^{\pm j}$ with $u = e^{\sigma\sqrt{T/n}}$, so this requires

$$\frac{\ln(S_0/H)}{\sigma\sqrt{T/n}} = m \quad \text{for an integer } m \quad \implies \quad n = \left\lfloor \frac{m^2 \sigma^2 T}{\ln^2(S_0/H)} \right\rfloor$$

m	Steps	Tree	Error
4	90	9.111965	+0.00074
6	202	9.119059	+0.00784
8	360	9.113142	+0.00192
9	456	9.111153	−0.00007
12	810	9.113370	+0.00215
16	1441	9.112680	+0.00146

At 90 steps with the layer on the barrier, the error is 0.0007. At 100 steps without it, the error is 0.833—more than a thousand times larger, with more steps.

The lesson generalizes well beyond barriers. A numerical method that resolves the payoff's discontinuities converges; one that does not, does not—and refining a grid that misses the discontinuity refines the wrong problem. Section 20.3's control variate worked for the same underlying reason, and Section 20.4's trinomial tree exists to give the free parameter that makes this placement possible for any n .

The practical rule is simply stated. Never let a user choose the step count on a barrier product. The pricing library should choose it from the barrier.

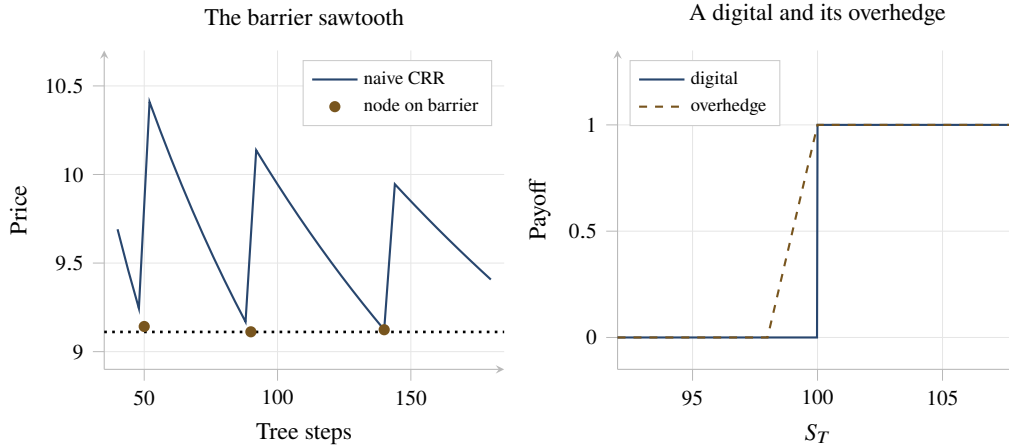


Figure 34.2: Left: Example 34.2; the dotted line is the analytic value. Right: a one-sided call spread dominating a digital.

In Figure 34.2's left panel, the sawtooth is not noise: each descending run is the effective barrier drifting up toward H as n increases, and each jump is a node crossing the barrier and resetting it.

Or choose a method whose grid is free. The binomial tree's trouble is that it has no spare parameter: $u = e^{\sigma\sqrt{\Delta t}}$ is fixed once σ and Δt are, so the node levels are determined and the barrier falls where it falls. A trinomial tree carries a space step that can be set independently of the time step, so choosing $\Delta x = \ln(S_0/H)/m$ for an integer m puts a layer of nodes on the barrier at every step count rather than at lucky ones. In $\ln S$ that tree is the explicit finite-difference scheme of Section 20.12, and an explicit scheme pays for the freedom with a stability condition, $\Delta x \geq \sigma\sqrt{\Delta t}$: at 50 time steps it allows only two space steps between the spot and the barrier, so hitting the barrier costs a coarse grid. An implicit scheme has no such condition. There the barrier is made the lower boundary of the grid, where the knock-out is the exact condition $V = 0$ rather than a test applied at interior nodes, and the space step is refined as far as wanted independently of the time step. That is the sense in which these are the natural methods for barriers: the discontinuity is put on the mesh by construction instead of being hunted for. On Example 34.2's contract the largest error over 50, 51, 100, 200 and 400 steps falls from **1.383** to **0.017**, and the jump from 50 steps to 51 that cost 1.35 becomes 0.02. This is why barrier books are priced on grids rather than on binomial trees.

34.5 STATIC REPLICATION

Section 18.13 argued that where gamma diverges, a static hedge is the only workable answer. Barriers are the main place that argument pays off.

Example 34.3 A barrier hedged with vanillas and never touched again

When the drift is zero— $r = q$ —a **reflection** argument gives an exact static replication. A down-and-in call with strike K and barrier $H < K$ has the same value as K/H puts struck at H^2/K .

The argument has two cases. If the barrier is never touched, the call never comes alive, and the puts expire worthless too: they are struck at H^2/K , below H , and the price has stayed above H . If the barrier is touched, then at that moment K/H puts struck at

H^2/K are worth exactly one call struck at K —a symmetry of the lognormal distribution that holds when $r = q$ —so the puts are sold and the call bought at no cost, and the hedge has become the option.

Nothing has to be computed to confirm it, because (34.11) contains it. A down-and-in call with $H \leq K$ is the single block C , and at $r = q$ the exponent is $\lambda = 1/2$, so the two weights collapse to H/S_0 and S_0/H :

$$C = e^{-rT} \left[H N(y_1) - K \frac{S_0}{H} N(y_1 - \sigma\sqrt{T}) \right], \quad y_1 = \frac{\ln(H^2/(S_0K))}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2}$$

Now price the puts. One struck at H^2/K has $d_1 = \ln(S_0K/H^2)/(\sigma\sqrt{T}) + \sigma\sqrt{T}/2$, so $-d_1 = y_1 - \sigma\sqrt{T}$ and $-d_2 = y_1$, and K/H of them are worth

$$\frac{K}{H} \left[\frac{H^2}{K} e^{-rT} N(y_1) - S_0 e^{-rT} N(y_1 - \sigma\sqrt{T}) \right] = e^{-rT} \left[H N(y_1) - K \frac{S_0}{H} N(y_1 - \sigma\sqrt{T}) \right]$$

term for term the same expression. The replication is therefore exact rather than close, at any common level of $r = q$: the bracket contains no r , so raising the carry only discounts the same number. The condition is zero *drift*, not zero rates.

Now break the condition. At $r = 5\%$, $q = 0$:

Down-and-in call	3.224778
Replicating puts	1.861532
Static hedge error	−42.3%

The replication is not robust to drift, and 42% is not a rounding error. Refinements exist—replicating with a strip of vanillas at several maturities rather than one, which trades exactness for robustness—and they are what desks actually use.

But the reason to want a static hedge survives the error. A static hedge is a portfolio of vanillas held to expiration. It requires *no rebalancing*, so it has no transaction costs and no gamma at the barrier; and it does not depend on the surface dynamics that Section 19.12 said barrier prices depend on. A hedge that is wrong by a known amount and never has to be adjusted can be better than one that is right in a model and must be adjusted precisely where the model is least reliable.

34.6 DISCRETE MONITORING

Every formula so far assumes the barrier is monitored continuously. *No contract is*: barriers are observed at a daily close, sometimes weekly or monthly. The gap between the two is larger than it looks and has a clean correction.

Example 34.4 What a daily close is worth

Broadie, Glasserman, and Kou showed that a barrier monitored m times behaves, to first order, like a *continuous* barrier shifted *away* from the spot:

$$H_{\text{eff}} = H e^{\pm\beta\sigma\sqrt{T/m}}, \quad \beta = -\frac{\zeta(1/2)}{\sqrt{2\pi}} = 0.5826 \quad (34.12)$$

where ζ is the Riemann zeta function, $\zeta(1/2) = -1.4604$, and the sign in the exponent is chosen so the barrier moves further from the current spot—down for a down-barrier, up

for an up-barrier.

Applied to Example 34.1's down-and-out call, whose continuous value is 9.111221:

Monitoring	m	Effective H	Price	vs. continuous
Daily	252	89.18	9.527218	+4.57%
Weekly	52	88.20	9.968231	+9.41%
Monthly	12	86.29	10.675121	+17.16%
Quarterly	4	83.68	11.365166	+24.74%

Even daily monitoring is worth 4.6% of the option, and monthly monitoring 17%. A knock-out priced as continuous is undervalued, because a path that dips below the barrier between closes is not caught.

Verify it independently. Simulating the daily-monitored contract directly with 400,000 antithetic paths:

	Value	Against Monte Carlo
Monte Carlo (daily)	9.523044	(s.e. 0.028437)
Corrected formula (34.12)	9.527218	+0.15 standard errors
Uncorrected (continuous)	9.111221	−14.5 standard errors

The correction lands within a fifth of a standard error and the uncorrected formula is fourteen and a half away. The correction looks like a fudge and is not: β comes from the expected overshoot of a random walk past a level, and the same constant appears wherever a discretely observed process is compared with a continuous one.

Three consequences for practice follow.

The correction is not optional at low monitoring frequencies. Quarterly observation moves the price by a quarter, which no bid–ask absorbs.

It interacts with Section 34.4. A tree with a node layer on the *stated* barrier prices the continuous contract; to price the discrete one, the layer goes on H_{eff} . Getting one right and the other wrong produces two errors that do not cancel, because they have the same sign.

And it matters most for the products of Chapter 35. An autocallable observes its knock-in barrier daily and its call barrier on a small number of dates, so a single instrument carries two different monitoring frequencies and needs two different corrections.

34.7 BARRIERS AND THE VOLATILITY SURFACE

A barrier price depends on the volatility surface in a way a vanilla price does not, and Section 21.11 measured by how much. Two models—local volatility and a stochastic-volatility model—were calibrated to the *same* vanilla surface and verified to reproduce it. They then priced a one-year up-and-out call, spot and strike at 100, barrier at 110, monitored daily: 0.8402 under the stochastic-volatility model against 0.4244 under local volatility, a difference of **49.5%** on an option worth less than a point.

The design of that experiment is Hull and Suo's (2002): fit the competing models to a common surface, *verify* that each reproduces it, and only then price something the surface does not determine. What the spread measures is not an error in either model—both are right about every quoted vanilla—but the size of the question the market has left unanswered.

The reason is Section 19.12's. A vanilla depends on the terminal distribution and a barrier depends on the path, and two models can agree on every terminal distribution while disagreeing about how the underlying gets there. Calibration to vanillas is therefore necessary and not

sufficient, and a desk that reports a barrier price to four decimals from a calibrated model is reporting the model.

Three practical responses follow, in ascending order of what they give up.

Price with the model, reserve against the range. Compute under two or three models and hold the spread as a reserve. Section 21.11's argument for reporting model risk applies exactly.

Hedge statically where possible, which converts the model question into a residual and a known error, as Section 34.5 described.

Or refuse the trade. A reverse knock-out at a barrier where the two models differ by half the premium is a trade whose price the desk does not know, and there are market conditions in which the correct response is to say so.

34.8 DIGITALS AND THE OVERHEDGE

This section treats digitals at more length than their own market would justify, because a digital is the simplest payoff with a jump in it. Everything that makes a discontinuous payoff hard to hedge dynamically appears here with nothing else in the way: a delta that grows without bound as expiration approaches, a hedge that has to be overhedged rather than replicated, and a bid-ask spread set by what the overhedge costs. A barrier option carries the same jump at its barrier, which is why Example 34.6 returns to the argument for the reverse knock-out.

Section 12.8 showed a digital has no finite replicating portfolio. Here is what that costs.

The closed forms are the two halves of Black-Scholes taken apart. A **cash-or-nothing** call pays Q if $S_T > K$; an **asset-or-nothing** call pays S_T in the same event. Their values are

$$V_{\text{cash}} = Qe^{-rT}N(d_2), \quad V_{\text{asset}} = S_0e^{-qT}N(d_1) \quad (34.13)$$

with d_1 and d_2 the usual Black-Scholes arguments, and the put versions replacing $N(\cdot)$ with $N(-\cdot)$. With $Q = 1$ the vanilla call is $V_{\text{asset}} - K V_{\text{cash}}$, which is Section 15.4's reading of $N(d_1)$ and $N(d_2)$ restated as two instruments rather than two integrals.

The formula is also the limit of the call spreads the example below prices, writing $c(K)$ for the value of a call struck at K :

$$\frac{c(K - \varepsilon/2) - c(K + \varepsilon/2)}{\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} -\frac{\partial c}{\partial K} = e^{-rT}N(d_2) \quad (34.14)$$

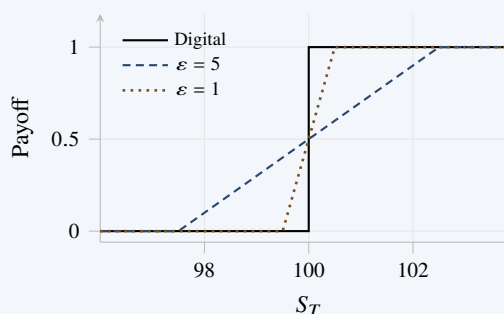
A digital is the strike-derivative of the vanilla, one differentiation short of Section 19.2's implied density, which is the second. The price is therefore a derivative of quoted prices, and a derivative of a quoted price is not itself quoted.

Example 34.5 What certainty costs

A cash-or-nothing call paying \$1 if $S_T > 100$, with $S_0 = 100$, $r = 5\%$, $\sigma = 25\%$, $T = 1$ and no dividends: worth **0.504049**.

Replicate with a call spread of width ε straddling the strike—long $K - \varepsilon/2$, short $K + \varepsilon/2$ —held in quantity $1/\varepsilon$. Each spread pays at most ε , so $1/\varepsilon$ of them pay nothing below $K - \varepsilon/2$, the digital's full dollar above $K + \varepsilon/2$, and rise linearly between them. That is (34.14) evaluated at finite ε :

ε	Spreads	Value
5.00	0.2	0.504159
1.00	1.0	0.504054
0.10	10.0	0.504050
0.01	100.0	0.504049



The value converges and the position diverges. The plot shows why the value converges—the payoff approaches the step—and also why this spread is not yet a hedge: it crosses the digital, paying more below the strike and less above it.

What the dealer has sold has no finite greeks. With τ the time remaining to expiration, differentiating $e^{-r\tau}N(d_2)$ once and twice, and collecting $d_1 = d_2 + \sigma\sqrt{\tau}$,

$$\Delta = \frac{e^{-r\tau}N'(d_2)}{S\sigma\sqrt{\tau}}, \quad \Gamma = -\frac{e^{-r\tau}N'(d_2)d_1}{S^2\sigma^2\tau}$$

At the money $d_1 = 1.3\sigma\sqrt{\tau}$, since $(r - q + \sigma^2/2)/\sigma^2 = 1.3$ here, so both greeks grow like $1/\sqrt{\tau}$: each doubles whenever the time remaining is quartered. Away from the money it is the d_1 in the numerator that carries the gamma's sign, so the gamma changes sign at the delta's peak, $Ke^{-(r-q+\sigma^2/2)\tau}$. That point closes on the strike as expiration approaches—three cents below it with a day to run, under a hundredth of a cent with a minute—while $|\Gamma|$ five cents to either side of it reaches 32. The delta is 0.25 with a day to run, 0.65 with an hour, 5.0 with a minute and 38.8 with a second.

Holding $1/\varepsilon$ spreads caps the delta at $1/\varepsilon$, which is the whole reason for holding spreads rather than the digital; and even at that cap, with $\varepsilon = 0.10$, ten shares of stock have to be traded for every dollar of payoff while the price sits inside the window in the last moments before expiration, with no time to trade them. That is what Section 18.13's “gamma diverges” means when someone has to do it.

So dealers overhedge. Sell the digital and buy $1/\varepsilon$ call spreads struck at $K - \varepsilon$ and K , so that the upper strike stays at the digital's strike and the whole window lies *below* it. That position already pays the full dollar at K and something below it, so it dominates the digital's payoff everywhere. The table widens the window and compares each cost with the digital's 0.504049:

ε	Strikes	Cost	Premium over the digital
1	[99, 100]	0.511635	+1.5%
2	[98, 100]	0.519249	+3.0%
5	[95, 100]	0.542210	+7.6%

That premium is what certainty costs. A dealer who will not carry the residual cannot sell the digital below the dominating spread, and buying it cannot pay above the mirror image—a spread struck entirely *above* the strike, which the digital dominates. At a one-point window those bounds are 0.4965 and 0.5116 around a digital worth 0.5040, so no two-way price narrower than **3.0%** is hedged on both sides. What a desk actually quotes is wider: under a skew the digital's own value moves (Example 19.4 differentiates along the smile), a jump through the strike can leave the overhedge short, every one of the

$1/\varepsilon$ spreads is bought at the offer, and the overhedged book ties up capital. None of these is in the model here; this is the floor the payoff alone puts under the quote. Figure 34.2's right panel shows the dominance.

The width is a business decision, not a modeling one. A tighter spread is a better price, a worse hedge, and a higher cap on the position the desk may have to trade—one share at a point wide, ten at a tenth, a hundred at a hundredth. The dealer chooses where on that trade-off to sit based on how much of the position they expect to hold to expiration.

And this generalizes to every discontinuous payoff in the chapter. A reverse knock-out is a digital in disguise—it pays a positive amount right up to the barrier and nothing beyond it—which is why Section 34.3 called it the most difficult vanilla-like instrument in the market. The overhedge for it is a barrier struck slightly further away, and the premium is correspondingly larger—much larger, as Example 34.6 shows.

Who trades digitals. Very little of the market trades a digital on its own. Most digital risk sits inside other products as a coupon trigger: each observation date on an autocallable pays its coupon only if the underlying is above a level, and a range accrual pays a day's interest only if a rate stays inside a band, so both are strips of digitals, which Chapter 35's investors are short and the issuing dealers long. The one active standalone market is in foreign exchange, where European digitals and their path-dependent relatives—the **one-touch**, which pays if the rate trades at a level at any time before expiration, and the **no-touch**, which pays if it never does—are standard dealer products. Macro hedge funds buy them to take a view on a level for a fraction of a vanilla's premium.

Professionals have little need for a listed digital, because a tight call spread already is one. Cboe listed binary options on the S&P 500 and the VIX in 2008, and they drew almost no volume. Binaries found their buyers among retail investors instead, sold by offshore platforms that were the counterparty to every trade, usually at maturities of minutes or hours. The European Securities and Markets Authority prohibited their sale to retail clients in 2018, and Israel closed its binary-options industry outright.

The growth since 2024 has come from **event contracts**: cash-or-nothing contracts on an inflation release, a central-bank decision or an election, listed on exchanges regulated by the Commodity Futures Trading Commission and traded offshore on prediction markets. They are digitals in the strict sense of this section, but on outcomes that no vanilla option spans. There is no call spread to overhedge them with, so Example 34.5's argument does not reach them: the price is a probability set by the order book, and the market maker carries the jump at the event rather than hedging it.

Example 34.6 The instrument that is short its own underlying

An up-and-out call struck at 100 with the barrier at 120, one year, 25% volatility, and $r = q = 0$: zero carry, so the reflection results of Section 34.5 apply exactly. Its value is the vanilla A less the up-and-in call $B - C + D$ from the third row of Section 34.3's table, by in-out parity. The vanilla is worth 9.9476; the knock-out is worth **0.6709**, 6.7% of it. Almost the whole premium has been given away, because the option is extinguished in precisely the states where a call is valuable.

Its delta is the striking part:

Spot	Value	Delta
100	0.6709	−0.0149
110	0.4081	−0.0359
115	0.2132	−0.0414
118	0.0863	−0.0430

The delta is negative from a spot of 94.3 upward. This is a call option whose holder loses money when the underlying rises, because approaching 120 removes more value through the knock-out than it adds through moneyness. Its holder hedges by buying stock, as though short a call—the reverse of every intuition built up in Chapter 18—and the sign flips back the moment the barrier is removed.

The overhedge is where the difficulty becomes a price. Dominating the payoff means selling a knock-out with the barrier further away, and the cost rises steeply:

Barrier used	Value	Cost over the contract
122	0.8947	+33.4%
124	1.1513	+71.6%
126	1.4377	+114.3%

Shifting the barrier by 1.7% of spot costs a third of the option’s value, against 7.6% for the digital’s much wider shift in Example 34.5. That ratio is the whole of why reverse knock-outs quote in percentage points while down-and-outs quote in basis points, and it is a property of the payoff rather than of any model: the cliff is 20 points high on an option worth 0.67.

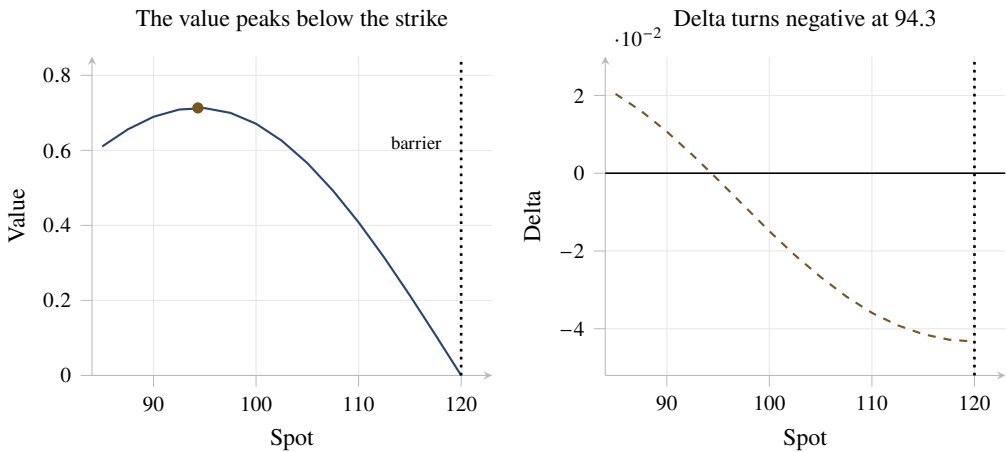


Figure 34.3: Example 34.6’s up-and-out call: $K = 100$, $H = 120$, $\sigma = 25\%$, $T = 1$, $r = q = 0$.

Figure 34.3 plots it. The value is a hump whose peak sits at 94.3, six points *below* the strike, because above that level a rise in spot removes more value by raising the probability of a knock-out than it adds by moving the call further into the money—and the delta in the right panel is simply the slope of the left one, crossing zero exactly at the peak. A call option with an interior maximum is not a call option in any sense Chapter 18 would recognize, and the hedging difficulty is that the maximum is inside the trading range rather than at the edge of it.

Figure 34.4 takes the same option with a few weeks to run rather than a year, and adds the gamma. The delta climbs to 0.67 by a spot of 104, turns down, crosses -1 at 115.7, and reaches -1.46 at the barrier, where a vanilla's delta is $+1$; with a year to run it never left $[-0.043, 0.020]$. The gamma is negative above 104.4, where a vanilla's is positive everywhere, and at -0.23 near a spot of 113 its magnitude is three times the largest gamma the vanilla ever has.

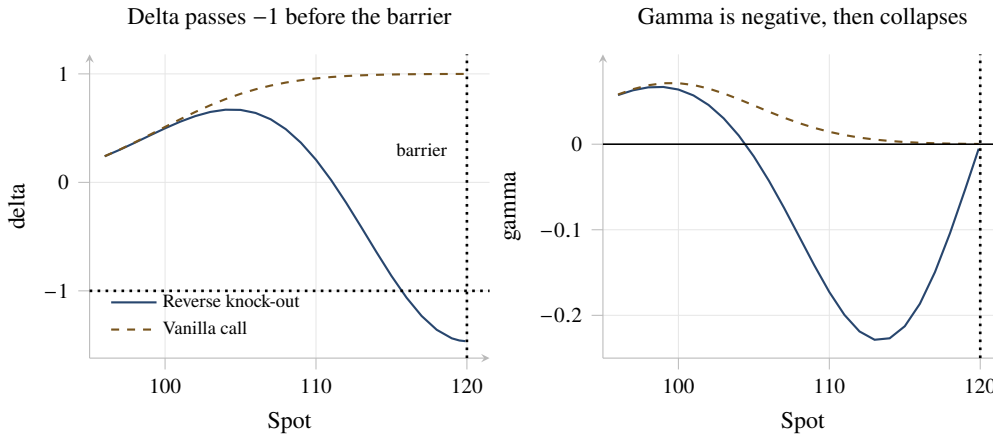


Figure 34.4: Up-and-out call against a vanilla call: $K = 100$, $H = 120$, $\sigma = 25\%$, $T = 0.05$, $r = q = 0$. The dotted vertical line is the barrier, the dotted horizontal one $\Delta = -1$.

The hedging problem is in the last few points of the right-hand panel. Approaching the barrier the value becomes nearly linear in the spot—0.146 at a spot of 119.9, against 0.146 for the straight line through the barrier—so the gamma collapses to -0.005 while the delta does not: it is still -1.46 . A hedge of 1.46 shares against an option worth 0.15 means a tenth of a point of spot moves the hedge by the option's entire value, and at the barrier the option ceases to exist, so that delta has to come off at once. Gamma, which normally warns a hedger that rebalancing is about to become expensive, is closest to zero exactly where the position is most dangerous.

34.9 LOOKBACKS AND ASIANS

This section compares two path-dependent payoffs, one cheaper than the vanilla and one dearer. Both have closed forms, which come first.

Two averages are in use. The traded contract takes the arithmetic average of its fixings, $\frac{1}{n} \sum_{i=1}^n S_{t_i}$, the discrete form of the A_T in (34.3). The **geometric average** multiplies the fixings and takes the n th root,

$$G_T = \left(\prod_{i=1}^n S_{t_i} \right)^{1/n} = \exp\left(\frac{1}{n} \sum_{i=1}^n \ln S_{t_i}\right),$$

which becomes $G_T = \exp\left(\frac{1}{T} \int_0^T \ln S_t dt\right)$ when the averaging is continuous, and which never exceeds the arithmetic average of the same prices. Its logarithm is an average of normal variables and is therefore normal, so the geometric-average Asian call, paying $(G_T - K)^+$, is exactly lognormal: it is Black–Scholes with two substitutions. Averaging continuously over

$[0, T]$,

$$c_G = S_0 e^{(b_G - r)T} N(d_1) - K e^{-rT} N(d_2), \quad \sigma_G = \frac{\sigma}{\sqrt{3}}, \quad b_G = \frac{1}{2} \left(r - q - \frac{\sigma^2}{6} \right) \quad (34.15)$$

with $d_1 = [\ln(S_0/K) + (b_G + \sigma_G^2/2)T] / (\sigma_G \sqrt{T})$ and $d_2 = d_1 - \sigma_G \sqrt{T}$. The $1/\sqrt{3}$ is Section 33.8's variance ratio; b_G is the drift the geometric average inherits, which is below $r - q$ because a geometric mean sits below an arithmetic one.

The arithmetic-average Asian—the one actually traded—has no closed form, because a sum of lognormals is not lognormal. Two devices cover the gap. Matching the average's first two moments to a lognormal and pricing with Black's formula gives 6.883 for the contract in Example 34.7, against 6.85 by simulation: about half a percent high, which is the size of error the whole approximation literature is arguing about, though it grows with volatility, to 1.4% at 50%. The arithmetic version is worth 5% more than the geometric one priced in that example, the gap between an arithmetic and a geometric mean, and that gap rises roughly in proportion to volatility, from 2% at a 10% volatility to 11% at 50%. And (34.15) serves as a control variate for simulation: the two payoffs correlate at 0.9994, and subtracting the geometric payoff one for one cuts the variance by a factor of 330, or about 850 with the regression-optimal coefficient. The instrument that is traded is the one without a formula, and it barely matters—which is the reverse of the situation for every barrier in this chapter.

The floating-strike lookback call paying $S_T - S_{\min}$, where $S_{\min}$ is the running minimum observed so far, is worth

$$c_{\text{lb}} = S_0 e^{-qT} N(a_1) - S_0 e^{-qT} \kappa N(-a_1) - S_{\min} e^{-rT} [N(a_2) - \kappa e^{Y_1} N(-a_3)], \quad \kappa = \frac{\sigma^2}{2(r - q)} \quad (34.16)$$

where

$$a_1 = \frac{\ln(S_0/S_{\min}) + (r - q + \sigma^2/2)T}{\sigma \sqrt{T}}, \quad a_3 = \frac{\ln(S_0/S_{\min}) + (-r + q + \sigma^2/2)T}{\sigma \sqrt{T}},$$

$a_2 = a_1 - \sigma \sqrt{T}$, and $Y_1 = -2(r - q - \sigma^2/2) \ln(S_0/S_{\min})/\sigma^2$. The floating-strike put, paying $S_{\max} - S_T$ with $S_{\max}$ the running maximum, is the mirror image: the maximum takes the place of the minimum and the drift terms change sign. The formula is singular at $r = q - \kappa$ diverges—though the value is not, the limit being finite; an implementation that does not special-case zero drift will return nonsense on precisely the parameter set Section 34.5 cares about.

Example 34.7 Averaging destroys optionality; looking back manufactures it

All at $S_0 = K = 100$, $\sigma = 25\%$, $r = 5\%$, $T = 1$:

	Value	Relative to vanilla
Vanilla at-the-money call	12.3360	1.00
Floating-strike lookback call	20.5522	1.67
Geometric-average Asian call	6.5298	0.53

Same underlying, same volatility, and a factor of three between the two exotics.

The Asian is cheap because an average has less variance than an endpoint. Section 33.8 computed the ratio: $(n+1)(2n+1)/6n^2$, tending to $1/3$, so the effective volatility is about $1/\sqrt{3}$ of the true one. An Asian option is a vanilla on a less volatile quantity, and that explains most of the discount: at $25\%/\sqrt{3} = 14.4\%$ an at-the-money call is worth

8.38, about two-thirds of the vanilla. The rest comes from the average's lower forward, the drift b_G of (34.15), which takes the ratio to 0.53.

The lookback is expensive because it removes the timing decision. A floating-strike lookback call pays $S_T - \min S$, so the holder buys at the best price achieved. Nothing is left to get wrong, which is worth 67% more than the vanilla.

Figure 34.5 repeats the comparison across volatilities, and it separates what is optionality from what is not. At zero volatility the vanilla and the lookback are worth the same 4.88: the path then rises steadily at the 5% rate, its minimum is the starting price, and both pay only the drift. The Asian starts lower, at 2.41, because the average of a rising path sits below its end, and that gap is carry rather than optionality. Above zero, each volatility point adds about 0.61 to the lookback, 0.38 to the vanilla and 0.18 to the Asian.

Three practical notes follow.

Asians dominate commercial usage and lookbacks barely exist. A corporate hedging a year of purchases has an average exposure and wants an Asian; no one has an economic exposure to the minimum of a price path. Lookbacks are taught more than they are traded.

The Asian's model risk is low and the lookback's is high. An average smooths the path, so an Asian is less sensitive to the path distribution than the vanilla is to the terminal one; a lookback is a bet on the extremes and is very sensitive to it.

And discrete monitoring matters in both, in opposite directions. A monthly-averaged Asian is worth more than a continuously-averaged one—fewer observations, more variance—while a daily-monitored lookback is worth less than a continuous one, because a discretely observed minimum is higher. Section 34.6's continuity correction is the same adjustment, and the sign follows from which way discretization moves the observed extreme.

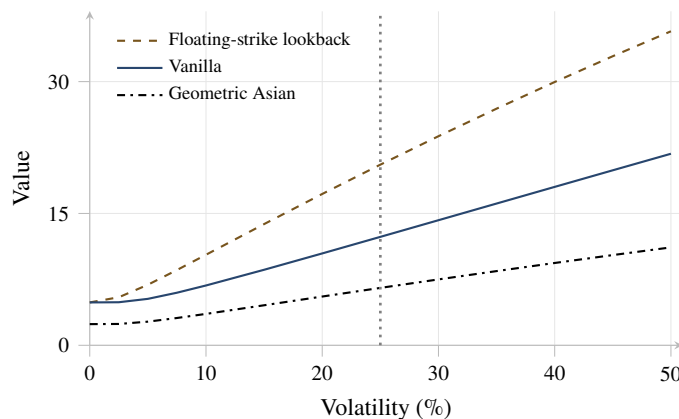


Figure 34.5: At-the-money calls against volatility: $S_0 = K = 100$, $r = 5\%$, $T = 1$, no dividends. The dotted line marks Example 34.7's 25%.

34.10 BASKETS, RAINBOWS, AND EXCHANGE OPTIONS

Multiple underlyings are the second source of genuine difficulty, and their new input is correlation.

Example 34.8 What correlation is worth

Two assets at 100 with volatilities 25% and 30%, $T = 1$, $r = 5\%$, no dividends.

A Margrabe exchange option—the right to give up asset 2 and receive asset 1—has the closed form derived in Section 28.11 as equation (28.14):

$$V_0 = S_1 e^{-q_1 T} N(d_1) - S_2 e^{-q_2 T} N(d_2), \quad \hat{\sigma} = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad (34.17)$$

with $d_1 = [\ln(S_1 e^{-q_1 T} / S_2 e^{-q_2 T}) + \hat{\sigma}^2 T / 2] / (\hat{\sigma} \sqrt{T})$ and $d_2 = d_1 - \hat{\sigma} \sqrt{T}$. Correlation enters only through $\hat{\sigma}$, so the table below is a table of Black–Scholes–Merton prices at five volatilities:

ρ	Effective volatility	Value
−0.50	47.70%	18.8495
0.00	39.05%	15.4808
+0.50	27.84%	11.0703
+0.90	13.23%	5.2737
+0.99	6.32%	2.5227

The value falls by a factor of seven across the correlation range. No interest rate appears, for the reason Section 28.11 gives: in units of asset 2 there is nothing to finance.

A *basket call* on $(S_1 + S_2)/2$ struck at 100, against a *basket of calls*, $(c_1 + c_2)/2$, half a 100-strike call on each asset:

ρ	Basket call	Basket of calls	Ratio
−0.50	8.3481	13.2836	0.628
0.00	10.3260	13.2836	0.777
+0.50	11.9088	13.2836	0.896
+0.90	13.0178	13.2836	0.980
+1.00	13.2792	13.2836	1.000

An option on a basket is always cheaper than a basket of options, with equality only at perfect correlation and equal volatilities—which is why the last row still shows a gap of 0.0044. The gap is the diversification the basket option does not let the holder keep: an option on each asset pays on whichever rises, while an option on the average pays only if the average rises.

And unlike the exchange option, the basket has no closed form, for the reason the arithmetic Asian does not: a weighted sum of lognormals is not lognormal. The workhorse approximation matches the first two moments of the basket $B = \sum_i w_i S_i$, with w_i units of asset i , to a lognormal,

$$M_1 = \sum_i w_i F_i, \quad M_2 = \sum_i \sum_j w_i w_j F_i F_j e^{\rho_{ij} \sigma_i \sigma_j T}, \quad \sigma_B = \sqrt{\frac{1}{T} \ln \frac{M_2}{M_1^2}} \quad (34.18)$$

where $F_i = S_i e^{(r-q_i)T}$, and then prices with Black's formula at forward M_1 and volatility σ_B . The approximation degrades exactly where the option is most interesting—against the exact values in the table it overstates by 0.06% at $\rho = 0.9$ and by 0.75% at $\rho = -0.5$,

because a negatively correlated basket is the least lognormal. The geometric basket $\prod_i S_i^{w_i}$, by contrast, is exactly lognormal and has a closed form, which makes it the natural control variate here just as (34.15) is for the Asian.

That gap is what Section 22.11's dispersion trade, short the index option and long the components, trades: it earns the difference between implied and realized correlation, and the ratio column shows how much is at stake.

The modeling difficulty is that ρ has to come from somewhere. With two assets and a liquid option on each plus a liquid basket option, it can be implied. With fifty assets it cannot, because one basket price cannot determine 1,225 pairwise correlations, and the standard device—assume a single correlation for all pairs—is a rank-one approximation whose error is exactly Section 32.9's factor-reduction problem in a different market.

34.11 WORST-OF OPTIONS

One multi-asset payoff is the single most common exotic sold to retail investors, and its correlation exposure runs the opposite way to a basket's.

With two assets it has a closed form and the derivation is one identity. Since $\min(S_1, S_2) = S_1 - \max(S_1 - S_2, 0)$, a forward on the minimum is a long position in asset 1 against a short exchange option (34.17), and the strike- K put on the minimum follows from

$$p_{\min}(K) = Ke^{-rT} - \underbrace{[S_1 e^{-q_1 T} - V_0^{\text{exch}}]}_{\text{PV of receiving } \min(S_1, S_2)} + c_{\min}(K) \quad (34.19)$$

Stulz's formula supplies the last term in bivariate normal probabilities:

$$c_{\min}(K) = S_1 e^{-q_1 T} M(y_1, -d_1; -\rho_1) + S_2 e^{-q_2 T} M(y_2, d_2; -\rho_2) - Ke^{-rT} M(y_1 - \sigma_1 \sqrt{T}, y_2 - \sigma_2 \sqrt{T}; \rho) \quad (34.20)$$

where $\hat{\sigma}$, d_1 , and d_2 are the exchange option's (34.17), and

$$y_i = \frac{\ln(S_i e^{-q_i T} / Ke^{-rT}) + \sigma_i^2 T / 2}{\sigma_i \sqrt{T}}, \quad \rho_1 = \frac{\sigma_1 - \rho \sigma_2}{\hat{\sigma}}, \quad \rho_2 = \frac{\sigma_2 - \rho \sigma_1}{\hat{\sigma}}$$

The exchange option appears twice—once as an instrument in (34.19) and once as the correlation structure inside M —which is the sense in which every two-asset payoff in this chapter is built from Margrabe.

Three assets is where it stops. The same construction extends to n assets—Johnson's formula—but the bivariate M becomes an n -variate normal distribution function, and quadrature for those is slow above three dimensions and unreliable above four. The three-asset worst-of that dominates the retail market is therefore simulated, which is why Example 34.9 reports a table rather than a formula, and why the note desk's risk numbers are Monte Carlo numbers with all the noise that implies.

Example 34.9 Why the coupon is high

Three assets, each at 100 with 28% volatility, $T = 1$, $r = 5\%$. A **worst-of put** struck at 100 pays on the *minimum* of the three. The benchmark is a vanilla put on any one of them, also struck at 100, which is worth 8.5955. The worst-of values are simulated, with 4 million antithetic pairs at each correlation; every standard error is below 0.005.

ρ	$E[\min S_T]$	Worst-of put	Relative to single-asset put
0.20	83.87	17.64	2.05
0.40	86.59	16.26	1.89
0.60	89.86	14.71	1.71
0.80	94.23	12.81	1.49
0.95	99.62	10.64	1.24

The worst-of put is worth more as correlation falls—exactly the opposite of the basket in Example 34.8, and for a reason that needs care.

A basket averages, so low correlation *reduces* dispersion and cheapens the option. A worst-of takes a minimum, so low correlation *increases* the chance that at least one asset falls a long way. At a correlation of 0.20 the expected minimum at one year is 83.87 against a starting level of 100—the drift is upward and the minimum still falls sixteen percent, purely from picking the loser three times.

A **worst-of note** is a bond whose principal is repaid in full only if the worst of the three assets finishes above the strike; otherwise the investor bears that asset's loss. The investor has therefore sold the issuer a worst-of put, and the issuer pays for it as a coupon above its own funding rate. At $\rho = 0.20$ the put is worth 17.64 per 100 of notional, roughly 18 percentage points of coupon on a one-year note, and twice what a note on any single one of the assets could pay.

So the investor in such a note is short correlation: the lower the correlation, the more the put is worth and the more the investor has given away. In a crash correlation rises, which on its own would lower the put's value, but it rises because everything is falling together, and with all three assets below the strike the put is deep in the money whatever the correlation. The investor therefore carries two losses for one coupon: the one that low correlation pays for, a single asset collapsing while the others hold up, and the market-wide fall, against which the rise in correlation offers no protection.

The practical warning is simpler than the analysis. Adding assets to a worst-of always raises the coupon and always raises the risk, and a note advertised as “linked to three blue-chip stocks” is riskier than one linked to any of them individually. Chapter 35 returns to this as the central example of the gap between how these products are described and what they are.

Figure 34.6 puts this example and Example 34.8 on one axis. Both ratios come to one at perfect correlation, where the assets move as one, and they leave it in opposite directions as correlation falls: the basket call loses value against its components, while the worst-of put gains it. A worst-of note's investor is short the upper curve, which is why the coupon is highest exactly where the assets are least alike.

34.12 CLIQUETS, AND THE CAP THAT REVERSES THE SIGN

Section 34.2 called a cliquet the purest instrument for trading forward volatility. The version actually sold is not, and the reason is instructive.

Example 34.10 What the cap does

A three-year annual-reset cliquet paying $\sum_i \max(r_i, 0)$, where r_i is the year's return, each period capped at 8%. Under Black–Scholes each period is a forward-start at-the-money call spread, so its value depends only on that year's forward volatility. By the scaling

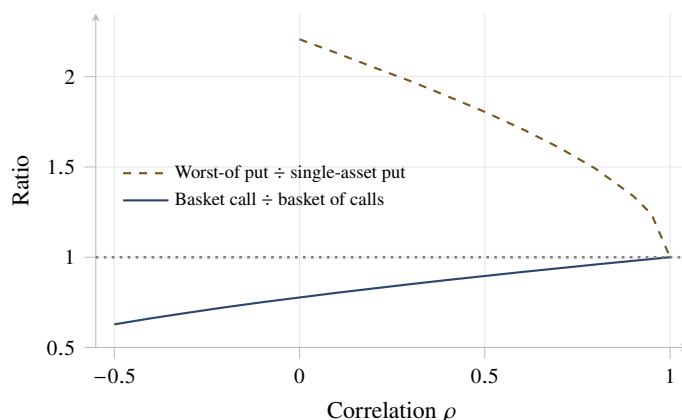


Figure 34.6: Examples 34.8 and 34.9 across correlation. The worst-of values are simulated.

in (34.9), the period payoff $\min(\max(r_i, 0), C)$ is worth, at its own reset date,

$$V = c(1, 1, \tau) - c(1, 1 + C, \tau) \quad (34.21)$$

per unit of notional, where $c(S, K, T)$ is the Black–Scholes–Merton call of Section 34.2, evaluated here at a spot of 1 and the period’s forward volatility, $C = 8\%$ is the cap, and $\tau = t_i - t_{i-1}$ is the length of a period, one year here. That is a number, not a function of the level, so the whole contract is $\sum_i e^{-rt_{i-1}} V$. The table reports the undiscounted sum $3V$, which isolates the volatility comparison; discounting the second and third periods scales the total by about 0.95.

Forward volatility	Per period	Three-year cliquet
15%	0.03800	0.11401
20%	0.03670	0.11009
25%	0.03561	0.10682
30%	0.03463	0.10388
35%	0.03372	0.10115

The value falls as volatility rises. An uncapped cliquet would be long volatility; the cap makes this one short it, because the 8% cap sits close to the one-year forward and higher volatility helps the short leg of the spread more than the long one. That reversal is what lets it be sold at an attractive headline rate: it is marketed on upside participation and is, in volatility terms, a *short*-volatility position.

And the model risk is severe for a reason Section 21.3 supplies. The value sits entirely in the forward smile: the second and third periods are forward-starting, so what matters is the smile prevailing a year and two years from now. Section 21.3 measured what local volatility predicts. The skew of its local-volatility function—the slope of volatility against log-moneyness at the money—decays from -0.981 at three months to -0.231 at three years, so the model expects the smile a few years out to be far flatter than today’s, while the market’s forward smile is roughly stationary. A local-volatility model prices the later periods off a smile it has flattened, and a cliquet’s value is concentrated exactly there.

This is the clearest case in the chapter where calibrating to today’s surface is not merely insufficient but actively misleading, because the model reproduces every vanilla

and gets the one quantity the product depends on structurally wrong.

Figure 34.7 shows the reversal against the contract the cap was applied to. The two slope in opposite directions, and an investor told they are buying participation in rising markets is holding the line that falls when volatility arrives. Both lines are close to straight because both options involved, the at-the-money call and the call at the cap, are struck near the forward, where an option's vega barely changes with volatility.

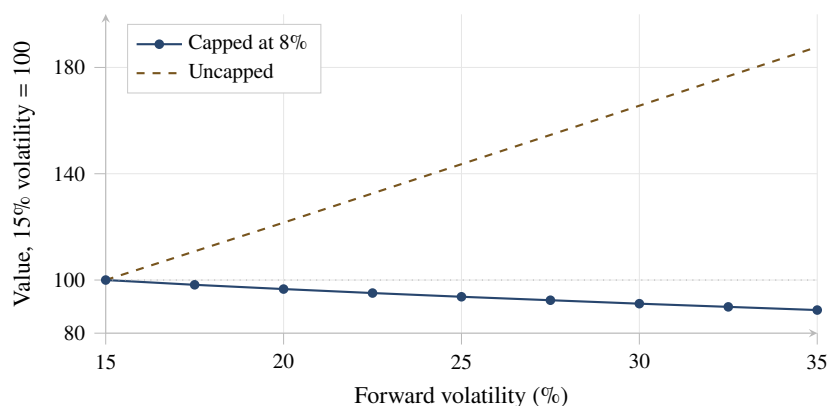


Figure 34.7: Three-year annual-reset cliquet of Example 34.10, each indexed to its own value at 15% volatility. In levels the capped contract falls from 0.11401 to 0.10115.

34.13 WHAT THEY HAVE IN COMMON

The chapter's material sorts on one question, put as a test rather than a list.

Can the payoff be replicated by traded instruments?

If it can be replicated statically—a package, a chooser, a digital as a limit of call spreads—then the vanilla market determines the price, model risk is small, and the residual is a hedging cost rather than a valuation uncertainty.

If it can only be replicated dynamically, then the price depends on the assumed dynamics, and calibration to vanillas constrains but does not determine it. Barriers, cliquets, lookbacks, and everything multi-asset live here.

The test is more useful than the taxonomy of Section 34.1 because it predicts the right thing. A compound option sounds harder than a barrier and is easier; a forward-start option sounds trivial and carries as much model risk as anything in the book.

And it is a different question from whether a formula exists, which is worth putting side by side because the two are so easily confused.

Instrument	Closed form under Black–Scholes	Model risk
Package, chooser	Exact, (34.8)	None
Compound	(34.7), plus one root search	Low
Digital	(34.13)	Low; the cost is hedging
Barrier, continuous	(34.11), all eight	Severe
Barrier, discrete	(34.11) with (34.12)	Severe
Lookback	(34.16)	High
Asian, geometric	(34.15)	Low
Asian, arithmetic	None ; moment match or simulate	Low
Exchange	(34.17)	Correlation only
Two-asset max/min	(34.20)	Correlation
Basket, worst-of ($n \geq 3$)	None ; simulate	Correlation, unhedgeable
Forward-start, cliquet	(34.9), (34.21)	Severe

The two columns are almost uncorrelated. Barriers, lookbacks, and cliquets all have closed forms and carry the worst model risk in the chapter; the arithmetic Asian has no formula at all and carries almost none. A closed form is a statement about Black–Scholes, not about the market: it returns the same number every time it is called, which is precision, not accuracy. The formulas are worth having because they are fast, exact within the model, and show where the model enters, not as evidence that the price is known.

And it explains a market structure fact that otherwise looks arbitrary. The exotics with static hedges trade at narrower spreads than those without, and the ordering follows the test rather than the complexity of the payoff description.

Case Study 34.1 The structure that sold a discount and bought a multiple

The **accumulator** was sold across private banks in Hong Kong and Singapore through 2007 and into 2008, and it is this chapter’s mechanism assembled into a single retail contract.

The terms sound like a discount. The client agrees to buy a fixed quantity of a stock on every trading day for a year, at a strike set a few percent below the spot price at inception. Nothing is paid at the outset. Two clauses complete it: the contract *knocks out* if the stock trades a few percent above the initial spot, and the daily quantity *doubles* if the stock is below the strike.

Decomposed, it is three of this chapter’s instruments. A strip of daily forwards at a discount; short a strip of up-and-out calls, which is what the knock-out clause is; and short a leveraged strip of puts, which is what the doubling clause is. The discount is the premium of the options the client has sold, and the client is short volatility, short the downside at twice the size, and capped on the upside at a few percent.

What happened is what the structure specifies. Hong Kong and mainland equities peaked in October 2007 and fell through 2008. Because prices fell, the knock-out never triggered, so the contracts ran their full term rather than terminating early as they did in the case the salesman described. The doubling clause then required the client to buy twice the daily quantity, every day, at a strike far above the market, with the obligation growing precisely as the losses grew. The instrument acquired a nickname that was a pun on its own name: *I kill you later*.

The failure was not mathematical. Nothing here needed a model the seller did not have. The knock-out was disclosed, the doubling was disclosed, and both behaved exactly as written. What was not conveyed is the asymmetry the two clauses create together — a payoff described by its best state and sized by its worst.

Three lessons connect it to the rest of the book. A barrier that removes the upside is worth more than it looks when it is one of two clauses rather than the only one. A zero-premium structure has a premium; Chapter 35's treatment of a note's estimated initial value gives the machinery for finding it, and here it is the discount itself. And leverage written into a contract's quantity is harder to see than leverage written into its notional, because nothing in the term sheet is labeled with a multiple.

34.14 WHO TRADES EXOTICS, AND WHY

Corporate hedgers are the economically motivated end of the market. An Asian option matches an averaged exposure and a barrier matches a hedge that is only wanted in some states—an importer who needs FX protection but not if the currency has already collapsed past the point where the business is unviable. These are exotics bought because the vanilla is the wrong shape.

Structured-note issuers are the volume. Chapter 35 is about them, and almost every payoff in this chapter appears inside a retail note: barriers as knock-in protection, digitals as coupon triggers, baskets as “worst-of” underlyings, cliquets as annual ratchets.

Dealers intermediate, and the pattern is one-directional. The client is almost always short the exotic optionality and the dealer long it, because the exotic feature is what cheapens the note or the hedge. So dealers run a structural long-exotic, short-vanilla book, which they hedge in the vanilla market—and Section 19.11's skew and Section 22.11's implied correlation both reflect that flow.

Hedge funds take the other side selectively, and the trades named in this book—dispersion in Section 22.11, forward-volatility positions against cliquets—are mostly bets that the dealer's hedging flow has pushed a price away from fair.

The recurring structure is worth naming once. An exotic option transfers a risk the client does not want to price to a dealer who does, and the dealer's compensation is the difference between what the exotic can be hedged for and what it was sold for. Sections 34.5 and 34.13 determine how large that difference has to be, which is the practical reason the test in this chapter matters.

SUMMARY

Exotics divide by path dependence, multiple underlyings, and non-standard exercise, but the useful division asks whether the payoff can be replicated by traded instruments. Statically replicable payoffs are priced by the vanilla market with small model risk; dynamically replicable ones depend on assumed dynamics, which calibration to vanillas constrains without determining.

A down-and-out call with $H = 90$ is worth **9.111** against a 12.336 vanilla, and a reverse knock-out loses value as the underlying rises toward its barrier. Barriers are hard to compute as well as to hedge: a CRR tree gives 9.142 at 50 steps and **10.494** at 51 unless a node layer sits on the barrier, and correcting for daily rather than continuous monitoring is worth +4.57%. The reflection hedge is exact only when $r = q$ and errs by -42.3% at $r = 5\%$, and two models calibrated to the same vanilla surface priced one barrier **49.5%** apart.

A digital is the limit of call spreads whose value converges while the position diverges, so dealers overhedge with a spread that dominates the payoff, at **7.6%** over the digital for a five-point window. Every discontinuous payoff in the chapter is priced this way.

A lookback is worth **1.67** times the vanilla and a geometric Asian **0.53** times: an Asian is a vanilla on a less volatile quantity, and a lookback removes the timing decision.

A worst-of put on three assets is worth **2.05 times** a single-asset put at $\rho = 0.20$, because a

minimum rewards dispersion where a basket's average punishes it, and that premium is what pays a worst-of note's coupon. A capped cliquet is short forward volatility, falling from 0.11401 at 15% to 0.10115 at 35%, and its value sits in the forward smile, which today's surface does not pin down.

A closed form is a statement about Black–Scholes, not about the market. Barriers, lookbacks and cliquets have exact formulas and the worst model risk in the chapter; the arithmetic Asian has no formula and almost none.

Corporates buy exotics because the vanilla is the wrong shape, and structured-note issuers supply the volume. Clients are almost always short the exotic optionality, so dealers run a long-exotic, short-vanilla book hedged in the vanilla market.

Companion notebook `ch34-exotic-options.ipynb`

FURTHER READING

Zhang's *Exotic Options* and Haug's *Complete Guide to Option Pricing Formulas* are the reference works for closed forms; Haug states each formula with a worked number, which makes it the place to check an implementation against, and both should be read knowing that a closed form exists only under assumptions the market violates. Derman, Ergener and Kani on static replication of barriers is the paper behind Section 34.5 and develops the strip-of-maturities refinement this chapter only mentions. Carr and Chou extend the reflection argument considerably further. Boyle and Lau (1994) is the source of Example 34.2's node-placement rule. Broadie, Glasserman and Kou on discrete barrier monitoring supplies the continuity correction of Section 34.6. On the multi-asset material, Margrabe's original paper is short and is the source of the exchange option; Stulz (1982) is the source of (34.20), and states put prices through the identity (34.19); and Johnson (1987) extends the result to n assets. The literature on basket approximations is large, and its methods differ mainly in accuracy, so decide what accuracy is needed before sampling it. For the model-risk question of Section 34.7, the practitioner literature on barrier reserves is thin and mostly internal; the academic treatment under uncertain volatility is the public substitute.

PRACTICE QUESTIONS

- 34.1** State the replication test that organizes this chapter, and use it to explain why a compound option carries less model risk than a forward-start.
- 34.2** Using Example 34.1, state the fraction of the vanilla premium the barrier removes, and explain why knock-out deltas can exceed one.
- 34.3** Using Example 34.2, state the tree prices at 50 and 51 steps, explain the difference, and state the general lesson in a form that does not mention barriers.
- 34.4** Explain why grid-averaging and Richardson extrapolation both make the barrier tree error worse rather than better.
- 34.5** Derive the step count that places a node layer on the barrier, and state the error it achieves against an unaligned grid with more steps.
- 34.6** Using Example 34.3, state the static hedge error when $r = 5\%$ and $q = 0$, and explain why the condition for exactness is zero drift rather than zero rates.
- 34.7** Explain why a static hedge that is wrong by a known amount can be preferable to a dynamic hedge that is right in a model.

- 34.8** Explain why two models calibrated to the same vanilla surface can price a barrier 49.5% apart, and state the three defensible responses.
- 34.9** Using Example 34.5, state the cost of the [95, 100] overhedge relative to the digital, and explain why the replicating call spread's notional diverges while its value converges.
- 34.10** Explain why an Asian option is worth about half a vanilla and a lookback about 1.67 times it, and why Asians dominate commercial usage while lookbacks do not.
- 34.11** Using Example 34.4, state the value of daily and of monthly monitoring relative to continuous, and explain why the continuity correction and node placement compound rather than offset.
- 34.12** Using Example 34.9, state the worst-of put's value relative to a single-asset put at $\rho = 0.20$, and explain why the correlation dependence runs opposite to a basket's.
- 34.13** Using Example 34.8, state the basket call's value relative to the basket of calls at $\rho = -0.5$ and at $\rho = 1$, explain what the gap represents, and explain why fifty assets cannot be handled with a single correlation.
- 34.14** Explain why a capped cliquet is short volatility, why its value sits entirely in the forward smile, and why that makes calibration to today's surface actively misleading rather than merely insufficient.
- 34.15** Explain why dealers run a structural long-exotic, short-vanilla book, and why exotics with static hedges trade at narrower spreads than others.
- 34.16** Using (34.11), price the down-and-in call of Example 34.1 from the single expression C , and explain why the up-and-in call with $H < K$ is worth the vanilla.
- 34.17** Name three instruments in this chapter with closed forms and severe model risk, and one with no closed form and very little. Explain why the two properties are nearly unrelated.

FURTHER QUESTIONS

- 34.18** Set $r = q$ in (34.11) and show that the down-and-in call's C is exactly K/H puts struck at H^2/K . Then show where the argument fails when $r \neq q$, and relate the failure to the exponents 2λ and $2\lambda - 2$.
- 34.19** Example 34.2 shows a thousandfold error reduction from node placement. Derive the effective barrier as a function of n , show that the resulting price error is first-order in the barrier displacement, and use that to explain the sawtooth's shape between jumps.
- 34.20** Derive the reflection result of Example 34.3 from the symmetry of the lognormal density under $r = q$. Then determine how the static hedge error scales with $(r - q)T$, and design a multi-maturity strip that reduces it, stating what it costs.
- 34.21** Section 34.7 reports a 49.5% model spread on a barrier. Propose a reserving policy: specify which models define the range, how the reserve is released over the trade's life, and how you would prevent the policy from being gamed by choosing convenient models.
- 34.22** Example 34.5 shows the overhedge premium rising with the spread's width. Derive the dealer's optimal width as a trade-off between the premium given up and the expected hedging loss, and state how it depends on the expected holding period.

- 34.23** Example 34.8 shows a basket option below a basket of options at all $\rho < 1$. Prove the inequality in general, determine the conditions for equality, and connect the result to the convexity argument of Section 11.3.
- 34.24** This chapter argues that spreads follow replicability rather than payoff complexity. Test the claim against the exotics in this chapter: rank them by replicability, rank them by the spreads you would expect, and identify any instrument where the two orderings disagree—then explain the disagreement.

Chapter 35

Structured Products and Autocallables

This is where most people meet derivatives. Far more households own a structured note than have ever traded an option, and the note is sold by a bank branch or a private banker rather than by a broker.

Chapter 12 said structured products are packaged strategies and deferred the packaging. Chapter 34 priced the components. This chapter puts them together and asks the question the previous two could not: what does the buyer actually own, and is the price fair? Section 35.10 answers with a number—a \$1,000 note worth \$978 the day it is sold, the \$22 sitting inside the coupon rather than beside it.

The chapter is deliberately blunt about a conflict. Every instrument here is legitimate and most serve a real purpose. But the products are sold on a headline number—a participation rate, a coupon—and the risk sits in terms that do not appear in the headline. That gap is the subject.

Section 35.1 takes the common notes apart into a bond and options, and Sections 35.2 and 35.3 price the two that dominate, the reverse convertible and the autocallable, and show what the autocallable's return distribution looks like. Sections 35.4 and 35.5 cover the secondary market and the remaining families. Sections 35.6 to 35.8 turn to the issuer's side: how the embedded option is hedged, what happens when many notes share a barrier, and why most notes reference the worst of several underlyings. Sections 35.9 to 35.11 are about what the buyer pays for—the issuer's credit and the fee—and give a procedure for evaluating a note; Section 35.12 asks what disclosure can and cannot fix, and Section 35.13 who buys these products and why.

35.1 TAKING A NOTE APART

Every structured note decomposes into a bond and a derivative. Doing the decomposition is the single most useful skill in this chapter, because it converts a marketing document into a portfolio.

The decompositions are short, and the rest of the chapter refers back to them. Write F for the face amount, S_0 and S_T for the underlying at issue and at maturity, and $N = F/S_0$ for the number of option units a note of face F carries. In every case the redemption is a bond plus a position in options, and the only question is which options and on which side.

Principal-protected note. The buyer is promised the full face back at maturity whatever the underlying does, plus a share p —the participation rate—of any rise in the underlying over the note’s life. A note with 90% participation on an index that rises 20% repays 118 per 100 of face; if the index falls, it repays 100. At maturity the redemption is

$$V_T = F + p N (S_T - S_0)^+ \quad (35.1)$$

a zero-coupon bond plus pN calls struck at the initial level. The bond secures the guarantee, and whatever the bond leaves over buys the calls—which is why the participation rate is a residual rather than a promise, and why Example 35.1 finds it governed by the issuer’s funding cost.

Capped participation. The same promise, except that the share of the rise is paid only up to a ceiling: above a level C the redemption stops growing. In exchange the participation below the cap is higher, often above 100%. Capping the upside at C turns the call into a call spread:

$$V_T = F + p N [(S_T - S_0)^+ - (S_T - C)^+] \quad (35.2)$$

The sold call is what pays for the higher headline participation. A note advertising a participation above 100% has almost always sold something, and this is usually what.

Reverse convertible. The buyer receives a fixed coupon c well above deposit rates. At maturity the face is repaid in cash if the underlying stock or index ends at or above a barrier H ; if it ends below, the buyer receives shares, or their cash value, instead and so takes the whole fall from the initial level:

$$V_T = F(1 + c) - N (S_0 - S_T)^+ \mathbf{1}_{\{S_T < H\}} \quad (35.3)$$

a bond and a coupon, less a put the investor has written: struck at the initial level, but paying only when the underlying ends below H . Setting $H = S_0$ deletes the indicator and leaves a plain short put, which is the un-barriered version of the product. *The indicator is also the discontinuity*—the note pays $F(1 + c)$ just above H and $F(1 + c) - N(S_0 - H)$ just below it—and Section 35.6 is about what that does to the hedge.

Shark fin. Principal is protected and the buyer shares in the rise, usually at a high participation rate, but only as long as the underlying never trades above an upper level H during the note’s life. If it ever does, the participation is lost and the buyer receives a small fixed rebate R instead. The name comes from the shape of the payoff, drawn in the bottom-right panel of Figure 35.1: flat at par below the initial level, rising steeply along the participation line up to H , then dropping back to par plus the rebate—the outline of a shark’s dorsal fin. At maturity

$$V_T = F + p N (S_T - S_0)^+ \mathbf{1}_{\{\max_{t \leq T} S_t < H\}} + R \mathbf{1}_{\{\max_{t \leq T} S_t \geq H\}} \quad (35.4)$$

a bond, an up-and-out call, and the rebate. The knock-out is cheaper than the vanilla call of (35.1), so the same budget buys more of it: that, and not a better view, is where a headline participation above 100% comes from. Shark fins come in two forms. Some observe the barrier only at maturity, so the payoff depends on S_T alone and Figure 35.1 draws it exactly; others observe it daily or continuously, which is what (35.4) writes and Example 35.4 prices, and then a path that touches H and falls back is knocked out too.

Autocallable. The buyer receives a high coupon, but the note can end early. On each of a series of observation dates, if the underlying is at or above a trigger level, the note is *called*: it repays the face plus all the coupons accrued so far, and stops. If it is never called, it repays the face with all the coupons at maturity unless the underlying ends below a barrier, in which case the buyer is repaid in shares, takes the loss as in a reverse convertible, and forfeits the coupons. In symbols, take observation dates $t_1 < \dots < t_n = T$, an autocall trigger K_a , a coupon c per annum with each t_i measured in years, and a barrier H , observed here only at maturity for simplicity; in practice it is usually observed daily, as in Example 35.2. Let τ be the first date on which $S_{t_i} \geq K_a$, and $\tau = \infty$ if that never happens. Then

$$V = \sum_{i=1}^n F(1 + c t_i) \mathbf{1}_{\{\tau=i\}} + \left[F(1 + cT) \mathbf{1}_{\{S_T \geq H\}} + NS_T \mathbf{1}_{\{S_T < H\}} \right] \mathbf{1}_{\{\tau=\infty\}} \quad (35.5)$$

a strip of digital payments, each conditional on not having been called already, plus, if the note is never called, the written put of (35.3): below the barrier the buyer receives shares worth NS_T and forfeits the coupons. Two things follow that the term sheet does not say. *The investor is short volatility twice*: once in the digitals, which want the underlying quietly above K_a , and once in the put. And a reverse convertible is the case $n = 1$ with no trigger. The two share their downside, a written put that pays only below the barrier, and are sold to the same buyer on the same high coupon, but their risk profiles differ: early redemption cuts the autocallable's expected life well below its stated term, so the buyer collects the coupon for less time while carrying the same tail, and the outcome depends on the path as well as on where the underlying ends.

Dual-currency deposit. The buyer places a deposit in one currency and earns a rate above the ordinary deposit rate. At maturity the bank chooses the currency of repayment—the original one, or a second currency at a conversion rate fixed at the outset—and it will choose whichever is worth less to the depositor. Write X_T for the exchange rate at maturity, in units of the second currency per unit of the deposit currency, K for the conversion rate, and c for the interest, paid in the deposit currency whatever happens. Measured in the deposit currency, the redemption is

$$V_T = F(1 + c) - \frac{F}{X_T} (X_T - K)^+ \quad (35.6)$$

a deposit less a written call on the deposit currency struck at the conversion rate. When the deposit currency strengthens past K , the principal comes back converted at K and is worth only FK/X_T . Example 35.5 carries the arithmetic, and the enhancement is the option premium, as it is in every structure above.

One identity organizes all of them. At issue,

$$\text{issue price} = \underbrace{\text{PV(principal)} + \text{PV(coupons)}}_{\text{the bond}} + \underbrace{\text{options bought} - \text{options sold}}_{\text{the derivative}} + \text{fee} \quad (35.7)$$

Figure 35.1 draws the first four on one scale, and the shapes carry the argument. In each panel the dashed line is what the same 100 would return left on deposit, at 4% for a year, so the gap between it and the solid line is what the embedded options add or take away against simply earning interest. The principal-protected note repays 100 where the deposit pays 104, which is the interest given up to buy the participation; the reverse convertible pays 110 above its barrier, the extra 6 being the premium for the put the buyer has written, and falls far below the deposit once the barrier is breached. Where each solid line crosses the dashed one marks the range of

the underlying in which the buyer comes out ahead: above 104.4 for the principal-protected note, above 104 for the capped note, anywhere at or above the barrier of 70 for the reverse convertible, and only between 103.6 and the knock-out at 130 for the shark fin. Each note is a bet that the underlying ends in its own range, and the reverse convertible's range is the widest because its losses, when they come, are the largest. *Every term but the fee is quotable*, and that is what makes the decomposition worth doing: price the bond at the issuer's own funding curve, price the options in the market, and the residual is what the note costs to own.

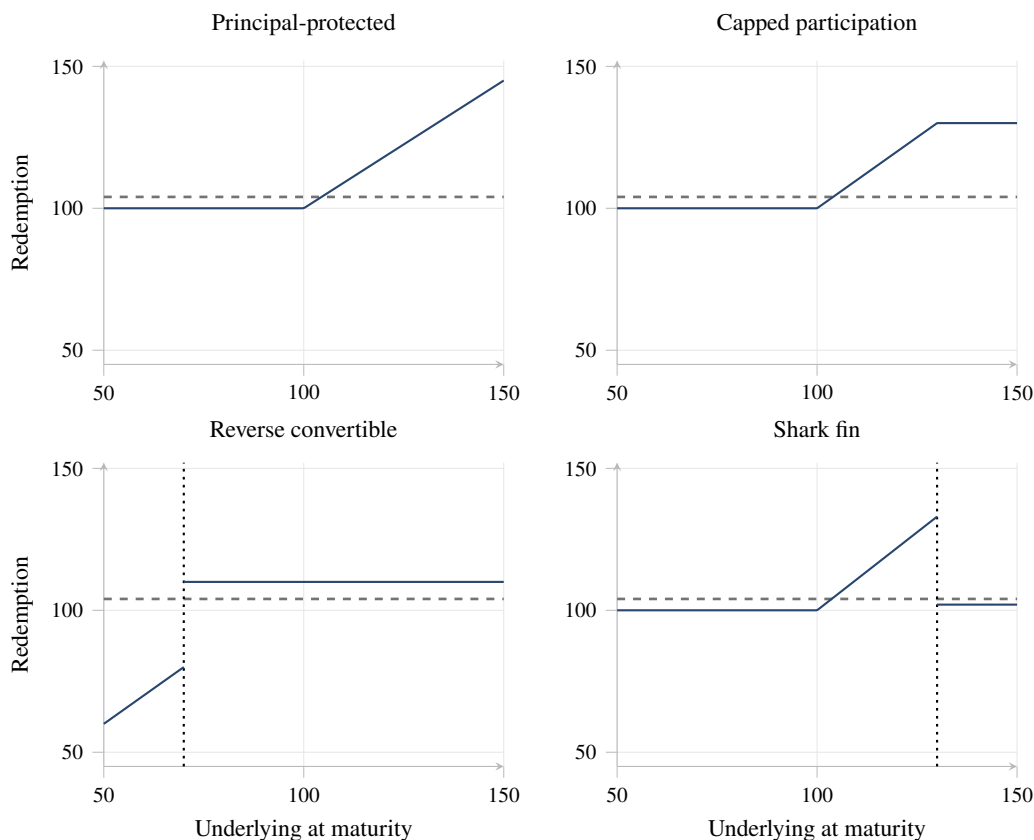


Figure 35.1: Redemption per 100 of face, initial level 100. Participation 90%, cap 130, barrier 70 with a 10% coupon, knock-out 130 with a 2% rebate. Dashed: 100 left on deposit at 4% for a year. Illustrative.

Example 35.1 Where the participation rate comes from

A five-year principal-protected note, \$1,000 face, linked to an equity index. The issuer funds at 4.5%, the index yields 2.0%, and its volatility is 20%. The issuer's 4.5% discounts both the bond and the option.

Zero-coupon bond to guarantee the \$1,000	798.52
Left to buy optionality	201.48
Cost of a 100% participation at-the-money call	209.60
Participation rate	96.1%

Now add what the brochure does not itemize—a 2% distribution fee:

Optionality budget after fees	181.48
Participation rate	86.6%

Ten points of participation, gone, and it appears nowhere in the term sheet.

Now the observation that matters most:

Issuer funding rate	Participation (no fee)	(2% fee)
2.0%	59.4%	46.9%
3.0%	77.7%	66.6%
4.5%	96.1%	86.6%
6.0%	107.3%	99.0%

The participation rate is mostly a statement about the issuer's funding cost, not about the index. It nearly doubles as funding goes from 2% to 6%, with the index unchanged.

Two consequences follow, and the second is uncomfortable.

A note looks better when rates are high, which is why principal-protected structures nearly vanished in the zero-rate years and returned after 2022. The product's attractiveness is a term-structure phenomenon, not an investment insight.

And across issuers at one moment, a higher participation rate signals a weaker credit. The issuer paying more to fund can buy more optionality. So the best-looking note in a comparison is systematically the one with the most credit risk, and Section 35.9 says what that has cost.

35.2 REVERSE CONVERTIBLES AND AUTOCALLABLES

The reverse convertible comes first, because an autocallable is a reverse convertible with early redemption added. The investor in a reverse convertible is short the put of (35.3), and the coupon is the premium: a note paying 10% when deposits pay 4% is not offering 6% of extra yield; it is paying 6% for a put the investor has written. The mis-selling risk is structural rather than incidental. The product is described in the vocabulary of fixed income—coupon, maturity, principal—and its risk is entirely equity, so a buyer who compares it with a bond is comparing the wrong thing; the right comparison is with selling puts, which the same buyer would very likely decline.

The autocallable is the dominant structured product globally, and the rest of this chapter is built on it.

The mechanics are simple to state and hard to price. The note pays a coupon and terminates early—"autocalls"—if the underlying is above a level on an observation date. If it never autocalls, principal is returned at maturity together with the accrued coupons *unless* the underlying has breached a lower knock-in barrier, in which case the investor takes the loss and forfeits the coupons.

Figure 35.2 is (35.5) drawn. The note is a sequence of chances to be redeemed early, each paying the coupons banked so far, and only the last one can end in shares.

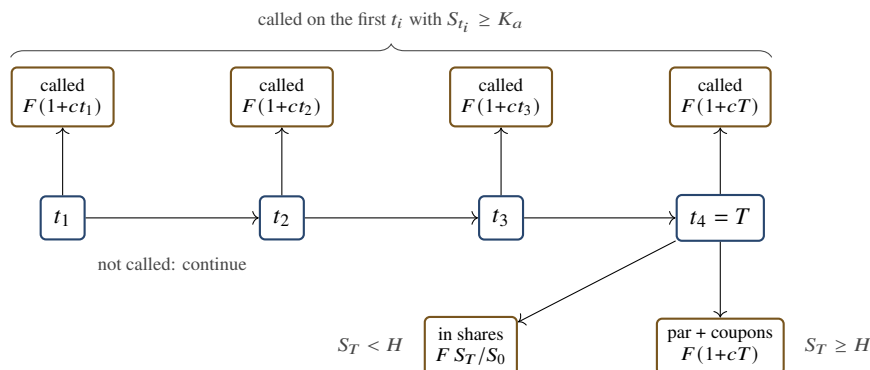


Figure 35.2: Each observation either redeems the note with the coupons accumulated to that point or passes it on. Only the final date can end below the barrier.

Example 35.2 Pricing one, and what the buyer owns

A two-year note on an index at 100. Autocall at 100% observed monthly after a two-month lockout; knock-in at 75% observed daily. The coupon accrues at an annual rate c : a note called at the end of month m repays $1 + c m/12$, and one that is never called and never knocked in repays $1 + 2c$ at maturity. If knocked in and never autocalled, the investor receives $\min(S_T/100, 1)$ per unit of face and no coupon. Volatility 25%, rate 2.5%; 400,000 antithetic paths. The coupon c is the unknown: it is set so that the note is worth exactly its issue price of 1.00, which makes it the *fair coupon*.

$P(\text{autocalls at some point})$	0.8296
$P(\text{knocks in at some point})$	0.4136
$P(\text{knocked in and never autocalled})$	0.1660
Fair coupon	18.63% per annum

An 18.6% coupon, and the note prices at par. Nothing is being given away: the coupon is exactly what the embedded short put is worth.

The probability structure is what sells it and what should worry the buyer. The note produces its advertised return about 83% of the time. That is a genuinely high win rate, and it is why these products are easy to sell and why holders report satisfaction for years at a stretch.

The remaining 17% is where all the loss is, and it is not a small loss—a knocked-in note pays $\min(S_T/100, 1)$, so the investor absorbs the full downside from par.

This is a negatively skewed payoff sold on its mean and its hit rate, which is the same structure as Section 22.8's inverse volatility products and Section 33.6's range accrual, and it recurs because it is what sells.

Now the exposure the buyer does not see:

Realized volatility	Note value at $c = 18.63\%$
15%	1.0532
20%	1.0210
25%	1.0005
30%	0.9856
40%	0.9579

The note is short volatility. It was priced at 25%; realized volatility above that makes it worth less than par, and the investor never agreed to a volatility view.

And it is short volatility in the worst way, because volatility rises when markets fall, which is exactly when the knock-in is approached. *The two risks are not independent*—this is Section 19.9's leverage effect and Section 27.9's wrong-way structure appearing in a retail product.

35.3 WHAT THE RETURN DISTRIBUTION LOOKS LIKE

A fair price says the mean is right. *It says nothing about the shape*, and the shape is what the buyer is actually choosing.

Example 35.3 The same expected return, a different distribution

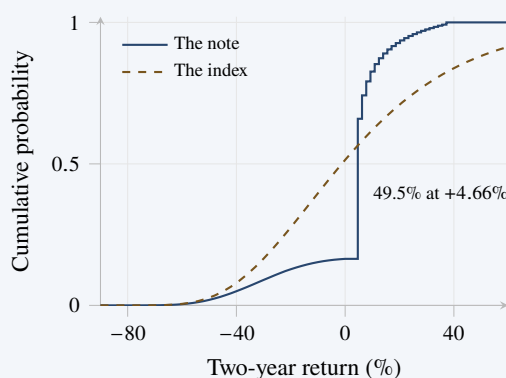
Simulate Example 35.2's note at its fair 18.63% coupon, 600,000 paths, and record the investor's total return—then do the same for simply holding the index for two years.

	The note	The index
Mean total return	+1.69%	+5.11%
Standard deviation	17.29%	38.30%
Skewness	−1.572	+1.149
Excess kurtosis	2.819	—
<i>P</i> (positive)	83.56%	48.60%
<i>P</i> (loss > 25%)	10.98%	—
5th percentile	−40.00%	−44.76%

Everything here is under the pricing measure, so both instruments earn the risk-free rate by construction. The note's lower mean is not a worse deal—it terminates early, after 0.70 years on average, so it earns that rate over a shorter life. $2.5\% \times 0.70 = 1.75\%$, against the simulated 1.69%, which is the internal check that the pricing is consistent.

With the mean neutralized, the comparison is purely about shape, and the shape is the product.

Median return	+4.66%
<i>E</i> [return positive]	+8.32%
<i>E</i> [return negative]	−32.00%
Average of the worst 1%	−60.19%



Wins average 8% and losses average 32%—a four-to-one asymmetry, at a five-to-one win rate. Those two ratios are the product, and they multiply out to a fair price.

Three observations follow.

The note's skewness is negative and the index's is positive. A buyer who moves from the index into the note has not reduced risk; they have exchanged a positively skewed

exposure for a negatively skewed one at the same expected return, which is the opposite of what “capital-efficient” and “defensive” suggest.

The 25th and 50th percentiles are identical, at +4.66%. Nearly half the distribution, 49.5%, is a single point—the note autocalled at the first opportunity—so every percentile from the 17th to the 65th is +4.66%, which is why holders experience these products as reliable. Most of the time the outcome is not merely good but identical, and repetition of an identical good outcome is unusually persuasive. The plot beside the second table draws both distributions cumulatively: the note’s single point is the 49.5-point vertical step at +4.66%. Below zero the note’s curve lies under the index’s, because a loss is far less likely—16.4% against 51.4%—but the note can never pay more than +37.26%, the level by which the index has reached only its 82nd percentile.

And the standard deviation is genuinely lower than the index’s, 17.3% against 38.3%. A risk measure that sees only the second moment will report this product as safer, which it is on that measure and is not on any measure that looks at the tail. Chapter 38’s critique of variance as a risk measure is not abstract; *this is the product it was written about*.

35.4 THE SECONDARY MARKET

A structured note is quoted daily and is not liquid, and the distinction matters more than it sounds.

The issuer is usually the only bid. There is no exchange, no competing market maker, and no obligation to quote. A buyer who wants out sells back to the party who sold it to them, at a price that party computes.

Three consequences follow.

The bid embeds a second spread. Having paid the issue premium of Section 35.10, an early seller pays again on the way out. Round-tripping a note in the first year can cost several percent of principal before any market move.

The quoted price is a model output. These are Level 2 or Level 3 valuations in the accounting sense, and Section 26.12’s observation about the crisis applies in miniature: when the only price comes from the party on the other side, the holder cannot distinguish a mark from a quote.

And illiquidity interacts with the payoff. An investor most wants to exit after a sharp fall—when the knock-in is close and the outlook has changed—and that is exactly when the bid is lowest and the embedded option’s value has moved most against them. *The exit is worst when it is most wanted*, which is this book’s recurring wrong-way structure appearing one last time.

Case Study 35.1 A market that learned this from its losses

South Korea developed the world’s largest retail market for autocallable notes, sold as **equity-linked securities** through bank branches to ordinary savers. A large share referenced the Hang Seng China Enterprises Index (HSCEI), chosen because its high volatility supported high coupons.

The structure worked as designed for years. Notes autocalled, holders were repaid early with the advertised coupon, and the proceeds were rolled into new notes—the 83% outcome of Example 35.3, repeated often enough to look like a rule rather than a probability.

Then the index fell a long way and stayed there. Notes issued in 2021 carried knock-in barriers set against 2021 levels; by the time those notes reached maturity in 2024, the index had not recovered, and the loss that Example 35.2 gives a 17% probability became

the realized outcome for a large cohort at once.

Four features of the episode are general and worth separating from the specifics.

The losses were synchronized because the issuance was. Notes sold during one period share a vintage, a barrier level, and a maturity. A product with a 17% loss probability sold to a hundred thousand people in the same quarter does not produce 17% of them losing independently; it produces all or none.

The distribution channel mattered as much as the product. The notes were sold through bank branches to savers, alongside deposits, by staff whose training and incentives were not designed for a negatively skewed derivative.

The regulatory response focused on sales practice rather than on the instrument—suitability assessment, recording requirements, limits on where such products may be sold—which is the right diagnosis if one accepts that the instrument is legitimate and the mismatch was between the payoff and the buyer.

And the mechanism was entirely visible in advance. Everything in Sections 35.2 through 35.3 could have been computed at issue from the term sheet. *Nothing about this episode required hindsight*, which is what distinguishes it from most of the case studies in this book.

35.5 OTHER FAMILIES

Four more structures account for most of the rest of the market, and each is a recombination of what has already been priced.

Shark-fin notes pay an enhanced participation up to a barrier and a small fixed rebate if the barrier is touched—an up-and-out call plus a rebate. *The buyer is selling the tail of the upside*, and the appeal is that giving up the potential of returns above 30% feels costless to someone who expects 10%. Example 35.4 prices what the feeling costs.

Example 35.4 Where the headline participation comes from

A one-year \$1,000 note at a 3% funding rate, on an index at 100 with 25% volatility. The zero-coupon bond costs 970.45, leaving **29.55** for the option and its rebate.

An at-the-money call is worth 11.3485 per 100 of index, so that budget buys **26.0%** participation—an unattractive number to print on a term sheet. Adding a knock-out at 130 drops the option to 2.1766, but a shark fin also pays a rebate, here 2% of face at maturity if the barrier is touched. The touch probability, from Section 34.3's formula, is 0.292, so the rebate costs $1,000 \times 2\% \times e^{-0.03} \times 0.292 = 5.68$, and the remaining 23.88 buys

$$\frac{23.88}{21.766} = \mathbf{109.7\% \text{ participation}}$$

The participation rate has more than quadrupled, and what has been given up is every path that touches 130 at any time before maturity—not only those that end above it.

The question the term sheet does not answer is how often those outcomes occur. That touch probability, **29.2%** under the pricing measure, is the answer: in nearly a third of scenarios the 109.7% participation is extinguished and the note pays its small rebate—and the knock-out happens in the states where the index rallied hardest. The buyer has not sold a remote tail; they have sold roughly the three strongest rallies in ten.

Two features make this harder to see than the autocallable's. When the barrier is monitored continuously, as here, a single intraday touch during a rally destroys the participation for the rest of the year even if the index closes the year at 128. And the

rebate is quoted as a percentage of face rather than as what it replaces: a 2% rebate against a forgone 109.7% participation on a 30% move is 2% against 32.9%. The instrument is Example 34.6's reverse knock-out sold retail, with the same discontinuity and none of the warning.

Dual-currency deposits pay an above-market rate in one currency, with the issuer holding the right to repay in another at a preset rate. *It is a deposit plus a short FX option*, and it is the most widely sold structured product in Asian private banking. The presentation as a “deposit” is the problem: the principal is at risk in currency terms and the vocabulary says otherwise. Example 35.5 shows that the entire enhancement is the option premium, to the basis point.

Example 35.5 The deposit that is not one

A three-month dollar deposit of \$100,000 with a conversion rate of 153 against a USD/JPY spot of 150. Dollar deposits pay 4.5%, yen 0.5%, and three-month volatility is 10%. The depositor has written a dollar call struck at 153: if the dollar strengthens past it, the bank repays in yen at 153. Interest is paid in dollars whatever happens; only the principal can be converted.

Chapter 16's Black–Scholes–Merton formula, with the yen rate for discounting and the dollar rate as the yield q (the form known as Garman–Kohlhagen), values that option at **1.2776** yen per dollar, which is 0.852% of the dollar notional over the quarter, or **3.41% annualized**. So:

$$4.5\% \text{ (deposit)} + 3.41\% \text{ (written option)} = \mathbf{7.91\%}$$

The advertised 7.9% rate is the deposit rate plus the premium, and nothing else. A depositor who understood the decomposition would ask whether 3.41% a year is a fair price for the option—which is the question the product's presentation prevents, because a deposit rate is not something one negotiates.

The outcomes are asymmetric in the familiar direction:

USD/JPY at maturity	Repaid in	Quarter's return
140	dollars	+1.98%
150	dollars	+1.98%
153	dollars	+1.98%
160	yen at 153	−2.40%
165	yen at 153	−5.30%
175	yen at 153	−10.59%

The gain is capped at one quarter's coupon and the loss is not capped at all, which is Section 35.2's shape in a three-month product and a different asset class. And the depositor is short exactly the move that would otherwise have helped: a dollar-based investor is converted into yen precisely when the dollar is strong, so the product removes the currency exposure the depositor wanted and leaves the one they did not.

Credit-linked notes embed the short CDS position of Section 25.14. The whole of Part IV's credit machinery applies, and the retail versions are the ones where the reference entity's identity is least understood by the buyer.

And range accruals appear as notes as well as swaps, with the coupon arithmetic and the short-vega exposure of Section 33.6 unchanged.

Hybrids attach a second condition from another asset class to the credit one—a coupon

paid only while an issuer survives *and* an index is above its strike—and they are the exception to this section’s arithmetic. The two legs cannot be priced separately and added, because the value depends on a correlation between default and the other factor that neither market quotes; Section 25.15 works the example, and the practical consequence for a buyer is that the usual decomposition is unavailable to them as well as to the dealer.

The pattern across the first four is the one this chapter began with. The enhanced headline term is paid for by an option the buyer has written, and the option is described by its trigger rather than by its risk.

35.6 ISSUER HEDGING, AND THE DISCONTINUITY

The investor’s position implies the issuer’s, and the issuer’s hedging is where these products stop being a private matter between two parties.

The investor is short a down-and-in put; the issuer is long it. A long put has negative delta, so *the issuer hedges by holding the index*, and the size of that holding is what matters.

Example 35.6 Why the hedge stabilizes, until it does not

The embedded option: a down-and-in put struck at 100 with a barrier at 75, one year, with Example 35.2’s 25% volatility and 2.5% rate and no dividends. Valued exactly with the down-and-in put row of Section 34.3’s barrier formulas. The note’s knock-in is observed daily, which Section 34.6’s correction handles by moving the barrier away from the spot, to 74.32; that lowers the value at 100 from 6.160 to 5.956 and leaves the shape of the delta, and its jump at the barrier, essentially unchanged, so the continuous barrier is used below.

Index	Option value	Delta
130	0.706	−0.055
110	3.128	−0.220
100	6.160	−0.399
90	11.365	−0.652
80	19.262	−0.921
76	23.124	−1.006
75.33	23.802	−1.017
74.67	24.411	−0.827
74	24.968	−0.836
70	28.414	−0.885

Above the barrier, delta steepens monotonically from −0.055 to −1.017 as the market falls, and it *exceeds one in magnitude*—a barrier feature Section 34.3 flagged.

The issuer, long this option, must therefore hold a growing long position in the index as the market falls. From 130 down to the barrier the hedge grows from 5.5% to 101.7% of notional. The issuer is buying into a falling market, continuously, and that flow is stabilizing.

Then the barrier is touched and the delta jumps from −1.017 to −0.827: a change of +0.19.

Two things happen at once and the second is the larger. The issuer must *sell* about 19% of notional immediately. And—more importantly—*the buying stops*: below the barrier the option is an ordinary put whose delta steepens far more slowly, so the persistent bid that had been supporting the market on the way down simply disappears.

The popular account of these episodes emphasizes the forced selling. The arithmetic

says the withdrawal of the bid is the bigger effect, and stating it that way changes what a regulator should monitor: not the size of the sell order at the barrier, but the size of the stabilizing flow that is about to vanish.

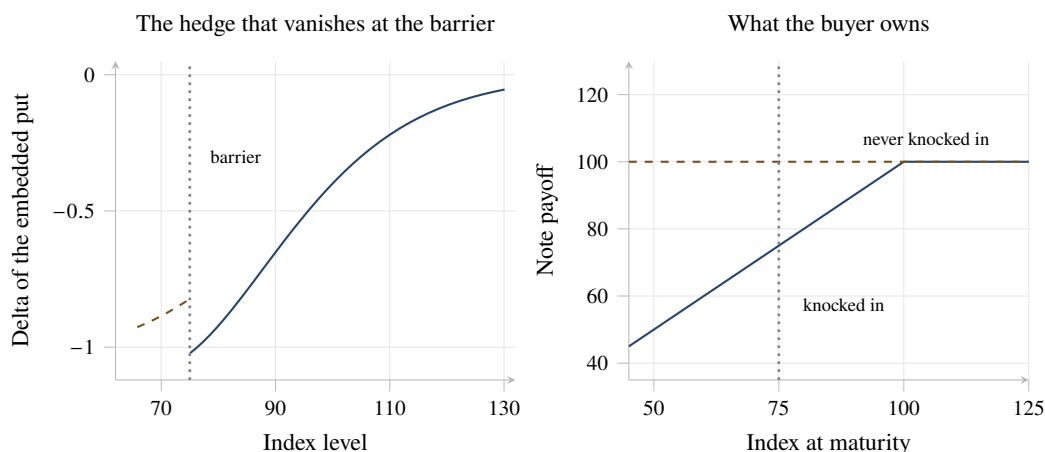


Figure 35.3: Left: Example 35.6's delta. Right: maturity payoff, coupon excluded.

Figure 35.3's left panel is the mechanism in one picture. *The gap at the barrier is the discontinuity*, and the steep approach to it on the left is the stabilizing flow that ends there.

35.7 WHEN MANY NOTES SHARE A BARRIER

One note's hedging is irrelevant to a market. *The problem is that barriers cluster.*

Notes are issued continuously against a rolling reference level, so a sustained decline of a given size brings a large cohort of barriers into range at once. The clustering is not coincidence but arithmetic: barriers are set at round fractions—75%, 80%—of the level prevailing at issue, so notes issued during any period of stable prices share nearly the same absolute barrier.

Three consequences follow, in order of increasing seriousness.

The stabilizing flow is larger than any one issuer knows. Each dealer sees its own book, which is Section 27.12's aggregation problem in a different market.

The withdrawal is synchronized. When the index passes the cluster, many issuers' bids disappear within the same few sessions.

And the feedback is self-reinforcing while it lasts. The buying on the way down slows the decline, which keeps the index above the barriers longer, which sustains the buying—until it does not. A stabilizer that works until it stops is not a stabilizer, and Section 23.10's saturating pin risk made the identical point about a different product.

Case Study 35.2 A cluster passed through: China, January–February 2024

China's structured-product market is built around the **snowball**, the domestic form of the autocallable, and its standard terms track Example 35.2 closely: a CSI 500 or CSI 1000 underlying, twelve to twenty-four months, monthly knock-out observation after a lockout of about three months, a knock-in observed daily at 70–80%, and coupons of 15–20% a year. The underlying is chosen for its volatility, which is what pays the coupon—the

same reason Korea's notes referenced an offshore China index—and that coupon range is the same order as the 18.63% of Example 35.2, because it is the same option at the same volatility.

One feature of the market differs from Korea's in the way that matters most. Snowballs are sold by securities firms acting as counterparty, wrapped in asset-management products with a minimum subscription around RMB 1 million, so the buyer is a qualified investor rather than a depositor. The second of the Korean episode's four general features—that the distribution channel matters as much as the payoff—had been addressed institutionally, and in advance.

What had not been addressed was the synchronization. Barriers were set at 75% or 80% of the index level prevailing at issue, so notes written through a comparatively stable 2023 shared nearly the same absolute knock-in level. The decline of January and February 2024 went straight through it.

Both of Example 35.6's effects are identifiable in what followed, and the popular account has their sizes the wrong way round. Dealers hedging the long down-and-in put had been buying index futures as the market fell, and that flow was supporting it. At the barrier the immediate sale is about 19% of notional, while the larger effect is that the bid supporting the market on the way down disappeared. So snowball hedging did not cause the fall; it was a mechanism that first slowed the decline and then stopped slowing it—and in a price series, “stopped slowing” looks a great deal like “caused.”

Attribution is harder still because a second leveraged chain was unwinding in the same week, on overlapping underlyings and intermediated by the same firms. Separating the two contributions requires dealer-level position data that is not public, so this section's first consequence—each dealer sees only its own book—obstructs the post-mortem as well as the risk management.

The regulatory response is the informative part, because most of it ran ahead of the event. Suitability requirements for snowball sales, an explicit instruction that the product is not principal-protected, and later limits on the size and concentration of dealers' snowball books were all in place before 2024. *None of them addressed the clustering*, because suitability governs who may buy, and the cluster is a fact about how many investors bought notes with nearly the same barrier level in the same quarter. A rule aimed at individual prudence does not solve a problem about an aggregate—Section 27.12's conclusion, reached again in a different market.

35.8 WORST-OF UNDERLYINGS

Section 34.11 priced the mechanism; here is what it does commercially.

Most autocallables reference the *worst* of two to four underlyings. *This roughly doubles the coupon*, and Example 34.9 showed why: at a correlation of 0.20 a worst-of put on three assets is worth 2.05 times a single-asset put.

The disclosure problem is acute and simple. A note “linked to three major indices” sounds diversified. *It is the opposite:* the investor is exposed to whichever of the three performs worst, so adding an underlying can only increase the risk and can only increase the coupon.

And the coupon's size is the tell. A coupon far above what a single-underlying note pays is not a better deal; it is a measure of how much extra risk was added, and it is the one number in the term sheet that reliably signals the risk it was designed to obscure.

35.9 THE ISSUER'S CREDIT

A structured note is an *unsecured obligation of the issuer*. The derivative is embedded in a bond, and if the bond defaults the derivative goes with it.

Principal protection means protection from the market, not from the issuer. That distinction was academic until September 2008, when retail holders of notes issued by a failed investment bank discovered that “100% principal protected” ranked alongside every other senior unsecured claim.

The pricing consequence is Example 35.1's uncomfortable corollary. The issuer's funding spread *is* the source of the note's attractive terms, so a better-looking note is a worse credit, mechanically and always. The investor is being paid for taking issuer credit risk, and in almost no case is that presented as the reason the terms are good.

Two developments have improved this and neither removes it. Disclosure rules now require the issuer's credit to be stated, and some jurisdictions require an estimated initial value—the note's fair value net of fees—to be printed on the term sheet. The estimated-value disclosure is the single most effective reform in this chapter's subject, because it makes Example 35.7's fee arithmetic visible without requiring the buyer to do it.

35.10 WHERE THE FEES ARE

The economics of the business are worth stating plainly, because they explain the product design.

A structured note carries no explicit management fee. The compensation is the difference between the note's issue price and its fair value—typically 1.5% to 3% of face for a simple structure and more for a complex one, earned at issue.

This creates a specific incentive: complexity is profitable because it is hard to price. A buyer can check a participation rate against a competitor's; the fair value of a worst-of autocallable on three underlyings with a monthly autocall and a daily knock-in is not something a buyer can check at all. The margin is largest where the price is hardest to verify, which is the same structure Section 33.13 identified in non-standard swaps and Section 34.13 in exotics.

The fee is not an estimate once the fair coupon is known, and Example 35.7 extracts it from the numbers already computed.

Example 35.7 Reading the fee out of the coupon

Example 35.2's note prices at par on a coupon of **18.63%**. Suppose the same structure is issued at 1,000 paying **12.65%**.

The factor that converts a coupon into value is not the 0.70-year average life of Section 35.3. A note that knocks in and is never called runs the full two years and pays no coupon at all, and those paths carry nearly half of that average; what remains is then discounted. Simulating the note gives the factor directly: each point of annual coupon is worth **0.366** points of face. So the value given up is

$$(18.63\% - 12.65\%) \times 0.366 = 2.19\% \text{ of face}$$

and the note's fair value at issue is about **978** against an issue price of 1,000. *The \$22 is the fee*, and it sits inside the coupon rather than beside it.

Three things follow that a term sheet does not say. The fee is paid at issue and cannot be recovered by holding to maturity—it is not an annual charge of 2.19% but a single 2.19% taken on day one, which on a note with a 0.70-year expected life is an annualized **3.1%**. *The buyer who sells early pays it twice*, once at issue and again in Section 35.4's

exit spread. And *the fee is invisible in every scenario table*, because those tables are computed from the stated coupon and the stated coupon is already net of it.

This is precisely what the estimated-initial-value disclosure repairs. Printing “estimated initial value: \$978 per \$1,000” converts an argument about fairness into a number a buyer can compare across issuers, without requiring anyone to understand a single thing about barrier options. That is why it is the most effective reform in this area: it does not attempt to explain the product, it prices it.

Illustrative magnitudes, scaled to complexity:

	Fair value	Gap to par
Principal-protected note, one index	0.9800	2.00%
Reverse convertible, one index	0.9750	2.50%
Autocallable, one index	0.9700	3.00%
Autocallable, worst-of three	0.9550	4.50%

The gap widens exactly as the buyer’s ability to check it narrows.

The defense, which deserves stating, is that the alternative is worse for the buyer. These structures cannot be assembled retail: the components are OTC, the sizes are institutional, and the hedging is continuous. A margin for manufacturing a payoff that could not otherwise be bought is a legitimate charge. The question is its size and whether it is visible, not whether it should exist.

35.11 HOW TO EVALUATE A NOTE

The chapter’s material assembles into a procedure. It is short, it requires no model the reader does not now have, and it answers the questions a term sheet is designed not to raise.

One. Decompose it. Every note is a bond plus one or more options. Identify the bond, identify each option, and—critically—identify which side of each option the buyer is on. In Example 35.1 the buyer is long a call; in Section 35.2’s reverse convertible and Example 35.2 the buyer is short a put. If the headline term is attractive, the buyer is short something.

Two. Price the bond leg at the issuer’s own curve. Example 35.1 showed that an attractive participation rate comes mostly from the issuer’s funding spread: an issuer that pays more to borrow spends less on the zero-coupon bond and has more left to buy options. A note from an AA issuer and one from a BBB issuer are not comparable on their headline terms, and the difference between them is credit compensation, not investment skill.

Three. Value the option leg and compare with par. The gap is the fee. Where an estimated initial value is disclosed, this step is done for the reader; where it is not, the table in Section 35.10 gives the order of magnitude, and the rule that it widens with complexity is reliable.

Four. Simulate the distribution, not the scenarios. A term sheet presents three or four illustrative outcomes. Example 35.3 shows what the term sheet omits: the skewness, the conditional loss given a loss, and the fact that the median and the 25th percentile can be the same number. A buyer who sees only $E[\text{return}]$ and $P(\text{positive})$ has seen the two statistics the product was engineered to optimize.

Five. Ask what the exit is worth. Section 35.4: the issuer is the only bid, the spread applies again, and the bid is lowest when the need to sell is greatest. A note should be evaluated as a hold-to-maturity instrument, because that is what it is.

Applied to Example 35.2’s note, the five steps produce a description no marketing document contains: a two-year unsecured loan to the issuer, bundled with a written down-and-in put on an index, sold, at Example 35.7’s 12.65% coupon, about 2.2% above fair value, returning a

positive amount with 83% probability—most often 4.7%, on average 8.3%—and on average –32% with 17% probability, and not sellable except back to the issuer.

That description is neither hostile nor favorable. It is the instrument, and a buyer who wants it after reading it is making an informed choice—which is the entire objective.

35.12 SUITABILITY, AND WHAT DISCLOSURE CANNOT FIX

Disclosure is necessary and it is not sufficient, and this chapter's material shows precisely where it stops working.

What disclosure fixes well. Fees and fair value: the estimated-initial-value requirement converts an invisible cost into a printed number, and it works because the quantity is a single figure the issuer already computes. Section 35.10's gap is exactly the kind of thing disclosure is good at.

What it fixes poorly. A distribution cannot be disclosed into intuition. A buyer told that the product has a skewness of –1.57 has been given a true fact and no understanding, and a buyer shown three illustrative scenarios has been given an understanding that is wrong. The gap between Example 35.3's two columns is not a disclosure failure; it is a comprehension problem, and jurisdictions that have tried to solve it with more disclosure have generally produced longer documents and unchanged outcomes.

What it cannot fix at all. The correlation between the product's failure and everything else the buyer owns. An autocallable knocks in when equity markets fall, which is when the buyer's other assets have fallen, their employer is under pressure, and their appetite for risk has collapsed. No disclosure of the product in isolation reveals that, because the fact is about the portfolio and not about the note.

That points at the reform this chapter would actually endorse: not more disclosure, but a suitability test framed around the payoff's shape rather than the buyer's stated risk tolerance—because a buyer who says they want “income with capital preservation” has described a payoff the product does not have, and no amount of the buyer confirming that they understand the risks changes that mismatch.

Two limits on that recommendation need stating, so that the argument is fair. Suitability regimes are paternalistic and impose real costs, including on buyers who understand the product perfectly well. And *the demand is genuine*: the yield problem these products solve is real, and a regime that removes them without addressing that demand pushes it somewhere less regulated. This is a trade-off rather than a solved problem, which is why Section 35.13 names the buyers rather than dismissing them.

35.13 WHO BUYS THESE, AND WHY

Retail and private-bank clients are the volume, and the demand is genuine rather than manufactured. The dominant motive is yield in a world where deposits pay less than the buyer needs, and the second is a preference for defined outcomes over open-ended ones. Both are real preferences and neither is irrational.

Insurers and pension funds buy principal-protected structures for the reason Section 31.17 gave: their liabilities contain guarantees, and a note with a floor matches them better than an unhedged equity holding.

Corporates occasionally buy them as treasury investments, and the suitability question is sharper there than in retail, because the sums are larger and the governance is often thinner.

Issuers sell them for funding and for the margin. *The funding motive is underrated*: a structured note is term funding raised from a source that does not otherwise lend to banks, and it is stickier than wholesale funding because the buyer cannot easily exit.

And dealers—frequently a different institution from the issuer—take the hedging risk, which is where Section 35.6's flows originate.

The recurring conclusion of the last three chapters applies here in its sharpest form. The client is short the optionality and the dealer is long it, and here the client is a household. *That does not make the trade illegitimate*, but it does mean the disclosure question is not a formality—it is the whole of the consumer-protection issue, and Section 35.9's estimated-value requirement is the reform that addresses it most directly.

SUMMARY

Every structured note decomposes into a bond and a derivative, and doing the decomposition converts a marketing document into a portfolio.

A five-year principal-protected note at 4.5% funding spends **798.52** of \$1,000 on the zero-coupon bond and buys **96.1% participation** with the rest, **86.6%** after a 2% fee. Participation runs from 59.4% to 107.3% as issuer funding goes from 2% to 6% with the index unchanged, so the best-looking note on offer is systematically the weakest credit. A reverse convertible is a bond plus a short put that pays only below its barrier: a 10% coupon against 4% deposits is 6% paid for a written put, described in fixed-income vocabulary while its risk is entirely equity.

A two-year autocallable at 100% monthly with a 75% daily knock-in prices at par on an **18.63%** coupon. It pays its advertised return about 83% of the time and loses in the other 17%, with skewness -1.572 against the index's $+1.149$: wins average $+8.32\%$ and losses -32.00% . A buyer moving from the index into the note exchanges a positively skewed exposure for a negatively skewed one at the same expected return, and a variance-based risk measure reports the note as the safer of the two. It is also short volatility at the worst time, since volatility rises as the knock-in is approached.

Korea's retail market showed what this does at scale: the 83% outcome repeated until it looked like a rule, and then a single vintage's barriers, maturities, and losses arrived together. A 17% loss probability sold to a hundred thousand people in one quarter does not produce 17% of them losing independently; it produces all or none.

The issuer is long the embedded down-and-in put, whose delta steepens to -1.017 just above the barrier and jumps to -0.827 at it. The withdrawal of the issuer's bid matters more than the forced sale, and because barriers cluster at round fractions of the level at issue, the withdrawal is synchronized across issuers who each see only their own book. Worst-of underlyings roughly double the coupon, and a note "linked to three major indices" is the opposite of diversified.

A note is unsecured issuer debt, so principal protection is protection from the market, not from the issuer. The fee, typically 1.5–3% of face, is largest where the price is hardest to verify; disclosure fixes fees well and distributions poorly. The reform this chapter would endorse is a suitability test framed around the payoff's shape rather than the buyer's stated risk tolerance.

FURTHER READING

The regulatory literature is where this subject is treated, because supervisors have had to define what a fair disclosure looks like: the estimated-value rules and the various packaged-product disclosure regimes are worth reading as primary sources, and their differences across jurisdictions are informative about what each regulator thought the problem was. On the academic side, the work measuring issue prices against model fair values is consistent enough across markets and decades to be treated as settled: notes are issued at a premium to fair value, and the premium rises with complexity. Bergstresser and others on the retail structured-products market document the demand side. On autocallable hedging flows the public literature is thin and the useful material is in central-bank financial stability reports, which have covered the

topic increasingly since 2022. For the pricing machinery, the references of Chapter 34 apply unchanged, since an autocallable is a barrier product with an early-termination feature and nothing in it is new.

PRACTICE QUESTIONS

- 35.1** Using Example 35.1, compute the participation rate before and after a 2% fee, and explain why the same note's participation would be 59.4% at 2% issuer funding and 107.3% at 6%.
- 35.2** Explain why the best-looking note in a comparison across issuers at one moment is systematically the weakest credit.
- 35.3** Decompose a reverse convertible paying 10% when deposits pay 4%, and state what the 6% actually buys.
- 35.4** Using Example 35.2, state the fair coupon and the three probabilities, and explain why a payoff sold on its hit rate is negatively skewed.
- 35.5** Explain why an autocallable is short volatility in the worst possible way, being specific about when volatility rises.
- 35.6** State the note's and the index's skewness from the simulated distribution, and explain what a buyer switching between them is exchanging when both earn the risk-free rate.
- 35.7** State the average win, the average loss, and the worst one percent, and reconcile them with the 83.6% probability of a positive return.
- 35.8** Explain why the 25th and 50th percentiles of the note's return coincide, and why that makes the product unusually persuasive.
- 35.9** Explain why a variance-based risk measure reports this note as safer than the index, and what that says about variance as a risk measure.
- 35.10** Explain why the exit from a structured note is worst precisely when it is most wanted.
- 35.11** Explain why a product with a 17% loss probability sold to a hundred thousand people does not produce 17% of them losing independently.
- 35.12** Using Example 35.6, state the delta just above and just below the barrier and the jump, then explain why the withdrawal of the issuer's bid matters more than the forced sale at the barrier.
- 35.13** Explain why barriers cluster across issuers, why no issuer can observe the clustering, and what that implies about what should be monitored.
- 35.14** Explain why a note linked to the worst of three indices pays roughly double the coupon, and why that is a measure of risk rather than of value.
- 35.15** List the five evaluation steps of Section 35.11, state what each is designed to reveal, and apply them to Example 35.2's note in one sentence.

FURTHER QUESTIONS

- 35.16** Example 35.1 shows participation nearly doubling with the issuer's funding rate. Derive the participation rate as a function of funding, maturity, and volatility; determine which

of the three a buyer can most easily verify; and design a comparison metric that would rank notes across issuers on a like-for-like basis.

- 35.17** Example 35.2 finds a fair coupon of 18.63% with a 16.6% probability of loss. Compute the distribution of the investor's return rather than its mean, state its skewness, and use it to argue either that the product is or is not suitable for a buyer whose stated objective is income with capital preservation.
- 35.18** Example 35.6 shows the hedge delta jumping by 0.19 at the barrier. Aggregate the flow: assume a market-wide notional and a barrier distribution, derive the total hedging demand as a function of the index level, and identify the level at which the withdrawal of the bid is largest.
- 35.19** Section 35.7 argues that a stabilizer which works until it stops is not a stabilizer. Assess the claim rigorously: construct the case that autocallable hedging reduces realized volatility on average, construct the case that it redistributes volatility into the tail, and state what data would settle which effect dominates.
- 35.20** Section 35.10 argues the margin is largest where the price is hardest to verify. Test this across the products in Chapters 33–35: rank them by verifiability and by typical margin, and state whether the relationship holds and what explains any exception.
- 35.21** Section 35.4 argues the exit is worst when most wanted. Quantify it: model the issuer's bid as fair value less a spread that widens with volatility, and compute the loss an investor takes by exiting after a 20% fall relative to holding to maturity.
- 35.22** Section 35.12 argues that a suitability test should be framed around payoff shape rather than stated risk tolerance. Design such a test: specify the statistics it would require, the thresholds it would apply, and how it would classify each product in this chapter—then identify the product it misclassifies and say why.
- 35.23** The estimated-initial-value disclosure tells a buyer what the note is worth on the day it is issued. Determine what it does *not* tell them, propose one further disclosure that would close the largest remaining gap, and explain why an issuer would resist it.

Subject Index

- Accumulator, **23**
- Antithetic variates, 10, 20, 33
- Asian option, **2**, 16, 25
- At-the-money, 16
- Autocallable, 29, 33, 42

- Barrier option, **2**, 6
- Basket / rainbow option, **2**, 19
- Binomial tree, 8
- Broadie–Glasserman–Kou, 10

- Cap, 21
- Cap / floor, 21
- Cash-or-nothing, **2**, **12**
- Cash-or-nothing / asset-or-nothing, **2**, **12**
- Chooser option, **3**
- Cliquet, **2**, **21**, 23, 25
- Closed form, 6, 16, 19, **23**
- Collar, 3
- Compound option, **3**
- Control variate, 8, 16, 19
- Conversion / reversal (parity trades), 21, 29
- Cox–Ross–Rubinstein tree, **8**
- Credit-linked note, 37

- Delta, 12, 39
- Discrete monitoring, **10**

- Exchange option, 2, 19, 20
- Exotic option, **2**
- Explicit / implicit scheme, 8

- Finite difference method, 8
- Forward smile, 21
- Forward-start option, **3**
- Funding rate, 29

- Geske formula, **3**

- Hang Seng China Enterprises Index (HSCEI),
36

- Knock-in / knock-out, 2, **6**, 10, 23, 29, 33, 36,
40

- Local / stochastic volatility, 11, 21
- Local volatility, 11, 21
- Lookback option, **2**

- Overhedge, 8, **12**

- Participation rate, **29**
- Principal-protected note, 29

- Reverse convertible, 29, 33, 42, 43
- Rho, 3, 20
- Running minimum / maximum, **2**, 16

- Shark-fin note, **37**
- Skewness, **35**
- Snowball, **40**
- Static replication, 9
- Stochastic volatility, 11
- Strip, 23
- Structured note, 25
- Suitability, 40, **44**

- Trinomial tree, 8

- Vanna, volga, 2
- Vega, 3
- Volatility surface, 11

- Worst-of note, **20**
- Worst-of option, **20**