Chapter 25

Solvability, Realism and What Is Left

The preceding chapters were written to be read in order and the argument they make is distributed across them, so this one assembles it. The claim is that the models in this book were chosen under a constraint that no longer binds, that the constraint which replaced it is not compute but identification, and that what solvability still buys is not speed but a bounded error — which makes the choice between a tractable model and a realistic one a trade of realism for auditability rather than a matter of taste. Along the way the same result appears from four directions: hedging a book better does not reduce its model risk, it concentrates it.

25.1 What the Book Did

The spine of these notes is two reductions, and chapter 14 names them.

The first turns a functional integral into a differential equation, by insisting that the state be finite. The second turns that equation into a system of ordinary ones, by insisting that the generator preserve a small family of functions. Neither is a fact about markets. Both are conditions imposed so that a calculation can be finished, and almost every model in this book is what one of them left standing.

Be blunt about how much was given up, because the chapters measured it rather than asserting it.

Restriction	**Imposed for**	**What it cost, measured**
one factor	a small Markov state	a 2s10s spread option at a quarter of its value
exponential volatility	a finite realisation at all	every non-quasi-exponential maturity profile
local volatility	an exact fit to vanillas	a forward skew a fifth of today's after a year
affine dynamics	a transform in closed form	a smile shape the market does not have at short expiry

Each row was established where it belongs — the first in chapter 12, the second there too, the third in chapter 10, the fourth in chapter 10's closing section — and each cost is large enough that nobody would accept it if the alternative were available.

25.2 The Constraint That Replaced Compute

The alternative is available. Chapter 19 sets out what a machine will now do, and the honest summary is that the first reduction is no longer needed for pricing and the second is no longer needed at all. A fifty-factor model can be simulated; a pricing equation in fifty dimensions can be solved to useful accuracy without a grid.

So the question is what binds instead, and the book answers it in three places without ever collecting the answer.

- Chapter 10 shows that a smile does not determine a model's dynamics: two models agreeing on every quoted price disagree on the hedge.

- Chapter 20 measures a SABR calibration in which the quotes fix the level of the smile to one per cent and leave the correlation free across most of its range, with the model exactly right and the data noiseless.

- Chapter 12 finds the same for mean reversion, and chapter 11 for the mixing weight.

These are not three difficulties. They are one, and it has a name.

Structure (Identification is the binding constraint)**.** A model is a map from parameters to prices. Calibration inverts it. The question of whether that inverse exists and is stable is not about optimisers or compute; it is about whether the instruments quoted carry enough information to distinguish the model from its neighbours, and generally they do not.

The situation is therefore the opposite of the one the classical models were built for. Then, a model that could not be solved could not be used, so tractability bound and data was ample relative to what a tractable model needed. Now a model can be solved at almost any complexity, and the data available to fit it has not grown at all — the same forty swaption quotes, the same handful of liquid expiries. A model with a million parameters fitted to forty quotes is the identification problem in its extreme form: a richer model does not extract information the market did not supply, it distributes the same information over more unknowns.

Chapter 20 makes the sharpest version of this quantitative, and it repays repeating because it is easy to file as a technicality. On a spread that genuinely mean reverts with a one-year half-life, five years of daily data gives a ninety per cent interval for the half-life spanning a factor of eight, and sampling ten times as often does not narrow it. The constraint is not the number of observations. It is the length of the record, and a desk cannot buy more history.

25.3 What Solvability Still Buys

If tractability is no longer necessary, the question is whether it is still worth anything, and the answer is yes for two reasons that have nothing to do with speed. Chapter 19 establishes both.

Prices that are consistent structurally rather than statistically. A closed form cannot produce a negative butterfly, because it is an expectation under a measure and an expectation of a convex function is convex. Once the expectation is estimated the guarantee must be earned: priced on common paths a butterfly inherits the payoff's convexity term by term and cannot be negative even at ten thousand paths, and priced on independent paths over a quarter of narrow triples invert. Arbitrage-freeness moves from a property of the model to a property of the estimator, and survives on three disciplines rather than on a theorem.

An error one can write down. A grid's error is $O(h^p)$ and two runs estimate it; a Monte Carlo standard error is a genuine confidence statement. A network fitted to a residual offers neither, because the map from residual to error runs through a stability constant unavailable in exactly the dimensions the method exists to handle. Over four dimensions the network is the more accurate method. It is not the more *measurable* one.

Structure (The trade is realism for auditability)**.** Putting those together gives the decision this book is ultimately about.

A solvable model is less realistic. That is not incidental — chapter 14 shows it was chosen for tractability, and the table above measures what the choice cost. A numerical model is more realistic and its error is not a quantity. So the choice is not between a good model and a bad one but between two different things one might want to be able to say about a number.

The reason this matters and is not merely tidy is chapter 24's: a valuation uncertainty has to be reserved against, and a reserve requires a number. An unbounded error cannot be reserved for. It can only be provisioned against by judgement, which is to say carried as model risk — so a more realistic model with an unmeasurable error has not removed a risk, it has moved it from the valuation to the capital, where it is less visible and harder to argue about.

Stated that way the trade becomes a decision with an owner rather than a habit with a history. A desk pricing a vanilla book wants the auditable answer. A desk pricing a fifty-factor structured note has no auditable answer available and should say so, rather than reporting a number with the same confidence as the vanilla one.

25.4 One Result, From Four Directions

The practical chapters were written independently and arrive at the same conclusion, which is the strongest evidence in the book that it is not an artefact of any one setting.

- Chapter 23 ends on the partial hedge, whose unhedged fraction is the transaction cost over the risk penalty. It is zero only if hedging is free, so a completely hedged book is not a well-run one.

- Chapter 24's attribution: rebalancing more often drives discretisation error down as the square root of the frequency and leaves model error untouched, so the share of remaining risk that is model error rises — from 1.15 times the correct-model risk at fifty rebalances a year to 2.64 at a thousand.

- Chapter 24's explain: hedging removes exposure to the factor the risk model understands and leaves the others alone, so the unexplained share of profit and loss rises with the quality of the hedge — under two per cent for an outright book, over ninety for the same book duration hedged, and all of it for a level-neutral curve trade.

- Chapter 22: a convergence trade's risk is not the direction, which may be certain, but the holding period, which depends on a parameter the data barely constrains.

Structure (Hedging concentrates model risk)**.** The four are one statement. Hedging removes the risks the model can see. It cannot remove the risks the model cannot see, because a hedge is computed from the model. So what survives hedging is, disproportionately, whatever the model is wrong about — and the better the hedging, the purer that residue.

This has a consequence for how a risk number should be read that no amount of care with ρ supplies. A well hedged book reports a small risk, and reports it in the model's own coordinates, so the report is small for the same reason it is unreliable. The unhedged book's number is larger and more nearly true.

It also gives a limit to how much hedging is worth doing, distinct from the transaction cost argument. Past the point where model error dominates, additional rebalancing reduces a term that is no longer binding: a desk hedging continuously against a surface marked to the nearest volatility point is being precise about the smaller half of its problem. The binding term is reduced by a better model, not by more trading, and knowing which of the two one is short of is the practical content of chapter 24's insistence on decomposition.

25.5 What To Do About It

Three habits, each of which the chapters arrived at by measuring something and finding it was not what was expected.

Decompose before reporting. A risk number without an attribution cannot be acted on, because the three sources — chosen, flowed, and imposed by the model — are reduced by three different actions and only one of them is free. Chapter 24 makes this its thesis and chapter 23 supplies the book that requires it.

Distinguish what cannot be hedged from what is not worth hedging. Both appear in a risk report as a residual and they are not the same thing: the first is a modelling constraint and no amount of spending reduces it, the second is a decision with a price and can be revisited. Chapter 23's projection makes the distinction precise.

Prefer a structural reason to a fitted parameter. Chapter 22 needs both a reason and a horizon, and chapter 20 shows that the horizon is the part the data cannot supply. A claim about a regulation or an index rule is a claim the evidence can support; a half-life from five years of history is a point estimate from a sample that could not distinguish six months from two years.

Remark (What will date and what will not)**.** The methods in chapter 19 will be superseded, probably soon, and the models in the middle chapters are already less realistic than what a machine can now handle. What will not date is the structure of the argument: that every tractable model is a restriction adopted for a reason, that the reason matters because it tells one where the model will fail, and that the failure is measurable before it is expensive.

That is why these notes spend as much effort on what the models get wrong as on what they compute. A reader who leaves knowing which compromise sits under each formula, and how to measure what it cost, is equipped for models that have not been invented yet. A reader who leaves knowing the formulae is equipped for the ones in here.