Lorenzo Bergomi

Stochastic Volatility Modeling

Chapman & Hall/CRC FINANCIAL MATHEMATICS SERIES

Stochastic Volatility Modeling

CHAPMAN & HALL/CRC
Financial Mathematics Series

Aims and scope:
The field of financial mathematics forms an ever-expanding slice of the financial sector. This series aims to capture new developments and summarize what is known over the whole spectrum of this field. It will include a broad range of textbooks, reference works and handbooks that are meant to appeal to both academics and practitioners. The inclusion of numerical code and concrete real-world examples is highly encouraged.

Published Titles

American-Style Derivatives; Valuation and Computation, *Jerome Detemple*

Analysis, Geometry, and Modeling in Finance: Advanced Methods in Option Pricing, *Pierre Henry-Labordère*

Commodities, *M. A. H. Dempster and Ke Tang*

Computational Methods in Finance, *Ali Hirsa*

Counterparty Risk and Funding: A Tale of Two Puzzles, *Stéphane Crépey and Tomasz R. Bielecki, With an Introductory Dialogue by Damiano Brigo*

Credit Risk: Models, Derivatives, and Management, *Niklas Wagner*

Engineering BGM, *Alan Brace*

Financial Mathematics: A Comprehensive Treatment, *Giuseppe Campolieti and Roman N. Makarov*

Financial Modelling with Jump Processes, *Rama Cont and Peter Tankov*

Interest Rate Modeling: Theory and Practice, *Lixin Wu*

Introduction to Credit Risk Modeling, Second Edition, *Christian Bluhm, Ludger Overbeck, and Christoph Wagner*

An Introduction to Exotic Option Pricing, *Peter Buchen*

Introduction to Risk Parity and Budgeting, *Thierry Roncalli*

Introduction to Stochastic Calculus Applied to Finance, Second Edition, *Damien Lamberton and Bernard Lapeyre*

Monte Carlo Methods and Models in Finance and Insurance, *Ralf Korn, Elke Korn, and Gerald Kroisandt*

Monte Carlo Simulation with Applications to Finance, *Hui Wang*

Nonlinear Option Pricing, *Julien Guyon and Pierre Henry-Labordère*

Numerical Methods for Finance, *John A. D. Appleby, David C. Edelman, and John J. H. Miller*

Option Valuation: A First Course in Financial Mathematics, *Hugo D. Junghenn*
Portfolio Optimization and Performance Analysis, *Jean-Luc Prigent*
Quantitative Finance: An Object-Oriented Approach in C++, *Erik Schlögl*
Quantitative Fund Management, *M. A. H. Dempster, Georg Pflug, and Gautam Mitra*
Risk Analysis in Finance and Insurance, Second Edition, *Alexander Melnikov*
Robust Libor Modelling and Pricing of Derivative Products, *John Schoenmakers*
Stochastic Finance: An Introduction with Market Examples, *Nicolas Privault*
Stochastic Finance: A Numeraire Approach, *Jan Vecer*
Stochastic Financial Models, *Douglas Kennedy*
Stochastic Processes with Applications to Finance, Second Edition, *Masaaki Kijima*
Stochastic Volatility Modeling, *Lorenzo Bergomi*
Structured Credit Portfolio Analysis, Baskets & CDOs, *Christian Bluhm and Ludger Overbeck*
Understanding Risk: The Theory and Practice of Financial Risk Management, *David Murphy*
Unravelling the Credit Crunch, *David Murphy*

Proposals for the series should be submitted to one of the series editors above or directly to:
CRC Press, Taylor & Francis Group
3 Park Square, Milton Park
Abingdon, Oxfordshire OX14 4RN
UK

This page intentionally left blank

Chapman & Hall/CRC FINANCIAL MATHEMATICS SERIES

Stochastic Volatility Modeling

Lorenzo Bergomi

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

Typeset in Linux Libertine

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20151105

International Standard Book Number-13: 978-1-4822-4407-6 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface

Pour soulever un poids si lourd
Sisyphe, il faudrait ton courage!
Bien qu'on ait du cœur à l'ouvrage,
L'Art est long et le Temps est court.

Baudelaire, *Les Fleurs du mal*

Tu quid? Quousque sub alio moveris? Impera et dic quod memoriae tradatur, aliquid et de tuo profer.

Seneca, *Letters to Lucilius, XXXIII*

C'est icy un livre de bonne foy, lecteur.

Montaigne, *Essais*

– This, Reader, is an honest book. It warns you at the outset that it is not a treatise on stochastic volatility.

Nor is it a mathematical finance textbook – there are treatises and textbooks galore on the shelves of bookshops and university libraries.

Rather, my intention has been to explain how stochastic volatility – and which kind of stochastic volatility – can be used to address practical issues arising in the modeling of derivatives.

Modeling in finance is an engineering field: while our task as engineers is to frame problems in mathematical terms and solve them using sophisticated machinery whenever necessary, the problems themselves originate in the form of embarrassingly practical trading questions.

I have been fortunate to have spent my career as a quant in an institution – Société Générale – and in a field – equity derivatives – that the derivatives community has come to associate. I have tried to convey the experience thus gained. My main objective has been to clearly state the motivation for each question I address and to represent the thought process one follows when designing or using a model.

A model's (ir)relevance is measured by its (in)ability to account for all nonlinearities of a derivative payoff with respect to hedging instruments, and to adequately reflect the former in the prices it produces.

Thus have I often lamented the propensity of "pricing quants" or "expectation calculators" to envision their task as that of producing (real) numbers auspiciously called prices, and to favor analytical tractability or computational speed at the expense of model relevance. What is the point in ultrafast mispricing?[1]

[1]Kind souls will allege it is still preferable to ultraslow mispricing.

This book is intended to be read in sequence. It assumes familiarity with the basic concepts and models of quantitative finance. The motivation for stochastic volatility is the subject of the first chapter; this is followed by a chapter on local volatility, both a special breed of stochastic volatility and a market model, which I survey as such.

I urge you, Reader, to read it fully, as I introduce notions and discuss issues that are referred to repeatedly throughout the book. After a warm-up with forward-start options we embark on stochastic volatility. Jump and Lévy models are briefly dealt with at the end of Chapter 10.

This work was written mostly at night, during hours normally devoted to rest and sleep. Thus, dear Reader, I beg your forgiveness. You will do me a great service by reporting typos, inaccuracies, downright errors, to lb.svbook@gmail.com.

For everything I have learned so far I have to be thankful to more people than my memory can remember. I wish to thank the practitioners and academics that I have met regularly at conferences and seminars, and also fellow workers at Société Générale: quants, traders, structurers.

I am especially indebted to my coworkers in the Quantitative Research team, past and present. They will find here a reflection of the very many discussions we have had over the years.

The generous help of colleagues – quants and traders – in proofreading the manuscript is also gratefully acknowledged, chief among them Florent Bersani and Pierre Henry-Labordère whose eagle eyes have helped clear the text of many imprecisions.

Finally, I would like to express my gratitude to Rena, Elisa and Chiara for their patience.

Thank you Rena for occasionally filling the loneliness of nightly writing sessions with music. This has resulted for the author in some unlikely pairings, e.g. local volatility and the piano part of the Kreutzer sonata, the Heston model and the Chopin Barcarolle, Rachmaninoff's Third and options on variance, etc.

Paris, Summer 2015

Cover art: in Albrecht Dürer's famous *Melencolia I* print (detail), a disgruntled quant sits amid the tools of his trade – some unfamiliar to the author. The scale and hourglass are symbols, respectively, for no-arbitrage and the convex order.

Rows and columns of the magic square all sum up to 34, as do numbers in the four quadrants, as well as in the center quadrant. The middle cells of the bottom row spell out the year the engraving was made: 1514.

Chapter 1

Introduction

Why would a trader use a stochastic volatility model? What for? Which issues does one address by using a stochastic volatility model? Why aren't practitioners content with just delta-hedging their derivative books? These are the questions we address in this introduction.

We begin our analysis by reviewing the Black-Scholes model and how it is used on trading desks. It may come as a surprise to many that, despite the widely publicized inconsistency between the actual dynamics of financial securities, as observed in reality, and the idealized lognormal dynamics that the Black-Scholes model postulates, it is still used daily in banks to risk-manage derivative books.

One may think that a model derives its legitimacy and usefulness from the accuracy with which it captures the historical dynamics of the underlying security – hence the scorn demonstrated by econometricians and econophysicists for the Black-Scholes model and its simplistic assumptions, upon first encounter. With regard to models, many are used to a normative thought process. Given the behavior of securities' prices – as specified for example by an xx-GARCH or xxx-GARCH model – this is what the price of a derivative should be. Models not conforming to such type of specification – or to some canonical set of *stylized facts* – are deemed "wrong".

This would be suitable if (a) the realized dynamics of securities benevolently complied with the model's specification and (b) if practitioners only engaged in delta-hedging. The dynamics of real securities, however, is not regular enough, nor can it be characterized with sufficient accuracy that the normative stance is appropriate. Moreover, such an approach skirts the issue of dynamical trading in options – at market prices – and of marking to market.

Rather than calibrating their favorite model to historical data for the spot process, and, armed with it and trusting its seaworthiness, endeavor to ride out the rough seas of financial markets, derivatives practitioners will be content with barely floating safely and making as few assumptions as possible about future market conditions.

Still, this requires some modeling infrastructure – hence this book: while they do not use the models' predictive power and may have little confidence in the reliability of the models' underlying assumptions, practitioners do need and make use of the models' *pricing equations.* This is an important distinction: while the Black-Scholes *model* is not used on derivatives desks, everybody uses the Black-Scholes *pricing equation.*

Indeed, a pricing equation is essentially an analytical accounting device: rather than predicting anything about the future dynamics of the underlying securities, a model's pricing equation supplies a decomposition of the profit and loss (P&L) experienced on a derivative position as time elapses and securities' prices move about. It allows its user to anticipate the sign and size of the different pieces in his/her P&L. We will illustrate this below with the example of the Black-Scholes equation, which could be motivated by elementary break-even *accounting* criteria.

More sophisticated models enable their users to characterize more precisely their P&L and the conditions under which it vanishes, for example by separating contributions from different effects that may be lumped together in simpler models. Again, the issue, from a practitioner's perspective, is not to be able to *predict* anything, but rather to be able to *differentiate* risks generated by these different contributions to his/her P&L and to ensure that the model offers the capability of *pricing* these different types of risk consistently across the book at levels that can be *individually controlled.*

It is then a trading decision to either hedge away some of these risks, by taking offsetting positions in more liquid – say vanilla – options or by taking offsetting positions in other exotic derivatives, or to keep these risks on the book.

1.1 Characterizing a usable model – the Black-Scholes equation

Imagine we are sitting on a trading desk and are tasked with pricing and risk-managing a short position in an option – say a European option of maturity T whose payoff at $t = T$ is $f(S_T)$, where S is the underlying.

The bank quants have coded up a pricing function: $P(t, S)$ is the option's price in the library model. Assume we don't know anything about what was implemented. How can we assess whether using the black-box pricing function $P(t, S)$ for risk-managing a derivative position is safe, that is whether the library model is usable?

We assume here that the underlying is the only hedging instrument we use. The case of multiple hedging instruments is examined next.

- The first sanity check we perform is set $t = T$ and check that P equals the payoff:

$$P(t = T, S) \;=\; f(S),\ \forall S. \tag{1.1}$$

Provided (1.1) holds, we proceed to consider the P&L of a delta-hedged position. For the purpose of splitting the total P&L incurred over the option's lifetime into pieces that can be ascribed to each time interval in between two successive delta rehedges, we can assume that we sell the option at time t, buy it back at $t + \delta t$ then start over again. δt is typically 1 day. Let Δ be the number of shares we buy at t as delta-hedge.

Our P&L consists of two pieces: the P&L of the option itself, of which we are short, which comprises interest earned on the premium received at t, and the P&L generated by the delta-hedge, which incorporates interest we pay on money we have borrowed to buy Δ shares, as well as money we make by lending shares out during δt:

$$P\&L \;=\; -\left[P\left(t+\delta t, S+\delta S\right)-P\left(t,S\right)\right]+rP\left(t,S\right)\delta t \;+\; \Delta(\delta S-rS\delta t+qS\delta t)$$

where δS is the amount by which S moves during δt. r is the interest rate and q the repo rate, inclusive of dividend yield.

How should we choose Δ? We pick $\Delta = \frac{dP}{dS}$ so as to cancel the first-order term in δS in the P&L above.

We now expand the P&L in powers of δS and δt. We would like to stop at the lowest non-trivial orders for δt and δS: order one in δt, and order two in δS, as the order one contribution is canceled by the delta-hedge. What about cross-terms such as $\delta S \delta t$?

In practice, this term, as well as higher order terms in δS, are smaller than δS^2 and δt terms. Indeed, to a good approximation, the variance of returns scales linearly with their time scale, thus $\left\langle \delta S^2 \right\rangle$ is of order δt and δS is of order $\sqrt{\delta t}$.[1] The contributions at order one in δt and order two in δS are then both of order δt while the cross-term $\delta S\delta t$ and terms of higher order in δS are of higher order in δt, thus become negligible as $\delta t \to 0$.[2]

We then get the following expression for our carry P&L – the standard denomination for the P&L of a hedged option position:

$$P\&L \;=\; -\left(\frac{dP}{dt} - \; rP \;+ \left(r-q\right) S\frac{dP}{dS}\right)\delta t \;-\frac{1}{2}S^2\frac{d^2P}{dS^2}\left(\frac{\delta S}{S}\right)^2 \tag{1.2}$$

- The first piece – called the theta portion – is deterministic. It is given by the time derivative of the option's price (sometimes theta is used to denote $\frac{dP}{dt}$ only), corrected for the financing cost/gain during δt of the delta hedge and the premium.

[1]The property that the variance of returns scales linearly with their time scale is equivalent to the property that returns have no serial correlation. Securities' returns do in fact exhibit some amount of serial correlation at varying time scales, of the order of several days down to shorter time scales and this is manifested in the existence of "statistical arbitrage" desks. Serial correlation in itself is of no consequence for the pricing of derivatives, however the measure of realized volatility will depend on the time scale of returns used for its estimation. As will be clear shortly, for a derivatives book, the relevant time scale is that of the delta-hedging frequency.

[2]How small should δt be so that this is indeed the case? The order of magnitude of δS is $S\sigma\sqrt{\delta t}$ where σ is the volatility of S. It turns out that for equities, volatility levels are such that for $\delta t = 1$ day, higher order terms can usually be ignored. There is nothing special about daily delta rebalancing; for much higher volatility levels, intra-day delta re-hedging would be mandatory.

- The second piece is random and quadratic in δS, as the linear term is cancelled by the delta position. $\frac{d^2P}{dS^2}$ is called "gamma". We usually prefer to work with the "dollar gamma" $S^2\frac{d^2P}{dS^2}$, as it has the same dimension as P.

Our daily P&L reads:

$$P\&L \;=\; -A\,(t,S)\,\delta t - B\,(t,S)\left(\frac{\delta S}{S}\right)^2 \tag{1.3}$$

where $A = \left(\frac{dP}{dt} - \;rP\; + (r-q)\,S\frac{dP}{dS}\;\right)$ and $B = \frac{1}{2}S^2\frac{d^2P}{dS^2}$. Because the second piece in the P&L is random we cannot demand that the P&L vanish altogether.

- What if $A \geq 0$ and $B \geq 0$? We lose money, regardless of the value of δS. This means P cannot be used for risk-managing our option. The initial price $P\,(t=0,S_0)$ we have charged is too low. We should have charged more so as not to keep losing money as we delta-hedge our option.
- What if $A \leq 0$ and $B \leq 0$? We make "free" money, regardless of δS. While less distressing than persistently losing money, the consequence is identical: P cannot be used for risk-managing our option. The initial price $P\,(t=0,S_0)$ we have charged is too high.
- The model is thus usable only if the signs of $A\,(t,S)$ and $B\,(t,S)$ are different, $\forall t, \forall S$. The values of $\frac{\delta S}{S}$ such that money is neither made nor lost are $\frac{\delta S}{S} = \pm\sqrt{-\frac{A(t,S)}{B(t,S)}}\sqrt{\delta t}$.

This condition is necessary, otherwise the model is unusable. We now introduce a further reasonable requirement.

While daily returns are random, empirically their squares average out over time to their realized variance. Let us call $\widehat{\sigma}$ the (lognormal) historical volatility of S: $\left\langle\,\left(\frac{\delta S}{S}\right)^2\,\right\rangle = \widehat{\sigma}^2\delta t$. Requiring that we do not lose or make money on average is a natural risk-management criterion – it reads: $A\,(t,S) = -\widehat{\sigma}^2 B\,(t,S)$, $\forall S,\ \forall t$.

- Replacing A and B with their respective expression yields the following identity that $P_{\widehat{\sigma}}$ ought to obey:

$$\frac{dP_{\widehat{\sigma}}}{dt} - rP_{\widehat{\sigma}} \;+\; (r-q)\,S\frac{dP_{\widehat{\sigma}}}{dS} \;=\; -\frac{\widehat{\sigma}^2}{2}S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2} \tag{1.4}$$

where subscript $\widehat{\sigma}$ keeps track of the dependence of P on the break-even level of volatility $\widehat{\sigma}$.

Plugging now in (1.2) the expression for $(\frac{dP}{dt} - \;rP\; + (r-q)\,S\frac{dP}{dS})$ in (1.4) yields:

$$P\&L \;=\; -\frac{S^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dS^2}\left(\frac{\delta S^2}{S^2} - \widehat{\sigma}^2\delta t\right) \tag{1.5}$$

The condition for the two pieces in the P&L to offset each other is then expressed very simply as a condition on the realized variance of S: the P&L will be positive or negative depending upon whether $\frac{\delta S^2}{S^2}$ is larger or smaller than $\widehat{\sigma}^2 \delta t$.

In the absence of a volatility market for S, $\widehat{\sigma}$ should be chosen as our best estimate of future realized volatility, weighted by the option's dollar gamma.[3]

For vanilla options that can be bought or sold at market prices we can define the notion of implied volatility – hence the hat: $\widehat{\sigma}$ is such that $P_{\widehat{\sigma}}$ is equal to the market price of the option considered.

(1.4) is in fact the Black-Scholes equation. Together with condition (1.1) it defines $P_{\widehat{\sigma}}(t, S)$.

Starting from expression (1.2) for our P&L and imposing the basic accounting criterion that the P&L vanish for $\left(\frac{\delta S}{S}\right)^2 = \widehat{\sigma}^2 \delta t$, at order one in δt and two in δS, a (gifted) trader would thus have obtained the Black-Scholes pricing equation (1.4), though he may not have known anything about Brownian motion and may have been reluctant to assume that real securities are lognormal. The Black-Scholes model is typical of the market models considered in this book:

- there exists a well-defined break-even level for $\left(\frac{\delta S}{S}\right)^2$ such that the P&L at order two in δS of a delta-hedged position vanishes,
- this break-even level does not depend on the specific payoff of the option at hand.

This last condition is important: should the gamma of an options portfolio vanish – that is the portfolio is locally riskless – then theta should vanish as well. If break-even levels were payoff-dependent, we could possibly run into one of the two absurd situations considered above, with $B = 0$ and $A \neq 0$, at the portfolio level.

A model not conforming to these criteria is unsuitable for trading purposes.[4]

Multiple hedging instruments

What if our pricing function is a function of several asset values: $P(t, S_1 \ldots S_n)$ where the S_i are market values of our hedge instruments – either different underlyings, or one underlying and its associated vanilla options?

---

[3] This is not exactly true. Equation (1.5) shows that the situation is different depending on whether our position is short gamma $\left(\frac{d^2P}{dS^2} > 0\right)$ or long gamma $\left(\frac{d^2P}{dS^2} < 0\right)$. In the short gamma situation, our gain is bounded while our loss is potentially unbounded – the reverse is true in the long gamma situation: our bid/offer levels for $\widehat{\sigma}$ will likely be shifted with respect to an unbiased estimate of future realized volatility.

[4] We may have more complex requirements, for example that our P&L vanishes on average, *inclusive* of P&Ls generated by stress-tests scenarios, or inclusive of a tax levied by the bank on our desk to cover losses generated by these stress test-scenarios. This leads to a different pricing equation than (1.4) – see Appendix A of Chapter 10, page 407.

Exceptions to the rule that break-even levels should not depend on the payoff occur if we explicitly demand that $\widehat{\sigma}$ be an increasing function of $S^2 \frac{d^2 P_{\widehat{\sigma}}}{dS^2}$, to ensure that, for larger gammas, the ratio of theta to gamma is increased, for the sake of conservativeness, with the deliberate consequence that the resulting model is non-linear. One example is the Uncertain Volatility Model, covered in Appendix A of Chapter 2.

Running through the same derivation that led to (1.3), the P&L in the multi-asset case reads:

$$P\&L \;=\; -A\,(t,S)\,\delta t - \frac{1}{2}\Sigma_{ij}\phi_{ij}(t,S)\frac{\delta S_i}{S_i}\frac{\delta S_j}{S_j} \tag{1.6}$$

where $\phi_{ij}(t,S) = S_i S_j \left.\frac{d^2P}{dS_i dS_j}\right|_{t,S}$ and S denotes the vector of the S_i.

Let us diagonalize ϕ, a real symmetric matrix, and denote by φ_k its eigenvalues and T_k the associated eigenvectors. Also denote by φ the diagonal matrix with the φ_k on its diagonal. We have:

$$\phi = T\varphi T^{\intercal}$$

The gamma portion of our P&L can be rewritten as:

$$\Sigma_{ij}\phi_{ij}\frac{\delta S_i}{S_i}\frac{\delta S_j}{S_j} \;=\; U^{\intercal}\phi U \;=\; U^{\intercal}T\varphi T^{\intercal}U \;=\; (T^{\intercal}U)^{\intercal}\varphi(T^{\intercal}U) \;=\; \Sigma_k\varphi_k\delta {z_k}^2$$

where $U_i = \frac{\delta S_i}{S_i}$ and $\delta z_k = T_k^{\intercal}U$. Our P&L now reads:

$$P\&L \;=\; -A\delta t - \frac{1}{2}\Sigma_k\varphi_k\delta {z_k}^2$$

which is a sum of P&Ls of the type in (1.3).

The δz_k are variations of particular baskets of the hedge instruments S_i. These baskets can be considered our effective hedge instruments, since the T_k form a basis.

$\delta {z_k}^2$ is always positive. As in the mono-asset case, the condition for our model to be usable is that there exist n positive numbers ω_k such that:

$$A \;=\; -\frac{1}{2}\Sigma_k\varphi_k\omega_k \tag{1.7}$$

so that our P&L reads:

$$P\&L \;=\; -\frac{1}{2}\Sigma_k\varphi_k\left(\delta {z_k}^2 - \omega_k\delta t\right)$$

Let us express A differently, so as to give our P&L in (1.6) a more symmetrical form. Denote by ω the diagonal matrix with the ω_k on the diagonal. We have:

$$\begin{aligned} A \;&=\; -\frac{1}{2}\Sigma_k\varphi_k\omega_k = -\frac{1}{2}\mathrm{tr}(\varphi\omega) = -\frac{1}{2}\mathrm{tr}(T^{\intercal}\phi T\omega) = -\frac{1}{2}\mathrm{tr}\left(\phi T\omega T^{\intercal}\right) = -\frac{1}{2}\mathrm{tr}\left(\phi C\right) \\ &=\; -\frac{1}{2}\Sigma_{ij}\phi_{ij}C_{ij} \end{aligned}$$

where $C = T\omega T^{\intercal}$ is a positive matrix by construction, as the ω_k are positive.

Our P&L then reads:

$$P\&L \;=\; -\frac{1}{2}\Sigma_{ij}\phi_{ij}\left(\frac{\delta S_i}{S_i}\frac{\delta S_j}{S_j} - C_{ij}\delta t\right) \tag{1.8}$$

Because C is a positive matrix, it can be interpreted as an (implied) covariance matrix; its elements are implied covariance break-even levels.

We have just shown that on the condition that our model is usable, there exists a positive break-even covariance matrix C such that our P&L reads as in (1.8).

In our construction C is given by: $C = T\omega T^{\intercal}$, based on expression (1.7) for A. Is this restrictive, or is P&L (1.8) guaranteed to be nonsensical, for *any* positive matrix C? The answer is yes.[5]

Conclusion

In the general case of multiple hedge instruments, the condition that our model is usable – no situation in which our carry P&L is systematically positive or negative – is that there exists a positive break-even covariance matrix $C(t, S)$, $\forall S, \forall t$, such that the model's theta and cross-gammas are related through:

$$A = -\frac{1}{2}\Sigma_{ij}\phi_{ij}C_{ij}$$

Again, a model not meeting this criterion is unsuitable for trading purposes. In the sequel, suitable models are also called *market models*.

The important thing here is that cross-gammas ϕ_{ij} involve derivatives with respect to values of *actual hedge instruments*, not model-specific state variables.

We will see in Chapter 2 that the local volatility model is a market model, in Chapter 7 that forward variance models are market models, and in Chapter 12 that most local-stochastic volatility models are not.

Specifying a break-even condition for the carry P&L at order 2 in δS leads to pricing equation (1.4). It so happens that the latter – a parabolic equation – has a probabilistic interpretation: the solution can be written as the expectation of the payoff applied to the terminal value of a stochastic process for S_t that is a diffusion: $dS_t = (r-q)S_t dt + \widehat{\sigma} S dW_t$.

The argument goes this way and not the other way around – modeling in finance *does not start* with the assumption of a stochastic process for S_t and has little to do with Brownian motion.

Expression (1.5) is a useful accounting tool – and the Black-Scholes equation (1.4) can be used to risk-manage options – despite the fact that real securities are not lognormal and do not exhibit constant volatility.

[5] Assume our P&L reads as in (1.8) with C an arbitrary positive matrix. We have:

$$A = -\frac{1}{2}\mathrm{tr}(\phi C) = -\frac{1}{2}\mathrm{tr}\left(T\varphi T^{\intercal} C\right) = -\frac{1}{2}\mathrm{tr}\left(\varphi T^{\intercal} C T\right) = -\frac{1}{2}\Sigma_k \varphi_k \left(T_k^{\intercal} C T_k\right) = -\frac{1}{2}\Sigma_k \varphi_k \alpha_k$$

where $\alpha_k = T_k^{\intercal} C T_k$ are positive numbers as C is positive. C can thus be any positive matrix.

1.2 How (in)effective is delta hedging?

Expression (1.5) quantifies the P&L of a short delta-hedged option position. The aim of delta-hedging is to reduce uncertainty in our final P&L – it removes the linear term in δS: is this sufficient from a practical point of view? How large is the gamma/theta P&L (1.5)? More precisely, how large is the average and standard deviation of the total P&L incurred over the option's life?

It can be shown – this is the principal result of the Black-Scholes-Merton analysis – that:

- if the underlying security indeed follows a lognormal process with the same volatility σ as that used for pricing and delta-hedging the option; that is, S follows the Black-Scholes *model* with volatility σ
- and if we take the limit of very frequent hedging: $\delta t \to 0$

then the sum of P&Ls (1.5) incurred over the option's life vanishes with probability 1.

In real life delta-hedging occurs discretely in time, typically on a daily basis, and real securities do not follow diffusive lognormal processes. Thus, the sum of P&Ls (1.5) over the option's life will not vanish. Already in the lognormal case, if S follows a lognormal process but with a different volatility – say higher – than the implied volatility $\widehat{\sigma}$, the sum of P&Ls (1.5) will not vanish in the limit $\delta t \to 0$.

Obviously, the condition that the final P&L vanishes on average requires that the implied volatility $\widehat{\sigma}$ used for pricing and risk-managing the option match on average the future realized volatility weighted by the option's dollar gamma over the option's life:

$$\left\langle \int_0^T e^{-rt} S^2 \frac{d^2 P_{\widehat{\sigma}}}{dS^2} \sigma_t^2 dt \right\rangle = \left\langle \int_0^T e^{-rt} S^2 \frac{d^2 P_{\widehat{\sigma}}}{dS^2} \widehat{\sigma}^2 dt \right\rangle$$

where σ_t is the instantaneous *realized* volatility defined by: $\sigma_t^2 \delta t = \frac{\delta S_t^2}{S_t^2}$ and the discount factor e^{-rt} is used to convert $P\&L$ generated at time t into $P\&L$ at $t = 0$.

Throughout this book, we use $\langle\ \rangle$ to denote either an average or a quadratic (co)variation – context should dispel any ambiguity as to which is intended.

Let us assume that this condition holds, so that our final P&L is not biased on average and let us concentrate on the dispersion – the standard deviation – of the final P&L. It vanishes in the Black-Scholes case with continuous hedging. How large is it, first in the Black-Scholes case with discrete hedging and then in the case of discrete hedging with real securities?

Assume that the option is delta-hedged daily at times t_i: $\delta t = 1$ day. The total P&L over the option's life, discounted at time $t = 0$, is:

$$P\&L \;=\; -\sum_i e^{-rt_i}\frac{S_i^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dS^2}\left(t_i, S_i\right)\left(r_i^2 - \widehat{\sigma}^2\delta t\right) \tag{1.9}$$

where r_i are daily returns, given by $r_i = \frac{S_{i+1}-S_i}{S_i}$. As expression (1.9) shows, at order 2 in δS and order 1 in δt, the total P&L is given by the sum of the differences between *realized* daily quadratic variation $\frac{\delta S_i^2}{S_i^2}$ and the *implied* quadratic variation $\widehat{\sigma}^2\delta t$, weighted by the prefactor $e^{-rt_i}\frac{S_i^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dS^2}\left(t_i, S_i\right)$, which is payoff-dependent and involves the gamma of the option. $\widehat{\sigma}$ is the implied volatility we are using to risk-manage our option position.

Let us make the approximation that the option's discounted dollar gamma $e^{-rt_i}S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}$ is a constant, equal to its initial value $S_0^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}\left(t_0, S_0\right)$ – this removes one source of randomness in the P&L.[6] The standard deviation of the total P&L depends on the variances of individual daily P&Ls as well as on their covariances. Let us write the daily return r_i as:

$$r_i \;=\; \sigma_i\sqrt{\delta t}z_i \tag{1.10}$$

where σ_i is the realized volatility for day i, and z_i is centered and has unit variance: $\left\langle z_i\right\rangle = 0$, $\left\langle z_i^2\right\rangle = 1$. Let us assume that the z_i are iid and are independent of the volatilities σ_i.

Because the z_i are independent, returns r_i have no serial correlation but are not independent, as daily volatilities σ_i may be correlated. Expression (1.10) allows separation of the effects of the scale σ_i of return r_i on one hand, and of the distribution of r_i – which up to a rescaling is given by that of z_i – on the other hand. Our total P&L now reads:

$$P\&L \;=\; -\frac{S_0^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dS^2}\left(t_0, S_0\right)\sum_i\left(\sigma_i^2 z_i^2 - \widehat{\sigma}^2\right)\delta t$$

Let us assume that the process for the σ_i is time-homogeneous so that, in particular, $\left\langle\sigma_i^2\right\rangle$ does not depend on i and let us take $\widehat{\sigma}^2 \;=\; \left\langle\sigma_i^2\right\rangle$. The variance of

[6]There exists actually a European payoff whose discounted dollar gamma is constant and equal to 1. It is called the log contract and pays at maturity $-2\ln S$; see Section 3.1.4.

$\sum_i \left(\sigma_i^2 z_i^2 - \widehat{\sigma}^2\right)\delta t$ is given by:

$$
\begin{aligned}
&\Big\langle \sum_{ij} (\sigma_i^2 z_i^2 - \widehat{\sigma}^2)\delta t \; (\sigma_j^2 z_j^2 - \widehat{\sigma}^2)\delta t \Big\rangle \\
&= \sum_i \left(\left\langle \sigma_i^4 z_i^4 \right\rangle \, + \widehat{\sigma}^4 - 2\widehat{\sigma}^4 \right) \delta t^2 \; + \sum_{i\neq j} \left\langle \sigma_i^2 \sigma_j^2 z_i^2 z_j^2 \; + \widehat{\sigma}^4 - 2\widehat{\sigma}^4 \right\rangle \delta t^2 \\
&= \sum_i \left(2+\kappa\right) \widehat{\sigma}^4 \delta t^2 \; + \quad \sum_{i\neq j} (\left\langle \sigma_i^2 \sigma_j^2 \right\rangle - \widehat{\sigma}^4) \delta t^2 \\
&= \sum_i \left(2+\kappa\right) \widehat{\sigma}^4 \delta t^2 \; + (\left\langle \sigma^4 \right\rangle - \widehat{\sigma}^4) \sum_{i\neq j} f_{ij} \delta t^2 \\
&= \widehat{\sigma}^4 \Big(\sum_i \left(2+\kappa\right) \delta t^2 \; + \Omega \sum_{i\neq j} f_{ij} \delta t^2 \Big)
\end{aligned}
\tag{1.11}
$$

where we have introduced the (excess) kurtosis κ of returns r_i and the variance/variance correlation function f defined by:

$$
\kappa \;=\; \frac{\left\langle \sigma_i^4 z_i^4 \right\rangle}{\widehat{\sigma}^4} - 3, \quad f_{ij} \;=\; \frac{\left\langle (\sigma_i^2 - \widehat{\sigma}^2)(\sigma_j^2 - \widehat{\sigma}^2) \right\rangle}{\sqrt{\left\langle \sigma_i^4 \; \right\rangle - \widehat{\sigma}^4} \sqrt{\left\langle \sigma_j^4 \right\rangle - \widehat{\sigma}^4}}
$$

and where the dimensionless factor Ω, which quantifies the variance of daily variances σ_i^2 is given by:

$$
\Omega \;=\; \frac{\left\langle \sigma^4 \right\rangle - \widehat{\sigma}^4}{\widehat{\sigma}^4} \;=\; \frac{\left\langle \sigma^4 \right\rangle - \left\langle \sigma^2 \right\rangle^2}{\left\langle \sigma^2 \right\rangle^2}
$$

We then get:

$$
\mathrm{StDev}\left(P\&L\right) \;=\; \left| \frac{S_0^2}{2} \frac{d^2 P_{\widehat{\sigma}}}{dS^2} \left(t_0, S_0\right) \right| \sqrt{\widehat{\sigma}^4 \Big(\sum_i \left(2+\kappa\right) \delta t^2 \; + \Omega \sum_{i\neq j} f_{ij} \delta t^2 \Big)}
$$

It is useful to measure the standard deviation of the final P&L in units of the option's vega, the sensitivity of the option's price to the implied volatility $\widehat{\sigma}$. In the Black-Scholes model, for European options the following relationship linking vega and gamma holds:

$$
\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}} \;=\; S^2 \frac{d^2 P_{\widehat{\sigma}}}{dS^2} \widehat{\sigma} T \tag{1.12}
$$

where T is the residual option's maturity – this is derived in Appendix A of Chapter 5, page 181. Using now the vega, the final expression for the standard deviation of the P&L is:

$$
\mathrm{StDev}\left(P\&L\right) \;=\; \left| \widehat{\sigma} \frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}} \right| \frac{1}{2T} \sqrt{\sum_i \left(2+\kappa\right) \delta t^2 \; + \Omega \sum_{i\neq j} f_{ij} \delta t^2} \tag{1.13}
$$

1.2.1 The Black-Scholes case

Let us first assume that S follows the lognormal Black-Scholes dynamics. σ_i is constant, equal to $\widehat{\sigma}$, hence $\Omega = 0$. Since $\sigma_i\sqrt{\delta t}$ is small (δt is one day and typically $\sigma_i = 20\%$, so that $\sigma_i\sqrt{\delta t} \simeq 0.01$), daily returns can be considered Gaussian: $\kappa = 0$. $\Sigma_i (2+\kappa)\,\delta t^2 = \frac{2T^2}{N}$, where T is the option's maturity and N is the number of delta rehedges: $N\delta t = T$. Expression (1.13) becomes:

$$\mathrm{StDev}\,(P\&L) = \frac{1}{\sqrt{2N}}\left|\widehat{\sigma}\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}}\right| \tag{1.14}$$

Thus, provided it is small, the standard deviation of our final P&L is equivalent to the impact on the option's price of a relative perturbation of $\widehat{\sigma}$ of size $\frac{1}{\sqrt{2N}}$.

Note that $\frac{\widehat{\sigma}}{\sqrt{2N}}$ is approximately the standard deviation of the historical volatility estimator. The standard variance estimator is given by:

$$\overline{\sigma}^2 = \frac{1}{N\delta t}\sum_i\left(\frac{S_{i+1}-S_i}{S_i}\right)^2$$

In the Black-Scholes case, for daily returns, $\frac{S_{i+1}-S_i}{S_i}$ is approximately Gaussian and we have:

$$\overline{\sigma}^2 \simeq \frac{\widehat{\sigma}^2}{N}\sum_i z_i^2$$

where z_i are standard normal random variables. The variance of $\overline{\sigma}^2$ is $\frac{2\widehat{\sigma}^4}{N}$, thus the *relative* standard deviation $\mathrm{StDev}\big(\overline{\sigma}^2\big)\,/\,\big\langle\overline{\sigma}^2\big\rangle$ is $\sqrt{\frac{2}{N}}$ and, if it is not too large, the *relative* standard deviation of the *volatility* estimator $\overline{\sigma}$ is approximately half of this, that is $\frac{1}{\sqrt{2N}}$.

The standard deviation observed on our final P&L is then approximately given by the option's vega multiplied by the standard deviation of the volatility estimator built on the same schedule as that of the delta rehedges.

Consider the example of a one-year at-the-money call option, with $\widehat{\sigma} = 20\%$, vanishing interest rates, repo and dividends, and $S = 1$. The option's price is then $P = 7.97\%$. There are about 250 trading days in one year, which gives $\frac{1}{\sqrt{2N}} \simeq 0.045$. An at-the-money option has the property that its price is approximately linear in $\widehat{\sigma}$ for short maturities: $P_{\widehat{\sigma}} \simeq \frac{1}{\sqrt{2\pi}}S\widehat{\sigma}\sqrt{T}$, thus $\widehat{\sigma}\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}} \simeq P$ (using this approximation yields a price of 7.98%).

We then get for the one-year at-the-money option: $\mathrm{StDev}(P\&L) \simeq 0.045P$: the standard deviation of our final P&L is about 5% of the option's price we charged at inception.

5% of the premium charged for the option – or equivalently 5% of the volatility – may seem a very reasonable risk to take. Alternatively, adjusting the option's price to account for one standard deviation of our final P&L would result in a relative bid/offer spread on the option price of about 10%.

1.2.2 The real case

In real life, in contrast to the Black-Scholes case, the second term in the square root in (1.13) does not vanish. It involves the variance/variance correlation function f_{ij}. We have made the (reasonable) assumption that the process for the σ_i is time-homogeneous: f_{ij} is then a function of the difference $j-i$, actually a function of $|j-i|$.

As δt is small compared to the option's maturity, we convert the sums in (1.11) into integrals:

$$\sum_{ij} f_{ij}\delta t^2 \;\simeq\; \int_0^T du \int_0^T dt f\left(t-u\right) \;=\; 2\int_0^T \left(T-\tau\right) f\left(\tau\right) d\tau$$

We now have from (1.13):

$$\begin{aligned}\mathrm{StDev}\left(P\&L\right) \;&\simeq\; \left|\widehat{\sigma}\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}}\right| \frac{1}{2T}\sqrt{\left(2+\kappa\right)\frac{T^2}{N} \;+\; 2\Omega\int_0^T \left(T-\tau\right) f\left(\tau\right) d\tau} \\ &= \left|\widehat{\sigma}\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}}\right| \sqrt{\frac{2+\kappa}{4N} \;+\; \frac{\Omega}{2T^2}\int_0^T \left(T-\tau\right) f\left(\tau\right) d\tau} \end{aligned} \tag{1.15}$$

Let us now examine the two contributions to StDev$(P\&L)$.

Imagine first that daily variances are constant: Ω vanishes and the first piece alone contributes to the standard deviation of the P&L. Just as in the Black-Scholes case (equation (1.14)), the variance of the final P&L scales like $1/N$, where N is the number of daily rehedges, which is natural as the final P&L is the sum of N identically distributed and independent daily P&Ls.

In contrast to the Black-Scholes case though, in which daily returns are approximately Gaussian, the effect of the tails of the distribution of daily returns appears through the kurtosis κ. By setting $\kappa = 0$ we recover result (1.14).

Consider now the second contribution in (1.15). The prefactor Ω quantifies the dispersion of daily variances while $f\left(\tau\right)$ quantifies how a fluctuation in daily variance σ_i^2 on day t_i impacts daily variances $\sigma_{i+\tau}^2$ on subsequent days. If f decays slowly, daily variances will be very correlated: in case one daily variance was higher than $\widehat{\sigma}$, daily variances for the following days are likely to be higher as well, resulting in daily gamma/theta P&Ls all having the same sign – thus generating strong correlation among daily P&Ls and increasing the variance of our final P&L.

For example, assume that daily variances are perfectly correlated: $f\left(\tau\right) = 1$. The second piece in (1.15) is then simply equal to $\frac{\Omega}{4}$. If Ω is small, the contribution of this term is then equivalent to the impact of a relative displacement of $\widehat{\sigma}$ by $\widehat{\sigma}\frac{\sqrt{\Omega}}{2}$, regardless of the number N of daily rehedges.[7]

[7]The case $f\left(\tau\right) = 1$ is unrealistic in that daily variances are random, but are all identical: the underlying security follows a lognormal dynamics with a constant volatility whose value is drawn randomly at inception.

Estimating $f(\tau), \Omega, \kappa$

Consider now the dynamics of daily variances σ_i in the case of real securities. Separating in r_i the contributions from σ_i and z_i is difficult if the only daily data we have are daily returns. In what follows we have estimated daily volatilities σ_i using 5-minute returns: σ_i is given by the square root of the sum of squared 5-minute returns during the exchange's opening hours, plus the square of the close-to-open return. Figure 1.1 shows the autocorrelation function f averaged over a set of European financial stocks, evaluated on a two-year sample: [August 2008, August 2010].[8]

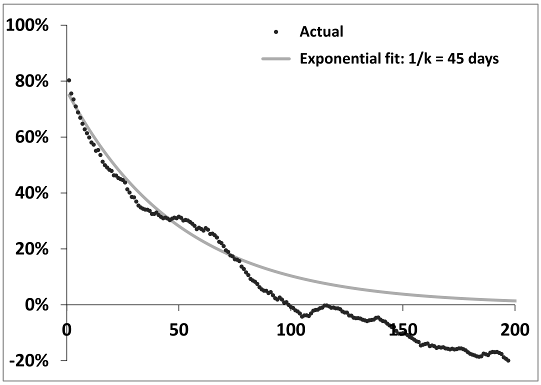

Figure 1.1: Correlation function $f(\tau)$ of daily volatilities evaluated on a basket of financial stocks. τ is in business days.

For $\tau = 0$, $f(\tau) = 1$. As is customary with correlation functions, however, $\lim_{\tau \to 0} f(\tau) \neq 1$, and the discontinuity in $\tau = 0$ quantifies the signal-to-noise ratio of our measurement of daily volatilities. As Figure 1.1 shows, this discontinuity is rather moderate and we get a robust estimation of the autocorrelation of daily volatilities up to time scales $\tau \simeq 100$ days.

For larger τ, Figure 1.1 displays negative autocorrelations: this is unphysical and most likely due to the fact that, over our historical sample (2 years), for $\tau > 100$ days, i and $i + \tau$ fall into two different regimes of respectively low and high volatilities.

We have also graphed in Figure 1.1 an exponential fit to f: $f(\tau) = \rho e^{-k\tau}$, with $\rho = 0.78$ and $\frac{1}{k} = 45$ days. The agreement of f with the exponential form is acceptable in the region $\tau < 100$ days, where our measurement is reliable.

Using this form for $f(\tau)$ yields our final expression for the standard deviation of the P&L:

$$\frac{\text{StDev}\,(P\&L)}{\left|\widehat{\sigma}\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}}\right|} \simeq \sqrt{\frac{2+\kappa}{4N} + \frac{\rho\Omega}{2}\,\frac{kT - 1 + e^{-kT}}{(kT)^2}} \tag{1.16}$$

[8] I am grateful to Benoît Humez for generating these data as well as estimates of Ω.

Ω quantifies the relative variance of daily variances σ_i^2. It varies appreciably, even among stocks of the same sector: a typical range for Ω is $[1.5,\ 4]$. Let us use the value $\Omega = 2$.

Estimating the unconditional kurtosis κ is also difficult, as the 4th order moment of daily returns converges slowly, so slowly that it is unreasonable to assume that the same regime of kurtosis holds throughout the historical sample used for its estimation: a typical order of magnitude is $\kappa = 5$.

1.2.3 Comparing the real case with the Black-Scholes case

We now use the typical values for Ω, κ, ρ, k estimated above in expression (1.16). Figure 1.2 shows the right-hand side of equation (1.16), that is the relative displacement of $\widehat{\sigma}$ that produces a variation of the option's price P equal to one standard deviation of the P&L. For an at-the-money option, whose price is approximately linear in $\widehat{\sigma}$, this number is also the ratio of one standard deviation of the P&L to the option's price itself.

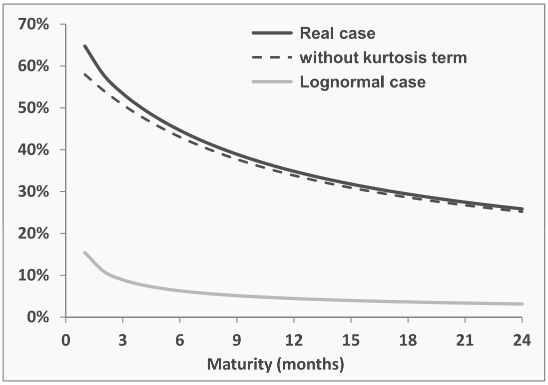

Figure 1.2: Right-hand side in equation 1.16 (darker line), as a function of maturity, compared to the same quantity, but without the kurtosis term (dashed line), and the lognormal case (lighter line).

Figure 1.2 also displays the same quantity, but without the term $\frac{2+\kappa}{4N}$, to remove the effect of the tails of the daily returns, as well as the standard deviation of the P&L in the lognormal, Black-Scholes case (1.14).

We can see that the standard deviation of the final P&L of a delta-hedged option in the real case is much larger than its estimation in the Black-Scholes case.

Consider again the example of a 1-year at-the-money option, with $\widehat{\sigma} = 0.2$, with $P = 7.97\%$. As Figure 1.2 shows, while in the Black-Scholes case, the standard deviation is 4.5% of the option's price, that is 0.35%, in the real case, for a 1-year maturity it is equal to 35% of the option's price, that is 2.8%.

Comparison of the dark and dashed lines in Figure 1.2 shows that, but for very short maturities, the dispersion of the P&L is mostly generated by correlation of daily volatilities rather than the thickness of the tails of daily returns.

Delta-hedging our one-year at-the-money option position exposes us to the risk of making or losing about one third of the option premium[9] – this is an unreasonable risk to take considering that, typically, commercial fees charged by banks on option transactions are much smaller than the option's value.

The conclusion is that, in real life, delta-hedging is not sufficient: while delta-hedging removes the linear term in δS in our daily P&L, the effect of the δS^2 term is still too large: the only way to remove it is to use other options – for example vanilla options – to offset the gamma of the option we are risk-managing.

This was expressed bluntly to the author upon starting his career in finance by Nazim Mahrour, an FX option trader: "options are hedged with options".

1.3 On the way to stochastic volatility

Let us then use other options to offset the gamma of the exotic option we are risk-managing: assume for simplicity that we use a single vanilla option, whose implied volatility is $\widehat{\sigma}_O$. The P&L of a delta-hedged position in the vanilla option O has the same form as in equation (1.5), except it involves the implied volatility $\widehat{\sigma}_O$:

$$P\&L_O \;=\; -\frac{S^2}{2}\frac{d^2O}{dS^2}\left(\frac{\delta S^2}{S^2} - \widehat{\sigma}_O^2\delta t\right) \tag{1.17}$$

The number λ of vanilla options O we are buying as gamma hedge is :

$$\lambda \;=\; \frac{1}{\frac{d^2O}{dS^2}}\frac{d^2P}{dS^2} \tag{1.18}$$

The gamma profiles of P and O are unlikely to be homothetic, thus this gamma hedge will be efficient only locally; as time elapses and S moves, we need to readjust the hedge ratio λ.

We could decide to risk-manage each option P and O with its own implied volatility $\widehat{\sigma}$ and $\widehat{\sigma}_O$, but this leads to incongruous carry P&Ls.

Indeed by selecting λ as specified in (1.18) we cancel the gamma of the hedged position. The P&Ls of options O and P are both of the form in (1.17). If $\widehat{\sigma} \neq \widehat{\sigma}_O$, the theta portion of our global P&L does not vanish, even though gamma vanishes, a situation as nonsensical as those encountered in Section 1.1, when A and B have the same sign – see also the discussion in Section 2.8 below.

[9]Remember that we have made the unrealistic assumption that we were able to predict the average realized volatility. Uncertainty about the future average level of realized volatility would push the standard deviation of our final P&L even higher.

We must thus choose $\widehat{\sigma} = \widehat{\sigma}_O$. We now have for P a pricing function that explicitly depends on two dynamical variables: S and $\widehat{\sigma}_O$:

$$P\left(t,\ S,\ \widehat{\sigma}_O\right)$$

which is natural as we are using two instruments as hedges.

This is an elementary instance of calibration: we *decide* to make our exotic option's price a function of other derivatives' prices. It is a trading decision.

In the unhedged case we were free to chose the implied volatility $\widehat{\sigma}$ as our best estimate of future realized volatility and kept it constant throughout: no P&L was generated by the variation of $\widehat{\sigma}$.

Unlike $\widehat{\sigma}$, however, $\widehat{\sigma}_O$ is a market implied volatility and cannot be kept constant. As S moves and time flows we readjust λ, thus buying or selling the vanilla option at prevailing market prices: $\widehat{\sigma}_O$ will move so as to reflect the market price O of the vanilla option. Daily P&Ls for O and P will include extra terms involving $\delta\widehat{\sigma}_O$. At second order in $\delta\widehat{\sigma}_O$:

$$P\&L_O = -\frac{S^2}{2}\frac{d^2O}{dS^2}\left(\frac{\delta S^2}{S^2} - \widehat{\sigma}_O^2\delta t\right) - \frac{dO}{d\widehat{\sigma}_O}\delta\widehat{\sigma}_O - \frac{1}{2}\frac{d^2O}{d\widehat{\sigma}_O^2}\delta\widehat{\sigma}_O^2 - \frac{d^2O}{dSd\widehat{\sigma}_O}\delta S\delta\widehat{\sigma}_O \tag{1.19}$$

The expansion of the P&L of the *hedged* position at second order in δS, $\delta\widehat{\sigma}_O$ and order 1 in δt reads:

$$\begin{aligned} P\&L &= -\left(\frac{dP}{d\widehat{\sigma}_O} - \lambda\frac{dO}{d\widehat{\sigma}_O}\right)\delta\widehat{\sigma}_O \qquad (1.20) \\ &\quad - \frac{1}{2}\left(\frac{d^2P}{d\widehat{\sigma}_O^2} - \lambda\frac{d^2O}{d\widehat{\sigma}_O^2}\right)\delta\widehat{\sigma}_O^2 - \left(\frac{d^2P}{dSd\widehat{\sigma}_O} - \lambda\frac{d^2O}{dSd\widehat{\sigma}_O}\right)\delta S\delta\widehat{\sigma}_O \end{aligned}$$

This is an accounting equation: the P&L generated by these three terms is no less real than the usual gamma/theta P&L – it is usually called *mark-to-market* P&L, while the gamma/theta P&L is typically called *carry* P&L.[10]

There is no contribution from δS^2 as $\frac{d^2P}{dS^2} = \lambda\frac{d^2O}{dS^2}$ by construction. Exotic options are typically path-dependent options: their final payoff is a function of values of S observed at discrete dates, specified in the option's term sheet. Between two observation dates, the pricing equation for P in the Black-Scholes framework is the same as that of a European option. Since P and O are given by a Black-Scholes pricing equation with the same implied volatility $\widehat{\sigma}_O$, cancellation of gamma implies cancellation of theta as well: there is no δt term in (1.20).

Consider the last two terms in $\delta\widehat{\sigma}_O^2$ and $\delta S\delta\widehat{\sigma}_O$ in (1.19) and (1.20). While their contributions to $P\&L_O$ and $P\&L$ look similar, they have a different status and have to be treated differently. Expression (1.19) is the P&L of a vanilla option position.

[10]The distinction between mark-to-market and carry P&L is somewhat arbitrary. Usually mark-to-market P&L refers to P&L generated by the variation of parameters that were supposed to stay constant in the pricing model: typically, in the Black-Scholes model a change in $\widehat{\sigma}$ generates mark-to-market P&L.

The extra terms that come in addition to the gamma/theta P&L do not warrant any adjustment to the price of the vanilla option: their contribution to the P&L is already priced-in in the market price of the vanilla option.

In expression (1.20), however, what appear as prefactors of $\delta\widehat{\sigma}_O^2$ and $\delta S \delta\widehat{\sigma}_O$ are the second-order sensitivities of the *hedged* position. We then need to adjust the price $P(t, S, \widehat{\sigma}_O)$ of our exotic option for the cost of these two contributions to the P&L.

What matters in the evaluation of extra-model cost is not so much the second-order sensitivities of the *naked* exotic option, but the residual sensitivities of the *hedged* position.

Three observations are in order:

- We now have a vega term in $\delta\widehat{\sigma}_O$. If P is a European option with the same maturity as O, the vega of a gamma-hedged position cancels out, owing to relationship (1.12) linking gamma and vega in the Black-Scholes model. A European payoff is statically hedged with a portfolio of vanilla options of the same maturity; it can hardly be called an exotic derivative.

  The situation we have in mind is that of real exotics that has no static hedge, whose hedge portfolio comprises vanilla options of different maturities: gamma cancellation does not imply vega cancellation. Depending on the relative sizes of the gamma and vega risks we may prefer to gamma-hedge or vega-hedge our exotic option: this is a trading decision. In practice an exotics book is a large caldron where mitigation of the gamma and vega risks of many different exotic and vanilla options takes place: gamma and vega hedging, unachievable on a deal-by-deal basis, can be reasonably achieved at the book level.[11]

- Our P&L does not involve realized volatility anymore. Instead, we have acquired sensitivity to $\widehat{\sigma}_O$. While in the unhedged case we were exposed to *realized* volatility, we are now exposed to the dynamics of the *implied* volatility $\widehat{\sigma}_O$.[12]

- Unlike in the unhedged case for the δS^2 term, no deterministic δt term is now offsetting the $\delta\widehat{\sigma}_O^2$ and $\delta S \delta\widehat{\sigma}_O$ terms: depending on their realized values and the signs of their prefactors, we may systematically make or lose money. This is a serious issue. While in the Black-Scholes pricing equation we had a parameter – the implied volatility – to control how the gamma and theta terms for the spot offset each other, we have no equivalent parameter at our disposal to control break-even levels for gammas on $\widehat{\sigma}_O$: no implied volatility of $\widehat{\sigma}_O$ and no implied correlation of S and $\widehat{\sigma}_O$. P and O should then be given by a different pricing equation than Black-Scholes', that explicitly includes

[11] Client preferences, pressure from the salesforce, unwillingness of other counterparties to take on exotic risks, may lead an exotics desk to pile up one-way risk. In normal circumstances, though, as exposure to a particular risk builds up, traders will be willing to quote aggressive prices for payoffs that offset this risk so as to keep the overall risk levels of the book under control.

[12] This is not exactly true – there remains a residual sensitivity to realized volatility in the covariance term $\delta S \delta\widehat{\sigma}_O$.

these new parameters so as to generate additional theta terms in the P&L: this is the general task of stochastic volatility models.[13]

The general conclusion is that by using options as hedges we lower – or cancel – our exposure to realized volatility, but acquire an exposure to the dynamics of implied volatilities. However, while the Black-Scholes pricing equation provides a theta term to offset the gamma term for S, no provision of a theta is made to offset the gamma P&Ls experienced on the variation of implied volatilities of options used as hedges.

This is not surprising as the notion of dynamic implied volatilities is alien to the Black-Scholes framework.

This is where stochastic volatility models are called for: their aim is not to model the dynamics of *realized* volatility, which is hedged away by trading other options, but to model the dynamics of *implied* volatilities, and provide their user with simple break-even accounting conditions for the P&L of a hedged position.

---

[13] The vanna-volga method – see [29] – once used on FX desks for generating FX smiles is a poor man's answer to this issue, with "exotic" option P a European option.

- Rather than using a single vanilla option O we use 3 of them, and find quantities λ_i so that the 3 sensitivities $\frac{d}{d\sigma}$, $\frac{d^2}{dSd\sigma}$, $\frac{d^2}{d\sigma^2}$ of the hedged position $P - \Sigma_{\iota=1}^{3}\lambda_i O_i$ vanish (a) in the Black-Scholes model for an implied volatility $\widehat{\sigma}_0$, (b) for current values of t, S. Cancellation of $\frac{d}{d\sigma}$ is equivalent to cancellation of $\frac{d^2}{dS^2}$, owing to the vega/gamma relationship in the Black-Scholes model – see Section A.1 of Chapter 5.
- The hedging options are bought/sold at market prices, at implied volatilities $\widehat{\sigma}_i$, thus the difference $O_i^{\mathrm{BS}}(\widehat{\sigma}_i) - O_i^{\mathrm{BS}}(\widehat{\sigma}_0)$ has to be passed on to the client as a hedging cost. We thus define the "market-adjusted" price P^{Mkt} of option P as:

$$P^{\mathrm{Mkt}} = P^{\mathrm{BS}}(\widehat{\sigma}_0) + \Sigma_i \lambda_i \left(O_i^{\mathrm{BS}}(\widehat{\sigma}_i) - O_i^{\mathrm{BS}}(\widehat{\sigma}_0) \right) \tag{1.21}$$

  The hedge portfolio is only effective for current values of t, S. It needs to be readjusted whenever either moves – the corresponding rehedging costs are not factored in P^{Mkt}.
- As observed in [29], the vanna-volga price in (1.21) can be written as:

$$\begin{aligned} P^{\mathrm{Mkt}} &= P^{\mathrm{BS}}(\widehat{\sigma}_0) + y_\sigma \left.\frac{dP^{\mathrm{BS}}}{d\widehat{\sigma}_0}\right|_{S,\widehat{\sigma}_0} + y_{\sigma^2} \left.\frac{d^2P^{\mathrm{BS}}}{d\widehat{\sigma}_0^2}\right|_{S,\widehat{\sigma}_0} + y_{S\sigma} \left.\frac{d^2P^{\mathrm{BS}}}{dSd\widehat{\sigma}_0}\right|_{S,\widehat{\sigma}_0} \\ &= P^{\mathrm{BS}}(\widehat{\sigma}_0) + y_{S^2} \left.\frac{d^2P^{\mathrm{BS}}}{dS^2}\right|_{S,\widehat{\sigma}_0} + y_{\sigma^2} \left.\frac{d^2P^{\mathrm{BS}}}{d\widehat{\sigma}_0^2}\right|_{S,\widehat{\sigma}_0} + y_{S\sigma} \left.\frac{d^2P^{\mathrm{BS}}}{dSd\widehat{\sigma}_0}\right|_{S,\widehat{\sigma}_0} \end{aligned} \tag{1.22}$$

  where the second line again follows from the vega/gamma relationship in the Black-Scholes model: $y_{S^2} = y_\sigma S^2 \widehat{\sigma}_0 T$. The interpretation of (1.22) is: we supplement the Black-Scholes price at implied volatility $\widehat{\sigma}_0$ with an estimation of future gamma P&Ls calculated (a) with current values of the gammas and cross-gammas, (b) values for $y_{S^2}, y_{S\sigma}, y_{\sigma^2}$ such that market prices for the three vanilla options O_i are recovered; $y_{S^2}, y_{S\sigma}, y_{\sigma^2}$ only depend on the $\widehat{\sigma}_i$, not on P. This underscores how local the vanna-volga adjustment is – it cannot replace a genuine model for pricing volatility-of-volatility risk.
- Historically, the vanna-volga method has been used for interpolating implied volatilities: pick a vanilla option of strike K and use (1.21) to generate the corresponding adjusted "market price" – hence implied volatility. There is obviously no guarantee that the resulting interpolation $\widehat{\sigma}^{\mathrm{Mkt}}(K, \widehat{\sigma}_0, \widehat{\sigma}_i)$ is arbitrage-free.

In practice, for liquid securities such as equity indexes, there are plenty of options available: rather than one implied volatility $\widehat{\sigma}_O$, one needs to model the dynamics of all implied volatilities $\widehat{\sigma}_{KT}$, where K and T are, respectively, the strikes and maturities of vanilla options. The two-dimensional set $\widehat{\sigma}_{KT}$ is known as the *volatility surface*.

While a stochastic volatility model should ideally offer maximum flexibility as to the range of dynamics of the volatility surface it is able to produce, we may not be able to build such a flexible model on one hand, and on the other hand we may not need so much versatility: some classes of exotic options are only sensitive to specific features of the dynamics of the volatility surface.

Before we delve into stochastic volatility models, we present two examples of exotic options whose type of volatility risk can be exactly pinpointed.

1.3.1 Example 1: a barrier option

Consider an option of maturity one year that pays at maturity 1 unless S_t hits the barrier $L = 120$, in which case the option expires worthless. The initial spot value is $S_0 = 100$. The pricing function $F(t, S)$ of this barrier option has to satisfy the terminal condition at maturity: $F(T, S) = 1$, for $S < L$ as well as the boundary condition $F(t, L) = 0$ for all $t \in [0, T]$.

How do we hedge this barrier option with vanilla options? Peter Carr and Andrew Chou show in [22] that, given a barrier option with payoff $f(S)$ and upper barrier L, it is possible to find a European payoff $g(S)$ of maturity T such that in the Black-Scholes model its value $G(t, S)$ exactly equals that of the barrier option, $F(t, S)$ for $S \leq L$, at all times.

The condition that $G(t, S) = F(t, S)$ at $t = T$ implies that $g(S) = f(S)$ for $S < L$. For $S > L$, f is not defined, but we have to find $g(S)$ such that $G(t, S = L)$ vanishes for all $t < T$.

Imagine we are able to find g such that this condition is satisfied. Then we have a European payoff that: (a) has the same final payoff as the barrier option, (b) satisfies the same boundary condition for $S = L$ and (c) solves the same pricing equation over $[0, L]$: this implies that $F(t, S) = G(t, S)$ for all $S \in [0, L]$, $t \in [0, T]$: the barrier option is statically hedged by the European payoff G.

Carr and Chou give the following explicit expression for g, in the Black-Scholes model:

$$S < L \quad g(S) = f(S) \tag{1.23a}$$

$$S > L \quad g(S) = -\left(\frac{L}{S}\right)^{\frac{2r}{\sigma^2}-1} f\left(\frac{L^2}{S}\right) \tag{1.23b}$$

where r is the interest rate and σ the volatility. Let us assume vanishing interest rates. The replicating European payoff for our barrier options is:

$$\begin{aligned} S<L \quad g(S) &= 1 \\ S>L \quad g(S) &= -\frac{S}{L} = -1-\frac{1}{L}(S-L)^+ \end{aligned}$$

This static hedge thus consists of two European digital options struck at L, each of which pays 1 if $S_T < L$ and 0 otherwise, minus (a) one zero-coupon bond that pays $1, \forall S_T$, and (b) $\frac{1}{L}$ call options of strike L. S_T is the value of S at maturity.

Equations (1.23a), (1.23b) for $g(S)$ show that if $f(L) \neq 0$, g has a discontinuity in $S = L$ whose magnitude is twice that of f. The replicating European payoff includes a digital option whose role is instrumental in replicating the sharp variation of F in the vicinity of L.

Let us consider for simplicity that we are only using the double European digital option: it pays 1 at T if $S_T < L$ and -1 if $S_T > L$. Even though European digitals are not liquid, they can be synthesized just like any European payoff by trading an appropriate set of vanilla options, in our case a very tight put spread, that is the combination of $\frac{1}{2\varepsilon}$ puts struck at $L+\varepsilon$ minus $\frac{1}{2\varepsilon}$ puts struck at $L-\varepsilon$.

The values of the barrier option, F, and of the double European digital option – minus the zero-coupon bond – are shown as a function of S at $t = 0$ on the left-hand side of Figure 1.3 while the right-hand side shows the dollar gamma for both options. We have used $\sigma = 20\%$.

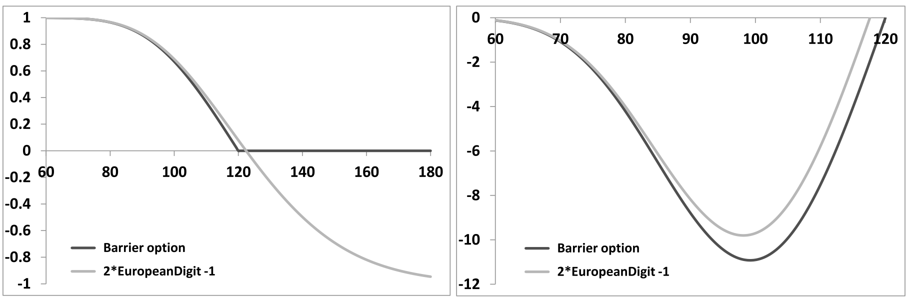

Figure 1.3: Value (left) and dollar gamma (right) of barrier option and double European digital option.

Had we used the exact static European hedge, curves would have overlapped exactly, in both graphs, by construction. Simply using the double European digital option still provides an acceptable hedge. Let us assume that we have sold at $t = 0$ the barrier option and have simultaneously purchased the double European digital as a hedge.

Which price do we quote for the barrier option? We are using as hedge a double European digital option whose market price will likely differ from its Black-Scholes price. The price we charge must thus be equal to the Black-Scholes price of the barrier option augmented by the difference between market and Black-Scholes prices of the double European digital: this extra charge covers the cost of actually purchasing the European hedge.[14]

If we reach maturity without hitting the barrier $L = 120$, the payoffs of the barrier option and the static European hedge exactly match: the hedge is perfect.

What if instead S hits the barrier? When S hits L at time τ, the barrier option expires worthlessly and we need to unwind our static European hedge. By construction, in the Black-Scholes model, its value for $S = L$ approximately vanishes.[15]

How about in reality? In reality, the value of our European static hedge will depend on market implied volatilities at time τ for European options of maturity T and will likely not vanish.

Let us make this dependence more explicit: the value D of the double European digital is given by:

$$D = 2\frac{\mathcal{P}_{L+\varepsilon} - \mathcal{P}_{L-\varepsilon}}{2\varepsilon} - 1 = 2 \left.\frac{d\mathcal{P}_K}{dK}\right|_L - 1$$

where $\mathcal{P}_K$ denotes the value of a put option of strike K, which is given by the Black-Scholes formula for put options, using the implied volatility for strike K : $\mathcal{P}_K = \mathcal{P}_K^{BS}(\widehat{\sigma} = \widehat{\sigma}_K)$. We have:

$$\begin{aligned} \frac{d\mathcal{P}_K}{dK} &= \frac{d\mathcal{P}_K^{BS}(\widehat{\sigma}_K)}{dK} = \frac{d\mathcal{P}_K^{BS}}{dK} + \frac{d\mathcal{P}_K^{BS}}{d\widehat{\sigma}}\frac{d\widehat{\sigma}_K}{dK} \\ &= \mathcal{D}_K^{BS}(\widehat{\sigma}_K) + \frac{d\mathcal{P}_K^{BS}}{d\widehat{\sigma}}\frac{d\widehat{\sigma}_K}{dK} \end{aligned}$$

where $\mathcal{D}^{BS}$ is the value of a (single) European digital option, which pays 1 if $S_T < L$ and 0 otherwise, in the Black-Scholes model. We get the following value for the double European digital:

$$\mathcal{D} = 2\left(\mathcal{D}_L^{BS}(\widehat{\sigma}_L) + \frac{d\mathcal{P}_L^{BS}}{d\widehat{\sigma}} \left.\frac{d\widehat{\sigma}_K}{dK}\right|_L\right) - 1 \tag{1.24}$$

$\mathcal{D}^{BS}$ is evaluated for $S = L$; as can be checked numerically, $\mathcal{D}^{BS}$ for $S = L$ is almost equal to 50% and has little sensitivity to the implied volatility $\widehat{\sigma}_L$. Expression (1.24) shows, though, that the value of the double European digital is very sensitive to $\left.\frac{d\widehat{\sigma}_K}{dK}\right|_L$ which is the at-the-money skew at the time S hits L. Take the example of a one-year ATM digital option; while $\mathcal{D}_L^{BS}(\widehat{\sigma}_L)$ is about 50%, the size of the correction term $\frac{d\mathcal{P}_L^{BS}}{d\widehat{\sigma}} \left.\frac{d\widehat{\sigma}_K}{dK}\right|_L$ for an equity index is typically about 8%: this is not a small effect.

[14] Black-Scholes prices are computed with the volatility σ that we choose to risk-manage the barrier option.

[15] It would vanish exactly had we used the exact static European hedge.

Thus, as we unwind our static hedge the magnitude of the then-prevailing at-the-money skew will determine whether we make or lose money. The Black-Scholes price of the barrier option has then to be adjusted manually to include an estimation of this gain or loss.

The lesson of this example is that the price of a barrier option is mostly dependent on the dynamics of the at-the-money skew conditional on S hitting the barrier.[16] A stochastic volatility model for barrier options would need to provide a direct handle on this precise feature of the dynamics of the volatility surface so as to appropriately reflect its P&L impact in the option price.

1.3.2 Example 2: a forward-start option

Forward-start options – also called cliquets[17] – involve the ratio of a security's price observed at two different dates – they are considered in detail in Chapter 3. Let T_1 and T_2 be two dates in the future and consider the case of a simple call cliquet whose payoff at T_2 is given by

$$\left(\frac{S_{T_2}}{S_{T_1}} - k\right)^+ \tag{1.25}$$

Let us choose $k = 100\%$ – this is called a forward-start at-the-money call. The price P of this option in the Black-Scholes model, because of homogeneity, does not depend on S and only depends on volatility. Assuming zero interest rates for simplicity, for $k = 100\%$, the Black-Scholes price of our cliquet is approximately given by:

$$P \simeq \frac{1}{\sqrt{2\pi}}\sigma\sqrt{T_2 - T_1} \tag{1.26}$$

The fact that P does not depend on S is worrisome: the only instrument whose dynamics is accounted for in the Black-Scholes model is S, yet S is not appearing in the pricing function.

P is only a function of volatility σ – σ is in fact the real underlying of the cliquet option.

A cliquet is an option on volatility, more precisely on forward implied volatility, that is the future implied volatility observed at T_1 for maturity T_2. At $t = T_1$, the cliquet becomes a vanilla option of maturity T_2, in our case a call option struck at kS_{T_1}. A suitable hedging strategy needs to generate at time T_1 the money needed to purchase a call option of maturity T_2 struck at kS_{T_1}.

While payoff formula (1.25) suggests that the cliquet is an option of maturity T_2 on S observed at T_1 and T_2, it is in fact an option of maturity T_1 whose underlying

[16] Besides the forward-skew risk we have just analyzed, the price of the barrier option needs to be adjusted for gap risk. Unwinding the European hedge – or unwinding the delta – cannot be done instantaneously as S crosses L. In our case, the delta of the barrier option we have sold is negative: we will need to buy stocks (or sell the double digital option) at a spot level that is presumably larger than L, thus incurring a loss. We must thus adjust the price charged for the barrier option to cover, on average, this loss.

[17] Ratchet, in French.

is the at-the-money implied volatility for maturity T_2, observed at T_1. This is the quantity whose dynamics a stochastic volatility model ought to provide a handle on.

1.3.3 Conclusion

Running an exotics book entails trading options dynamically to hedge other options. Vanilla options should be considered as hedging instruments in their own right and their dynamics modeled accordingly; as such the task of a stochastic volatility model is to model the joint dynamics of the underlying security and its associated volatility surface.

Chapter's digest

▶ Delta hedging removes the order-one contribution of δS to the P&L of an option position. Specifying a break-even condition for the lowest-order portion – the second order in δS – of the residual P&L leads to the Black-Scholes pricing equation – a parabolic equation. The latter has a probabilistic interpretation: the solution can be expressed as the expectation of the payoff under a density which is generated by a diffusion for S_t.

The argument goes this way and not the other way around – modeling does not start with the assumption of a diffusion for S_t and has little to do with Brownian motion; in this respect we refer the reader to Section 4.2 of [53].[18] For alternative break-even criteria that involve higher-order terms in δS see Chapter 10.

When there are multiple hedge instruments, the suitability of a model depends on the existence of a – possibly state- and time-dependent – break-even covariance matrix for hedge instruments that ensures gamma/theta cancellation.

▶ Delta hedging is not adequate for reducing the standard deviation of the P&L of an option position to reasonable levels. The sources of the dispersion of this P&L are: (a) the tails of returns, (b) the volatility of realized volatility and the correlation of future realized volatilities – see (1.15). Except for very short options, the latter effect prevails, because of the long-ranged nature of volatility/volatility correlations.

▶ Using options for gamma-hedging immunizes us against realized volatility. Dynamical trading of vanilla options, however, exposes us to uncertainty as to future levels of implied volatilities. Stochastic volatility models are thus needed for modeling the dynamics of implied volatilities, rather than that of realized volatility.

▶ Exotic options often depend in a complex way on the dynamics of implied volatilities. Some specific classes of options, such as barrier options, or cliquets, are such that their volatility risk can be pinpointed, enabling an easier assessment of the suitability of a given model.

[18] This is not to mean we can write down just *any* pricing equation. It has to comply with the basic requirements that (a) given two payoffs f and g, if $g(S) \geq f(S)\ \forall S$ then g should be more expensive than f – this expresses absence of arbitrage, and for a linear pricing equation implies the existence of a (risk-neutral) density and (b) that it obeys the convex order condition – see Section 2.2.2, page 29.

Chapter 2

Local volatility

This chapter covers the simplest and most widely used stochastic volatility model: the local volatility model. Local volatility [37], [40] was introduced as an extension of the Black-Scholes model that can be exactly calibrated to the whole volatility surface $\widehat{\sigma}_{KT}$.

While its proponents did not have stochastic volatility in mind, local volatility is a particular breed of stochastic volatility. It is also the simplest *market model*.

2.1 Introduction – local volatility as a market model

Market models aim at treating vanilla options on the same footing as the underlying itself: vanilla option prices observed at $t = 0$ are initial values of hedge instruments to be used as inputs in the model.[1] A stochastic volatility model should be able to accommodate as initial condition any configuration of these asset values, provided it is not nonsensical: for example, call option prices for a given maturity should be a decreasing function of the option's strike. We will see in Chapter 4 that this basic capability is difficult to achieve – most stochastic volatility models cannot be calibrated to the volatility surface exactly.

The enduring popularity of the local volatility model lies in its ability – as a market model – to take as input an arbitrary volatility surface provided it is free of arbitrage. Because any European option can be synthesized using call and put options of the European option's maturity (see Section 3.1.3, page 106), the local volatility model prices European options exactly.

It is a peculiar market model, however, because calibration on the market smile fully determines the model. Such frugality comes at a price: the dynamics it generates for the volatility surface is fully fixed by the vanilla smile used for calibration, it is not explicit, and must be extracted *a posteriori*. Mathematically, it is a market model that possesses a Markov representation in terms of t, S_t.

[1] We can use a subset of vanilla options in our market model or other types of options. The models of Chapter 7 are market models for log-contracts, or for a term structure of vanilla options of an arbitrary moneyness. In Section 4.3 of Chapter 4 we show how power payoffs can be used as well.

Historically, the local volatility model has been published and presented as a variant of the Black-Scholes model such that the instantaneous volatility is a deterministic function of t, S: $\sigma(t, S)$.

The local volatility function $\sigma(t, S)$, however, only serves an ancillary purpose and has no physical significance. It is a by-product of the fact that the model has a one-dimensional Markov representation in terms of t, S_t.

Traditional presentations of the local volatility model make $\sigma(t, S)$ a central object: the local volatility function is calibrated at $t = 0$ on the market smile and kept frozen afterwards. This contravenes the typical trading practice of recalibrating the local volatility function on a daily basis – which then seems to amount to an improper use of the model.

We will see instead in Section 2.7 that:

- this is how the local volatility model should be used,
- the resulting carry P&L has the standard expression in terms of offsetting spot/volatility gamma/theta contributions with well-defined and payoff-independent break-even levels – the trademark of a market model.

Our aim in the following sections is to characterize the dynamics generated for implied volatilities and then discuss the issue of the delta and the carry P&L. We will first need to establish the relationships linking local and implied volatilities.

2.1.1 SDE of the local volatility model

In the local volatility model, all assets have a one-dimensional representation in terms of t, S. The stochastic differential equation (SDE) for S_t is:

$$dS_t \ = \ (r - q)\, S_t dt \ + \ \sigma\left(t, S_t\right) S_t dW_t \tag{2.1}$$

where r is the interest rate and q the repo rate inclusive of the dividend yield. The pricing equation is identical to the Black-Scholes equation (1.4), except $\sigma\left(t, S\right)$ now replaces $\widehat{\sigma}$:

$$\frac{dP}{dt} + (r - q)\, S \frac{dP}{dS} + \frac{\sigma\left(t, S\right)^2}{2} S^2 \frac{d^2P}{dS^2} \ = \ rP \tag{2.2}$$

Given a particular local volatility function $\sigma\left(t, S\right)$ one can get prices of vanilla options by setting $P\left(t = T, S\right)$ equal to the option's payoff and solving equation (2.2) backwards from T to t, to generate $P\left(t, S\right)$. Conversely, given a configuration of vanilla options' prices, can we find a function $\sigma\left(t, S\right)$ such that, by solving equation (2.2) they are recovered? What is the condition for the existence of a $\sigma\left(t, S\right)$?

The expression of $\sigma\left(t, S\right)$,which we derive below, was found by Bruno Dupire [40]:

$$\sigma\left(t, S\right)^2 \ = \ 2 \ \left. \frac{\frac{dC}{dT} + qC \ + \ (r - q)\, K \frac{dC}{dK}}{K^2 \frac{d^2C}{dK^2}} \right|_{\substack{K=S \\ T=t}} \tag{2.3}$$

where $C(K,T)$ is the price of a call option of strike K and maturity T. This equation expresses the fact that the local volatility for spot S and time t is reflected in the differences of option prices with strikes straddling S and maturities straddling t.

2.2 From prices to local volatilities

2.2.1 The Dupire formula

Consider the following diffusive dynamics for S_t:

$$dS_t \;=\; (r-q)\, S_t dt + \sigma_t S_t dZ_t$$

where σ_t is for now an arbitrary process. By only using vanilla option prices, how precisely can we characterize σ_t?

The price of a call option is given by:

$$C(K,T) \;=\; e^{-rT} E[(S_T-K)^+]$$

The dynamics of S_t on the interval $[T, T+dT]$ determines how much prices of options of maturities T and $T+dT$ differ. Let us write the Itô expansion for $(S_T-K)^+$ over $[T, T+dT]$:

$$\begin{aligned}
& d\,(S_T-K)^+ \\
&= \frac{d(S_T-K)^+}{dS_T}\big(\,(r-q)\, S_T dT + \sigma_T S_T dZ_T\big) \\
&\quad + \frac{1}{2}\frac{d^2(S_T-K)^+}{dS_T^2}\sigma_T^2 S_T^2 dT \\
&= \theta(S_T-K)\,\big(\,(r-q)\, S_T dT + \sigma_T S_T dZ_T\big) + \frac{1}{2}\delta(S_T-K)\,\sigma_T^2 S_T^2 dT \qquad (2.4)
\end{aligned}$$

where $\theta(x)$ is the Heaviside function: $\theta(x)=1$ for $x>0$, $\theta(x)=0$ for $x<0$, and δ is the Dirac delta function.

For simplicity let us switch temporarily to *undiscounted* option prices $\mathcal{C}(K,T)$. Taking derivatives with respect to K of the identity: $\mathcal{C}(K,T) \;=\; E[(S_T-K)^+]$ we get:

$$E[\theta(S_T-K)] \;=\; -\frac{d\mathcal{C}}{dK}, \quad E[\delta(S_T-K)] \;=\; \frac{d^2\mathcal{C}}{dK^2} \qquad (2.5)$$

The second equation expresses the well-known property that the second derivative of undiscounted call or put prices with respect to their strike yields the pricing (or risk-neutral) density of S_T.

From the identity

$$\begin{aligned}
\mathcal{C} \;&=\; E[(S_T-K)^+] \;=\; E[(S_T-K)\,\theta(S_T-K)] \\
&=\; E[S_T\theta(S_T-K)] \;-\; K E[\theta(S_T-K)]
\end{aligned}$$

we get:

$$E[S_T\theta(S_T-K)] \;=\; \mathcal{C}-K\frac{d\mathcal{C}}{dK}$$

Now take the expectation of both sides of equation (2.4). In the left-hand side, $E[d\,(S_T-K)^+] \;=\; dE[(S_T-K)^+]$, that is the difference of the undiscounted prices of two call options of strike K expiring at T and $T+dT$: this is equal to $\frac{d\mathcal{C}}{dT}dT$.

$$\frac{d\mathcal{C}}{dT}dT \;=\; (r-q)\,\left(\mathcal{C}-K\frac{d\mathcal{C}}{dK}\right)dT \;+\; \frac{K^2}{2}E\left[\sigma_T^2\delta(S_T-K)\right]dT$$

yields:

$$E\left[\sigma_T^2\delta(S_T-K)\right] \;=\; \frac{2}{K^2}\left(\frac{d\mathcal{C}}{dT}-(r-q)\left(\mathcal{C}-K\frac{d\mathcal{C}}{dK}\right)\right)$$

Dividing the left-hand side by $E[\delta\,(S_T-K)]$ and the right-hand side by $\frac{d^2\mathcal{C}}{dK^2}$, which are equal, we get:

$$\frac{E[\sigma_T^2\delta(S_T-K)]}{E[\delta(S_T-K)]} \;=\; 2\frac{\frac{d\mathcal{C}}{dT}-(r-q)\left(\mathcal{C}-K\frac{d\mathcal{C}}{dK}\right)}{K^2\frac{d^2\mathcal{C}}{dK^2}}$$

and reverting back to discounted option prices: $C \;=\; e^{-rT}\mathcal{C}$:

$$E\left[\sigma_T^2|S_T=K\right] \;=\; \frac{E[\sigma_T^2\delta(S_T-K)]}{E[\delta(S_T-K)]} \;=\; 2\frac{\frac{dC}{dT}+qC+(r-q)\,K\frac{dC}{dK}}{K^2\frac{d^2C}{dK^2}} \tag{2.6}$$

This identity, known as the Dupire equation, expresses a general relationship linking the expectation of the instantaneous variance conditional on the spot price to the maturity and strike derivatives of vanilla option prices.

It holds in diffusive models for S_t: knowledge of vanilla options prices is not sufficient to pin down the process σ_t, but characterizes the class of diffusive processes that yield the same vanilla option prices. Two processes σ_t, σ_t' generate the same vanilla smile if $E[\sigma_T^2|S_T=K]=E[\sigma_T'^2|S_T=K]$ for all K,T.[2]

A stochastic volatility model aiming to reproduce at time $t=0$ the market smile has to satisfy this condition. The simplest way of accommodating this constraint is to take for process σ_t a deterministic function of t,S:

$$\sigma_t \;\equiv\; \sigma(t,S)$$

The conditional expectation of the instantaneous variance on the left-hand side of (2.6) is then simply $\sigma\,(t,S)^2$ and we get the Dupire formula (2.3). The Dupire

[2]The general result that the marginals of an arbitrary diffusive process with instantaneous volatility σ_t are exactly recovered by using an effective local volatility model whose local volatility is given by $\sigma^2(t,S)=E\left[\sigma_t^2|S\right]$ is due to Gyöngy – see [54].

equation can also be used to compute call and put option prices for a known local volatility function $\sigma(t,S)$ – let us rewrite it as:

$$\frac{dC}{dT}+(r-q)K\frac{dC}{dK}-\frac{\sigma^2(t=T,S=K)}{2}K^2\frac{d^2C}{dK^2} \;=\; -qC \tag{2.7}$$

This is called the forward equation. Unlike the usual pricing equation, which is a backward equation and provides prices of a single call option with given maturity and strike for a range of initial spot prices, the forward equation, with initial condition $C(K,T=0) \;=\; (S_0-K)^+$, supplies prices for a single value of the spot price S_0, but for all K and T: this makes it attractive in situations when derivatives with respect to S_0 are not needed. Put option prices are obtained by changing the initial condition to $(K-S_0)^+$.

In the derivation above, we have made the assumption of a diffusive process for S_t. Consider now a given market smile – does there exist a local volatility function $\sigma(t,S)$ that is able to reproduce it? Choosing $\sigma(t,S)$ as specified by (2.3) will do the job, but what if the numerator or denominator in the right-hand side of (2.3) are negative? We now prove that this cannot be the case unless vanilla option prices are arbitrageable.

2.2.2 No-arbitrage conditions

Strike arbitrage

The denominator in (2.6) involves the second derivative of the call price with respect to K:

$$\frac{d^2C(K,T)}{dK^2} \;=\; \lim_{\varepsilon\to 0}\frac{C(K-\varepsilon,T)-2C(K,T)+C(K+\varepsilon,T)}{\varepsilon^2}$$

Consider the European payoff consisting of $\frac{1}{\varepsilon^2}$ calls of strike $K-\varepsilon$, $\frac{1}{\varepsilon^2}$ calls of strike $K+\varepsilon$ and $-\frac{2}{\varepsilon^2}$ calls of strike K – this is known as a butterfly spread.

The payout at maturity as a function of S_T has a triangular shape whose surface area is unity: it vanishes for $S_T \le K-\varepsilon$ and $S_T \ge K+\varepsilon$ and is equal to $\frac{1}{\varepsilon}$ for $S_T = K$. For $\varepsilon \to 0$ it becomes a Dirac delta function. It either vanishes or is strictly positive depending on S_T: its price at inception must be positive.

Options' markets are arbitraged well enough that butterfly spreads do not have negative prices:[3] the denominator in the Dupire formula (2.3) is positive.

In a model, $\frac{d^2C(K,T)}{dK^2}$ is related to the probability density of S_T through:

$$\frac{d^2C(K,T)}{dK^2} \;=\; e^{-rT}E[\delta(S_T-K)] \tag{2.8}$$

[3]Bid/offer spreads of options are usually not negligible: arbitrage opportunities may appear more attractive than they really are.

thus is positive by construction. The condition $\frac{d^2C(K,T)}{dK^2} > 0$ is equivalent to requiring that the market implied density be positive. Violation of the positivity of the denominator of (2.6) is called a strike arbitrage.

Maturity arbitrage

What about the numerator in (2.3)? It can be rewritten as:

$$e^{-qT}\frac{d}{dT}[e^{qT}C(Ke^{(r-q)T},T)]$$

For it to be positive, e^{qT} times the price of a call option struck at a strike that is a fixed proportion of the forward $F_T = Se^{(r-q)T}$ – that is $K = kF_T$ – must be an increasing function of maturity. For $T_1 \leq T_2$:

$$e^{qT_1}C\left(kF_{T_1},T_1\right) \;\leq\; e^{qT_2}C\left(kF_{T_2},T_2\right) \tag{2.9}$$

Imagine that this condition is violated: there exist two maturities $T_1 < T_2$ and k such that:

$$e^{qT_1}C\left(kF_{T_1},T_1\right) \;>\; e^{qT_2}C\left(kF_{T_2},T_2\right)$$

Set up the following strategy: buy one option of maturity T_2, strike kF_{T_2} and sell $e^{-q(T_2-T_1)}$ options of maturity T_1, strike kF_{T_1}: we pocket a net premium at inception. At T_1 take the following Δ position on S:

$$\begin{aligned}&\text{if } S_{T_1} < kF_{T_1}: \quad \Delta = 0\\ &\text{if } S_{T_1} > kF_{T_1}: \quad \Delta = -1\end{aligned}$$

Our P&L at T_2 comprises the payout of the T_2 option which we receive, the payout of the T_1 option which we pay, capitalized up to T_2, and the P&L generated by the delta position entered at T_1, which we unwind at T_2 – inclusive of financing costs. Its expression is:

$$\begin{aligned}&(S_{T_2}-kF_{T_2})^+ - e^{r(T_2-T_1)}e^{-q(T_2-T_1)}\left(S_{T_1}-kF_{T_1}\right)^+ + \Delta\Big(S_{T_2}-\frac{F_{T_2}}{F_{T_1}}S_{T_1}\Big)\\ &= \;(S_{T_2}-kF_{T_2})^+ - \left[\frac{F_{T_2}}{F_{T_1}}\left(S_{T_1}-kF_{T_1}\right)^+ + \mathbf{1}_{S_{T_1}>kF_{T_1}}\Big(S_{T_2}-\frac{F_{T_2}}{F_{T_1}}S_{T_1}\Big)\right]\\ &= \;(S_{T_2}-kF_{T_2})^+ - \left[\left(S^\star_{T_1}-kF_{T_2}\right)^+ + \mathbf{1}_{S^\star_{T_1}>kF_{T_2}}\left(S_{T_2}-S^\star_{T_1}\right)\right]\end{aligned}$$

where $S^\star_{T_1} = \frac{F_{T_2}}{F_{T_1}}S_{T_1}$. The last equation reads:

$$f\left(S_{T_2}\right) - \left[f\left(S^\star_{T_1}\right) + \frac{df}{dx}\left(S^\star_{T_1}\right)\left(S_{T_2}-S^\star_{T_1}\right)\right]$$

with $f(x) = (x-kF_{T_2})^+$. Since f is convex this is positive. Our strategy not only produces strictly positive P&L at inception; it also generates positive P&L at T_2.

Real markets are sufficiently arbitraged that arbitrage opportunities of this type do not exist: market prices of vanilla options are such that the numerator in the Dupire equation (2.3) is always positive.

In a model, the numerator in (2.3) is positive by construction. Writing the price of an option of maturity T_2 as an expectation and conditioning with respect to S_{T_1} at T_1 we get, using Jensen's inequality:

$$
\begin{aligned}
& e^{qT_2} C\left(kF_{T_2}, T_2\right) \\
& = e^{qT_2} e^{-rT_2} E\big[(S_{T_2} - kF_{T_2})^+\big] \ = \ e^{-(r-q)T_2} E\big[E[(S_{T_2} - kF_{T_2})^+ | S_{T_1}]\big] \\
& \geq e^{-(r-q)T_2} E\big[(\frac{F_{T_2}}{F_{T_1}} S_{T_1} - kF_{T_2})^+\big] \ = \ e^{-(r-q)T_2} \frac{F_{T_2}}{F_{T_1}} E\big[(S_{T_1} - kF_{T_1})^+\big] \\
& \geq e^{qT_1} C\left(kF_{T_1}, T_1\right)
\end{aligned}
$$

Violation of the positivity of the numerator of (2.6) is called a maturity arbitrage.

Conclusion

In conclusion, a violation of (2.9) can be arbitraged and the local volatility given by the Dupire equation is well-defined for any arbitrage-free smile.

Contrary to a frequently heard assertion, the steep skews observed for short-maturity equity smiles are no evidence that jumps are needed to generate them – as long as they are non-arbitrageable, local volatility will be happy to oblige.

See also Section 8.7.2 for an example of how a two-factor stochastic volatility model is also able to generate the typical term structures of ATMF skews observed for equity indexes.

2.2.2.1 Convex order condition for implied volatilities

What does the convex order condition (2.9) for prices mean for implied volatilities?

Consider a call option of maturity T for a strike $K = kF_T$, whose implied volatility we denote by $\widehat{\sigma}_{kT}$. We have:

$$
\begin{aligned}
e^{qT} C_{BS}\left(kF_T, T, \widehat{\sigma}_{kT}\right) \ &= \ e^{-(r-q)T} E\big[(S_T - kF_T)^+\big] \\
&= \ S_0 E\big[(U_{\tau(T)} - k)^+\big] \ = \ S_0 f\left(k, \tau\right)
\end{aligned}
\tag{2.10}
$$

where we have introduced U_τ defined by: $U_\tau = e^{-\frac{\tau}{2} + W_\tau}$, $\tau(T) = \widehat{\sigma}^2_{kT} T$ and f is defined by:

$$
f\left(k, \tau\right) \ = \ E\big[(U_\tau - k)^+\big] \tag{2.11}
$$

U_τ is a martingale: for $\tau_1 \leq \tau_2$ $E[U_{\tau_2} | U_{\tau_1}] = U_{\tau_1}$. We could use Jensen's inequality exactly as above: for $\tau_1 \leq \tau_2$: $E[(U_{\tau_2} - k)^+] = E[E[(U_{\tau_2} - k)^+ | U_{\tau_1}]] \geq E[(U_{\tau_1} - k)^+]$ thus

$$
\tau_1 \leq \tau_2 \ \Rightarrow \ f(k, \tau_1) \leq f(k, \tau_2) \tag{2.12}
$$

What we need, however, is the reverse implication.

From (2.11) $f(\tau, k)$ is the price of call option of strike k in the Black-Scholes model where the underlying U starts from $U_0 = 1$ and has a constant volatility equal to 1. It obeys the following forward PDE:

$$\frac{df}{d\tau} = \frac{1}{2}k^2\frac{d^2f}{dk^2} \tag{2.13}$$

with initial condition $f(k, \tau = 0) = (1-k)^+$.

From (2.8) $\frac{d^2f}{dk^2}$ is proportional to the risk-neutral density of U_τ, which in a lognormal model with constant volatility, is *strictly* positive. Thus, from (2.13) $\frac{df}{d\tau} > 0$.

We now have property (2.12), with a *strict* inequality: $\tau_1 < \tau_2 \Rightarrow f(k, \tau_1) < f(k, \tau_2)$. This, together with (2.12) yields the following equivalence:

$$\tau_1 \le \tau_2 \Leftrightarrow f(k, \tau_1) \le f(k, \tau_2)$$

which, using (2.10), translates into:

$$e^{qT_1}C_{BS}(kF_{T_1}, T_1, \widehat{\sigma}_{kT_1}) \le e^{qT_2}C_{BS}(kF_{T_2}, T_2, \widehat{\sigma}_{kT_2}) \Leftrightarrow T_1\widehat{\sigma}^2_{kT_1} \le T_2\widehat{\sigma}^2_{kT_2} \tag{2.14}$$

Thus, in an arbitrage-free smile, the integrated variance corresponding to any given moneyness k is an increasing function of maturity:

$$T_1\widehat{\sigma}^2_{kF_{T_1},T_1} \le T_2\widehat{\sigma}^2_{kF_{T_2},T_2} \tag{2.15}$$

2.2.2.2 Implied volatilities of general convex payoffs

The notion of implied volatility is not a privilege of hockey-stick payoffs. One can show that, in the absence of arbitrage, the notion of (lognormal) implied volatility can be defined for any convex payoff. Moreover, consider a family of European options such that the payoff $f(S_T)$ for maturity T is given by:

$$f(S_T) = h(x) \quad \text{with } x = \frac{S_T}{F_T} \text{ and } h \text{ convex.} \tag{2.16}$$

It is shown in [81] that:

- there exists one single Black-Scholes implied volatility $\widehat{\sigma}_T$ that matches a given market price for payoff f.
- no-arbitrage in market prices for maturities T_1, T_2 implies that the following convex order condition holds:

$$T_2\widehat{\sigma}^2_{T_2} \ge T_1\widehat{\sigma}^2_{T_1} \tag{2.17}$$

Vanilla options are but a particular case of convex payoffs – the payoffs of maturities T_1, T_2 used above to derive (2.15) are of type (2.16), with $h(x) = (x-k)^+$.

We will consider in Section 4.3 the particular class $h(x) = x^p$ and will focus on the special case $p \to 0$.

A note on "arbitrage" arguments

In all fairness, the type of arbitrage strategy we have outlined – entering a position and keeping it until maturity to pocket the (positive) arbitrage profit – is a bit unrealistic as it does not take into account mark-to-market P&L and the discomfort that comes with it, in the case of a large position.[4]

Imagine we bought yesterday a butterfly spread that had negative market value and today's market value is even more negative: we have lost money on yesterday's position. Our management may demand that we cut our position – at a loss – despite our plea that we will eventually make money if allowed to hold on to our position, that the arbitrage has actually become more attractive, and that we should in fact increase the size of our position.

2.3 From implied volatilities to local volatilities

The Dupire equation (2.3) expresses the local volatility as a function of derivatives of call option prices. Let us assume that there are no dividends or, less strictly, that dividend amounts are expressed as fixed yields applied to the stock value at the dividend payout date.[5] The dividend yield can then be lumped together with the repo and we can use the Black-Scholes formula to express call option prices as a function of implied volatilities. Let us use the parametrization $f(t, y)$ with:

$$y = \ln\left(\frac{K}{F_t}\right) \tag{2.18a}$$

$$f(t, y) = (t - t_0)\,\widehat{\sigma}^2_{Kt} \tag{2.18b}$$

where F_t is the forward for maturity t: $F_t = S_0 e^{(r-q)(t-t_0)}$. Replacing C in the Dupire equation (2.3) with the Black-Scholes formula with implied volatility $\widehat{\sigma}_{KT}$, computing analytically all derivatives of C, and using f and y rather than $\widehat{\sigma}$ and K yields the following formula:

$$\sigma(t, S)^2 = \left.\frac{\frac{df}{dt}}{\left(\frac{y}{2f}\frac{df}{dy} - 1\right)^2 + \frac{1}{2}\frac{d^2 f}{dy^2} - \frac{1}{4}\left(\frac{1}{4} + \frac{1}{f}\right)\left(\frac{df}{dy}\right)^2}\right|_{y=\ln\left(\frac{S}{F_t}\right)} \tag{2.19}$$

---

[4] Also note that, as we take advantage of a maturity arbitrage, we make a bet on the repo level prevailing at T_1 for maturity T_2 – which could turn sour.

[5] While this is reasonable for dividends far into the future, it is a poor assumption for nearby dividends whose cash amount is usually known, either because it has been announced, or through analysts' forecasts. As a result, equities are probably the only asset class for which even vanilla options cannot be priced in closed form.

As mentioned above, option markets typically do not violate the no-arbitrage conditions of Section 2.2.2. Market prices, however, are only available for discrete strikes and maturities: prior to using equation (2.19) we need to build an interpolation in between discrete strikes and maturities – and an extrapolation outside the range of market-traded strikes – of market implied volatilities that comply with no-arbitrage conditions.

The latter take a particularly simple form in the (y, t) coordinates. Let T_i be the discrete maturities for which implied volatilities are available and set $f_i(y) = f(T_i, y)$. The convex order condition (2.15) translates into $\frac{df}{dt} \geq 0$, thus implies the simple rule: $f_{i+1}(y) \geq f_i(y)$.

Once each $f_i(y)$ function has been created by interpolating $\widehat{\sigma}^2(K, T_i)T_i$ as a function of $\ln(K/F_{T_i})$ the simple rule that the f_i profiles should not cross ensures the positivity of the numerator in the right-hand side of (2.19).

$f(t, y)$ for $y \in [T_i, T_{i+1}]$ is then generated by affine interpolation:

$$f(t, y) = \frac{T_{i+1} - t}{T_{i+1} - T_i} f_i(y) + \frac{t - T_i}{T_{i+1} - T_i} f_{i+1}(y) \tag{2.20}$$

Though rustic, interpolation (2.20) ensures that the convex order condition holds over $[T_i, T_{i+1}]$ and that local volatilities $\sigma(t, S)$ for $t \in [T_i, T_{i+1}]$ only depend on implied volatilities for maturities T_i, T_{i+1}. Otherwise – in case a spline interpolation was used, for example – a European option expiring at $T \in [T_i, T_{i+1}]$ would be sensitive to implied volatilities for maturities longer than T_{i+1}, an incongruous and unintended consequence of the interpolation scheme.

As we turn to extrapolating $f(t, y)$ for values of y corresponding to strikes that lie beyond the lowest/highest market-quoted strikes, care must be taken not to create strike arbitrage. Typically an affine extrapolation is used: $f_i(y) = a_i y + b_i$. It is easy to check that $|a_i| \leq 2$ is a necessary condition for positivity of the denominator in (2.19) for large values of y.

Finally, there may be situations – for illiquid underlyings – when one needs to build from scratch a volatility surface; we refer the reader to [49] for a popular example of a parametric volatility surface that, under certain conditions, is arbitrage-free: the SVI formula.

2.3.1 Dividends

In the presence of cash-amount dividends, while the Dupire formula (2.3) with option prices is still valid, its version (2.19) expressing local volatilities directly as a function of implied volatilities cannot be used as is, as option prices are no longer given by the Black-Scholes formula.

We first present an exact solution then an accurate approximate solution.

2.3.1.1 An exact solution

The exact solution is taken from [58] and [19]. It relies on the mapping of S to an asset X that does not jump on dividend dates.

Let us assume that dividends consist of two portions: a fixed cash amount and a proportional part. The dividend d_i falling at time t_i is given by:

$$d_i = y_i S_{t_i^-} + c_i$$

When looking for a security that does not experience dividend jumps the forward naturally comes to mind. However, we would have to pick an arbitrary maturity T for the forward – the local volatility function would change whenever an option with maturity longer than T was priced.

Let us instead use a driftless process X which starts with the same value as S: $X_{t=0} = S_{t=0}$ and define X_t as:

$$S_t = \alpha(t) X_t - \delta(t) \tag{2.21}$$

with $\alpha(t), \delta(t)$ given by:

$$\begin{aligned}\alpha(t) &= e^{(r-q)t} \prod_{t_i<t} (1-y_i) \\ \delta(t) &= \sum_{t_i<t} c_i e^{(r-q)(t-t_i)} \prod_{t_i<t_j<t} (1-y_j)\end{aligned}$$

One can check that X_t is driftless and does not jump across dividend dates. Because the relationship of S to X is affine, the price of a vanilla option on X is a multiple of the price of a vanilla option on S, with a shifted strike. We then have all implied volatilities for X and can use equation (2.19) to get the local volatility function for X: $\sigma_X(t, X)$. The local volatility for S is then given by:

$$\sigma(t, S) = \frac{S + \delta(t)}{S} \sigma_X(t, X(S, t)) \tag{2.22}$$

Across dividend dates σ_X is continuous, but σ is not, as $\delta(t)$ jumps. Those taking local volatility seriously may object to this. Consider, however, that just before a dividend date, the portion of S which is the cash dividend is frozen and has no volatility: the volatility of S only comes from the volatility of $S - c$. Consequently, as one crosses the dividend date, it is natural that the lognormal volatility of S jumps, in a fashion that is exactly expressed by (2.22).

Equation (2.21) seems to imply that S can go negative. This would be the case, for example, if X were lognormal. In reality, it does not happen, as the implied volatilities of X are derived from the smile of S which – if extrapolated properly – ensures that S_T cannot go negative, hence X_T cannot go below $\delta(T)/\alpha(T)$. For a typical negatively skewed smile for S, the smile of X will have a similar shape, except implied volatilities for low strikes, of the order of $\delta(T)/\alpha(T)$, will fall off.