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 non-linearities 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. $\delta t$ is typically 1 day. Let $\Delta$ 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 $\Delta$ shares, as well as money we make by lending shares out during $\delta t$:

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

where $\delta S$ is the amount by which $S$ moves during $\delta t$. $r$ is the interest rate and $q$ the repo rate, inclusive of dividend yield.

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

We now expand the P&L in powers of $\delta S$ and $\delta t$. We would like to stop at the lowest non-trivial orders for $\delta t$ and $\delta S$: order one in $\delta t$, and order two in $\delta S$, 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 $\delta S$, are smaller than $\delta S^2$ and $\delta t$ 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 $\delta t$ and $\delta S$ is of order $\sqrt{\delta t}$.[1] The contributions at order one in $\delta t$ and order two in $\delta S$ are then both of order $\delta t$ while the cross-term $\delta S \delta t$ and terms of higher order in $\delta S$ are of higher order in $\delta t$, 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:

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

- 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 $\delta t$ 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 $\delta t$ be so that this is indeed the case? The order of magnitude of $\delta S$ is $S\sigma\sqrt{\delta t}$ where $\sigma$ 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 $\delta S$, 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:

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

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 $\delta S$. 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 $\delta S$. 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:

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

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:

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

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 $\delta t$ and two in $\delta S$, 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 $\delta S$ 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 $S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}$, 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 $\phi$, a real symmetric matrix, and denote by $\varphi_k$ its eigenvalues and $T_k$ the associated eigenvectors. Also denote by $\varphi$ the diagonal matrix with the $\varphi_k$ 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 $\delta z_k$ 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 $\omega_k$ 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 $\omega$ the diagonal matrix with the $\omega_k$ 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 $\omega_k$ 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 $\phi_{ij}$ 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 $\delta S$ 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:

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

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 $\delta S$: 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 $\sigma$ as that used for pricing and delta-hedging the option; that is, $S$ follows the Black-Scholes *model* with volatility $\sigma$

- 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 $\sigma_t$ 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:

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

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 $\delta S$ and order 1 in $\delta t$, 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 $e^{-rt_i}\frac{S_i^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dS^2}(t_i, S_i)$, 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 $S_0^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}(t_0, S_0)$ – 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 $\sigma_i$ is the realized volatility for day $i$, and $z_i$ is centered and has unit variance: $\langle z_i\rangle = 0$, $\langle z_i^2\rangle = 1$. Let us assume that the $z_i$ are iid and are independent of the volatilities $\sigma_i$.

Because the $z_i$ are independent, returns $r_i$ have no serial correlation but are not independent, as daily volatilities $\sigma_i$ may be correlated. Expression (1.10) allows separation of the effects of the scale $\sigma_i$ 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:

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

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

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

where we have introduced the (excess) kurtosis $\kappa$ 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 $\Omega$, which quantifies the variance of daily variances $\sigma_i^2$ 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:

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

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:

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

### 1.2.1 The Black-Scholes case

Let us first assume that $S$ follows the lognormal Black-Scholes dynamics. $\sigma_i$ is constant, equal to $\widehat{\sigma}$, hence $\Omega = 0$. Since $\sigma_i\sqrt{\delta t}$ is small ($\delta t$ 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 $\sigma_i$ is time-homogeneous: $f_{ij}$ is then a function of the difference $j-i$, actually a function of $|j-i|$.

As $\delta t$ is small compared to the option's maturity, we convert the sums in (1.11) into integrals:

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

We now have from (1.13):

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

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

Imagine first that daily variances are constant: $\Omega$ 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 $\kappa$. By setting $\kappa = 0$ we recover result (1.14).

Consider now the second contribution in (1.15). The prefactor $\Omega$ quantifies the dispersion of daily variances while $f(\tau)$ quantifies how a fluctuation in daily variance $\sigma_i^2$ 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: $f(\tau) = 1$. The second piece in (1.15) is then simply equal to $\frac{\Omega}{4}$. If $\Omega$ 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 $f(\tau) = 1$ 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 $\sigma_i$ in the case of real securities. Separating in $r_i$ the contributions from $\sigma_i$ and $z_i$ is difficult if the only daily data we have are daily returns. In what follows we have estimated daily volatilities $\sigma_i$ using 5-minute returns: $\sigma_i$ 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]

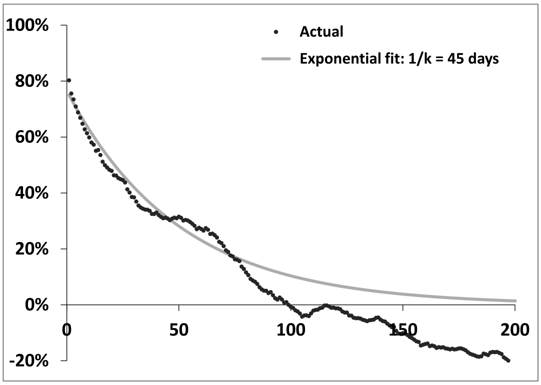



**Figure 1.1**: Correlation function $f(\tau)$ of daily volatilities evaluated on a basket of financial stocks. $\tau$ is in business days.

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 $\tau$, 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:

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

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

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

Estimating the unconditional kurtosis $\kappa$ 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 $\Omega$, $\kappa$, $\rho$, $k$ 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.

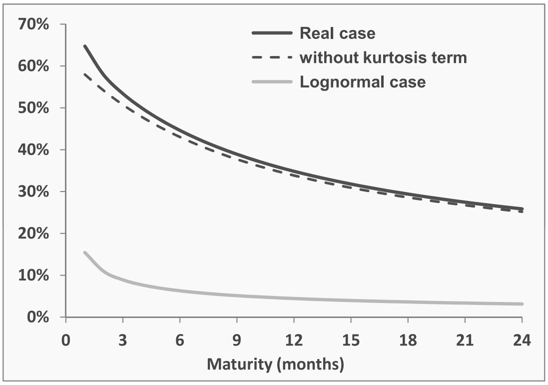



**Figure 1.2**: Right-hand side in equation 1.16 (darker line), as a function of maturity, compared to the same quantity, but without the kurtosis term (dashed line), and the lognormal case (lighter line).

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 $\delta S$ in our daily P&L, the effect of the $\delta S^2$ 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 $\lambda$ 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 $\lambda$.

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 $\lambda$ 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 $\lambda$, 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 $\delta S$, $\delta\widehat{\sigma}_O$ and order 1 in $\delta t$ 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 $\delta S^2$ 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 $\delta t$ 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 $\delta S^2$ term, no deterministic $\delta t$ 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 $\lambda_i$ 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 $P^{\mathrm{Mkt}}$ 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 $P^{\mathrm{Mkt}}$.
- 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 $\sigma$ 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 $\sigma = 20\%$.

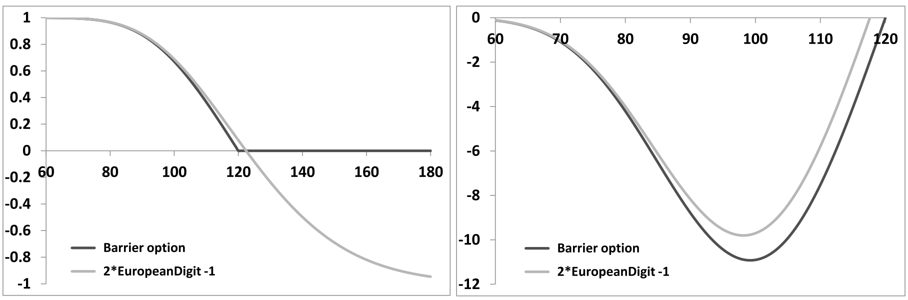



**Figure 1.3**: Value (left) and dollar gamma (right) of barrier option and double European digital option.

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 $\tau$, 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 $\tau$ 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 $\sigma$ 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 $\sigma$ – $\sigma$ 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 $\delta S$ to the P&L of an option position. Specifying a break-even condition for the lowest-order portion – the second order in $\delta S$ – 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 $\delta S$ 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:

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

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 $\sigma\,(t, S)$ now replaces $\widehat{\sigma}$ :

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

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

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

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

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

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 $[T, T+dT]$ :

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

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

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

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

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

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 $\sigma_t$, but characterizes the class of diffusive processes that yield the same vanilla option prices. Two processes $\sigma_t$, $\sigma'_t$ 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 $\sigma_t$ 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 $\sigma_t$ 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 $T_1 \le T_2$:

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

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 $\Delta$ 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:

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

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

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

$U_\tau$ 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

$$\tau_1 \leq \tau_2 \;\Rightarrow\; f(k, \tau_1) \leq f(k, \tau_2) \qquad (2.12)$$

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 $U_\tau$, 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:

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

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 $\sigma_X$ is continuous, but $\sigma$ 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.

#### 2.3.1.2 An approximate solution

We really would like to use an expression relating local volatilities to implied volatilities directly, similar to (2.19). Because of the presence of cash-amount dividends, the definition of $y$ in (2.18a) has to change.

The ingredient in (2.19) is $f(t,y)$, that is a parametrization of the implied volatility surface. When there are dividends, a suitable parametrization must ensure that appropriate matching conditions hold across dividend dates. Consider a dividend $d$ falling at time $\tau$, part cash amount, part yield:

$$S_{\tau^+} \;=\; (1-z)S_{\tau^-} - c$$

and a call option of strike $K$, maturity $\tau^+$. Its payoff can be written as a function of $S_{\tau^-}$:

$$\begin{aligned}(S_{\tau^+}-K)^+ \;&=\; ((1-z)S_{\tau^-}-c-K)^+\\ &= (1-z)\left(S_{\tau^-}-\frac{c+K}{1-z}\right)^+\end{aligned}$$

This option's payoff is proportional to that of a vanilla option of strike $\frac{c+K}{1-z}$, maturity $\tau^-$. Their implied volatilities are thus identical:

$$\widehat{\sigma}_{K\tau^+} \;=\; \widehat{\sigma}_{\frac{K+c}{1-z}\,\tau^-}$$

Equivalently:

$$\widehat{\sigma}_{K\tau^-} \;=\; \widehat{\sigma}_{(1-z)K-c\;\tau^+} \tag{2.23}$$

How can we alter the definition of $y$ in (2.18a) so that (2.23) holds? Our inspiration comes from an approximation for vanilla option prices in the Black-Scholes model when cash-amount dividends are present.

Proportional dividends are readily converted in an adjustment of the initial spot value. With regard to the cash-amount portion of dividends, one typically uses an approximation that condenses them into a smaller number of effective cash-amount dividends.

The most well-known is that published by Michael Bos and Stephen Vandermark – see [16] – where dividends are replaced with just two effective dividends. Each cash-amount dividend is split into two pieces, one falling at $t=0$, resulting in a negative adjustment $\delta S$ of the initial spot value, the other at maturity $T$, which translates into a positive shift $\delta K$ of the strike. For vanishing interest rate and repo, the proportions are, respectively, $\frac{T-t}{T}$ and $\frac{t}{T}$ where $t$ is the time the dividend falls.

Let $y_i$ and $c_i$ be the yield and cash-amount of the dividend falling at time $t_i$:$S_{t_i^+} = (1-y_i)S_{t_i^-} - c_i$. Define the functions $\alpha(T),\ \delta S(T),\ \delta K(T)$ as:

$$\begin{cases} \alpha(T) & = & \prod_{t_i<T}(1-y_i) \\ \delta S(T) & = & \sum_{t_i<T} \frac{T-t_i}{T}\, c_i^*\, e^{-(r-q)t_i} \\ \delta K(T) & = & \sum_{t_i<T} \frac{t_i}{T}\, c_i^*\, e^{(r-q)(T-t_i)} \end{cases} \tag{2.24}$$

with the effective cash amounts $c_i^*$ given by:

$$c_i^* \ = \ c_i \prod_{t_i<t_j<T} (1-y_j)$$

The price of a vanilla option of strike $K$, maturity $T$ is given approximately by the Black-Scholes formula with rate $r$ and repo $q$, with the initial spot value $S$ and strike $K$ replaced, respectively, by $\alpha(T)S - \delta S(T)$ and $K + \delta K(T)$:

$$C\left(K,T\right) \ = \ P_{BS}\left(t_0,\ \alpha(T)S_0 - \delta S(T),\ K+\delta K(T),\ \widehat{\sigma}_{KT}\right) \tag{2.25}$$

When there are no dividends, $\alpha = 1$, $\delta S = 0$, $\delta K = 0$.

Directly using the Bos & Vandermark approximation for vanilla option prices in the Dupire formula (2.3) does not work. Indeed, (2.3) expresses the square of the local volatility as the ratio of two quantities, each of which becomes very small when $S \ll S_0$ or $S \gg S_0$. An approximation of $C(K,T)$ has to be such that its derivatives with respect to $K,T$ are very accurate and remain so even when they are very small – this is too much to ask from (2.25).

For example consider a flat implied volatility surface: $\widehat{\sigma}_{KT} \equiv \sigma_0$. In order to recover $\sigma_0$ as the (constant) local volatility out of (2.3), $C\left(K,T\right)$ needs to obey the forward equation (2.7). It is easy to check that expression (2.25) for $C\left(K,T\right)$ with $\widehat{\sigma}_{KT} \equiv \sigma_0$ does not fulfill this condition.

While we will not use (2.25), we make use of the expressions of $\alpha(T),\ \delta S(T),\ \delta K(T)$. Consider the following amended definition of $y$ and parametrization $f(t,y)$ of the volatility surface:

$$\begin{cases} y & = & \ln\left(\frac{K+\delta K(t)}{\alpha(t)S_0-\delta S(t)}\right) - (r-q)(t-t_0) \\ f(t,y) & = & (t-t_0)\,\widehat{\sigma}_{Kt}^2 \end{cases} \tag{2.26}$$

where $f(t,y)$ is continuous across dividend dates.

Consider a dividend falling at time $\tau$, with cash-amount $c$ and yield $z$ and a vanilla option of strike $K$, maturity $\tau^-$. Let us check that condition (2.23) holds.

Since $f$ is continuous across $\tau$, $\widehat{\sigma}_{K\tau^-} = \widehat{\sigma}_{K'\tau^+}$, where $K, K'$ are such that they correspond to the same value of $y$:

$$\frac{K' + \delta K(\tau^+)}{\alpha(\tau^+)S_0 - \delta S(\tau^+)} = \frac{K + \delta K(\tau^-)}{\alpha(\tau^-)S_0 - \delta S(\tau^-)} \tag{2.27}$$

From the definition of $\alpha, \delta S, \delta K$:

$$\begin{aligned} \alpha(\tau^+) &= (1-z)\alpha(\tau^-) \\ \delta S(\tau^+) &= (1-z)\delta S(\tau^-) \\ \delta K(\tau^+) &= (1-z)\delta K(\tau^-) + c \end{aligned}$$

which, once plugged in (2.27) yields:

$$K' = (1-z)K - c$$

Thus (2.23) is exactly obeyed: using a smooth function $f\,(t, y)$ with $y$ given by (2.26) automatically takes care of the matching conditions across dividend dates. Our final recipe is thus:

- Build a smooth interpolation of $f = (t - t_0)\,\widehat{\sigma}^2_{Kt}$ as a function of $(t, y)$ with $y$ defined in (2.26).

- Use formula (2.19) to generate the local volatility function.

In the author's experience this approximate technique is accurate for indexes (many small dividends) and stocks (few large dividends) alike – see Figure 2.1 for an example with the S&P 500 index. As an additional benefit, whenever we input a flat volatility surface – $\widehat{\sigma}_{KT} = \sigma_0, \forall\, K, T$ – we exactly recover a flat local volatility function: $\sigma\,(t, S) = \sigma_0, \forall\, t, S$.[6]

## 2.4 From local volatilities to implied volatilities

Expression (2.19) gives local volatilities as a function of implied volatilities. For the sake of analyzing the dynamics of the local volatility model, we need to study how, for a set local volatility function $\sigma(t, S)$, implied volatilities respond to a move of $S$. Rather than solving the forward equation (2.7) for call option prices, we will use an approximate formula that expresses implied volatilities as a function of the local volatility function directly.

We first derive a more general identity.

[6]The ratios $\frac{T-t_i}{T}$, $\frac{t_i}{T}$ in the definition of $\delta S$, $\delta K$, could be replaced by other functions of $\frac{t_i}{T}$, provided these vanish respectively for $t_i = T$ and $t_i = 0$, for the sake of ensuring condition (2.23). We leave this optimization to the reader.

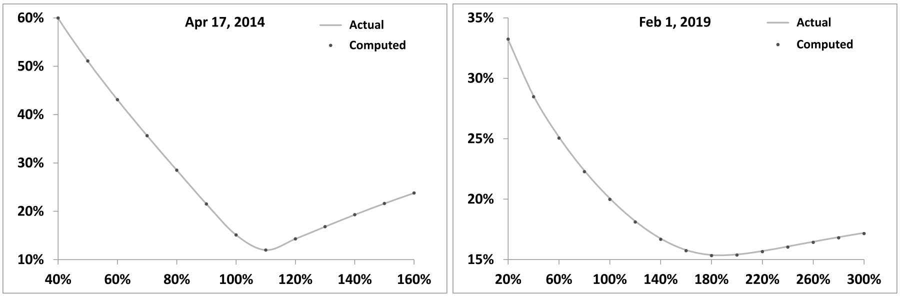



**Figure 2.1**: Comparison of market implied volatilities (solid line) with implied volatilities computed using the local volatility function generated by the approximate technique of Section 2.3.1.2 (dots), for the S&P 500 index, for two maturities. Market parameters as of February 1, 2014 have been used.

### 2.4.1 Implied volatilities as weighted averages of instantaneous volatilities

Consider a base model – denoted model I – a local volatility model with volatility function $\sigma_1(t,S)$. Denote by $P_1(t,S)$ the price of a vanilla option of strike $K$, maturity $T$ in this base model.

Consider another model – denoted model II – such that the dynamics of $S_t$ is given by:

$$dS_t = (r-q)\,S_t dt + \sigma_{2t} S_t dW_t \tag{2.28}$$

where the instantaneous volatility $\sigma_{2t}$ is for now an arbitrary process.

Imagine delta-hedging a short vanilla option position of maturity $T$ using the base model with the actual dynamics of $S_t$ given by the second model. Our final P&L is the sum of all gamma/theta P&Ls between successive delta rehedges, each one equal to: $-S_t^2 \left.\frac{d^2P_1}{dS^2}\right|_{S=S_t} \left(\sigma_{2t}^2 - \sigma_1(t,S_t)^2\right)dt$.

The price $P_2$ we should charge for this option is then the price in model I plus the opposite of an estimate of the sum of such gamma/theta P&Ls. The following derivation makes this trading intuition more precise.

Consider process $Q_t$ defined by:

$$Q_t = e^{-rt} P_1(t,S_t)$$

At $t=0$, $Q_{t=0}$ is simply the initial price in model I, for the initial spot value $S_0$. The variation during $dt$ of $Q_t$ reads, in the dynamics of model II:

$$\begin{aligned} dQ_t &= e^{-rt}\left[\left(-rP_1+\frac{dP_1}{dt}\right)dt+\frac{dP_1}{dS}dS_t+\frac{1}{2}\frac{d^2P_1}{dS^2}\left\langle dS_t^2\right\rangle\right] \\ &= e^{-rt}\left[\left(-rP_1+\frac{dP_1}{dt}\right)dt+\frac{dP_1}{dS}dS_t+\frac{1}{2}\sigma_{2t}^2S_t^2\frac{d^2P_1}{dS^2}dt\right] \end{aligned} \tag{2.29}$$

where $P_1$ and its derivatives are evaluated at $(t, S_t)$. Taking the expectation of (2.29) yields:

$$\begin{aligned} E_2[dQ_t|t,S_t] &= e^{-rt}\left(-rP_1+\frac{dP_1}{dt}+(r-q)\,S_t\frac{dP_1}{dS}+\frac{1}{2}\sigma_{2t}^2S_t^2\frac{d^2P_1}{dS^2}\right)dt \\ &= e^{-rt}\frac{S_t^2}{2}\frac{d^2P_1}{dS^2}\left(\sigma_{2t}^2-\sigma_1(t,S_t)^2\right)dt \end{aligned}$$

where we have made use of pricing equation (2.2) for $P_1$ and where the subscript 2 indicates that the expectation is taken over paths of $S_t$ generated by SDE (2.28). Integrating the above expression on $[0,T]$:

$$\begin{aligned} E_2[Q_T] &= Q_0+\int_0^T E_2[dQ_t] \\ &= P_1\left(0,S_0\right)+E_2\left[\int_0^T e^{-rt}\frac{S_t^2}{2}\frac{d^2P_1}{dS^2}\left(\sigma_{2t}^2-\sigma_1(t,S_t)^2\right)dt\right] \end{aligned}$$

Now $Q_T = e^{-rT}P_1\left(T,S_T\right) = e^{-rT}f\left(S_T\right)$, where $f$ is the European option's payoff, thus $E_2[Q_T] = e^{-rT}E_2[f\left(S_T\right)] = P_2\left(0,S_0,\bullet\right)$ where $P_2$ is the pricing function of model 2, which involves its own parameters in addition to $t, S$. We thus have:

$$P_2\left(0,S_0,\bullet\right) = P_1\left(0,S_0\right)+E_2\left[\int_0^T e^{-rt}\frac{S_t^2}{2}\frac{d^2P_1}{dS^2}\left(\sigma_{2t}^2-\sigma_1(t,S_t)^2\right)dt\right] \tag{2.30}$$

which expresses mathematically what trading intuition suggested.

The price of an option for an arbitrary dynamics of $S_t$ is given by its price in a given base model – here a local volatility model – supplemented with the expectation of the discounted gamma/theta P&Ls incurred as one delta-hedges the option using the base model until maturity.

Assume now that model I is the Black-Scholes model with a constant volatility equal to the implied volatility of the vanilla option at hand, backed out of model II price: $\sigma_1(t,S_t)\equiv\widehat{\sigma}_{KT}$. By definition of $\widehat{\sigma}_{KT}$:

$$P_1\left(0,S_0\right) = P_{\widehat{\sigma}_{KT}}\left(0,S_0\right) = P_2\left(0,S_0,\bullet\right)$$

where $P_{\widehat{\sigma}_{KT}}\left(t,S\right)$ denotes the Black-Scholes price with volatility $\widehat{\sigma}_{KT}$. (2.30) then yields:

$$\widehat{\sigma}_{KT}^2 = \frac{E_2\left[\int_0^T e^{-rt} S_t^2 \frac{d^2 P_{\widehat{\sigma}_{KT}}}{dS^2} \sigma_{2t}^2 dt\right]}{E_2\left[\int_0^T e^{-rt} S_t^2 \frac{d^2 P_{\widehat{\sigma}_{KT}}}{dS^2} dt\right]} \tag{2.31}$$

This is true for an arbitrary instantaneous volatility $\sigma_{2t}$. Specialize now to the case of a local volatility model: $\sigma_{2t} \equiv \sigma(t,S)$ and $\widehat{\sigma}_{KT}$ is the implied volatility corresponding to the local volatility function. We have:

$$\widehat{\sigma}_{KT}^2 = \frac{E_{\sigma(t,S)}\left[\int_0^T e^{-rt} S_t^2 \frac{d^2 P_{\widehat{\sigma}_{KT}}}{dS^2} \sigma\left(t,S\right)^2 dt\right]}{E_{\sigma(t,S)}\left[\int_0^T e^{-rt} S_t^2 \frac{d^2 P_{\widehat{\sigma}_{KT}}}{dS^2} dt\right]} \tag{2.32}$$

$\widehat{\sigma}_{KT}^2$ is thus the average value of $\sigma\left(t,S\right)^2$, weighted by the dollar gamma computed with the constant volatility $\widehat{\sigma}_{KT}$ itself, over paths generated by the local volatility $\sigma\left(t,S\right)$.

Going back to identity (2.30), consider instead that model I is the local volatility model – $\sigma_1\left(t,S\right) \equiv \sigma\left(t,S\right)$ – and that model II is the Black-Scholes model with volatility $\widehat{\sigma}_{KT}$. Again we have $P_1\left(0,S_0\right) = P_2\left(0,S_0\right)$, and now get:

$$\widehat{\sigma}_{KT}^2 = \frac{E_{\widehat{\sigma}_{KT}}\left[\int_0^T e^{-rt} S^2 \frac{d^2 P_{\sigma(S,t)}}{dS^2} \sigma\left(t,S\right)^2 \, dt\right]}{E_{\widehat{\sigma}_{KT}}\left[\int_0^T e^{-rt} \, S^2 \frac{d^2 P_{\sigma(S,t)}}{dS^2} dt\right]} \tag{2.33}$$

where the averaging is now performed using the density generated by a Black-Scholes model with constant volatility $\widehat{\sigma}_{KT}$ and $\sigma\left(t,S\right)^2$ is now weighted by the dollar gamma computed in the local volatility model.[7]

In this derivation we have used as base model the local volatility model, with a fixed volatility function – or the Black-Scholes model, a particular version of it – and have considered a delta-hedged vanilla option position.

In Section 8.4, page 316, we derive a different expression for European option prices in diffusive models by considering a delta-hedged, vega-hedged option position, using as base model a Black-Scholes model whose volatility is constantly recalibrated to the terminal VS volatility.

### 2.4.2 Approximate expression for weakly local volatilities

While the Dupire formula (2.19) explicitly expresses local volatilities as a function of implied volatilities, formulas (2.32) and (2.33) for $\widehat{\sigma}_{KT}$ are implicit, as the right-hand side depends on the unknown implied volatility $\widehat{\sigma}_{KT}$, either through the dollar gamma or through the density used for averaging the numerator and denominator.

[7] As far as I can remember, expression (2.32) for implied volatilities in the local volatility model was first presented by Bruno Dupire at a Global Derivatives conference in the late 90s.

In order to get a more usable expression, we will assume that $\sigma(t, S)$ is only weakly local.

We will use formula (2.32) with the Black-Scholes model with time-dependent volatility $\sigma_0(t)$ as base model.

Let us use the local variance $u(t, S) = \sigma(t, S)^2$ and assume that

$$u(t,S) = u_0(t) + \delta u(t,S)$$

where $u_0 = \sigma_0^2(t)$ and $\delta u$ is a small perturbation. If $\delta u = 0$, $\widehat{\sigma}_{KT} = \widehat{\sigma}_{0T}$ where $\widehat{\sigma}_{t_1 t_2}$ is defined as:

$$\widehat{\sigma}_{t_1 t_2}^2 = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \sigma_0^2(t) dt$$

Let us use expression (2.32) and expand $\widehat{\sigma}_{KT}^2$ at first order in $\delta u$: $\widehat{\sigma}_{KT}^2 + \delta(\widehat{\sigma}_{KT}^2)$. $u(t, S)$ appears explicitly in the numerator as well as implicitly in the density used for computing expectations in both numerator and denominator. For the sake of computing the order-one perturbation in $\delta u$, however, the contribution generated by the density vanishes: from equation (2.32) and using a compact notation:

$$\begin{aligned} \widehat{\sigma}_{KT}^2 + \delta(\widehat{\sigma}_{KT}^2) &= \frac{E_{u_0+\delta u}[(u_0 + \delta u)\bullet]}{E_{u_0+\delta u}[\bullet]} = u_0 + \frac{E_{u_0+\delta u}[\delta u \bullet]}{E_{u_0+\delta u}[\bullet]} \\ &= u_0 + \frac{E_{u_0}[\delta u \bullet]}{E_{u_0}[\bullet]} \end{aligned} \tag{2.34}$$

where $\bullet$ is computed at order zero in $\delta u$. We thus have:

$$\delta(\widehat{\sigma}_{KT}^2) = \frac{E_{\sigma_0}\left[\int_0^T e^{-rt}\, \delta u(t,S)\, S^2 \frac{d^2 P_{\sigma_0}}{dS^2} dt\right]}{E_{\sigma_0}\left[\int_0^T e^{-rt}\, S^2 \frac{d^2 P_{\sigma_0}}{dS^2} dt\right]} \tag{2.35}$$

The right-hand side of equation (2.35) now only requires the density and the dollar gamma of a call or put option, evaluated in the Black-Scholes model with deterministic volatility $\sigma_0(t)$ – both analytically known. The denominator in (2.35) involves the discounted dollar gamma of a European option, averaged over its lifetime. In the Black-Scholes model, $e^{-rt}\, S^2 \frac{d^2 P_{\sigma_0}}{dS^2}$ is a martingale – see a proof in Appendix A of Chapter 5, page 181. Thus:

$$E_{\sigma_0}\left[e^{-rt}\, S^2 \frac{d^2 P_{\sigma_0}}{dS^2}\right] = S_0^2 \left.\frac{d^2 P_{\sigma_0}}{dS^2}\right|_{t=0, S=S_0} \tag{2.36}$$

where $S_0$ denotes today's spot price. The denominator is then equal to $T\, S_0^2 \left.\frac{d^2 P_{\sigma_0}}{dS^2}\right|_{t=0,\, S=S_0}$.

Focus now on the numerator in (2.35). It reads:

$$\int_0^T dt \int_0^\infty dS\, \rho_{\sigma_0}(t,S)\, e^{-rt}\, \delta u(t,S)\; S^2 \frac{d^2 P_{\sigma_0}}{dS^2} \tag{2.37}$$

where $\rho_{\sigma_0}(t,S)$ is the lognormal density with deterministic volatility $\sigma_0(t)$. Define $x = \ln(S/F_t)$ where $F_t$ is the forward for maturity $t$: $F_t = S_0 e^{(r-q)t}$. $\rho_{\sigma_0}(t,S)$ and $S^2 \frac{d^2 P_{\sigma_0}}{dS^2}$ are given by:

$$\rho_{\sigma_0}(t,S) = \frac{1}{\sqrt{2\pi\omega_t}S} e^{-\frac{\left(x+\frac{\omega_t}{2}\right)^2}{2\omega_t}} \tag{2.38}$$

$$S^2 \frac{d^2 P_{\sigma_0}}{dS^2} = S\frac{F_T}{F_t} e^{-r(T-t)} \frac{1}{\sqrt{2\pi(\omega_T-\omega_t)}} e^{-\frac{\left(-x_K+x+\frac{(\omega_T-\omega_t)}{2}\right)^2}{2(\omega_T-\omega_t)}} \tag{2.39}$$

where $x_K = \ln(\frac{K}{F_T})$ and we have introduced the integrated variance $\omega_t$, defined as:

$$\omega_t = \int_0^t \sigma_0^2(\tau)d\tau$$

The numerator then reads:

$$F_T e^{-rT} \int_0^T dt \int_{-\infty}^{+\infty} dx \, \frac{\delta u(t, F_t e^x)}{\sqrt{2\pi(\omega_T-\omega_t)}\sqrt{2\pi\omega_t}} e^x e^{-\frac{\left(-x_K+x+\frac{(\omega_T-\omega_t)}{2}\right)^2}{2(\omega_T-\omega_t)}} e^{-\frac{\left(x+\frac{\omega_t}{2}\right)^2}{2\omega_t}}$$

Combining the exponentials, we can rewrite this expression as:

$$F_T e^{-rT} \int_0^T dt \int_{-\infty}^{+\infty} dx \frac{\delta u(t, F_t e^x)}{\sqrt{2\pi(\omega_T-\omega_t)}\sqrt{2\pi\omega_t}} e^{-\left(\frac{1}{2(\omega_T-\omega_t)}+\frac{1}{2\omega_t}\right)\frac{(\omega_T x - \omega_t x_K)^2}{\omega_T^2}} e^{-\frac{\left(x_K-\frac{\omega_T}{2}\right)^2}{2\omega_T}}$$

We now divide this by (2.36), the dollar gamma evaluated at $t=0$ and multiplied by $T$, which reads:

$$T F_T e^{-rT} \frac{1}{\sqrt{2\pi\omega_T}} e^{-\frac{\left(-x_K+\frac{\omega_T}{2}\right)^2}{2\omega_T}}$$

and get for the ratio in (2.35):

$$\frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dx \, \frac{\sqrt{\omega_T}}{\sqrt{\omega_t}} \frac{\delta u(t, F_t e^x)}{\sqrt{2\pi(\omega_T-\omega_t)}} e^{-\left(\frac{1}{2(\omega_T-\omega_t)}+\frac{1}{2\omega_t}\right)\frac{(\omega_T x - \omega_t x_K)^2}{\omega_T^2}}$$

Switching now from $x$ to a new coordinate $y$:

$$x = \frac{\omega_t}{\omega_T} x_K + \frac{\sqrt{(\omega_T-\omega_t)\omega_t}}{\sqrt{\omega_T}} y$$

yields our final formula for $\delta(\widehat{\sigma}_{KT}^2)$ at order one in $\delta u$:

$$\delta(\widehat{\sigma}_{KT}^2) = \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \, \delta u\left(t, F_t e^{\frac{\omega_t}{\omega_T} x_K + \frac{\sqrt{(\omega_T-\omega_t)\omega_t}}{\sqrt{\omega_T}} y}\right)$$

Noting that, at order zero, $\widehat{\sigma}^2_{KT} = \frac{1}{T}\int_0^T \sigma_0^2(t)dt = \frac{1}{T}\int_0^T u_0(t)dt$, this can be rewritten, at order one in $\delta u = \sigma^2(t,S) - \sigma_0^2(t)$, as:

$$\widehat{\sigma}^2_{KT} = \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy\, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, u\left(t, F_t e^{\frac{\omega_t}{\omega_T}x_K + \frac{\sqrt{(\omega_T-\omega_t)\omega_t}}{\sqrt{\omega_T}}y}\right) \tag{2.40}$$

This is the expression of $\widehat{\sigma}_{KT}$ at order one in the perturbation of $\sigma$ around a time-dependent volatility $\sigma_0(t)$. The square of the implied volatility is expressed as an integral of the square of the instantaneous volatility – thus (2.40) is exact when $u$ is a function of $t$ only.[8]

### 2.4.3 Expanding around a constant volatility

Consider as base case the Black-Scholes model with constant volatility $\sigma_0$: $\sigma_0(t) = \sigma_0$, thus $\omega_t = \sigma_0^2 t$ and write:

$$\sigma(t,S) = \sigma_0 + \delta\sigma(t,S)$$

thus $\delta u = 2\sigma_0\delta\sigma(t,S)$. Using (2.40) together with $\widehat{\sigma}^2_{KT} = \sigma_0^2 + 2\sigma_0\delta(\widehat{\sigma}_{KT})$ yields:

$$\delta\widehat{\sigma}_{KT} = \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy\, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, \delta\sigma\left(t, F_t e^{\frac{t}{T}x_K + \sigma_0\sqrt{\frac{(T-t)t}{T}}y}\right) \tag{2.41}$$

Adding $\sigma_0$ on both sides yields, at order one in $\delta\sigma(t,S)$:

$$\widehat{\sigma}_{KT} = \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy\, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, \sigma\left(t, F_t e^{\frac{t}{T}x_K + \sigma_0\sqrt{\frac{(T-t)t}{T}}y}\right) \tag{2.42}$$

Even though (2.40) may be slightly more accurate when the term-structure of volatilities is strong, we will use (2.42) in the sequel, since resulting expressions for the ATMF skew and the SSR are simpler.

### 2.4.4 Discussion

For $t = 0$, $F_t \exp\left(\frac{t}{T}\ln\frac{K}{F_T} + \frac{\sqrt{\sigma_0^2(T-t)t}}{\sqrt{T}}y\right) = S_0$. Thus values of $\sigma\,(t=0,S)$ for $S \neq S_0$ do not contribute in (2.42) to $\widehat{\sigma}_{KT}$.

Likewise, for $t=T$, $F_t \exp\left(\frac{t}{T}\ln\frac{K}{F_T} + \frac{\sqrt{\sigma_0^2(T-t)t}}{\sqrt{T}}y\right) = K$: values of $\sigma\,(t=T,S)$ for $S \neq K$ do not contribute either. This is natural, upon inspection of expression

[8] This is comforting but should not cause too much rejoicing – (2.40) is only an order-one approximation, besides we will see further below that for $T \to 0$ there is an exact expression of the *inverse* of $\widehat{\sigma}_{KT}$ as an average of the *inverse* of $\sigma\,(0,S)$.

(2.37): for $t = 0$, the density $\rho_{\sigma_0}$ vanishes unless $S = S_0$, while for $t = T$ the dollar gamma $S^2 \frac{d^2 P_{\sigma_0}}{dS^2}$ vanishes, unless $S = K$.

In (2.42), the largest weight is obtained for $y = 0$: this singles out a path for $\ln(S)$ which is a straight line starting at $\ln(S_0)$ for $t = 0$ and ending at $\ln(K)$ at $t = T$. Replacing the integral over $y$ by the value for $y = 0$ would give an approximation of $\widehat{\sigma}_{KT}$ as a uniform average of the local volatility along this line:

$$\widehat{\sigma}_{KT} \simeq \frac{1}{T} \int_0^T \sigma\left(t, F_t e^{\frac{t}{T} \ln \frac{K}{F_T}}\right) dt \tag{2.43}$$

Summing over values of $y \neq 0$, includes other paths in the average, with their ends pinned down at $S_0$ at $t = 0$ and at $K$ for $t = T$ by the factor $\sqrt{\sigma_0^2\,(T-t)\,t}$.

However appealing expression (2.42) for $\widehat{\sigma}_{KT}$ may be, market smiles on equity underlyings are strong enough, and bid-offer spreads on vanilla option prices are tight enough, that its numerical accuracy is not sufficient for practical trading purposes: an order one expansion in $\delta\sigma(t, S)$ is simply not adequate.[9]

Can we do better? Expression (2.42) for $\widehat{\sigma}_{KT}$ amounts to merely setting $\widehat{\sigma}_{KT} = \sigma_0$ and using the lognormal density with volatility $\sigma_0$ for computing both averages in the numerator and denominator of the right-hand side of equation (2.32).

A number of tricks have been proposed under the loose name of "most likely path" techniques for approximating the right-hand side of (2.32), using a lognormal density but with a different implied volatility for each time slice – still, their accuracies are not adequate.

The reason for this is that formula (2.32) seems to suggest that the main contribution of $\sigma\,(t, S)$ is embodied in the explicit gamma term in the numerator and that using an approximate density – for example lognormal – for computing both averages in the numerator and the denominator will do. This is not the case. In practice, taking into account the dependence of the density itself on $\sigma\,(t, S)$ is mandatory for achieving the accuracy needed in trading applications.[10] For realistic equity smiles, there isn't yet a computationally efficient alternative to numerically solving the forward equation (2.7).

Is formula (2.42) then of any use? While a good approximation of absolute volatility levels is hard to come by, the skew – which is the difference of two volatilities – is more easily approximated. We now use equation (2.42) to calculate the skew and curvature of the smile, as a function of parameters of the local volatility function.

---

[9] The cruder version (2.43) is even less usable.

[10] Julien Guyon and Pierre Henry-Labordère provide in [51] a nice summary and comparison of different "most likely path" approximations, along with a technique based on a short-time heat-kernel expansion for $\rho$.

### 2.4.5 The smile near the forward

Let us assume that the local volatility is smooth and given by:

$$\sigma(t,S) = \overline{\sigma}(t) + \alpha(t)\,x + \frac{\beta(t)}{2}x^2 \tag{2.44}$$

where $x$ – which we call moneyness – is given by $x = \ln(S/F_t)$.

We assume that $\alpha(t)$, $\beta(t)$ are small, and that $\overline{\sigma}(t)$ does not vary too much, so that the difference $\overline{\sigma}(t) - \sigma_0$ is small, where $\sigma_0$ is the constant volatility level around which the order-one expansion in (2.35) is performed.

We could as well perform the expansion around the time-deterministic volatility $\overline{\sigma}(t)$ – calculations are similar. For the sake of simplicity we carry out the expansion around a constant $\sigma_0$ – expressions for the more general case appear in (2.60), page 51.

Equation (2.42) gives :

$$\begin{aligned}\widehat{\sigma}_{KT} &= \frac{1}{T}\int_0^T \overline{\sigma}(t)\,dt \qquad (2.45)\\ &+ \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy\, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\left[\alpha(t)X(t,y) + \frac{\beta(t)}{2}X(t,y)^2\right]\end{aligned}$$

where

$$X(t,y) = \frac{t}{T}x_K + \frac{\sqrt{\sigma_0^2(T-t)\,t}}{\sqrt{T}}y \tag{2.46}$$

Doing the integrals over $y$ we get at order 1 in $\alpha, \beta$:

$$\begin{aligned}\widehat{\sigma}_{KT} &= \frac{1}{T}\int_0^T \overline{\sigma}(t)\,dt + \frac{\sigma_0^2 T}{2}\frac{1}{T}\int_0^T \frac{(T-t)\,t}{T^2}\beta(t)dt \qquad (2.47)\\ &+ \left(\frac{1}{T}\int_0^T \frac{t}{T}\alpha(t)dt\right)x_K + \frac{1}{2}\left(\frac{1}{T}\int_0^T \frac{t^2}{T^2}\beta(t)dt\right)x_K^2\end{aligned}$$

Thus, for a sufficiently smooth local volatility function, the skew and curvature of the smile near the forward are related to the skew and curvature of the local volatility function through:

$$\mathcal{S}_T = \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{F_T} = \frac{1}{T}\int_0^T \frac{t}{T}\alpha(t)dt \tag{2.48}$$

$$\left.\frac{d^2\widehat{\sigma}_{KT}}{d\ln K^2}\right|_{F_T} = \frac{1}{T}\int_0^T \left(\frac{t}{T}\right)^2 \beta(t)dt \tag{2.49}$$

where we have introduced $\mathcal{S}_T$ as a notation for the ATMF (at the money forward) skew.[11]

[11] These approximate formulas for the implied ATMF skew and curvature in the local volatility model can be obtained in a number of ways – see [78] for an alternative derivation of the "skew-averaging" expression (2.48).

#### 2.4.5.1 A constant local volatility function

Assume that $\alpha$ and $\beta$ are constant. We get:

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{F_T} = \frac{1}{T}\int_0^T \frac{t}{T}dt\,\alpha = \frac{\alpha}{2} \tag{2.50a}$$

$$\left.\frac{d^2\widehat{\sigma}_{KT}}{d\ln K^2}\right|_{F_T} = \frac{1}{T}\int_0^T \left(\frac{t}{T}\right)^2 \beta dt = \frac{\beta}{3} \tag{2.50b}$$

Thus, for a local volatility function of the form (2.44) with $\alpha$ and $\beta$ constant, at order one in $\alpha, \beta$, the ATMF skew of the implied volatility is half the skew of the local volatility function, while its curvature is one third the curvature of the local volatility function.

#### 2.4.5.2 A power-law-decaying ATMF skew

Let us assume a power-law form for $\alpha(t)$. To prevent divergence as $t \to 0$ we take:

$$\begin{cases} \alpha(t) = \alpha_0 \left(\frac{\tau_0}{t}\right)^\gamma & t > \tau_0 \\ \alpha(t) = \alpha_0 & t \le \tau_0 \end{cases} \tag{2.51}$$

where $\tau_0$ is a cutoff – typically $\tau_0 = 3$ months – and $\gamma$ is the characteristic exponent of the long-term decay of $\alpha(t)$. We get, from (2.48):

$$\begin{cases} \mathcal{S}_T = \frac{\alpha_0}{2} & T \le \tau_0 \\ \mathcal{S}_T = \frac{1}{2-\gamma}\alpha_0 \left(\frac{\tau_0}{T}\right)^\gamma - \frac{\gamma}{2(2-\gamma)}\alpha_0 \left(\frac{\tau_0}{T}\right)^2 & T \ge \tau_0 \end{cases} \tag{2.52}$$

$\gamma$ is typically smaller than 2. For (very) long maturities the second piece in (2.52) can be ignored and we get:

$$\mathcal{S}_T \simeq \frac{1}{2-\gamma}\alpha_0 \left(\frac{\tau_0}{T}\right)^\gamma \tag{2.53}$$

The long-term ATMF skew thus decays with the same exponent $\gamma$ as the local volatility function. For typical equity smiles, $\gamma \simeq \frac{1}{2}$ – see examples in Figure 2.3, page 58. With respect to the local volatility skew $\alpha(t)$, $\mathcal{S}_T$ is rescaled by a factor $\frac{1}{2-\gamma}$.

### 2.4.6 An exact result for short maturities

We consider here the case $T \to 0$, for which a particularly simple relationship linking local and implied volatilities exists, which we now derive.

Let us recall the definition of $y$ and $f$, which appear in expression (2.19):

$$y = \ln\left(\frac{K}{F_T}\right)$$
$$f(T,\ y) = (T - t_0)\,\widehat{\sigma}^2_{K=F_T e^y, T}$$

To lighten the notation we use $\widehat{\sigma}(T,y)$ instead of $\widehat{\sigma}_{K=F_Te^y,T}$ and take $t_0 = 0$. Expression (2.19) reads:

$$\sigma^2 = \frac{\widehat{\sigma}^2 + 2T\widehat{\sigma}\widehat{\sigma}_T}{\left(\frac{y}{\widehat{\sigma}^2}\widehat{\sigma}\widehat{\sigma}_y - 1\right)^2 + T\left(\widehat{\sigma}_y^2 + \widehat{\sigma}\widehat{\sigma}_{yy}\right) - \left(\frac{1}{4} + \frac{1}{T\widehat{\sigma}^2}\right)\widehat{\sigma}^2\widehat{\sigma}_y^2 T^2}$$

where $\widehat{\sigma}_y, \widehat{\sigma}_T$ denote derivatives of $\widehat{\sigma}$ with respect to $y$ and $T$. Take the limit $T \to 0$ and keep the leading term in the numerator and denominator in an expansion in powers of $T$, assuming that $\widehat{\sigma}$ is a smooth function of $T$ and $y$ as $T \to 0$. We get:

$$\sigma^2 = \frac{\widehat{\sigma}^2}{\left(\frac{y}{\widehat{\sigma}}\widehat{\sigma}_y - 1\right)^2} = \frac{1}{\left(\frac{y}{\widehat{\sigma}^2}\widehat{\sigma}_y - \frac{1}{\widehat{\sigma}}\right)^2}$$

which yields

$$\frac{1}{\sigma} = \pm\left(\frac{y}{\widehat{\sigma}^2}\widehat{\sigma}_y - \frac{1}{\widehat{\sigma}}\right) = \mp\left(y\left(\frac{1}{\widehat{\sigma}}\right)_y + \frac{1}{\widehat{\sigma}}\right) = \mp\left(\frac{y}{\widehat{\sigma}}\right)_y$$

Thus

$$\int_0^y \frac{du}{\sigma(T=0,\ Se^u)} = \mp\frac{y}{\widehat{\sigma}(T=0, Se^y)}$$

Following the usual convention of using positive volatilities:

$$\frac{1}{\widehat{\sigma}(T=0, Se^y)} = \frac{1}{y}\int_0^y \frac{du}{\sigma(T=0,\ Se^u)} \tag{2.54}$$

or, equivalently:

$$\frac{1}{\widehat{\sigma}(T=0, K)} = \frac{1}{\ln\frac{K}{S}}\int_S^K \frac{1}{\sigma(T=0,\ \mathcal{S})}\frac{d\mathcal{S}}{\mathcal{S}}$$

This result was first published by Henri Beresticki, Jérôme Busca and Igor Florent [6]. The fact that the inverse of $\widehat{\sigma}$ should be given by the average of the inverse of $\sigma$ may surprise at first.

The squared volatility that one may have expected appears usually in averages only because they are temporal averages, akin to a quadratic variation. In our case $T \to 0$ and there is no temporal averaging.

Then note that the harmonic average complies with the basic requirement that if the local volatility vanishes in a region between the initial spot level and the strike, the implied volatility for that strike should vanish, as for $T \mapsto 0$, the effect of the drift is immaterial and the spot would be unable to cross that region.[12]

[12]The motivation for the harmonic average is that it appears naturally in the density of $\ln(S_T)$ for short maturities as the change of variable $S \to z = \int_{S_0}^S \frac{dS}{S\sigma(S)}$ results in a process for $z_t$ which is Gaussian at short times. A derivation of (2.54) using the zeroth order of the heat-kernel expansion can be found in Section 5.2.2 of [56].

## 2.5 The dynamics of the local volatility model

Once the volatility function $\sigma(t, S)$ is set, expression (2.32) shows that, in the local volatility model, $\widehat{\sigma}_{KT}$ changes only if time advances or $S$ moves. Mathematically, this is a consequence of the fact that the local volatility model is a market model for spot and implied volatilities that has a one-dimensional Markovian representation in terms of $t, S$.

Thus, practically, to characterize the joint dynamics of spot and implied volatilities, we only need to analyze how implied volatilities respond to a move of $S$.

As will be made clear in Section 2.7 below, the volatilities of volatilities and spot/volatility covariances, or SSRs, that we derive below are exactly the break-even levels of the P&L of a delta and vega-hedged position – inclusive of recalibration of the local volatility function – an aspect that looks counter-intuitive at first.

### 2.5.1 The dynamics for strikes near the forward

How much do implied volatilities move when $S$ moves? Let us take the derivative of equation (2.42) with respect to $\ln(S_0)$, introducing the notation $S(t,y) = F_t \exp(\frac{t}{T}x_K + \frac{\sqrt{\sigma_0^2(T-t)t}}{\sqrt{T}}y)$ and remembering that $F_t$ and $x_K$ depend on $S_0$: $F_t = S_0 e^{(r-q)t}$; $x_K = \ln\left(\frac{K}{S_0 e^{(r-q)T}}\right)$.

$$
\begin{aligned}
\frac{d\widehat{\sigma}_{KT}}{d\ln S_0} &= \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \, \frac{d\sigma}{d\ln S_0}(t, S(t,y)) \\
&= \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \, \frac{d\sigma}{dS}(t, S(t,y)) \, \frac{dS(t,y)}{d\ln S_0} \\
&= \frac{1}{T}\int_0^T dt \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \, \frac{d\sigma}{dS}(t, S(t,y)) \left(1 - \frac{t}{T}\right) S(t,y) \\
&= \frac{1}{T}\int_0^T \left(1 - \frac{t}{T}\right) dt \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \, \frac{d\sigma}{d\ln S}(t, S(t,y))
\end{aligned}
$$

Using expression (2.44) for $\sigma(t, S)$ and the definition of $X(t,y)$ in (2.46) we get:

$$
\begin{aligned}
\frac{d\widehat{\sigma}_{KT}}{d\ln S_0} &= \frac{1}{T}\int_0^T \left(1 - \frac{t}{T}\right) dt \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \Big(\alpha(t) + \beta(t) X(t,y)\Big) \\
&= \frac{1}{T}\int_0^T \left(1 - \frac{t}{T}\right) \alpha(t) dt + \left[\frac{1}{T}\int_0^T \left(1 - \frac{t}{T}\right) \frac{t}{T} \beta(t) dt\right] x_K
\end{aligned}
$$

Let us consider the special case of the ATMF volatility, that is the implied volatility for a strike equal to the forward: $x_K = 0$. We get:

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln S_0}\right|_{K=F_T} = \frac{1}{T}\int_0^T \Big(1-\frac{t}{T}\Big)\alpha(t)dt$$

This formula quantifies how the implied volatility for a fixed strike equal to the forward $F_T$ moves when the spot moves. It resembles equation (2.48) except the weight is $1-\frac{t}{T}$ rather than $\frac{t}{T}$. This is natural, as when $S_0$ moves while $K$ stays fixed, for $t = T$ only the *value* $\sigma(T,\ S = K)$ contributes to formula (2.32), thus $\alpha(t = T)$ is immaterial.

Symmetrically, for calculating how the implied volatility changes with strike $K$ for a fixed spot $S_0$, knowledge of $\alpha(t)$ for $t = 0$ is not needed as only the *value* $\sigma\,(t = 0, S_0)$ is contributing – hence the vanishing weight for $\alpha\,(t = 0)$ in equation (2.48).

Consider now the motion of the ATMF implied volatility $\widehat{\sigma}_{K=F_T T}$ keeping in mind that, as $S_0$ moves, strike $K$ moves as well, so as to track the change in the forward $F_T$. We get the sum of two contributions:

$$\begin{aligned}\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} &= \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{K=F_T} + \left.\frac{d\widehat{\sigma}_{KT}}{d\ln S_0}\right|_{K=F_T} \\ &= \frac{1}{T}\int_0^T \frac{t}{T}\alpha(t)dt + \frac{1}{T}\int_0^T \Big(1-\frac{t}{T}\Big)\alpha(t)dt \\ &= \frac{1}{T}\int_0^T \alpha(t)dt \end{aligned} \tag{2.55}$$

Thus the rate at which the ATMF volatility varies as the spot moves is simply given by the uniform time average of the skew of the local volatility function at the forward. In practice, given a market smile, $\alpha\,(t)$ is not accessible, but $\mathcal{S}_T$ is. Inverting equation (2.48) gives:

$$\alpha\,(t) = \frac{d}{dt}\,(t\mathcal{S}_t) + \mathcal{S}_t \tag{2.56}$$

Inserting this expression in equation (2.55) yields:

$$\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} = \mathcal{S}_T + \frac{1}{T}\int_0^T \mathcal{S}_t dt \tag{2.57}$$

The rate at which the ATMF volatility moves when the spot moves is purely dictated by the term structure of the ATMF skew for maturities from 0 to $T$.

Let us assume that $\mathcal{S}$ is constant. Formula (2.57) then gives:

$$\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} = 2\mathcal{S}_T \tag{2.58}$$

We get the property that, for weak skews that do not depend on maturity, the rate at which the ATMF implied volatility moves as the spot moves is exactly twice the rate at which the implied volatility varies as a function of the strike, near the forward.

This is a fundamental feature of the dynamics of implied volatilities in the local volatility model: their dynamics is entirely determined by the implied smile to which the model has been calibrated.

Recalling (2.48), let us summarize the three key properties for implied volatilities near the forward that we have derived at order one in $\alpha(t)$, in an expansion around a constant volatility $\sigma_0$:

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{K=F_T} = \frac{1}{T}\int_0^T \frac{t}{T}\alpha(t)dt \tag{2.59a}$$

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln S_0}\right|_{K=F_T} = \frac{1}{T}\int_0^T \left(1-\frac{t}{T}\right)\alpha(t)dt \tag{2.59b}$$

$$\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} = \frac{1}{T}\int_0^T \alpha(t)dt = \mathcal{S}_T + \frac{1}{T}\int_0^T \mathcal{S}_t dt \tag{2.59c}$$

**Expanding around a time-dependent volatility**

What about expanding around a time-dependent volatility $\overline{\sigma}(t)$ rather than a constant volatility $\sigma_0$? Starting from expression (2.40), page 44, and local volatility function (2.44): $\sigma(t,S) = \overline{\sigma}(t) + \alpha(t)x + \frac{\beta(t)}{2}x^2$, and expanding around $\sigma_0(t) = \overline{\sigma}(t)$ yields the following expressions:

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{K=F_T} = \frac{1}{T}\int_0^T \frac{\widehat{\sigma}_t^2 t}{\widehat{\sigma}_T^2 T}\frac{\overline{\sigma}(t)}{\widehat{\sigma}_T}\alpha(t)dt \tag{2.60a}$$

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln S_0}\right|_{K=F_T} = \frac{1}{T}\int_0^T \left(1-\frac{\widehat{\sigma}_t^2 t}{\widehat{\sigma}_T^2 T}\right)\frac{\overline{\sigma}(t)}{\widehat{\sigma}_T}\alpha(t)dt \tag{2.60b}$$

$$\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} = \frac{1}{T}\int_0^T \frac{\overline{\sigma}(t)}{\widehat{\sigma}_T}\alpha(t)dt = \mathcal{S}_T + \frac{1}{T}\int_0^T \frac{\overline{\sigma}^2(t)}{\widehat{\sigma}_t\widehat{\sigma}_T}\mathcal{S}_t dt \tag{2.60c}$$

where $\widehat{\sigma}_\tau = \sqrt{\frac{1}{\tau}\int_0^\tau \overline{\sigma}^2(u)du}$. $\overline{\sigma}(t)$ is arbitrary – a natural choice is to calibrate $\overline{\sigma}(t)$ to the term structure of ATMF volatilities.

### 2.5.2 The Skew Stickiness Ratio (SSR)

It is useful to normalize $\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0}$ by the ATMF skew of maturity $T$, $\mathcal{S}_T$, thus defining a dimensionless number $\mathcal{R}_T$ which we call the Skew Stickiness Ratio (SSR) for maturity $T$:

$$\mathcal{R}_T = \frac{1}{\mathcal{S}_T}\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} \tag{2.61}$$

$\mathcal{R}_T$ quantifies how much the ATMF volatility for maturity $T$ responds to a move of the spot, *in units of the ATMF skew.*

$\mathcal{R}_T$ will be given in Chapter 9 a more general definition as the regression coefficient of the ATMF volatility on $\ln S$, normalized by the ATMF skew:

$$\mathcal{R}_T = \frac{1}{\mathcal{S}_T}\frac{\langle d\widehat{\sigma}_{F_TT} d\ln S_0\rangle}{\langle (d\ln S_0)^2\rangle} \tag{2.62}$$

$\mathcal{R}_T$ essentially quantifies the spot/volatility covariance in the model at hand.

In the local volatility model, $\widehat{\sigma}_{F_TT}$ is a *function* of $(t, S)$, hence expression (2.62) for $\mathcal{R}_T$ simplifies to (2.61).[13]

Practitioners routinely refer to two archetypical regimes:

- The "sticky-strike" regime corresponds to $\mathcal{R}_T = 1$. As the spot moves, implied volatilities *for fixed strikes* near the money stay frozen – the ATMF volatility slides along the smile.
- The "sticky-delta" regime corresponds to $\mathcal{R}_T = 0$. The whole smile experiences a translation alongside the spot: volatilities *for fixed log-moneyness* are frozen.

While the sticky-delta regime is observed for all $T$ in models with iid increments for $\ln S$ – such as jump-diffusion models – sticky-strike behavior is only observed as a limiting regime for long maturities for certain types of stochastic volatility models – see Section 9.5.

Using expression (2.57) for $\frac{d\widehat{\sigma}_{F_TT}}{d\ln S_0}$ we get:

$$\langle d\widehat{\sigma}_{F_TT} d\ln S_0\rangle = \left(\mathcal{S}_T + \frac{1}{T}\int_0^T \mathcal{S}_t dt\right)\left\langle (d\ln S_0)^2\right\rangle dt \tag{2.63}$$

which yields the following approximate expression for the SSR in the local volatility model, at order one in $\delta\sigma(t, S)$:

$$\mathcal{R}_T = 1 + \frac{1}{T}\int_0^T \frac{\mathcal{S}_t}{\mathcal{S}_T} dt \tag{2.64}$$

For strong skews (2.64) typically overestimates the SSR – see the examples in Figure 2.4, page 59. This is due to the omission of higher-order contributions from $\alpha(t)$ in the expansion.

**Expanding around a time-dependent volatility**

Expanding around a time-dependent volatility $\overline{\sigma}(t)$, rather than a constant, yields the following expression for $\mathcal{R}_T$:

$$\mathcal{R}_T = 1 + \frac{1}{T}\int_0^T \frac{\overline{\sigma}^2(t)}{\widehat{\sigma}_t\widehat{\sigma}_T}\frac{\mathcal{S}_t}{\mathcal{S}_T} dt \tag{2.65}$$

This follows directly from (2.60c) – we use the same notations.[14]

[13]See Section 9.2 for a study of the SSR in stochastic volatility models.

[14]Expression (2.65), with $\overline{\sigma}(t)$ calibrated to the term structure of ATMF volatilities, is more accurate than its counterpart (2.64). (2.64) is, however, already a good approximation. It owes its robustness to

#### 2.5.2.1 The $\mathcal{R} = 2$ rule

In case $\mathcal{S}$ does not depend on maturity, or equivalently when $\alpha(t)$ is constant, (2.57) – or (2.64) – yields:

$$\mathcal{R}_T \;=\; 2 \tag{2.66}$$

for all $T$. This is also true in the limit $T \to 0$ if $\alpha\,(t)$ is smooth:

$$\lim_{T\to 0} \frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} \;=\; 2\lim_{T\to 0} \left.\frac{d\widehat{\sigma}_{KT}}{d\ln S_0}\right|_{K=F_T} \;=\; 2\lim_{T\to 0} \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{K=F_T} \tag{2.67}$$

Thus for short maturities, in the local volatility model:

$$\lim_{T\to 0} \mathcal{R}_T \;=\; 2 \tag{2.68}$$

We will see in Chapter 9 that this property is shared by stochastic volatility models.

### 2.5.3 The $\mathcal{R} = 2$ rule is exact

While we have just derived them using approximation (2.42), we now show that the properties

- $\mathcal{R}_T = 2\ \forall T$, if $\alpha(t)$ is a constant
- $\lim_{T\to 0} \mathcal{R}_T \;=\; 2$

are in fact exact.

#### 2.5.3.1 Time-independent local volatility functions

Imagine that the local volatility function is a function of $\frac{S}{F_t}$ only:

$$\sigma\,(t, S) \;\equiv\; \sigma\left(\frac{S}{F_t}\right) \tag{2.69}$$

with $F_t = S^{\star} e^{(r-q)t}$, where $S^{\star}$ is some fixed reference spot level. The time dependence is embedded in the moneyness and $\sigma$ has no explicit time dependence: we call this a time-independent local volatility function.

Let $C(tS; KT)$ be the price of a call option of maturity $T$ and strike $K$, computed at time $t$ and spot value $S$. $C$ solves the following usual backward equation:

$$\frac{dC}{dt} + (r-q)\,S\frac{dC}{dS} + \frac{1}{2}\sigma^2\left(\frac{S}{F_t}\right)S^2\frac{d^2C}{dS^2} \;=\; rC$$

with terminal condition: $C(t = T, S; KT) = (S - K)^+$. Consider now the change of variables: $\tau = T - t,\ \ s = S/F_t,\ \ k = K/F_T$ and let $f\,(\tau s; k)$ be the solution of the following *forward* equation:

the fact that it does not involve $\sigma_0$, the constant volatility around which the order-one expansion is performed – see Figure 2.4, page 59.

$$\frac{df}{d\tau} = \frac{1}{2}\sigma^2(s)\, s^2 \frac{d^2 f}{ds^2} \tag{2.70}$$

with initial condition: $f(\tau = 0, s; k) = (s-k)^+$. $C$ can be expressed as:

$$C(tS; KT) = e^{-r\tau} F_T f(\tau s; k) \tag{2.71}$$

Let now $P(tS; KT)$ be the price of a *put* option of maturity $T$ and strike $K$, computed at time $t$ and spot value $S$ and let us now express the fact that $P$ solves the forward equation:

$$\frac{dP}{dT} + (r-q)\, K\frac{dP}{dK} - \frac{1}{2}\sigma^2\left(\frac{K}{F_T}\right) K^2 \frac{d^2 P}{dK^2} = -qP$$

with initial condition $P(tS; K, T = t) = (K-S)^+$. $P$ can be written as:

$$P(tS; KT) = e^{-r\tau} F_T f(\tau k; s) \tag{2.72}$$

Notice how the right-hand sides of (2.71) and (2.72) are identical, except $s$ and $k$ are exchanged. For a constant $\sigma$, the Black-Scholes solution of (2.70) is denoted $f_{BS}(\tau s; k; \sigma)$. Given a general solution of (2.70) with initial condition $(s-k)^+$ let us denote $\Sigma_{k\tau}(s)$ its Black-Scholes implied volatility; $\Sigma_{k\tau}(s)$ is such that:

$$f(\tau s; k) = f_{BS}(\tau s; k; \Sigma_{k\tau}(s))$$

(2.71) and (2.72) can be rewritten as:

$$C(tS; KT) = e^{-r\tau} F_T f_{BS}(\tau s; k; \Sigma_{k\tau}(s)) \tag{2.73}$$

$$P(tS; KT) = e^{-r\tau} F_T f_{BS}(\tau k; s; \Sigma_{s\tau}(k)) \tag{2.74}$$

The following identity holds for $f_{BS}$:

$$f_{BS}(\tau s; k; \Sigma) = (s-k) + f_{BS}(\tau k; s; \Sigma) \tag{2.75}$$

Using (2.75) and the call/put parity, we derive from (2.74) the following expression for the value of the *call* option:

$$C(tS; KT) = e^{-r\tau} F_T f_{BS}(\tau s; k; \Sigma_{s\tau}(k)) \tag{2.76}$$

The right-hand sides of equations (2.73) and (2.76) are Black-Scholes formulas for the price of a call option with the same strike and maturity, computed for the same initial spot value. Their implied volatilities are then identical:

$$\Sigma_{k\tau}(s) = \Sigma_{s\tau}(k) \tag{2.77}$$

Standard implied volatilities $\widehat{\sigma}_{KT}(S)$ are given by: $\widehat{\sigma}_{KT}(S) = \Sigma_{\frac{K}{F_T}, T}\left(\frac{S}{F_t}\right)$. Using (2.77) we get our final result:[15]

$$\widehat{\sigma}_{S\frac{F_T}{F_t}, T}\left(K\frac{F_t}{F_T}\right) = \widehat{\sigma}_{KT}(S) \tag{2.78}$$

[15] I am grateful to Julien Guyon for pointing out this symmetry property to me – see also exercise 9.1 in [56].

For zero interest rate and repo this simplifies to:

$$\widehat{\sigma}_{ST}(K) \ = \ \widehat{\sigma}_{KT}(S)$$

Thus knowledge of implied volatilities of all strikes for a given initial spot level $S$ (the right-hand side) supplies information on the implied volatility of a particular strike equal to the initial spot $S$, for all values of the spot level (the left-hand side).

**The $R = 2$ rule**

Taking the derivative of both sides of equation (2.78) with respect to $\ln(K)$ and setting $t = 0$, $K = F_T$ and $S = F_t = S_0$ then yields:

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{K=F_T} \ = \ \left.\frac{d\widehat{\sigma}_{KT}}{d\ln S_0}\right|_{K=F_T} \tag{2.79}$$

From this we derive the relationship linking the rate at which the ATMF implied volatility moves as the spot moves to the ATMF skew:

$$\frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} \ = \ \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{K=F_T} + \left.\frac{d\widehat{\sigma}_{KT}}{d\ln S_0}\right|_{K=F_T} \ = \ 2\left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{K=F_T}$$

Hence:

$$\mathcal{R}_T \ = \ 2$$

This is an important result. The rule that the rate at which the ATMF implied volatility moves when the spot moves is twice the ATMF skew – or that the SSR equals 2 – is in fact exact for local volatilities that are a function of $S/F_t$ only.

The reason why the order-one expansion of $\widehat{\sigma}_{KT}$ in equation (2.42) yields this result is that it complies with the symmetry condition (2.78). Remember that in (2.42), $F_t$ is the forward associated to $S$, not the reference spot $S^\star$. For a local volatility of the form (2.69), equation (2.42) reads:

$$\widehat{\sigma}_{KT}(S) \ = \ \frac{1}{T-t}\int_t^T du \int_{-\infty}^{+\infty} dy \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, \sigma\left(\frac{S}{S^*} e^{\frac{u-t}{T-t}\ln\left(\frac{KF_t}{SF_T}\right) + \frac{\sqrt{\sigma_0^2 (T-u)(u-t)}}{\sqrt{T-t}} y}\right)$$

where we have set the initial time equal to $t$. One can check that, replacing in this expression $S$ with $K\frac{F_t}{F_T}$ and $K$ with $S\frac{F_T}{F_t}$ and making the change of variables $u \to T + t - u$ leaves the integrand unchanged and yields:

$$\widehat{\sigma}_{S\frac{F_T}{F_t},T}\left(K\frac{F_t}{F_T}\right) \ = \ \widehat{\sigma}_{KT}(S)$$

The diligent reader will have noticed that the backward-forward symmetry condition that yields equations (2.73) and (2.74) still holds if the local volatility function $\sigma$ is allowed to depend on $t$ such that $\sigma(t, s)$ is symmetric on $[0, T]$ with respect to $\frac{T}{2}$. Again one can check that if this holds, expression (2.42) yields identity (2.79).

#### 2.5.3.2 Short maturities

Consider the case $T \to 0$. Implied volatilities are then given by the exact formula 2.54. One can check using (2.54) that the identity (2.79) holds, hence:

$$\mathcal{R}_T \;=\; 2$$

holds, for any local volatility function: for $t \to 0$ the local volatility function becomes in effect "time-independent".

### 2.5.4 SSR for a power-law-decaying ATMF skew

Let us use the power-law benchmark (2.53) for $\mathcal{S}_T$, with characteristic exponent $\gamma$ and vanishing cutoff: $\tau_0 = 0$. (2.64) yields the following maturity-independent value of $\mathcal{R}_T$:

$$\frac{2-\gamma}{1-\gamma} \tag{2.80}$$

That $\mathcal{R}_T$ does not depend on $T$ is due to our assumption of a vanishing cutoff. In practice $\mathcal{S}_T$ does not diverge as $T \to 0$.

Assume that $\mathcal{S}_T$ is given by (2.52), which is derived from expression (2.51) for $\alpha(t)$ with characteristic exponent $\gamma$ and cutoff $\tau_0$. Evaluation of the integral in (2.64) is straightforward. The resulting profile of $\mathcal{R}_T$ appears in Figure 2.2 for $\gamma = \frac{1}{2}$ and $\tau_0 = 0.05$ and $0.25$.

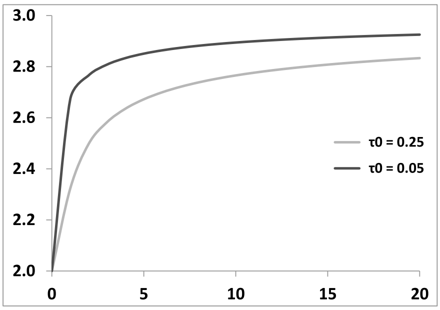



**Figure 2.2**: $\mathcal{R}_T$ as a function of $T$ (years) as given by formula (2.64), using expression (2.52), page 47, for $\mathcal{S}_T$, with $\gamma = \frac{1}{2}$ and two values of $\tau_0$. $\mathcal{R}_\infty = 3$.

$\mathcal{R}_T$ is very sensitive to $\tau_0$. The limiting value (2.80) for long maturities

$$\mathcal{R}_\infty \;=\; \frac{2-\gamma}{1-\gamma} \tag{2.81}$$

may thus be reached for outrageously large maturities only – the limiting value in Figure 2.2 is $\mathcal{R}_\infty = 3$ ($\gamma = \frac{1}{2}$). Stated differently, $\mathcal{R}_T$ is very dependent on the ATMF skew for short maturities.

What happens when $\gamma = 1$? From (2.81) $\mathcal{R}_\infty = \infty$. How fast does $\mathcal{R}_T$ diverge? Going back to expression (2.64) we can see that, for large $T$, $\mathcal{R}_T \propto \ln T$. Again, the precise value of $\mathcal{R}_T$ depends on the short end of the smile – see the example in Section 2.5.6 below and Figure 2.6, page 61, for a smile whose ATMF skew decays like $\frac{1}{T}$ for long maturities.

### 2.5.5 Volatilities ofvolatilities

In the local volatility model implied volatilities are a function of $S_t$.

$$d\widehat{\sigma}_{F_T T} \;=\; \frac{d\widehat{\sigma}_{F_T T}}{d\ln S}\, d\ln S_t \;+\; \bullet\, dt$$

From (2.61):

$$d\widehat{\sigma}_{F_T T} \;=\; \mathcal{R}_T \mathcal{S}_T \, d\ln S_t \;+\; \bullet\, dt \tag{2.82}$$

Let us now set $t = 0$ and note that the instantaneous volatility $\sigma(0, S_0)$ is equal to the short ATMF volatility $\widehat{\sigma}_{F_0 0}$: $\left\langle d\ln S^2 \right\rangle = \widehat{\sigma}^2_{F_0 0} dt$. The instantaneous (lognormal) volatility of $\widehat{\sigma}_{F_T T}$ is given by:

$$\mathrm{vol}(\widehat{\sigma}_{F_T T}) \;=\; \mathcal{R}_T \mathcal{S}_T \frac{\widehat{\sigma}_{F_0 0}}{\widehat{\sigma}_{F_T T}}$$

- Inserting expression (2.64) for $\mathcal{R}_T$:

$$\mathrm{vol}(\widehat{\sigma}_{F_T T}) \;=\; \left(\mathcal{S}_T + \frac{1}{T}\int_0^T \mathcal{S}_t dt\right) \frac{\widehat{\sigma}_{F_0 0}}{\widehat{\sigma}_{F_T T}} \tag{2.83}$$

- If instead we use expression (2.65) for $\mathcal{R}_T$:

$$\mathrm{vol}(\widehat{\sigma}_{F_T T}) \;=\; \left(\mathcal{S}_T + \frac{1}{T}\int_0^T \frac{\overline{\sigma}^2(t)}{\widehat{\sigma}_t \widehat{\sigma}_T} \mathcal{S}_t dt\right) \frac{\widehat{\sigma}_{F_0 0}}{\widehat{\sigma}_{F_T T}} \tag{2.84}$$

For short maturities $\mathcal{R}_T = 2$ and we get:

$$\mathrm{vol}(\widehat{\sigma}_{F_T T}) \;=\; 2\mathcal{S}_T \tag{2.85}$$

The (lognormal) volatility of a short volatility is just twice the ATMF skew.

For typical equity index smiles, whose ATMF skews decrease with $T$, the longer the maturity, the lower the instantaneous volatility of the ATMF volatility. For a power-law decay of the ATMF skew with a characteristic exponent $\gamma$, we get, for long maturities and ignoring the factor $\frac{\widehat{\sigma}_{F_0 0}}{\widehat{\sigma}_{F_T T}}$, which is only dependent on the term structure of ATMF volatilities:

$$\mathrm{vol}(\widehat{\sigma}_{F_T T}) \;=\; \frac{2-\gamma}{1-\gamma}\, \mathcal{S}_T \tag{2.86}$$

For long maturities the volatility of $\widehat{\sigma}_{F_T T}$ thus approximately decays as a function of $T$ with the same exponent as the ATMF skew.[16]

### 2.5.6 Examples and discussion

We now illustrate what we have just discussed with the example of two Euro Stoxx 50 smiles, then end with a remark on local volatility considered as a stochastic volatility model.

We use the Euro Stoxx 50 smiles of October 4, 2010 (a strong smile) and May 16, 2013 (a mild smile). Let us first consider implied volatilities for, respectively, September 16, 2011 and June 20, 2014 – roughly a 1-year maturity in both cases – and only use implied volatility data for this single maturity for calibrating the local volatility function.

As we have a single maturity, we take the local volatility function to be a function of $\frac{S}{F_t}$ only, so that it falls in the class of time-independent local volatilities. This is easily achieved by using the parametrization $f(t, y)$ in equations (2.18a) and (2.18b) where $f(T, y)$ is given for $T$, our single maturity. $f$ is defined for $t < T$ by $f(t, y) = \frac{t}{T} f(T, y)$. We are then in the setting of Section 2.5.3.1. Once we have calibrated a time-independent local volatility function, we move the initial spot value $S_0$ and reprice vanilla options.

The resulting smiles, along with the initial smile, are shown in Figure 2.3. In each smile the marker highlights the ATMF volatility.

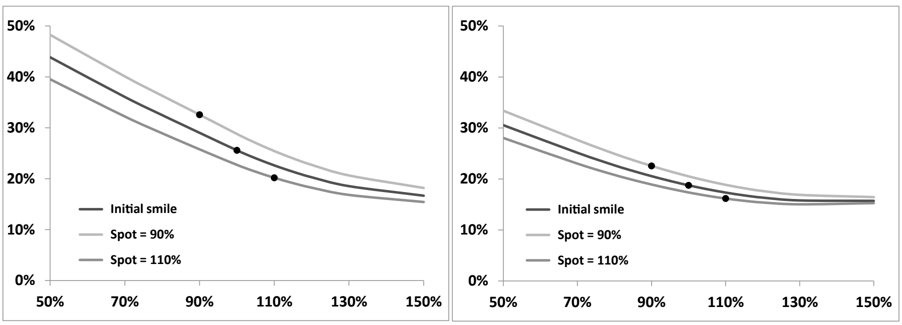



**Figure 2.3**: Smiles of the Euro Stoxx 50 index for a maturity $\simeq$ 1 year (see text), observed on October 4, 2010 (left) and May 16, 2013 (right), along with smiles produced by the local volatility model – calibrated on these initial smiles – for two other values of $S_0$.

[16]Typically the ATMF skews of index smiles decay like $\frac{1}{\sqrt{T}}$. (2.86) implies that in the local volatility model, vol($\widehat{\sigma}_{F_T T}$) decays approximately like $\frac{1}{\sqrt{T}}$ as well.

That the rate at which the ATMF volatility varies when $S_0$ varies is twice the ATMF skew – or equivalently that $\mathcal{R}_T = 2$ – is apparent to the eye.[17]

We now use implied volatilities for all of the available maturities. The local volatility function cannot be assumed to be time-independent anymore and the SSR will be different than 2. We use formula (2.64):

$$\mathcal{R}_T \;=\; 1 + \frac{1}{T}\int_0^T \frac{\mathcal{S}_t}{\mathcal{S}_T}dt \tag{2.87}$$

$\mathcal{R}_T$ as a function of $T$ is shown in Figure 2.4, together with the actual value of $\mathcal{R}_T$ obtained by shifting the spot value and repricing vanilla options.

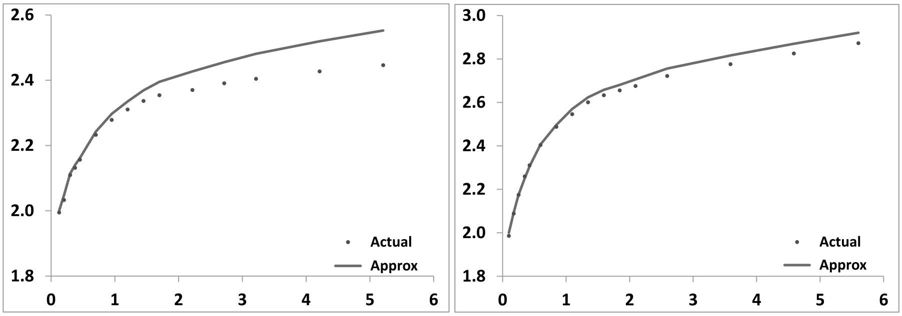



**Figure 2.4**: $\mathcal{R}_T$ for the Euro Stoxx 50 index as a function of $T$ computed: (a) directly in the local volatility model (actual), (b) using expression (2.87) (approx), for the smiles of October 4, 2010 (left) and May 16, 2013 (right).

Agreement is good except for the long end of the smile of October 4, 2010; (2.87) overestimates the SSR as the order-one expansion that leads to (2.87) ignores contributions from higher orders, which become material for strong skews.

Still the relative error in the estimation of the SSR – or equivalently in the level of volatility of volatility, or spot/volatility covariance – is about $5\%$.

For the smile of May 16, 2013, which displays an appreciable (increasing) term-structure of ATMF volatilities, using (2.65) rather than (2.87) results in a slightly higher value for $\mathcal{R}_T$ – about 0.05. For smiles with a strong term-structure of ATMF volatilities, that are not too steep, (2.65) is in practice more accurate than (2.87).

That ATMF skews of equity smiles are well captured by the power-law benchmark (2.52) is illustrated in Figure 2.5.

In our two examples the SSR values given by (2.87) using either the actual market ATMF skew or expression (2.52) are similar.

[17]We have used zero repo and interest rate for simplicity.

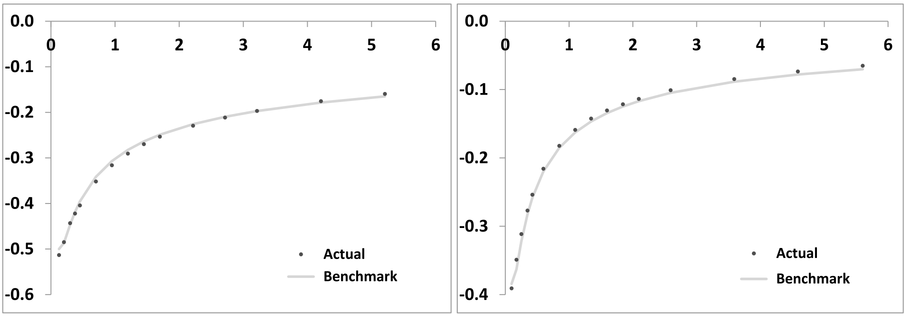



**Figure 2.5**: $\mathcal{S}_T$ for the Euro Stoxx 50 index as a function of $T$ (in years) as read off the market smile (actual) and as given by the power-law benchmark (2.52), for the smiles of October 4, 2010 (left: $\tau_0 = 0.15,\ \gamma = 0.37$) and May 16, 2013 (right: $\tau_0 = 0.12,\ \gamma = 0.52$).

The long-maturity value of the Âš is given by (2.81):

$$\mathcal{R}_\infty = \frac{2-\gamma}{1-\gamma}$$

$\mathcal{R}_\infty = 2.6$ for the October 4, 2010 smile and $\mathcal{R}_\infty = 3.1$ for the May 16, 2013 smile – we know that $\mathcal{R}_\infty$ is only reached for very long maturities.

In the local volatility model, because implied volatilities are a *function* of $(t, S)$ the SSR provides substantial information: it determines both the break-even levels of the spot/volatility cross-gamma *and* of the volatility gamma. Expressions (2.64) for $\mathcal{R}_T$ and (2.83) for $\mathrm{vol}(\widehat{\sigma}_{F_T T})$ are useful for sizing up these break-even levels and comparing them to realized levels.

Remember that the SSR involves the ratio of the spot/volatility covariance to the ATMF skew. Large values of the SSR may only be due to weak or vanishing ATMF skews and may not be a signal of particularly large volatilities of volatilities or spot/volatility covariances.

This applies to the following example of a fast-decaying ATMF skew.

**A $\frac{1}{T}$ decay for the ATMF skew**

We now consider the case of a smile whose long-term ATMF skew decays like $\frac{1}{T}$. This is the case of stochastic volatility models of Type I – see Section 9.5 in Chapter 9.

For $\gamma = 1$ formula (2.81) yields $\mathcal{R}_\infty = \infty$. From (2.87), it can be checked that if $\mathcal{S}_T \propto \frac{1}{T}$ for large $T$, then $\mathcal{R}_T \propto \ln T$.

Figure 2.6 shows that this is indeed the case. The local volatility function is calibrated on a smile generated by a one-factor stochastic volatility model of the type discussed in Chapter 7, with $k = 6.0$; for $T \gg \frac{1}{k}$ the resulting ATMF skew decays like $\frac{1}{T}$.[18]

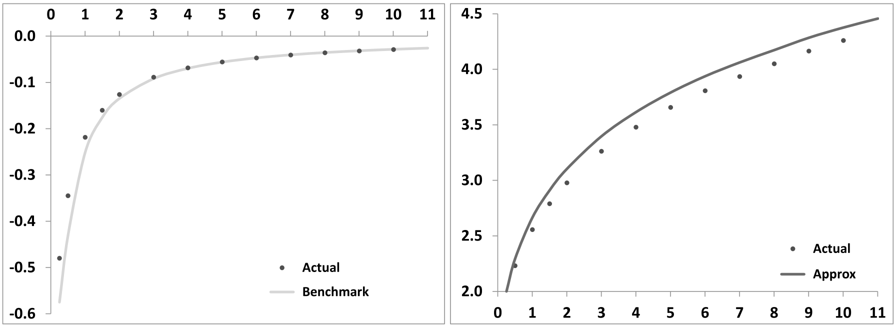



**Figure 2.6**: Left: $\mathcal{S}_T$ as a function of $T$, as read off the smile used as input (actual) and as given by the power-law benchmark (2.52), with $\gamma = 0.999$). Right: $\mathcal{R}_T$ as a function of $T$ computed: (a) directly in the local volatility model (actual), (b) using expression (2.87) (approx). The smile used as input has been generated by a stochastic volatility model of Type I – its ATMF skew decays like $\frac{1}{T}$ for large $T$.

### 2.5.7 SSR in local and stochastic volatility models

Consider the instantaneous covariance of $\ln S_t$ and $\widehat{\sigma}_{F_T T}$ – denoted more compactly by $\widehat{\sigma}_T(t)$ – observed at $t$. Because $\widehat{\sigma}_T(t)$ is a function of $\ln S_t$ we have simply:

$$\begin{aligned}\frac{\langle d\ln S_t d\widehat{\sigma}_T(t)\rangle}{dt} &= \frac{d\widehat{\sigma}_T(t)}{d\ln S_t}\sigma^2(t, S_t)\\ &= \widehat{\sigma}_t^2\Big(\frac{1}{T-t}\int_t^T \alpha(\tau)d\tau\Big)\end{aligned}$$

where we have used expression (2.59c) for $\frac{d\widehat{\sigma}_T}{d\ln S}$, applied to time $t$ rather than 0, and the fact that $\sigma(t, S_t) = \widehat{\sigma}_t$.

We are using formulas at order one in $\alpha(t)$, perturbing around a constant volatility $\sigma_0$. At order one in $\alpha(t)$, we can use the order-zero value for $\widehat{\sigma}_{F_t t}$, that is $\sigma_0$. We thus have:

$$\frac{\langle d\ln S_t d\widehat{\sigma}_T(t)\rangle}{dt} = \sigma_0^2\Big(\frac{1}{T-t}\int_t^T \alpha(\tau)d\tau\Big) \tag{2.88}$$

[18] We are discussing the SSR of the local volatility model as calibrated on the smile generated by a stochastic volatility model. The SSR of the stochastic volatility model is different as the instantaneous spot/volatility covariance is different. Type I models, whose ATMF skew decays like $\frac{1}{T}$ or faster are such that $\mathcal{R}_T \to 1$ for $T \to \infty$ – see Section 9.5, page 361.

Multiplying both sides by $(T-t)$ and integrating with respect to $t$ on $[0,T]$ yields:

$$\begin{aligned}\int_0^T (T-t)\frac{\langle d\ln S_t d\widehat{\sigma}_T(t)\rangle}{dt}dt &= \sigma_0^2\int_0^T dt\int_t^T \alpha(\tau)d\tau \\ &= \sigma_0^2\int_0^T \tau\alpha(\tau)d\tau \\ &= \sigma_0^2 T^2\mathcal{S}_T\end{aligned}$$

where the last line follows from the order-one expression of the ATMF skew in (2.59a). We thus get:

$$\mathcal{S}_T = \frac{1}{\widehat{\sigma}_T^2 T}\int_0^T \frac{T-t}{T}\frac{\langle d\ln S_t\, d\widehat{\sigma}_T(t)\rangle}{dt}\,dt \tag{2.89}$$

where we have replaced in the denominator $\sigma_0^2$ with $\widehat{\sigma}_T^2$, still preserving the order-one accuracy in $\alpha(t)$.

As a formula for $\mathcal{S}_T$ (2.89) is useless – we may just as well use (2.59a). What it expresses though – that the ATMF skew for maturity $T$ is given by the integrated instantaneous covariance of $\ln S_t$ and the ATMF volatility for the residual maturity, $\widehat{\sigma}_T(t)$, weighted by $\frac{T-t}{T}$ – has wider relevance.

We have derived it here in the context of local volatility at order one in $\alpha(t)$ but this result is more general.[19]

As will be proven in Section 8.4 of Chapter 8, page 316, formula (2.89) holds for any diffusive model, at order one in volatility of volatility, whenever the instantaneous spot/variance covariation $\langle d\ln S_t d\widehat{\sigma}_T(t)\rangle$ does not depend on $S_t$. This is the case for a local volatility function linear in $\ln S$, hence (2.89).

Expression (2.89) accounts for why local volatility and stochastic volatility models calibrated to the same smile may have different SSRs, or, equivalently, generate different break-even levels for the spot/volatility cross-gamma.

From (2.89), the ATMF skew sets the integrated value of the covariance of $\ln S$ and $\widehat{\sigma}_T(t)$, the implied ATMF volatility for the residual maturity. Changing the distribution of this covariance on $[0,T]$ without changing its integrated value leaves the ATMF skew unchanged, but changes $\langle d\ln S_t d\widehat{\sigma}_T(t)\rangle_{t=0}$ – which sets the value of the SSR.

With respect to time-homogeneous stochastic volatility models, the local volatility model tends to generate larger covariances at short times – and consequently smaller covariances at future times. This translates into:

- larger SSRs than in time-homogeneous stochastic volatility models
- weaker future skews

See also related discussions in Section 9.11.1 of Chapter 9, page 379, and Section 12.6 of Chapter 12, page 482.

[19]Had we started from the order-one expansion (2.40) around a deterministic volatility $\sigma_0(t)$ rather than around a constant volatility $\sigma_0$, we would have obtained the exact same formula.

## 2.6 Future skews and volatilities of volatilities

In Section 1.3.1 we showed how the price of a barrier digital option is mostly determined by the magnitude of the local skew at the barrier for the residual maturity, as generated by the model used for pricing. To assess how a given model prices barrier options, one needs to investigate the ATMF skews generated by a model for a given residual maturity, at future dates, and for different future spot levels. We do this now, for the local volatility model.

Imagine we are using a local volatility model calibrated to the market smile, with a local volatility function given by (2.44).

Let us assume that we are sitting at the forward date $\tau > 0$ with spot $S_\tau$. What is the ATMF skew generated by the local volatility model? Expression (2.48) for the skew was derived for $t = 0$ and a local volatility function given by (2.44) where $F_t$ is the forward at time $t$ for the initial spot level: $F_t = S_0 e^{(r-q)t}$.

We can reuse the results above, but first need to express the local volatility function as a function of $y = \ln\left(\frac{S}{F_t(S_\tau)}\right)$, where $F_t(S_\tau) = S_\tau e^{(r-q)(t-\tau)}$. We have:

$$\begin{aligned} x &= \ln\left(\frac{S}{F_t}\right) = \ln\left(\frac{S}{F_t(S_\tau)}\right) + \ln\left(\frac{F_t(S_\tau)}{F_t}\right) \\ &= y + x_\tau \end{aligned}$$

where $x_\tau = \ln\left(\frac{S_\tau}{F_\tau}\right)$. Sitting at time $\tau$ and using $S_\tau$ as reference spot level, the local volatility function for $t > \tau$ is given by:

$$\begin{aligned} \sigma(t,S) &= \overline{\sigma}(t) + \alpha(t)(y + x_\tau) + \frac{\beta(t)}{2}(y + x_\tau)^2 \\ &= \overline{\sigma}_\tau(t) + \alpha_\tau(t)\, y + \frac{\beta_\tau(t)}{2} y^2 \end{aligned}$$

with

$$\left\{ \begin{array}{lcl} \overline{\sigma}_\tau(t) &=& \overline{\sigma}(t) + \alpha(t)\, x_\tau + \frac{\beta(t)}{2} x_\tau^2 \\ \alpha_\tau(t) &=& \alpha(t) + \beta(t) x_\tau \\ \beta_\tau(t) &=& \beta(t) \end{array} \right.$$

If $\beta(t) \neq 0$, since $\alpha_\tau(t)$ depends on $x_\tau$, the ATMF skew at time $\tau$ will depend on the spot level $S_\tau$. Let us, however, set $x_\tau = 0$ – that is $S_\tau = F_\tau$ – and focus instead on how the ATMF skew at $\tau$ for a given residual maturity $\theta$ depends on $\tau$; or equivalently consider that $\beta(t) = 0$.

Using (2.48), the ATMF skew at time $\tau$ for a residual maturity $\theta$ – that is for maturity $\tau + \theta$ – is given by:

$$\mathcal{S}_\theta(\tau) = \left.\frac{d\widehat{\sigma}_{K\tau+\theta}(S_\tau,\tau)}{d\ln K}\right|_{K=F_{\tau+\theta}(S_\tau)} = \frac{1}{\theta}\int_\tau^{\tau+\theta} \frac{t-\tau}{\theta}\alpha(t)dt \qquad (2.90)$$

Using now expression (2.56), page 50, for $\alpha(t)$ we get the following expression of the foward-starting ATMF skew as a function of the term structure of the ATMF skew read off the vanilla smile used for calibration:

$$\mathcal{S}_\theta(\tau) \;=\; \mathcal{S}_{\tau+\theta} \;-\; \frac{\tau}{\theta}\left(\frac{1}{\theta}\int_\tau^{\tau+\theta}\mathcal{S}_t dt - \mathcal{S}_{\tau+\theta}\right) \tag{2.91}$$

where $\mathcal{S}_t$ is the spot-starting ATMF skew for maturity $t$.

- Formula (2.91) for $\mathcal{S}_\theta(\tau)$ involves the ATMF skew of the initial vanilla smile, but only for maturities in $[\tau, \tau+\theta]$. Thus there is no reason that $\mathcal{S}_\theta(\tau)$ should bear any resemblance with $\mathcal{S}_\theta$.

- The second piece in (2.91) involves $\mathcal{S}_{\tau+\theta}$ minus the average of $\mathcal{S}_t$ on the interval $[\tau, \tau+\theta]$. For a decreasing term structure of the ATMF skew – which is typical – the latter is larger, in absolute value. We then have the property that:

$$|\mathcal{S}_\theta(\tau)| \;\leq\; |\mathcal{S}_{\tau+\theta}| \;\ll\; |\mathcal{S}_\theta|$$

  The future skew is weaker than the spot-starting skew for maturity $\tau+\theta$, thus (much) weaker than the spot-starting skew for the same residual maturity.

- Consider the case of a power-law decaying skew with $\alpha(t)$ given by (2.51), without any cutoff for simplicity: $\alpha(t) \;\propto\; \left(\frac{\tau_0}{t}\right)^\gamma$. Focus on the case of a very short residual maturity $\theta$. From (2.90) for $\theta$ small, we have:

$$\mathcal{S}_\theta \;=\; \frac{1}{2-\gamma}\alpha_0\left(\frac{\tau_0}{\theta}\right)^\gamma$$
$$\mathcal{S}_\theta(\tau) \;=\; \frac{\alpha(\tau)}{2} \;=\; \frac{\alpha_0}{2}\left(\frac{\tau_0}{\tau}\right)^\gamma$$

  thus

$$\mathcal{S}_\theta(\tau) \;\propto\; \left(\frac{\theta}{\tau}\right)^\gamma \mathcal{S}_\theta \tag{2.92}$$

  For typical equity smiles $\gamma \simeq \frac{1}{2}$. Thus, in the local volatility model, future ATMF skews will be much weaker than spot-starting skews for the same residual maturity: $\mathcal{S}_\theta(\tau) \ll \mathcal{S}_\theta$. This is apparent in Figure 2.7 which shows ATMF skews for the two smiles considered in the examples of Section 2.5.6, for different forward dates, calculated using (2.91).

- In the local volatility model, volatilities of ATMF volatilities are determined by the ATMF skew – see formula (2.83), page 57. Thus low levels of future ATMF skews translate also into low future levels of volatility of volatility.

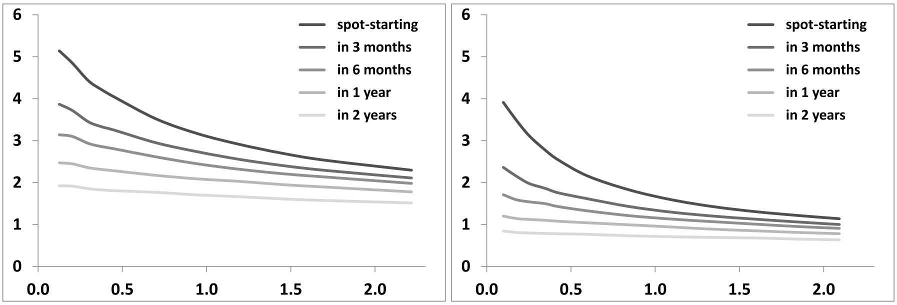



**Figure 2.7**: $\mathcal{S}_\theta(\tau)$ as a function of residual maturity $\theta$ for different future dates: $\tau = 0$ (forward-starting), 3 months, 6 months, 1 year and 2 years. $\mathcal{S}_\theta(\tau)$ is evaluated using (2.91) and multiplied by $\ln(95/105)$ to convert it into the difference of implied volatilities for the 95% and 105% moneynesses. Smiles of the Euro Stoxx 50 index on October 4, 2010 (left) and May 16, 2013 (right) have been used.

### 2.6.1 Comparison with stochastic volatility models

Our analysis has concentrated on the price as generated by the local volatility model at $t = 0$. We derive in the following section the expression of the daily P&L of a delta and vega-hedged position. As we risk-manage our forward-start or barrier option together with its vanilla hedge, the daily P&L of the hedged position reads as in (2.105). In case gammas and cross-gammas of the hedged position are sizeable, these P&Ls will be large and unpredictable, randomly polluting our final P&L – this is the price we pay for using the local volatility model.

In contrast, the stochastic volatility models of Chapter 7 are time-homogeneous. Future skews and future volatilities of volatilities are determined by the model's parameters and are commensurate with their values at $t = 0$ – they are not altered by recalibration to the term structure of VS volatilities.

When using the local volatility model, it is then essential to realize that the model price at $t = 0$ incorporates assumptions on future skews and break-even levels of spot/volatility gammas and cross-gammas that cannot be locked and will change as the model is recalibrated to future market smiles.

We refer the reader to Section 3.2 of the following chapter, page 119, where we continue our investigation of forward-start options in the local volatility model; the case of a forward-start call is covered in detail.

See also Section 12.6.1 of Chapter 12 where future smiles generated by local, stochastic and local volatility models are compared.

## 2.7 Delta and carry P&L

We now consider two practically important issues that need to be answered before we can even consider using a model for trading purposes:

- Which delta should we trade?
- What is the carry P&L of a hedged position?

We will show that the local volatility model is indeed a legitimate market model.

### 2.7.1 The "local volatility delta"

The price $P^{\mathrm{LV}}$ of a derivative in the local volatility model is given by:

$$P^{\mathrm{LV}}(t, S, \sigma)$$

where $\sigma$ denotes the local volatility function.

Imagine using the "local volatility delta", that is the delta computed with a fixed local volatility function:

$$\Delta^{\mathrm{LV}} = \left.\frac{dP^{\mathrm{LV}}}{dS}\right|_{\sigma} \tag{2.93}$$

The pricing equation (2.2) implies that the P&L during $\delta t$ of a delta-hedged position reads – see expression (1.5):

$$P\&L \;=\; -\frac{1}{2}S^2\frac{d^2P^{\mathrm{LV}}}{dS^2}\left(\frac{\delta S^2}{S^2} - \sigma^2(t,S)\,\delta t\right) \tag{2.94}$$

In the local volatility model, implied volatilities of vanilla options are functions of $t$ and $S$. For strike $K$ and maturity $T$:

$$\widehat{\sigma}_{KT}(t,S) \;\equiv\; \Sigma^{\mathrm{LV}}_{KT}(t,S,\sigma) \tag{2.95}$$

where $\sigma$ denotes the local volatility function used.

As $S$ moves by $\delta S$ during $\delta t$, only if market implied volatilities $\widehat{\sigma}_{KT}$ move as prescribed by (2.95) does the P&L of the delta-hedged position read as in (2.94). Expression (2.94) has thus little usefulness.

In reality, implied volatilities will move however they wish. Changes in market implied volatilities $\delta\widehat{\sigma}_{KT}$ will be arbitrary, thus the local volatility function calibrated to the market smile at $t + \delta t$ differs from that calibrated at time $t$: our P&L will include additional terms reflecting this change.

Practically, we will be using the local volatility model in a way that it wasn't meant to be used – that is, recalibrating daily the local volatility function on market smiles. Is this nonsensical? Is it nevertheless possible to express simply our carry P&L? Do payoff-independent break-even levels for volatilities of implied volatilities and correlations of spot and implied volatilities exist?

### 2.7.2 Using implied volatilities – the sticky-strike delta $\Delta^{\mathrm{SS}}$

Denote by $P(t, S, \widehat{\sigma}_{KT})$ the option price given by the local volatility model, at time $t$, for a spot value $S$ and implied volatilities $\widehat{\sigma}_{KT}$:

$$P(t, S, \widehat{\sigma}_{KT}) \equiv P^{\mathrm{LV}}(t, S, \sigma[t, S, \widehat{\sigma}_{KT}]) \tag{2.96}$$

where the notation $\sigma[t, s, \widehat{\sigma}_{KT}]$ signals that the local volatility function is calibrated at time $t$, spot $S$, to the volatility surface $\widehat{\sigma}_{KT}$. Our state variables are thus $S$ and the $\widehat{\sigma}_{KT}$. Equivalently, we could use option prices $O_{KT}$ rather than implied volatilities – see the following section in this respect.

$P(t, S, \widehat{\sigma}_{KT})$ is our pricing function. Assume we are short the option; the P&L during the interval $[t, t+\delta t]$ of our option position − without its delta hedge − is simply:

$$P\&L = -\left(P(t+\delta t,\ S+\delta S,\ \widehat{\sigma}_{KT}+\delta\widehat{\sigma}_{KT}) - (1+r\delta t)P(t, S, \widehat{\sigma}_{KT})\right)$$

Expanding at order one in $\delta t$ and two in $\delta S$ and $\delta\widehat{\sigma}_{KT}$:

$$\begin{aligned} P\&L = {} & rP\delta t \\ & -\frac{dP}{dt}\delta t - \frac{dP}{dS}\delta S - \frac{dP}{d\widehat{\sigma}_{KT}} \bullet \delta\widehat{\sigma}_{KT} \\ & -\left(\frac{1}{2}\frac{d^2P}{dS^2}\delta S^2 + \frac{d^2P}{dS\,d\widehat{\sigma}_{KT}} \bullet \delta\widehat{\sigma}_{KT}\delta S + \frac{1}{2}\frac{d^2P}{d\widehat{\sigma}_{KT}d\widehat{\sigma}_{K'T'}} \bullet \delta\widehat{\sigma}_{KT}\delta\widehat{\sigma}_{K'T'}\right) \end{aligned} \tag{2.97}$$

The notation $\bullet$ stands for:[20]

$$\frac{df}{d\widehat{\sigma}_{KT}} \bullet \delta\widehat{\sigma}_{KT} \equiv \iint dK\,dT\ \frac{\delta f}{\delta\widehat{\sigma}_{KT}}\ \delta\widehat{\sigma}_{KT}$$

We have made no model assumption so far – (2.97) is a basic accounting statement.

The derivatives $\frac{dP}{dS}$, $\frac{dP}{dt}$ are computed keeping the $\widehat{\sigma}_{KT}$ fixed − the underlying local volatility function is *not* fixed. Let us call $\frac{dP}{dS}$ the sticky-strike delta and denote it by $\Delta^{\mathrm{SS}}$

$$\Delta^{\mathrm{SS}} = \left.\frac{dP}{dS}\right|_{\widehat{\sigma}_{KT}} \tag{2.98}$$

[20] The finicky reader will observe – with reason – that the $\widehat{\sigma}_{KT}$ cannot be considered as independent variables as shifting by a finite amount one single point of the volatility surface creates arbitrage. Similarly, for $T \to 0$, in-the-money options are redundant with respect to $S$.

In practice, implied volatilities $\widehat{\sigma}_{K_iT_j}$ for discrete sets of strikes and maturities are used, out of which the full volatility surface $\widehat{\sigma}_{KT}$ is interpolated/extrapolated. Our pricing function indeed takes as inputs a finite number of parameters.

$\frac{df}{d\widehat{\sigma}_{KT}} \bullet \delta\widehat{\sigma}_{KT}$ should thus really be understood as $\Sigma_{ij}\frac{df}{\widehat{\sigma}_{K_iT_j}}\delta\widehat{\sigma}_{K_iT_j}$.

Let us now utilize the fact that our "black box" valuation function $P$ is in fact the local volatility price. We will express the derivatives of $P^{\text{LV}}$ with respect to $t, S$ – that is with a fixed local volatility function – in terms of derivatives of $P$ and use the pricing equation of the local volatility model to derive an identity involving derivatives of $P$.

By definition of $\Sigma^{\text{LV}}_{KT}(t,S,\sigma)$, $P\left(t,S,\widehat{\sigma}_{KT}=\Sigma^{\text{LV}}_{KT}(t,S,\sigma)\right)$ is the local volatility price:

$$P^{\text{LV}}(t,S,\sigma) \;=\; P\left(t,S,\widehat{\sigma}_{KT}=\Sigma^{\text{LV}}_{KT}(t,S,\sigma)\right)$$

We will not carry the three arguments of $\Sigma^{\text{LV}}_{KT}$ anymore, unless necessary. Taking the derivative of this expression with respect to $t$ and $S$ yields:

$$\frac{dP^{\text{LV}}}{dt} \;=\; \frac{dP}{dt}+\frac{dP}{d\widehat{\sigma}_{KT}}\bullet\frac{d\Sigma^{\text{LV}}_{KT}}{dt} \tag{2.99}$$

$$\frac{dP^{\text{LV}}}{dS} \;=\; \frac{dP}{dS}+\frac{dP}{d\widehat{\sigma}_{KT}}\bullet\frac{d\Sigma^{\text{LV}}_{KT}}{dS} \tag{2.100}$$

Taking once more the derivative of $\frac{dP^{\text{LV}}}{dS}$ with respect to $S$ we get:

$$\begin{aligned}\frac{d^2P^{\text{LV}}}{dS^2} = &\left(\frac{d^2P}{dS^2}+2\frac{d^2P}{dSd\widehat{\sigma}_{KT}}\bullet\frac{d\Sigma^{\text{LV}}_{KT}}{dS}+\frac{d^2P}{d\widehat{\sigma}_{KT}d\widehat{\sigma}_{K'T'}}\bullet\frac{d\Sigma^{\text{LV}}_{KT}}{dS}\frac{d\Sigma^{\text{LV}}_{K'T'}}{dS}\right)\\ &+\frac{dP}{d\widehat{\sigma}_{KT}}\bullet\frac{d^2\Sigma^{\text{LV}}_{KT}}{dS^2}\end{aligned} \tag{2.101}$$

We now use (2.99) to express $\frac{dP}{dt}$ in terms of $\frac{dP^{\text{LV}}}{dt}$ and then use the pricing equation of the local volatility model:

$$\frac{dP^{\text{LV}}}{dt} \;+\; (r-q)S\frac{dP^{\text{LV}}}{dS} \;+\; \frac{1}{2}\sigma^2(t,S)\,S^2\frac{d^2P^{\text{LV}}}{dS^2} \;=\; rP^{\text{LV}} \tag{2.102}$$

to express $\frac{dP^{\text{LV}}}{dt}$ as a function of $\frac{dP^{\text{LV}}}{dS}$, $\frac{P^{\text{LV}}}{dS^2}$. We then use (2.101) and (2.100) to write everything in terms of derivatives of $P$. This yields the following expression of $\frac{dP}{dt}$:

$$\begin{aligned}\frac{dP}{dt} \;=\;& rP \;-\; (r-q)S\frac{dP}{dS} \;-\; \frac{dP}{d\widehat{\sigma}_{KT}}\bullet\mu_{KT}\\ -\frac{1}{2}\sigma^2(t,S)\,S^2&\left(\frac{d^2P}{dS^2}+2\frac{d^2P}{dSd\widehat{\sigma}_{KT}}\bullet\frac{d\Sigma^{\text{LV}}_{KT}}{dS}+\frac{d^2P}{d\widehat{\sigma}_{KT}d\widehat{\sigma}_{K'T'}}\bullet\frac{d\Sigma^{\text{LV}}_{KT}}{dS}\frac{d\Sigma^{\text{LV}}_{K'T'}}{dS}\right)\end{aligned}$$

where $\mu_{KT}$ is given by:

$$\mu_{KT} \;=\; \frac{d\Sigma^{\text{LV}}_{KT}}{dt}+\frac{1}{2}\sigma^2(t,S)\,S^2\frac{d^2\Sigma^{\text{LV}}_{KT}}{dS^2}+(r-q)S\frac{d\Sigma^{\text{LV}}_{KT}}{dS} \tag{2.103}$$

Inserting now this expression of $\frac{dP}{dt}$ in (2.97) yields the following expression for our P&L:

$$
\begin{aligned}
P\&L \;=\; & -\frac{dP}{dS}\left(\delta S-(r-q)S\delta t\right) \;-\; \frac{dP}{d\widehat{\sigma}_{KT}}\bullet\left(\delta\widehat{\sigma}_{KT}-\mu_{KT}\delta t\right)\\
+\frac{1}{2}\sigma^2(t,S)\,S^2 & \left(\frac{d^2P}{dS^2}+2\frac{d^2P}{dSd\widehat{\sigma}_{KT}}\bullet\frac{d\Sigma^{\mathrm{LV}}_{KT}}{dS}+\frac{d^2P}{d\widehat{\sigma}_{KT}d\widehat{\sigma}_{K'T'}}\bullet\frac{d\Sigma^{\mathrm{LV}}_{KT}}{dS}\frac{d\Sigma^{\mathrm{LV}}_{K'T'}}{dS}\right)\delta t\\
& -\left(\frac{1}{2}\frac{d^2P}{dS^2}\delta S^2+\frac{d^2P}{dSd\widehat{\sigma}_{KT}}\bullet\delta\widehat{\sigma}_{KT}\delta S\;+\frac{1}{2}\frac{d^2P}{d\widehat{\sigma}_{KT}d\widehat{\sigma}_{K'T'}}\bullet\delta\widehat{\sigma}_{KT}\delta\widehat{\sigma}_{K'T'}\right)
\end{aligned}
\tag{2.104}
$$

In the first line of (2.104), the linear contributions of $\delta S$ and $\delta\widehat{\sigma}_{KT}$ to the P&L are accompanied by their respective financing costs – or risk-neutral drifts. While the financing cost of $S$ is model-independent, $\mu_{KT}$ is not as it depends on the dynamics of implied volatilities assumed by the model – in our case the local volatility model – hence the LV superscript.

This is not an issue – it only happens because $\widehat{\sigma}_{KT}$ is not a tradeable asset; when we use option prices rather than implied volatilities, the model-dependence of the drift disappears – see Section 2.7.3 below.

The P&L in (2.104) can be rewritten so as to make the break-even volatilities and correlations of $\delta S$ and $\delta\widehat{\sigma}_{KT}$ apparent. Denote by $\nu_{KT}$ the instantaneous (lognormal) volatility of $\widehat{\sigma}_{KT}$ in the local volatility model – that is of $\Sigma^{\mathrm{LV}}_{KT}$:

$$
\nu_{KT} \;=\; \frac{1}{\Sigma^{\mathrm{LV}}_{KT}}\frac{d\Sigma^{\mathrm{LV}}_{KT}}{dS}S\sigma(t,S)
$$

Our P&L during $\delta t$ can be rewritten as:

$$
P\&L \;=
$$

$$
-\frac{dP}{dS}\left(\delta S-(r-q)S\delta t\right) \;-\; \frac{dP}{d\widehat{\sigma}_{KT}}\bullet\left(\delta\widehat{\sigma}_{KT}-\mu_{KT}\delta t\right) \tag{2.105a}
$$

$$
-\frac{1}{2}S^2\frac{d^2P}{dS^2}\left[\frac{\delta S^2}{S^2}-\sigma^2(t,S)\,\delta t\right] \tag{2.105b}
$$

$$
-\frac{d^2P}{dSd\widehat{\sigma}_{KT}}\bullet S\widehat{\sigma}_{KT}\left[\frac{\delta S}{S}\frac{\delta\widehat{\sigma}_{KT}}{\widehat{\sigma}_{KT}}-\sigma(t,S)\,\nu_{KT}\delta t\right] \tag{2.105c}
$$

$$
-\frac{1}{2}\frac{d^2P}{d\widehat{\sigma}_{KT}d\widehat{\sigma}_{K'T'}}\bullet\widehat{\sigma}_{KT}\widehat{\sigma}_{K'T'}\left[\frac{\delta\widehat{\sigma}_{KT}}{\widehat{\sigma}_{KT}}\frac{\delta\widehat{\sigma}_{K'T'}}{\widehat{\sigma}_{K'T'}}-\nu_{KT}\nu_{K'T'}\delta t\right] \tag{2.105d}
$$

$\frac{dP}{dS}$ in (2.105a) is the sticky-strike delta $\Delta^{\mathrm{SS}}$. This is our total P&L – (2.105) incorporates the P&L generated by the change in local volatility function over $[t,t+\delta t]$ – up to second order – except it is expressed in terms of changes in implied volatilities.

### 2.7.3 Using option prices – the market-model delta $\Delta^{\text{MM}}$

While we have used implied volatilities $\widehat{\sigma}_{KT}$ as state variables, there is nothing special about them. The reader can check that the fact that $\widehat{\sigma}_{KT}$ is an implied volatility has played no part in the derivation leading to (2.105).

We could have used a different representation of the vanilla smile – for example straight option prices. Let us then replace $\widehat{\sigma}_{KT}$ with $O_{KT}$, the price of the vanilla option (say, a call) of strike $K$, maturity $T$ and replace $\Sigma^{\text{LV}}_{KT}(t,S,\sigma)$ with $\Omega^{\text{LV}}_{KT}(t,S,\sigma)$, the price of the same vanilla option in the local volatility model, as a function of time $t$, spot $S$, and the local volatility function $\sigma$.

Denote by $\mathcal{P}(t,S,O_{KT})$ the price of our exotic option, now a function of vanilla option prices, rather than implied volatilities. $\mathcal{P}$ is related to $P$ through:

$$P(t,S,\widehat{\sigma}_{KT}) \;=\; \mathcal{P}\big(t,S,O_{KT} = P^{\text{BS}}_{KT}(t,S,\widehat{\sigma}_{KT})\big) \tag{2.106}$$

The expression of our P&L is similar to (2.105), with $\widehat{\sigma}_{KT}$ replaced with $O_{KT}$. Drift $\mu_{KT}$, according to expression (2.103), is now given by:

$$\mu_{KT} \;=\; \frac{d\Omega^{\text{LV}}_{KT}}{dt} + \frac{1}{2}\sigma^2(t,S)\,S^2\frac{d^2\Omega^{\text{LV}}_{KT}}{dS^2} + (r-q)S\frac{d\Omega^{\text{LV}}_{KT}}{dS}$$

Because $\Omega^{\text{LV}}_{KT}$ is the price in the local volatility model, with a fixed local volatility function, it obeys (2.102), thus we have:

$$\mu_{KT} \;=\; r\Omega^{\text{LV}}_{KT} \;=\; rO_{KT}$$

– which we knew in the first place: $O_{KT}$ is the price of an asset, thus its drift is model-independent and is simply its financing cost.

The expression of our P&L during $\delta t$ using option prices thus reads:

$$P\&L =$$

$$-\frac{d\mathcal{P}}{dS}\,(\delta S - (r-q)S\delta t) \;-\; \frac{d\mathcal{P}}{dO_{KT}} \bullet (\delta O_{KT} - rO_{KT}\delta t) \tag{2.107a}$$

$$-\frac{1}{2}\frac{d^2\mathcal{P}}{dS^2}\left[\delta S^2 - \sigma^2(t,S)\,S^2\delta t\right] \tag{2.107b}$$

$$-\frac{d^2\mathcal{P}}{dSdO_{KT}} \bullet \left[\delta S\delta O_{KT} - \sigma^2(t,S)\,S^2\frac{d\Omega^{LV}_{KT}}{dS}\delta t\right] \tag{2.107c}$$

$$-\frac{1}{2}\frac{d^2\mathcal{P}}{dO_{KT}dO_{K'T'}} \bullet \left[\delta O_{KT}\delta O_{K'T'} - \sigma^2(t,S)\,S^2\frac{d\Omega^{LV}_{KT}}{dS}\frac{d\Omega^{LV}_{K'T'}}{dS}\delta t\right] \tag{2.107d}$$

Three observations are in order:

- Expression (2.107) for our carry P&L is typical of a market model – remember our discussion in Section 1.1. Second-order gamma contributions involving prices of all hedge instruments – the spot and vanilla options – appear together with their offsetting thetas. The corresponding break-even levels are payoff-independent.

- (2.107) makes plain that the hedge ratios of our exotic option are simply given by $\frac{d\mathcal{P}}{dS}$ and $\frac{d\mathcal{P}}{dO_{KT}}$. In particular, the delta – that is the sensitivity of $\mathcal{P}$ to a move of $S$ that is not offset by the vanilla option hedge – is given by $\frac{d\mathcal{P}}{dS}$. This is obtained by moving $S$ while keeping vanilla option prices fixed. This is the natural delta as generated in any market model: move one asset value keeping all others unchanged. We call it the market-model delta, denoted by $\Delta^{\text{MM}}$:

$$\Delta^{\text{MM}} = \left.\frac{d\mathcal{P}}{dS}\right|_{O_{KT}} \tag{2.108}$$

- In a market model – which local volatility is – the delta of a vanilla option is an irrelevant notion. $S$ and $O_{KT}$ are prices of two different assets, which are both hedge instruments. The incongruity of asking a model to output a hedge ratio of one hedge instrument on another is made manifest in the two-asset example of Section 2.7.6 below.

The local volatility delta of vanilla options, $\frac{d\Omega_{KT}^{\text{LV}}}{dS}$, only appears in the expressions of the break-even levels as option prices in the local volatility model are functions of $(t, S)$. The instantaneous covariance of two securities $O_1, O_2$ in the model – be they the underlying or vanilla options – is then given by $\frac{d\Omega_1^{\text{LV}}}{dS}\frac{d\Omega_2^{\text{LV}}}{dS}\sigma^2(t,S)\,S^2\delta t$.

### 2.7.4 Consistency of $\Delta^{\text{SS}}$ and $\Delta^{\text{MM}}$

$\Delta^{\text{MM}}$ is the "real" delta of the local volatility model: it is the number of units of $S$ that need to be held in the hedge portfolio, alongside a position $\frac{d\mathcal{P}}{dO_{KT}}$ in vanilla options of strike $K$, maturity $T$.

$\Delta^{\text{SS}}$ on the other hand is tied to a specific representation of vanilla option prices – in terms of Black-Scholes lognormal implied volatilities. Should we use it? What is its connection to $\Delta^{\text{MM}}$?

Our hedge portfolio $\Pi$ comprises $\Delta^{\text{MM}} = \frac{d\mathcal{P}}{dS}$ units of the underlying and $\frac{d\mathcal{P}}{dO_{KT}}$ vanilla options of strike $K$, maturity $T$:

$$\Pi = \frac{d\mathcal{P}}{dS}S + \frac{d\mathcal{P}}{dO_{KT}} \bullet O_{KT}$$

Typically, the convention on an exotic desk is to use delta-hedged – rather than naked – vanilla options, where the delta hedge is the vanilla option's Black-Scholes delta computed using the option's implied volatility $\widehat{\sigma}_{KT}$. The composition of our hedge portfolio can thus be rewritten differently:

$$\Pi = \left[\frac{d\mathcal{P}}{dS} + \frac{d\mathcal{P}}{dO_{KT}} \bullet \frac{dP^{\text{BS}}_{KT}}{dS}\right] S + \frac{d\mathcal{P}}{dO_{KT}} \bullet \left[O_{KT} - \frac{dP^{\text{BS}}_{KT}}{dS} S\right]$$

What does $\frac{d\mathcal{P}}{dS} + \frac{d\mathcal{P}}{dO_{KT}} \bullet \frac{dP^{\text{BS}}_{KT}}{dS}$ correspond to? It is the sensitivity of $\mathcal{P}$ to a simultaneous move of $S$ and a variation of vanilla option prices generated by their Black-Scholes deltas. Saying that a vanilla option's price varies by its Black-Scholes delta times the variation of $S$ is equivalent to saying that its Black-Scholes implied volatility stays fixed. This is the sticky-strike delta: $\frac{d\mathcal{P}}{dS} + \frac{d\mathcal{P}}{dO_{KT}} \bullet \frac{dP^{\text{BS}}_{KT}}{dS} = \Delta^{\text{SS}}$, which we can rewrite as:

$$\Delta^{\text{MM}} + \frac{d\mathcal{P}}{dO_{KT}} \bullet \frac{dP^{\text{BS}}_{KT}}{dS} = \Delta^{\text{SS}}$$

Thus $\Delta^{\text{SS}}$ is equal to the market-model delta $\Delta^{\text{MM}}$ augmented by the Black-Scholes deltas of the hedging vanilla options. Once Black-Scholes deltas of the hedging options are accounted for, we recover $\Delta^{\text{MM}}$ as the aggregate delta in our hedge portfolio.

Had we used a different representation of vanilla option prices,[21] we would have obtained a different "sticky-strike" delta. Yet, once the hedge porfolio is broken down into underlying + *naked* vanilla options the delta is always equal to $\Delta^{\text{MM}}$.

$\Delta^{\text{SS}}$ thus has no special status. It is simply the delta an exotic desk should trade if – as is customary – the delta hedges of delta-hedged vanilla options used as vega hedges are their Black-Scholes deltas.

### 2.7.5 Local volatility as the simplest market model

Expression (2.107) for the carry P&L is so natural from a trading point of view that the derivation in Section 2.7.2 seems unnecessarily cumbersome on one hand and on the other hand does not shed much light on why things pan out so neatly – indeed we have obtained (2.107) by using the local volatility model in an unorthodox manner, as we recalibrate the local volatility function at $t + \delta t$.

Consider a stochastic volatility market model[22] defined by the following joint SDEs for the spot and vanilla option prices:

$$dS_t = (r - q) S_t dt + \sigma_t S_t dW^S_t \tag{2.109a}$$

$$dO_{KT,t} = rO_{KT,t} dt + \lambda_{KT,t} dW^{KT}_t \tag{2.109b}$$

and the following condition:

$$O_{KT,t=T} = (S_T - K)^+ \tag{2.110}$$

[21] For example the implied intensity of a Poisson process.

[22] By stochastic volatility market model, we mean here a market model driven by diffusive processes.

together with the initial values $S_{t=0}$, $O_{KT,t=0}$. Processes $\sigma_t$ (resp. $\lambda^t_{KT}$) are the instantaneous lognormal (resp. normal) volatilities of $S_t$ (resp. $O_{KT}$). Imagine we are able to build such a model – that is a solution to (2.109) that satisfies (2.110). Then, the carry P&L during $\delta t$ in such a model is exactly of the form in (2.107), with the respective break-even levels given by:

- $S_t^2\sigma_t^2$ for the spot/spot gamma
- $\sigma_t S_t \lambda_{KT,t}\rho_{S,KT,t}$ for spot/vanilla option cross-gammas
- $\lambda_{KT,t}\lambda_{K'T',t}\rho_{KT,K'T',t}$ for the option/option cross-gammas.

$\rho_{S,KT,t}$ and $\rho_{KT,K'T',t}$ are, respectively, the instantaneous correlations of $S_t$ and $O_{KT,t}$ and of $O_{KT,t}$ and $O_{K'T',t}$.

Building market models from scratch is difficult. Furthermore only market models possessing a Markov representation in terms of a (small) finite number of state variables can realistically be considered. Failing that, pricing requires simultaneous simulation of $S_t$ as well as all of the $O_{KT}$, a task which is unfeasible numerically.

Denote by $\sigma[O_{KT}, t, S]$ the local volatility function calibrated at time $t$ using vanilla option prices $O_{KT}$ and spot value $S$, and denote by $\sigma[O_{KT}, t, S](\tau, \mathcal{S})$ its value for time $\tau$ and spot value $\mathcal{S}$. Denote by $\Omega^{\mathrm{LV}}_{KT}(t, S, \sigma)$ the local volatility price of a vanilla option of strike $K$, maturity $T$, as a function of time $t$, spot $S$ and local volatility function $\sigma$.

Set:

$$
\begin{aligned}
W_t^{KT} &\equiv W_t && \text{(2.111a)}\\
\sigma_t &= \sigma[O_{KT,t}, t, S_t]\,(t, S_t) && \text{(2.111b)}\\
\lambda_{KT,t} &= \sigma_t S_t \left.\frac{d\Omega^{\mathrm{LV}}_{KT}}{dS}\right|_{t,\ S=S_t,\ \sigma=\sigma[O_{KT,t},t,S_t]} && \text{(2.111c)}
\end{aligned}
$$

SDEs (2.109) together with (2.111) define a market model that starts from the initial condition $S_0, O_{KT,0}$.

$\sigma_t$ and $\lambda_{KT,t}$ are functions of $S_t, O_{KT,t}$ only – information available at time $t$. Our model is Markovian in these state variables.

It is in fact the local volatility model and we know that it has a Markov representation in terms of $t, S_t$. We know that in the model defined by (2.109) and (2.111) – which is the local volatility model – the local volatility function $\sigma[O_{KT,t}, t, S_t]$ is in fact constant and equal to $\sigma[O_{KT,0}, 0, S_0]$. $O_{KT,t}$ can thus be written as:

$$O_{KT,t} = \Omega^{\mathrm{LV}}_{KT}\left(t, S_t, \sigma[O_{KT,0}, 0, S_0]\right)$$

An alternative definition of local volatility is thus that it is a[23] diffusive market model that has a one-dimensional Markov representation in terms of $(t, S)$. Because

[23]Presumably the only one.

the local volatility model is a diffusive market model – that is its SDEs are of the form in (2.109) – the carry P&L is automatically of the form in (2.107) – no need to go through the rigmarole of Section 2.7.2.

In Chapter 12 we examine local-stochastic volatility models. They have a Markov representation in terms of $t, S$, plus a few other state variables – $X_t$ and $Y_t$ if one uses the two-factor model of Chapter 7. This additional flexibility can be exploited to produce different break-even levels for gamma/theta P&Ls.

As we will see, only a few of them are market models, that is can actually be used to risk-manage a derivatives book.

### 2.7.6 A metaphor of the local volatility model

We now illustrate the irrelevance of the local volatility delta $\Delta^{\mathrm{LV}}$ and the inconsequentiality of the "recalibration" of the local volatility function, using the simple example of a basket option on two underlyings $S_1, S_2$ – say the Euro Stoxx 50 and S&P 500 indexes. In our analogy with the local volatility model, $S_1$ is the actual underlying while $S_2$ is a vanilla option.

Consider the example of an ATM call option on an equally weighted basket of $S_1,\ S_2$ in a Black-Scholes model with identical volatilities and correlation $\rho$. We assume that the initial values $S_{1,\tau_0}, S_{2,\tau_0}$ of both underlyings are equal. The option price is $P\,(t, S_1, S_2)$.

The delta of an ATM option is about $50\%$ thus:

$$\Delta_1 \;\simeq\; 25\%, \quad \Delta_2 \;\simeq\; 25\%$$

Imagine now raising $\rho$ until it reaches $100\%$.[24] For $\rho = 100\%$, we still have $\Delta_1 \simeq 25\%$, $\Delta_2 \simeq 25\%$.

For $\rho = 100\%$, however, rather than solving the two-dimensional PDE for the option price, we can express $S_2$ as a function of $(t, S_1)$. If the volatilities of $S_1, S_2$ are equal, $S_2 = \left(\frac{S_{2,\tau_0}}{S_{1,\tau_0}}\right) S_1$. Our model has a Markov representation in terms of $(t, S_1)$ – the counterpart of the fact that the local volatility model has a one-dimensional Markov representation in terms of $(t, S)$.

We can then solve a one-dimensional equation for the option price which we denote by $P^{\mathrm{LV}}\,(t, S_1)$.

- In the local volatility model $P^{\mathrm{LV}}$ does not depend explicitly on $\widehat{\sigma}_{KT}$ anymore, because $\widehat{\sigma}_{KT}$ is a function of $t, S$: $\widehat{\sigma}_{KT} = \Sigma_{KT}\,(t, S, \sigma)$.
- Likewise, in our example $P^{\mathrm{LV}}$ does not depend on $S_2$ because $S_2$ is a function of $S_1$: $S_2 = \left(\frac{S_{2,\tau_0}}{S_{1,\tau_0}}\right) S_1$ – hence the notation $P^{\mathrm{LV}}$. The ratio $\frac{S_{2,\tau_0}}{S_{1,\tau_0}}$ plays the role of the local volatility function.

[24] We may use such high correlation for setting conservatively the break-even level of the cross-gamma $\frac{d^2P}{dS_1 dS_2}$.

We now compute deltas as derivatives of $P^{\mathrm{LV}}(t, S_1)$. We have:

$$\Delta_1^{\mathrm{LV}} = \frac{dP^{\mathrm{LV}}}{dS_1} \simeq 50\%, \quad \Delta_2^{\mathrm{LV}} = \frac{dP^{\mathrm{LV}}}{dS_2} = 0$$

- In the local volatility model, hedge ratios computed by taking the derivatives of $P^{\mathrm{LV}}$ imply that the only hedge instrument is $S$ and delta is $\Delta^{\mathrm{LV}} = \frac{dP^{\mathrm{LV}}}{dS}$. Vanilla options are not needed as hedges.

- Likewise, in our example, $\Delta_2 = 0$ means there's no need for $S_2$ in our hedge portfolio.

$\Delta_1^{\mathrm{LV}}$, $\Delta_2^{\mathrm{LV}}$ are obviously ludicrous – in particular any move of $S_2$ generates a change in option price that our (nonexistent) delta $\Delta_2^{\mathrm{LV}}$ is unable to offset.

Likewise, in the local volatility model, as discussed on page 66, $\Delta^{\mathrm{LV}}$ is useless since any move of $\widehat{\sigma}_{KT}$ that is not equal to that specified by $\Sigma_{KT}(t, S, \sigma)$ is not hedged.

What about the significance of recalibrating the local volatility function at $\tau_0 + \delta\tau$ to take into account the fact that, at $\tau_0 + \delta\tau$, $\widehat{\sigma}_{KT} \neq \Sigma_{KT}(\tau_0 + \delta\tau, S + \delta S, \sigma)$?

In our example, at $\tau + \delta\tau$, $S_{2,\tau_0+\delta\tau}$ will likely not be equal to $\left(\frac{S_{2,\tau_0}}{S_{1,\tau_0}}\right) S_{1,\tau_0+\delta\tau}$. Thus by using the actual value $S_{2,\tau_0+\delta\tau}$ of $S_2$, we "recalibrate" the ratio $\left(\frac{S_{2,\tau_0}}{S_{1,\tau_0}}\right)$ which, in our one-dimensional Markov representation, relates $S_2$ to $S_1$.

Yet, this "recalibration" of $S_2$ at $\tau_0 + \delta\tau$ is not an issue: we just enter at $\tau_0 + \delta\tau$ the new values of $S_1, S_2$ in our pricing function $P(t, S_1, S_2)$. The fact that we are using $\rho = 100\%$ in our model is of no consequence, other than that of setting the break-even level for the P&L generated by the cross-gamma $\frac{d^2 P}{dS_1 dS_2}$.

It is important to stress that (a) the delta, (b) the covariance structure of the model, are unrelated issues – a point we make again in Section 7.3.3.

The purpose of delta-hedging is to immunize a position at first order against arbitrary moves of the underlying assets, not just those allowed by the covariance structure of the model.

### 2.7.7 Conclusion

- The local volatility model is a diffusive market model – a somewhat special one as (a) it is a one-factor model, (b) it possesses a Markov representation in terms of $t, S$.

  It can be used for risk-managing options, recalibrating on a daily basis the local volatility function to market smiles. The gamma/theta P&L is well-defined: break-even levels for volatilities and correlations of $S$ and $\widehat{\sigma}_{KT}$ exist and are *payoff-independent.*

Obviously, one may wish that break-even correlations were different than $100\%$ and that volatilities of implied volatilities could be controlled exogenously and not be dictated by the smile used for calibration, but this is how much we can get with a one-factor model.

- The fact that volatilities of implied volatilities, as generated by the model, are neither chosen by the user, nor even deterministic but depend at each point in time on the then-prevailing market smile is however an issue. There is no guarantee that these levels will be adequate with respect to realized levels. Moreover, they will vary unpredictably: in case future market smiles happen to be flat, so that $\frac{d\Sigma_{KT}^{\mathrm{LV}}}{dS} \simeq 0$, hence $\nu_{KT} = 0$, we will find ourselves risk-managing our exotic option with a model that is locally pricing vanishing volatilities of implied volatilities.

  In a stochastic volatility model of the type studied in Chapter 7, in contrast, these levels depend on parameters that are set at inception.

- In practice, unlike delta hedging, vega hedging is typically not performed on a daily basis, because of larger bid/offer costs. When rehedging frequencies for delta and vega hedges differ, what gamma/theta P&L do we materialize? This question is answered in Section 9.11.3, page 383.

- The "local volatility delta" $\Delta^{\mathrm{LV}}$, computed with a fixed local volatility function, has no special significance or usefulness.

  The delta of the local volatility model is the market-model delta $\Delta^{\mathrm{MM}}$ in (2.108), that is the sensitivity of the option's price to a move of the spot, with fixed vanilla option prices: $\frac{d\mathcal{P}}{dS}$. The hedge then consists of a position in $\frac{d\mathcal{P}}{dS}$ shares and $\frac{d\mathcal{P}}{dO_{KT}}$ *naked* vanilla options of strike $K$, maturity $T$.

  If, as is customary, we use implied volatilities $\widehat{\sigma}_{KT}$ as a representation of market prices of vanilla options, rather than prices $O_{KT}$, we need to trade the sticky-strike delta $\Delta^{\mathrm{SS}}$ defined in (2.98) computed by moving $S$ and keeping the $\widehat{\sigma}_{KT}$ fixed. This is not the total delta in our hedge, as we also need to delta-hedge the vanilla options used as vega hedges. Their deltas are computed in the Black-Scholes model using their respective implied volatilities. The aggregate delta is thus:

$$\Delta^{\mathrm{SS}} - \frac{d\mathcal{P}}{dO_{KT}} \bullet \frac{dP_{KT}^{\mathrm{BS}}}{dS} = \left[\frac{d\mathcal{P}}{dS} + \frac{d\mathcal{P}}{dO_{KT}} \bullet \frac{dP_{KT}^{\mathrm{BS}}}{dS}\right] - \frac{d\mathcal{P}}{dO_{KT}} \bullet \frac{dP_{KT}^{\mathrm{BS}}}{dS}$$
$$= \frac{d\mathcal{P}}{dS} = \Delta^{\mathrm{MM}}$$

  that is, we recover as the aggregate delta of our hedge portfolio the market-model delta $\Delta^{\mathrm{MM}}$.

- The very notion of a vanilla option's delta *does not make sense* in the context of a market model, such as the local volatility model. A market model takes as inputs vanilla option prices, in addition to the spot. The former are treated as

> hedge instruments, on the same footing as the underlying itself. Just as the notion of the delta of one asset with respect to a different asset in a multi-asset model makes no sense, the delta of a vanilla option in a market model is an irrelevant notion.

We refer the reader to Section 2.9 for the practical numerical calculation of vega hedge ratios $\frac{\delta P}{\delta \widehat{\sigma}_{KT}}$.

### 2.7.8 Appendix – delta-hedging only

Consider a position that is only delta-hedged. The P&L of the delta-hedged position is the P&L in (2.105) supplemented with the contribution from the delta position, equal to $\Delta\,(\delta S - (r-q)S\delta t)$ where $\Delta$ is the delta we trade.

Trading a delta given by $\Delta = \frac{dP}{dS}$ leaves us with a residual vega position. At order one in $\delta\widehat{\sigma}_{KT}, \delta S, \delta t$ the P&L or our delta-hedged position is:

$$P\&L \;=\; -\,\frac{dP}{d\widehat{\sigma}_{KT}} \bullet \left(\delta\widehat{\sigma}_{KT} - \mu^{\mathrm{LV}}_{KT}\delta t\right) \tag{2.112}$$

Consider instead trading the local volatility delta $\Delta^{\mathrm{LV}}$ – computed with a fixed local volatility function. This is the delta generated by the local volatility model *for a fixed volatility function*. From (2.100):

$$\Delta^{\mathrm{LV}} \;=\; \frac{dP}{dS} + \frac{dP}{d\widehat{\sigma}_{KT}} \bullet \frac{d\Sigma^{\mathrm{LV}}_{KT}}{dS}$$

The P&L at order one of our delta-hedged position now reads:

$$P\&L \;=\; -\frac{dP}{d\widehat{\sigma}_{KT}} \bullet \left(\delta\widehat{\sigma}_{KT} - \frac{d\Sigma^{\mathrm{LV}}_{KT}}{dS}\left(\delta S - (r-q)S\delta t\right) - \mu^{\mathrm{LV}}_{KT}\delta t\right) \tag{2.113}$$

In the dynamics of the local volatility model, $S$ and $O_{KT}$ – or equivalently $S$ and $\widehat{\sigma}_{KT}$ – are perfectly correlated. One can be written as a function of the other: $\widehat{\sigma}_{KT} = \Sigma^{\mathrm{LV}}_{KT}(t,S)$. The SDE for $\widehat{\sigma}_{KT}$ reads:

$$\begin{aligned} d\widehat{\sigma}_{KT} \;&=\; \frac{d\Sigma^{\mathrm{LV}}_{KT}}{dS}dS \;+\; \left(\frac{1}{2}\frac{d^2\Sigma^{\mathrm{LV}}_{KT}}{dS^2}\sigma^2(t,S)S^2 + \frac{d\Sigma^{\mathrm{LV}}_{KT}}{dt}\right)dt \\ &=\; \frac{d\Sigma^{\mathrm{LV}}_{KT}}{dS}\left(dS - (r-q)S\delta t\right) + \mu^{\mathrm{LV}}_{KT}dt \end{aligned}$$

thus P&L (2.113) vanishes with probability one.

This property, however, has no practical relevance. In real life $S$ and $\widehat{\sigma}_{KT}$ are separate instruments that will move about freely.

- Just as in the case of two equity underlyings – see the example in Section 2.7.6 – it makes no sense to try to offset P&L (2.112) by spreading a vega position against a delta position. Doing so using the ratio $\frac{d\Sigma_{KT}^{\mathrm{LV}}}{dS}$ prescribed by the local volatility model is even less defensible: the realized regression coefficient of $\delta\widehat{\sigma}_{KT}$ on $\delta S$ will likely not match the regression coefficient implied by the local volatility model $\beta_{KT}^{\mathrm{LV}} = \frac{d\Sigma_{KT}^{\mathrm{LV}}}{dS}$.

- This is seen in the fact that realized values of the SSR (see for example Figure 9.3, page 367) are lower than their value in the local volatility model (see Figure 2.4, page 59 – typically $\mathcal{R} > 2$ for equity smiles) – and also higher than what the sticky-strike delta implies ($\mathcal{R} = 1$). Neither sticky-strike nor local-volatility deltas are good proxies for hedging the residual vega position.

- Local and stochastic volatility models, and more generally diffusive models, share the property that the SSR for short maturities – thus the value of $\beta_{KT}$ – is model-independent: $\mathcal{R}_{T\to 0} = 2$. This implies that the *implied* spot/ATMF volatility instantaneous covariance can be read off the market smile in model-independent fashion, just as a short ATM option's implied volatility supplies the implied instantaneous break-even variance of $\delta \ln S$. However, this is irrelevant to delta-hedging: the way we compute deltas has nothing to do with the covariance structure – the break-even levels of correlations and volatilities – of the particular model at hand. We also refer the reader to the discussions in Sections 7.3.3, page 225, and 9.11.1, page 379.

- In case we are adamant about delta-hedging the residual vega exposure (2.112) we should use the delta that minimizes the standard deviation of the order-one contribution to the P&L:

$$\Delta = \frac{dP}{dS} + \frac{dP}{d\widehat{\sigma}_{KT}} \bullet \beta_{KT}$$

  where $\beta_{KT}$ is the *historical* – rather than *implied* – regression coefficient of $\delta\widehat{\sigma}_{KT}$ on $\delta S$.

### 2.7.9 Appendix – the drift of $V_t$ in the local volatility model

Local volatility is but a special kind of stochastic volatility. Let $V_t$ be the instantaneous variance. Two features single out the local volatility model:

- $V_t$ is 100% correlated with $S_t$.
- $V_t$ is not just a functional of the path of $S_t$, it is actually a function of $S_t$.

In the local volatility model $V_t = \sigma^2(t, S_t)$. The dynamics of $S_t, V_t$ then reads:

$$\begin{aligned} dS_t &= (r-q)\, S_t dt + \sqrt{V_t} S_t dW_t \\ dV_t &= \left( (r-q)\, S_t \frac{d\sigma^2}{dS} + \frac{d\sigma^2}{dt} + \frac{S_t^2}{2} \frac{d^2\sigma^2}{dS^2} \sigma^2 \right) dt + S_t \frac{d\sigma^2}{dS} \sqrt{V_t} dW_t \end{aligned}$$

where $\sigma^2$, $\frac{d\sigma^2}{dS}$, $\frac{d^2\sigma^2}{dS^2}$ are evaluated with arguments $t, S_t$. Notice again that the instantaneous volatility of $V_t$ is determined by the skew of the local volatility function, $S\frac{d\sigma}{dS}$.

Also note how complicated the drift of $V_t$ is. Yet, while the drift of $V_t$ – historically called the "market price of risk" – has been the subject of much ado in stochastic volatility papers, no one seems to get quite as concerned about the drift of $V_t$ in the local volatility model.

We come back to the issue of the drift of $V_t$ and its significance in Section 6.3.

---

## 2.8 Digression – using payoff-dependent break-even levels

We have stressed in the previous section that the local volatility model is a market model: break-even levels of volatilities and correlations of hedging instruments – spot and vanilla options – are payoff-independent. This has an important consequence: if the gamma of a position locally vanishes, so does its theta.

What if we do not use a real model?

Imagine for example being short a one-year 90/110 call spread: short a call option struck at $K_1 = 90$ and long a call option struck at $K_2 = 110$ and assume that the initial value of $S$ is 100. Let us also make the assumption that we are free to risk-manage these options until maturity without remarking them to market at intermediate times.

Figure 2.8 shows the price of this call spread in the Black-Scholes model as a function of volatility for zero interest rate and repo.

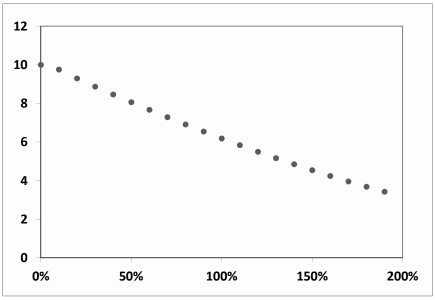



**Figure 2.8**: Black-Scholes price of a one-year 90/110 call spread as a function of volatility, for $S = 100$, zero interest rate and repo.

Notice that the price is maximum for zero volatility. For other configurations of $K_1, K_2$ it would have peaked for a different volatility: because of the varying sign of the payoff's convexity and unlike vanilla options, the Black-Scholes price of a call spread is not necessarily a monotonic function of volatility.

The market price of the call spread is simply $P_{BS}(tS, K_1T, \widehat{\sigma}_{K_1T}) - P_{BS}(tS, K_2T, \widehat{\sigma}_{K_2T})$. The smile of equity underlyings is usually strong enough ($\widehat{\sigma}_{K_1T}$ larger than $\widehat{\sigma}_{K_2T}$) that the market price of the call spread lies above the highest price attainable in the Black-Scholes model: the notion of an implied volatility vanishes. For example, taking $\widehat{\sigma}_{K_1T} = 22.5\%$, $\widehat{\sigma}_{K_2T} = 17.5\%$ gives a price equal to 11.04.

Thus we cannot risk-manage our short call spread position in the Black-Scholes model using a single implied volatility – what about delta-hedging each vanilla option using its own implied volatility? This gives rise during $\delta t$ to the following P&L:

$$P\&L = -\frac{1}{2}S^2\frac{d^2P^1_{BS}}{dS^2}\left(\frac{\delta S^2}{S^2} - \widehat{\sigma}^2_{K_1T}\delta t\right) + \frac{1}{2}S^2\frac{d^2P^2_{BS}}{dS^2}\left(\frac{\delta S^2}{S^2} - \widehat{\sigma}^2_{K_2T}\delta t\right) \tag{2.114}$$

Now, consider risk-managing this position with the local volatility model, keeping the local volatility function fixed, that is using the delta $\Delta^{\text{LV}}$ in (2.93). There is no contradiction with our discussion in the previous section: since we are not marking our call spread to market, we are free to keep the local volatility function fixed. The P&L over $[t,\ t+\delta t]$ is then given by (2.94) rather than (2.105). Our P&L reads:

$$\begin{aligned} P\&L &= -\frac{1}{2}S^2\frac{d^2P^1_{\sigma}}{dS^2}\left(\frac{\delta S^2}{S^2} - \sigma(t,S)^2\delta t\right) + \frac{1}{2}S^2\frac{d^2P^2_{\sigma}}{dS^2}\left(\frac{\delta S^2}{S^2} - \sigma(t,S)^2\delta t\right) \\ &= -\frac{1}{2}S^2\left(\frac{d^2P^1_{\sigma}}{dS^2} - \frac{d^2P^2_{\sigma}}{dS^2}\right)\left(\frac{\delta S^2}{S^2} - \sigma(t,S)^2\delta t\right) \end{aligned} \tag{2.115}$$

In the Black-Scholes model, for a vanilla option, $\frac{d^2P_{BS}}{dS^2}$ is positive and peaks in the vicinity of the option's strike. There is then a particular value of $S$ such that $\frac{d^2P^1_{BS}}{dS^2} - \frac{d^2P^2_{BS}}{dS^2} = 0$. For this value of $S$, the P&L in equation (2.114) becomes:

$$P\&L = \frac{1}{2}S^2\frac{d^2P^1_{BS}}{dS^2}\left(\widehat{\sigma}^2_{K_1T} - \widehat{\sigma}^2_{K_2T}\right)\delta t$$

Our delta strategy pays us "free" money ($\widehat{\sigma}_{K_1T} > \widehat{\sigma}_{K_2T}$) in a region of spot prices where there is no gamma risk. Note that this is not the case if we risk-manage both options in the same model – see (2.115) – as cancellation of gamma implies cancellation of the associated theta as well, which is much more reasonable. Note that prices of the call spread in both delta-hedging strategies are identical. Depending on which one is used, however, the theta is distributed differently.

It seems more judicious to use a hedging strategy that pays more theta in regions where gamma is large and pays no theta wherever gamma vanishes – which is what the local volatility model, or any model for that matter, does – rather than squander theta in regions of $S$ where there is little or no risk. We examine this issue further below in Appendix A in the context of the Uncertain Volatility Model.

## 2.9 The vega hedge

In the Black-Scholes model there is only one vega, as there is only one volatility parameter: any option can be used to vega-hedge any other option. The situation improves somewhat in the Black-Scholes model with deterministic time-dependent volatility $\sigma(t)$. The relationship linking $\sigma(t)$ to implied volatilities $\widehat{\sigma}_T$ is:

$$\sigma^2(t) \;=\; \frac{d}{dT}\left(T\widehat{\sigma}_T^2\right)\Big|_{T=t}$$

Thus an option's sensitivity to $\delta\sigma(t)$ can be hedged – within the model – by trading narrow calendar spreads of vanilla options. Equivalently, given the sensitivity of an option to $\sigma(t)$ for all $t$ up to its maturity, we can derive the maturity distribution of vanilla options to be used as hedges.

In the local volatility model, the price of an exotic option is a functional of the whole volatility surface $\widehat{\sigma}_{KT}$. Alongside a position in the underlying, the hedge portfolio consists of $\frac{d\mathcal{P}}{dO_{KT}}$ vanilla options of strike $K$, maturity $T$. Equivalently, the vanilla option hedge immunizes the global position at order one against fluctuations of the local volatility function $\sigma(t,S)$.

How do we calculate $\frac{d\mathcal{P}}{dO_{KT}}$? This question was first tackled by Bruno Dupire [41]. What follows is based on his and Pierre Henry-Labordère's work – see [59].

### 2.9.1 The vanilla hedge portfolio

Let $P(t,S,\,\bullet)$ be the price of an exotic option, where $\bullet$ stands for all path-dependent variables, whose number usually varies during the option's life.[25] Path-dependent variables only change discontinuously at times when $S$ is observed, as mandated by the term sheet of our exotic payoff. While the option price is continuous across each of these dates, the *function* $P$ changes discontinuously so as to absorb the discontinuity in the path-dependent variables, but in between obeys the usual pricing equation:

$$\frac{dP}{dt} + (r-q)\,S\frac{dP}{dS} + \frac{\sigma^2(t,S)\,S^2}{2}\frac{d^2P}{dS^2} - rP \;=\; 0 \tag{2.116}$$

Let us perturb $\sigma^2(t,S)$ by $\delta\sigma^2(t,S)$ and let us call $\delta P$ the resulting perturbation for $P$ – working with variances is equivalent to working with volatilities and lightens the notation. Replacing $\sigma^2$ with $\sigma^2+\delta\sigma^2$ and $P$ with $P+\delta P$ in (2.116) and expanding at order one in $\delta\sigma^2$ yields the following equation for $\delta P$:

$$\frac{d\delta P}{dt} + (r-q)\,S\frac{d\delta P}{dS} + \frac{\sigma^2(t,S)\,S^2}{2}\frac{d^2\delta P}{dS^2} - r\delta P \;=\; -\frac{1}{2}S^2\frac{d^2P}{dS^2}\delta\sigma^2(t,S)$$

[25] In the case of payoff $\left(S_{T_2}/S_{T_1}-1\right)^+$ for example, $P$ is a function of $t,S$ for $t<T_1$ and of $t,S,S_{T_1}$ for $t\in[T_1,T_2]$.

which is similar to (2.116) except it has a source term. At maturity $\delta P = 0$. Application of the Feynman-Kac theorem yields:

$$\delta P \;=\; \frac{1}{2}E_\sigma\left[\int_0^T dt\, e^{-rt} S^2\frac{d^2P}{dS^2}\,(t,S,\;\bullet)\,\delta\sigma^2(t,S)\right]$$

where $E_\sigma$ denotes the expectation taken over paths of $S_t$ generated by the local volatility $\sigma\,(t,S)$. Conditioning now with respect to the value of $S$ at time $t$:

$$\begin{aligned}\delta P \;&=\; \frac{1}{2}\int_0^T dt\, e^{-rt}\int_0^\infty dS\,\rho\,(t,S)\,E_\sigma\left[S^2\frac{d^2P}{dS^2}\,(t,S,\;\bullet)\,|S,t\right]\delta\sigma^2(t,S)\\ &=\; \frac{1}{2}\int_0^T dt\, e^{-rt}\int_0^\infty dS\,\rho\,(t,S)\,\phi\,(t,S)\,\delta\sigma^2(t,S) \end{aligned} \tag{2.117}$$

where $\rho\,(t,S)$ is the density of $S$ at time $t$ and $\phi\,(t,S)$ is the expectation of the dollar gamma conditional on the underlying having the value $S$ at time $t$:

$$\phi\,(t,S) \;=\; E_\sigma\left[S^2\frac{d^2P}{dS^2}\,(t,S,\;\bullet)\,|S,t\right] \tag{2.118}$$

Equation (2.117) expresses $\delta P$ as an average of $\delta\sigma^2$, weighted by the product of the density and $\phi$. We now look for a portfolio $\Pi$ of call options of all strikes and maturities:

$$\Pi \;=\; \int_0^T d\tau\int_0^\infty dK\,\mu\,(\tau,K)\,C_{K\tau} \tag{2.119}$$

that hedges our exotic option at order one against any perturbation $\delta\sigma^2(t,S)$. $\mu\,(\tau,K)$ is the density of vanilla options of strike $K$, maturity $\tau$. Equation (2.117) implies:

$$\phi_\Pi\,(t,S) \;=\; \phi\,(t,S)$$

where $\phi_\Pi$ is the dollar gamma of portfolio $\Pi$. How can we choose $\mu$ so that the resulting dollar gamma is $\phi$? Imagine that the exotic option was in fact a straight vanilla option – could we tell by just looking at $\phi$?

First note that, because a vanilla option is European, it is not path-dependent and $\phi$ is simply the dollar gamma:

$$\phi\,(t,S) \;=\; S^2\frac{d^2P}{dS^2}\,(t,S)$$

Starting from the pricing equation:

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

and applying the operator $S^2\frac{d^2}{dS^2} \equiv \left(S\frac{d}{dS}\right)^2 - S\frac{d}{dS}$:

$$\frac{d\big(S^2\frac{d^2P}{dS^2}\big)}{dt} + (r-q)\,S\frac{d\big(S^2\frac{d^2P}{dS^2}\big)}{dS} + \frac{1}{2}S^2\frac{d^2}{dS^2}\left(\sigma^2(t,S)\,S^2\frac{d^2P}{dS^2}\right) \;=\; r\left(S^2\frac{d^2P}{dS^2}\right)$$

yields:

$$\frac{d\phi}{dt} + (r-q)\,S\frac{d\phi}{dS} + \frac{1}{2}S^2\frac{d^2}{dS^2}\left(\sigma^2(t,S)\,\phi\right) \;=\; r\phi$$

Let us define operator $\mathcal{L}$ as:

$$\mathcal{L}f = \frac{df}{dt} + (r-q)\,S\frac{df}{dS} + \frac{1}{2}S^2\frac{d^2}{dS^2}\left(\sigma^2(t,S)\,f\right) - rf \tag{2.120}$$

We get the property that $\mathcal{L}\phi = 0$ for a European option. Consider a discrete portfolio $\Pi$ of vanilla options of strikes $(K_i, \tau_i)$, in quantities $\mu_i$, and let the corresponding dollar gamma be $\phi_\Pi$. For $t \in ]\tau_{i-1},\ \tau_i[$, $\mathcal{L}\phi_\Pi = 0$. However, as we cross – forward – time $\tau_i$, $\phi_\Pi$ is discontinuous since for $t > \tau_i$ it does not include the dollar gamma of option $i$ anymore. This discontinuity contributes to $\mathcal{L}\phi_\Pi$, through the $\frac{d}{dt}$ operator in $\mathcal{L}$. The dollar gamma of option $i$ vanishes at $t = \tau_i^-$ except for $S = K_i$. It is then equal to $\mu_i K_i^2 \delta\,(S - K_i)$ where $\delta$ denotes the Dirac distribution: this is the distinguishing feature of call and put options. For our discrete portfolio we then have:

$$(\mathcal{L}\phi_\Pi)\,(t,S) \;=\; -\sum_i \mu_i K_i^2 \delta(t-\tau_i)\,\delta(S-K_i)$$

Consider now a portfolio consisting of a continuous density of call options as in (2.119):

$$\mathcal{L}\phi_\Pi\,(\tau,K) \;=\; -K^2\mu\,(\tau,K)$$

We then have our final result. $\mu$ is simply given by:

$$\mu\,(\tau,K) \;=\; -\frac{1}{K^2}\mathcal{L}\phi\,(\tau,K) \tag{2.121}$$

To be able to use (2.121) we need an estimate for $\phi$ that is sufficiently smooth so that we can apply operator $\mathcal{L}$. Practically $\phi$ will be evaluated on a grid, in a Monte Carlo simulation, using Malliavin techniques as it is a conditional expectation: this is numerically delicate.

Our task is made a lot easier if we simply carry out the perturbation analysis around a flat local volatility function, as the forward transition densities are all analytically known and $S$ is simulated in the Black-Scholes model. In practice, for exotic options that do not depend explicitly on realized variance, the nature of the vega hedge will not depend much on the precise shape of the local volatility function around which perturbation is performed.

Consider a path-dependent payoff $f\,(\mathbf{S})$ where $\mathbf{S}$ is the vector of spot observations $S_i \equiv S_{t=t_i}$. Consider a time $t$ and let $t_{k-1}, t_k$ be spot observation dates such that $t\ \in ]t_k, t_{k+1}[$. $\phi\,(t,S)$ is given by:

$$\phi\,(t,S) \;=\; \frac{\displaystyle\int \prod_{i<k}(p_{i-1,i}dS_i)\,p\,(t_{k-1}S_{k-1}, tS)\,S^2\tfrac{d^2P}{dS^2}\,(t,S,\bullet)}{p\,(t_0 S_0, tS)}$$

$$= \frac{\int \prod_{i<k} \left(p_{i-1,i} dS_i\right) p\left(t_{k-1}S_{k-1}, tS\right) S^2 \frac{d^2}{dS^2} \left[p\left(tS, t_kS_k\right) dS_k \prod_{j>k} \left(p_{j-1,j} dS_j\right) f(\mathbf{S})\right]}{p\left(t_0S_0, tS\right)}$$

$$= \frac{\int \prod_{i\neq k} \left(p_{i-1,i} dS_i\right) p\left(t_{k-1}S_{k-1}, tS\right) S^2 \frac{d^2p}{dS^2} \left(tS, t_kS_k\right) dS_k \; f(\mathbf{S})}{p\left(t_0S_0, tS\right)} \tag{2.122}$$

where $p_{i,i+1}$ is a shorthand notation for the transition density $p\left(t_iS_i, t_{i+1}S_{i+1}\right)$ in the Black-Scholes model:

$$p\left(t_iS_i, t_{i+1}S_{i+1}\right) = \frac{1}{S_{i+1}\sqrt{2\pi\sigma_0^2 \left(t_{i+1} - t_i\right)}} e^{-\frac{\left(\ln\left(S_{i+1}/S_i\right) - \left(r - q - \frac{\sigma_0^2}{2}\right)\left(t_{i+1} - t_i\right)\right)^2}{2\sigma_0^2\left(t_{i+1} - t_i\right)}}$$

In equation (2.122), $S^2 \frac{d^2}{dS^2}$ only acts on $p\left(tS, t_kS_k\right)$: the calculation can be done analytically and we get our final expression for $\phi$:

$$\phi\left(t, S\right) = \frac{\int \prod_{i\neq k} \left(p_{i-1,i} dS_i\right) p\left(t_{k-1}S_{k-1}, tS\right) w\left(tS, t_kS_k\right) p\left(tS, t_kS_k\right) dS_k f(\mathbf{S})}{p\left(t_0S_0, tS\right)}$$

$$= \frac{1}{p\left(t_0S_0, tS\right)} E\left[\frac{p\left(t_{k-1}S_{k-1}, tS\right) w\left(tS, t_kS_k\right) p\left(tS, t_kS_k\right)}{p\left(t_{k-1}S_{k-1}, t_kS_k\right)} f(\mathbf{S})\right] \tag{2.123}$$

where $w$ is given by:

$$\begin{aligned} w\left(tS, t_kS_k\right) &= \frac{1}{\sigma_0^2\left(t_k - t\right)} \left(\frac{z^2}{\sigma_0^2\left(t_k - t\right)} - 1 - z\right) \\ z &= \ln\left(S_k/S\right) - \left(r - q - \frac{\sigma_0^2}{2}\right)\left(t_k - t\right) \end{aligned}$$

Equation (2.123) expresses $\phi$ as an expectation of the option's payoff multiplied by a weight that involves $S_{k-1}, S, S_k$. We recognize in $w$ the classical expression of the weight for computing gamma in a Monte Carlo simulation, in the Black-Scholes model.

As is well known, in practice it provides noisy estimates of gamma, especially when $t_k - t$ is small, as the variance of $w$ blows up. In our context this will be the case whenever the spot observation dates of our exotic option are closely spaced. This issue is compounded by the fact that in (2.123) $w$ is sandwiched in between two transition densities that contribute their fair share of the variance of our estimator for $\phi$. Getting an accurate estimate for $\phi$ is then computationally expensive but presents no special difficulty.

### 2.9.2 Calibration and its meaningfulness

What do we do once we have the hedge portfolio $\Pi$? Can we use it in practice?

Consider a constant volatility $\widehat{\sigma}_0$ and call $P^0$ the corresponding Black-Scholes price. Let us assume that the market smile is not too strong so that $\widehat{\sigma}_{K\tau} - \widehat{\sigma}_0$ is small. Using the above results and expanding at order one in $\widehat{\sigma}_{K\tau} - \widehat{\sigma}_0$:

$$P = P^0 + \int_0^T d\tau \int_0^\infty dK\, \mu(\tau, K)\left(C_{K\tau} - C^0_{K\tau}\right) \tag{2.124}$$

where $C^0_{K\tau}$ is the Black-Scholes price with volatility $\widehat{\sigma}_0$ and $C_{K\tau}$ the market price of the vanilla option of strike $K$, maturity $\tau$.

(2.124) can be interpreted as expressing the following:[26]

- Choose an implied volatility $\widehat{\sigma}_0$ for risk-managing the exotic option. This generates price $P^0$.
- Determine the quantities $\mu(\tau, K)$ of vanilla options to be used as hedges.
- Seting up the hedging portfolio entails paying market prices $C_{K\tau}$ rather than model prices $C^0_{K\tau}$ for the hedging vanilla options. The corresponding mismatch is passed on to the client as a hedging cost – the price we quote for the exotic option is given by (2.124).

The conclusion is that the price produced by a calibrated model is as credible as the hedge it implies. Is the latter really a statement on the exotic option or does it reflect model-specific features? This issue needs to be assessed on a case-by-case basis.

The impatient reader can jump to Section 3.2.4 where the case of a forward-start call is analyzed in detail.

---

## 2.10 Markov-functional models

In the local volatility model $S_t$ is generally a function of the path of $W_t$ up to time $t$, where $W_t$ is the driving Brownian motion. Are there special forms of the local volatility function such that $S_t$ can be written as a function of $t$ and $W_t$, hence can be simulated without any time-stepping? The Black-Scholes model is one example:

$$S_t = S_0 e^{\left(r-q-\frac{\sigma^2}{2}\right)t+\sigma W_t}$$

[26] This is how trading desks use to price exotic options in the second half of the '90s, before models were available and/or (mis)understood.

Imagine there exists $f(t,x)$ such that

$$S_t = f(t, W_t) \tag{2.125}$$

The condition that the drift of $S$ be $r-q$ translates into the following PDE for $f$:

$$\frac{df}{dt} + \frac{1}{2}\frac{d^2 f}{dx^2} = (r-q) f \tag{2.126}$$

and the instantaneous (lognormal) volatility of $S$ is given by:

$$\sigma(t,S) = \left.\frac{d\ln f}{dx}\right|_{x=f^{-1}(t,S),t} \tag{2.127}$$

which makes it clear that it is a local volatility model.

For $\sigma(t,S)$ to be well-defined, $f$ has to be a monotonic function of $x$. Markov-functional Models (MFM) for equities were first introduced by Peter Carr and Dilip Madan – see [23] – who provide some analytic non-trivial solutions to (2.126).

Given now a market smile, is it possible to find a function $f$ such that market prices of vanilla options are recovered?

If $f$ is known for a given time $T$, equation (2.126) generates $f$ for times $t \leq T$, thus determining smiles for maturities less than $T$. This implies that we can at most calibrate the smile for *one* maturity $T$. Smiles for maturities shorter than $T$ are dictated by the smile at $T$.

We have shown at the beginning of this chapter that, given a full vanilla smile that is free of arbitrage, there exists *one* local volatility function $\sigma(t,S)$ that is able to generate it. If instead we only have a volatility smile for a single maturity, there exist generally many different local volatility functions that are able to recover it. What we have just shown is that one of them corresponds to a Markov-functional model: the process for $S_t$ in this particular local volatility model can be simulated without any time-stepping by simply drawing $W_t$ and setting $S_t = f(t, W_t)$.

Assume we are given the market smile for maturity $T$: we can price digital options for all strikes, which gives access to the cumulative distribution function of $S_T$, $\mathcal{F}(S)$. Denoting $\mathcal{N}$ the cumulative distribution of the centered normal distribution, $f(t=T,\ x)$ is given by:

$$f(T,x) = \mathcal{F}^{-1}\left(\mathcal{N}\left(\frac{x}{\sqrt{T}}\right)\right) \tag{2.128}$$

$f(t=T,\ x)$ is monotonic by construction and so is $f(t,\ x)$ for $t<T$.[27]

[27] Assume $f(T,x)$ is monotonic in $x$ – say increasing: $\frac{df}{dx}(T,x) \geq 0$, $\forall x$. Take the derivative of both sides of (2.126) with respect to $x$: $\frac{df}{dx}$ obeys the same PDE as $f$, thus $\frac{df}{dx}(t,x) = e^{-(r-q)(T-t)} E\left[\frac{df}{dx}(T, W_T) | W_t = x\right]$ where $W_t$ is a Brownian motion. $\frac{df}{dx}(T,x) \geq 0$, $\forall x$ then implies $\frac{df}{dx}(t,x) \geq 0$, $\forall x$.

Note that by using other processes than a straight Brownian motion in (2.125) one can generate different smiles for intermediate maturities, for example by taking $S_t = f(t, Z_t)$ where $dZ_t = \sigma(t)\, dW_t$.

In the context of equities MFMs are very rarely used: one usually needs to calibrate a set of maturities simultaneously. In fixed income markets, on the other hand, MFMs are natural, as cap/floors on LIBOR rates, swaptions, have maturities that match the fixing date of the underlying rate – see [64] for Markov-functional interest rate models. This is also the case of futures in commodity markets and VIX futures – see Section 7.8.2 for an example of an MFM in this context.

### 2.10.1 Relationship of Gaussian copula to multi-asset local volatility prices

MFMs can be used for European options on a basket of equities $S^i$. Let us call $T$ the option's maturity: one draws the (correlated) Gaussian random variables $W_T^i$, applies the mapping in (2.128) and evaluates the payoff. This is exactly equivalent to using the marginal densities supplied by the market smile for maturity $T$ for each asset, and then using a Gaussian copula function to generate the multivariate density for the $S_T^i$.

This is worth noting as, usually, given a multivariate density $\rho$ generated by an arbitrary copula function, one is unable to characterize the dynamics that underlies $\rho$: it is not even clear that there exists a diffusive process that is able to generate $\rho$ – one then has no idea of what the gamma/theta break-even levels of his/her position are: the model is unusable.

Because MFMs are a particular instance of local volatility, in the case of a multi-asset European option, pricing with a Gaussian copula thus exactly boils down to using a particular[28] multi-asset local volatility model calibrated on implied volatilities of the option's maturity, with constant correlations, equal to the correlations of the Gaussian copula.

---

## Appendix A – the Uncertain Volatility Model

Treating the Uncertain Volatility Model (UVM) as a local volatility model is not quite natural, as it is typically used in situations when there are no market implied volatilities. We still cover it as it is a (very) special and useful instance of local volatility: in its basic version the local volatility function is not determined by the market smile, but set by a trading criterion.

[28]Because we only calibrate implied volatilities of maturity $T$, there exist many other local volatility functions that achieve exact calibration. Prices generated by these other volatility functions will differ from the Gaussian copula price.

Imagine selling an option on an underlying for which no option market exists – typically a fund share. For example, consider a short position in a call option, whose dollar gamma is always positive. We should typically sell it for a Black-Scholes implied volatility $\widehat{\sigma}$ that is sufficiently higher than the expected realized volatility $\sigma_r$ to ensure that our gamma/theta P&L is mostly positive. Conversely, we would buy this option for a value of $\widehat{\sigma}$ lower than $\sigma_r$. What about an option whose gamma can be positive or negative, depending on $S, t$?

For example, imagine being short a call spread: we sell the call struck at $K_1$ and buy the call struck at $K_2$ ( $K_2 > K_1$), and assume that we price the $K_1$ call with $\widehat{\sigma}_1$ and the $K_2$ call with $\widehat{\sigma}_2$, with $\widehat{\sigma}_1 > \widehat{\sigma}_2$. Our P&L during $\delta t$ is:

$$P\&L \;=\; -\frac{\Gamma_1}{2}\left(\sigma_r^2 - \widehat{\sigma}_1^2\right) + \frac{\Gamma_2}{2}\left(\sigma_r^2 - \widehat{\sigma}_2^2\right) \tag{2.129}$$

where $\Gamma_1$,$\Gamma_2$ are the (positive) dollar gammas of both calls.

Whenever $\sigma_r$ is such that $\widehat{\sigma}_1 < \sigma_r < \widehat{\sigma}_2$, both contributions in (2.129) are positive. For $S \ll K_1$, $\Gamma_2 \ll \Gamma_1$ and $P\&L \simeq -\Gamma_1\left(\sigma_r^2 - \widehat{\sigma}_1^2\right)/2$. Similarly, for $S \gg K_2$, $P\&L \simeq \Gamma_2\left(\sigma_r^2 - \widehat{\sigma}_2^2\right)/2$: our gamma/theta break-even levels are $\widehat{\sigma}_1$ (resp. $\widehat{\sigma}_2$) for very low (res. high) values of $S$. Now, as discussed in Section 2.8, for $S$ such that the residual gamma $\Gamma_1 - \Gamma_2$ vanishes, our P&L is uselessly positive. Can we use a pricing and hedging scheme such that this P&L is redistributed to regions where $\Gamma$ is sizeable, thus improving our gamma/theta break-even levels?

This is exactly what the Uncertain Volatility Model (UVM), introduced by Marco Avellaneda, Arnon Levy, Antonio Paras in [3] and Terry Lyons in [71], does.

The UVM is a local volatility model where the instantaneous volatility $\sigma$ is a function of the dollar gamma of the option being priced. In the original version of the UVM, $\sigma$ can take two values: $\sigma_{\min}$, $\sigma_{\max}$, depending on the sign of the dollar gamma: $\sigma(t,S) = \sigma_{\max}$ if $\frac{d^2P}{dS^2} > 0$, $\sigma(t,S) = \sigma_{\min}$ if $\frac{d^2P}{dS^2} < 0$ . Our P&L thus reads:

$$\begin{aligned} P\&L \;&=\; -\frac{\Gamma}{2}\left(\sigma_r^2 - \sigma_{\max}^2\right) \;\text{ if } \Gamma > 0 \\ &=\; -\frac{\Gamma}{2}\left(\sigma_r^2 - \sigma_{\min}^2\right) \;\text{ if } \Gamma < 0 \end{aligned}$$

The pricing equation for $P$ is non-linear as $\sigma(t,S)$ is a function of $P$ :

$$\frac{dP}{dt} + (r-q)\,S\frac{dP}{dS} \;+\; \frac{1}{2}\,\sigma^2\left(\frac{d^2P}{dS^2}\right)S^2\frac{d^2P}{dS^2} = rP$$

which can be written as:

$$\frac{dP}{dt} + (r-q)\,S\frac{dP}{dS} \;+\; \frac{1}{2}\;\max_{\sigma=\sigma_{\min},\sigma_{\max}}\left(\sigma^2 S^2\frac{d^2P}{dS^2}\right) = rP \tag{2.130}$$

Equation (2.130) is the Hamilton-Jacobi-Bellman equation for the following stochastic control problem:

$$P = \max_{\sigma_t \in [\sigma_{\min}, \sigma_{\max}]} E[f(S_T)] \tag{2.131}$$

where $f$ is the option's payoff, the maximum is taken over all processes for the instantaneous volatility $\sigma_t$ such that $\sigma_t \in [\sigma_{\min}, \sigma_{\max}]$ and the expectation is taken over paths of $S_t$.

This implies that for an option whose payoff is a sum of two payoffs, the UVM price $P$ satisfies the inequality $P \leq P_1 + P_2$, where $P_1, P_2$ are the respective UVM prices of both payoffs. Going back to our example, the price of the call spread in the UVM with $\sigma_{\max} = \widehat{\sigma}_1$, $\sigma_{\min} = \widehat{\sigma}_2$ is lower than the difference of the Black-Scholes prices of each call, computed with volatilities $\widehat{\sigma}_1, \widehat{\sigma}_2$.

Equivalently, we can find better levels $\sigma_{\max} > \widehat{\sigma}_1$ and $\sigma_{\min} < \widehat{\sigma}_2$ that still allow us to match the market price. Risk-managing our call spread with the UVM pays zero theta wherever gamma vanishes but provides better break-even volatility levels elsewhere.

For a payoff whose dollar gamma has varying sign, there exist generally many different couples $(\sigma_{\min}, \sigma_{\max})$ such that the UVM price matches a given price level.

## A.1 An example

Take the example of a 3-year maturity call spread with $K_1 = 90\%$, $K_2 = 160\%$ with $\widehat{\sigma}_1 = 38\%$, $\widehat{\sigma}_2 = 31\%$. Examples of $(\sigma_{\min}, \sigma_{\max})$ couples that match the Black-Scholes price of the call spread are given in Table 2.1.

| | | | | | |
|---|---|---|---|---|---|
| $\sigma_{\min}$ | 20% | 22% | 24% | 26% | 28% |
| $\sigma_{\max}$ | 28% | 32% | 36% | 41% | 46% |

**Table 2.1**: Examples of $(\sigma_{\min}, \sigma_{\max})$ couples yielding the same price for a 3-year 90%/160% call spread, in the UVM.

From a trading point of view, for the same price charged, it is more reasonable to not have any theta P&L whenever gamma vanishes and have break-even levels $26\%/41\%$, for example, than getting positive theta P&L in a region of vanishing gamma and having break-even levels $31\%/38\%$ in regions where gamma is sizeable.

In practice, rather than taking $\sigma = \sigma_{\max}\theta(\Gamma) + \sigma_{\min}(1 - \theta(\Gamma))$, we can use smoother functions of $\Gamma$, for example requiring more comfortable break-even levels as the dollar gamma increases.

It is however necessary to ensure that $\sigma(\Gamma)^2\Gamma$ is an increasing function of $\Gamma$ to preclude arbitrage, that is the possibility that given two payoffs $u(S)$, $v(S)$ such that $u(S) \geq v(S)$, $u$ might be cheaper than $v$.

## A.2 Marking to market

The UVM was originally designed for underlyings for which no volatility market exists. Is it suited to underlyings for which implied volatilities exist?

### A.2.1 An unhedged position

Consider the case of a large trade in a call spread. While the liquidity of vanilla options is not sufficient for us to hedge ourselves in the market, it may be sufficient enough that we decide[29] to mark our position to market.

With respect to the previous situation, the benefits of the wider break-even levels are wiped out because of the mark-to-market constraint: $\sigma_{\min}, \sigma_{\max}$ have to be moved throughout time so that the UVM price of the call spread always matches its market value.

Let us assume that, during the option's lifetime, implied volatilities $\widehat{\sigma}_1, \widehat{\sigma}_2$ do not move: we start initially with $\sigma_{\max} > \widehat{\sigma}_1$ and $\sigma_{\min} < \widehat{\sigma}_2$. As we reach maturity, the dollar gammas $\Gamma_1, \Gamma_2$ become localized near their respective strikes and do not overlap anymore. Thus, at maturity $\sigma_{\max} = \widehat{\sigma}_1$ and $\sigma_{\min} = \widehat{\sigma}_2$. As time advances, $\sigma_{\max}$ (resp. $\sigma_{\min}$) will converge to $\widehat{\sigma}_1$ (resp. $\widehat{\sigma}_2$). This daily remarking of $\sigma_{\min}$, $\sigma_{\max}$ will generate additional theta.

By doing this, we extract more theta from the UVM than we need with the consequence that, near the option's maturity, when dollar gammas are largest, our break-even levels become identical to the Black-Scholes ones ($\sigma_{\max} = \widehat{\sigma}_1$, $\sigma_{\min} = \widehat{\sigma}_2$). This defeats the purpose of the UVM.

### A.2.2 A hedged position – the $\lambda$-UVM

Let us assume here that we have traded an exotic option $F$ that can be reasonably – but not perfectly – hedged with a portfolio of vanilla options. Pricing the exotic at hand in the UVM is a very conservative approach: rather than pricing the full gamma of $F$ with $\sigma_{\min}, \sigma_{\max}$, depending on its sign, it is preferable to first assemble a portfolio of vanilla options that best offsets the gamma profile of our exotic, and then price the package consisting of the exotic option minus its hedge in the UVM. This idea was first proposed in [4] under the name of $\lambda$-UVM model – or Lagrangian UVM; see also [47] for more recent work.

The price $\mathcal{P}(F)$ we quote for the exotic is then given by:

$$\mathcal{P}(F) \; = \; \mathcal{P}_{\text{UVM}}(F - \Sigma\lambda_i O_i) + \Sigma\lambda_i \mathcal{P}_{\text{Mkt}}(O_i) \tag{2.132}$$

where $F$ is the exotic option's payoff, $\mathcal{P}_{\text{UVM}}$ and $\mathcal{P}_{\text{Mkt}}$ denote, respectively, the UVM price and the market price, and $\lambda_i$ the quantities of vanilla options $O_i$ traded. The second piece represents the cost of buying the vanilla portfolio at market price.

How should we choose vector $\lambda$? The $\lambda_i$ should be chosen so that for given values of $\sigma_{\min}, \sigma_{\max}$, we quote the most competitive – i.e. lowest possible – price

[29] or that the Risk Control department decides.

$\overline{\mathcal{P}}(F)$ for our exotic option:

$$\lambda = \arg\min_{\lambda}\left[\mathcal{P}_{\mathrm{UVM}}(F-\Sigma\lambda_i O_i)+\Sigma\lambda_i\mathcal{P}_{\mathrm{Mkt}}(O_i)\right]$$
$$\overline{\mathcal{P}}(F) = \min_{\lambda}\left[\mathcal{P}_{\mathrm{UVM}}(F-\Sigma\lambda_i O_i)+\Sigma\lambda_i\mathcal{P}_{\mathrm{Mkt}}(O_i)\right] \tag{2.133}$$

**Optimality conditions**

Let us express that the derivative of $\overline{\mathcal{P}}(F)$ with respect to $\lambda_i$ vanishes. Consider equation (2.130) for the UVM price $P$ of the package $F-\Sigma\lambda_i O_i$ and assume a small perturbation of the $\lambda_i$ which results in a variation $\delta P$ of $P$ and a variation $\delta(\sigma^2)$ of $\sigma^2$:

$$\frac{d(P+\delta P)}{dt}+(r-q)S\frac{d(P+\delta P)}{dS}+\frac{\sigma^2+\delta(\sigma^2)}{2}S^2\frac{d^2(P+\delta P)}{dS^2} = r(P+\delta P) \tag{2.134}$$

where $\sigma$ is the local volatility function that maximizes the UVM price of $F-\Sigma\lambda_i O_i$. Expanding (2.134) at order one in $\delta P$ and $\delta(\sigma^2)$ gives for $\delta P$ the following PDE:

$$\frac{d\delta P}{dt}+(r-q)S\frac{d\delta P}{dS}+\frac{\sigma^2}{2}S^2\frac{d^2\delta P}{dS^2} = r\delta P - \frac{\delta(\sigma^2)}{2}S^2\frac{d^2P}{dS^2} \tag{2.135}$$

with the terminal condition $\delta P(T,S) = -\Sigma\delta\lambda_i O_i(S)$. The solution of (2.135) is given by:

$$\delta P = e^{-rT}E_\sigma[\delta P(T,S_T)] + E_\sigma\left[\int_0^T e^{-rt}S^2\frac{d^2P}{dS^2}\frac{\delta(\sigma^2)}{2}dt\right] \tag{2.136}$$

Consider the second piece of (2.136). It would be the only contribution to $\delta P$ if $\delta P(T,S)=0$, that is if the $\delta\lambda_i$ were vanishing. It represents the effect of a small perturbation of $\delta\sigma^2$ on $P$ – for an unchanged payoff. By definition, $\sigma$ is the local volatility function that maximizes the price $\mathcal{P}_{\mathrm{UVM}}(F-\Sigma\lambda_i O_i)$: at order one any perturbation $\delta(\sigma^2)$ leaves $P$ unchanged: the second piece in (2.136) vanishes.

We are then left with:

$$\delta P = e^{-rT}E_\sigma[\delta P(T,S_T)]$$

which expresses that $\delta P$ is given by a standard local volatility PDE where the local volatility is fixed, given by the solution of the UVM price for $F-\Sigma\lambda_i O_i$. In other words:

$$\delta\mathcal{P}_{\mathrm{UVM}}(F-\Sigma\lambda_i O_i) = -\Sigma\delta\lambda_i\mathcal{P}_{\mathrm{UVM}}^{F-\Sigma\lambda_i O_i}(O_i)$$

where $\mathcal{P}_{\mathrm{UVM}}^{F-\Sigma\lambda_i O_i}(G)$ is defined as the price of payoff $G$ calculated with a local volatility function $\sigma$ that maximizes the UVM price *of the package* $F-\Sigma\lambda_i O_i$.

Taking now the derivative of $\mathcal{P}(F)$ with respect to $\lambda_i$ yields:

$$\frac{d\mathcal{P}(F)}{d\lambda_i} = -\mathcal{P}_{\mathrm{UVM}}^{F-\Sigma\lambda_i O_i}(O_i)+\mathcal{P}_{\mathrm{Mkt}}(O_i)$$

Condition $\frac{d\mathcal{P}(F)}{d\lambda_i} = 0$ implies:

$$\mathcal{P}_{\text{UVM}}^{F-\Sigma\lambda_i O_i}(O_i) \;=\; \mathcal{P}_{\text{Mkt}}(O_i) \tag{2.137}$$

Thus, for $\lambda$ such that $\mathcal{P}\left(F\right)$ is extremal, prices of vanilla options used as hedges calculated using the local volatility that maximize the UVM price of $F - \Sigma\lambda_i O_i$ match their market prices. The $\lambda$-UVM thus ensures that if $F$ should collapse to a vanilla option $O_i$, $\overline{\mathcal{P}}\left(O_i\right)$ would match the market price $\mathcal{P}_{\text{Mkt}}(O_i)$. $\overline{\mathcal{P}}\left(F\right)$ can then be called a mark-to-market price.

Going back to the definition of $\overline{\mathcal{P}}\left(F\right)$ in (2.133) and using identity (2.137) yields:

$$\overline{\mathcal{P}}\left(F\right) \;=\; \mathcal{P}_{\text{UVM}}^{F-\Sigma\lambda_i O_i}(F)$$

Alternatively $\overline{\mathcal{P}}\left(F\right)$ can also be characterized as the solution of the following stochastic control problem:

$$\overline{\mathcal{P}}\left(F\right) \;=\; \max_{\substack{\sigma_t \,\in\, [\sigma_{\min},\sigma_{\max}] \\ E_\sigma[O_i] \,=\, \mathcal{P}_{\text{Mkt}}(O_i)}} E_\sigma[F] \tag{2.138}$$

which generalizes criterion (2.131): we have added the additional constraint that market prices of options used as hedges have to be matched.[30]

### A.2.3 Discussion

Note that, for given market prices $\mathcal{P}_{\text{Mkt}}\left(O_i\right)$, $\sigma_{\text{min}}$, $\sigma_{\text{max}}$ have to be chosen so that $\mathcal{P}_{\text{Mkt}}\left(O_i\right)$ is attainable in the UVM. For example, imagine that the market implied volatilities are all equal to $25\%$ and that we have chosen $\sigma_{\text{min}} = 10\%$, $\sigma_{\text{max}} = 20\%$. Obviously UVM prices of vanilla options are such that their implied volatilities cannot exceed $20\%$. Problem (2.138) has no solution. Practically this would manifest itself in the fact that $\mathcal{P}\left(F\right)$ in (2.132) can be made as negative as we wish by making $\lambda$ sufficiently negative.

The above derivation was carried out for the situation when we are selling payoff $F$: $\overline{\mathcal{P}}\left(F\right)$ is our offer price. If instead we are buying the exotic option we can follow a similar derivation, defining our bid price as $\underline{\mathcal{P}}\left(F\right)$ given by:

$$\underline{\mathcal{P}}\left(F\right) \;=\; \max_{\lambda}\left[\mathcal{P}_{\text{UVM}}(F - \Sigma\lambda_i O_i) + \Sigma\lambda_i \mathcal{P}_{\text{Mkt}}(O_i)\right]$$

$\underline{\mathcal{P}}\left(F\right)$ is also characterized as:

$$\underline{\mathcal{P}}\left(F\right) \;=\; \min_{\substack{\sigma_t \,\in\, [\sigma_{\min},\sigma_{\max}] \\ E_\sigma[O_i] \,=\, \mathcal{P}_{\text{Mkt}}(O_i)}} E_\sigma[F]$$

In practice, whenever $F$ can be suitably hedged with vanilla options – this is assessed by checking that for a comfortably wide interval $[\sigma_{\min}, \sigma_{\max}]$, $\overline{\mathcal{P}}\left(F\right)$ is

[30]The $\lambda_i$ introduced in (2.132) are Lagrange multipliers for these constraints.

not inconsiderately expensive, or better that our bid/offer spread $\overline{\mathcal{P}}(F) - \underline{\mathcal{P}}(F)$ is not too large – the $\lambda$-UVM model is an effective tool for pricing and risk-managing exotic options. Unfortunately this isn't often the case, which is testament to the fact that most exotic risks are of a different nature than vanilla risks. We refer the reader to our experiment with forward-start options in Section 3.1.7.

When using vanilla options as hedges, $\overline{\mathcal{P}}(F) - \underline{\mathcal{P}}(F)$ provides an indication of how vanilla-like the risk of our exotic option is. We can, however, also use exotic options as hedges – for example cliquets. $\overline{\mathcal{P}}(F) - \underline{\mathcal{P}}(F)$ then supplies a measure of the kinship of exotic risks of the hedged and hedging options. Practically though, solving the stochastic control problem (2.138) in the situation of an exotic option hedged with other exotic options is typically not possible, as the high degree of path-dependence of exotic options results in the high dimensionality of equation (2.130).

Note that, as time elapses and we update $\lambda$ so as to keep $\overline{\mathcal{P}}(F)$ in (2.133) minimal, we never lose any money, as, by construction, shifting from a previously optimized to a currently optimal vector $\lambda$ lowers $\overline{\mathcal{P}}(F)$.

What if we set $\sigma_{\min} = 0$, $\sigma_{\max} = +\infty$? $\underline{\mathcal{P}}(F)$ (resp. $\overline{\mathcal{P}}(F)$) are then model-independent lower (resp. upper) bounds for the price of payoff $F$, given market prices of payoffs $O_i$ – see [47] for more on this and the connection with the dual problem of model-independent sub- (resp. super-) replication.

## A.3 Using the UVM to price transaction costs

Consider selling a call option on a security $S$, risk-managed in the Black-Scholes model with an implied volatility $\widehat{\sigma}$, but assume that bid/offer costs are incurred as we adjust our delta – say on a daily basis. Assume the relative bid-offer spread is $k$: we pay $\left(1+\frac{k}{2}\right)S$ to buy one unit of the security, and receive $\left(1-\frac{k}{2}\right)S$ when we sell it.

The P&L over $[t, t+\delta t]$ of a delta-hedged short option position reads as in (1.5), page 4, with the additional contribution of bid/offer costs:

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

where $\delta\Delta$ is the variation of our delta during $\delta t$; we use an absolute value as the impact of bid/offer is always a cost. $\delta\Delta$ is generated by (a) time advancing by $\delta t$, (b) $S$ moving by $\delta S$.

Since $\delta S$ if or order $\sqrt{\delta t}$ we keep the latter contribution only: $\delta\Delta = \frac{d^2P}{dS^2}\delta S$.

$$\begin{aligned} P\&L \;&=\; -\frac{S^2}{2}\frac{d^2P}{dS^2}\left(\frac{\delta S^2}{S^2} - \widehat{\sigma}^2\delta t\right) - \frac{k}{2}S\left|\frac{d^2P}{dS^2}\delta S\right| \\ &=\; -\frac{S^2}{2}\frac{d^2P}{dS^2}\left(\frac{\delta S^2}{S^2} - \widehat{\sigma}^2\delta t + \varepsilon_\Gamma k\left|\frac{\delta S}{S}\right|\right) \end{aligned}$$

where $\varepsilon_\Gamma$ is the sign of $\frac{d^2P}{dS^2}$.

If the realized volatility of $S$ is indeed $\widehat{\sigma}$, the gamma and theta contributions cancel out and the third piece will make our P&L persistently negative. Can we find an implied volatility $\widehat{\sigma}^*$ such that risk-managing the option position at $\widehat{\sigma}^*$ generates, on average, zero P&L? $\widehat{\sigma}^*$ must be such that:

$$\left\langle \frac{\delta S}{S^2}^2 \right\rangle - \widehat{\sigma}^{*2}\delta t + \varepsilon_\Gamma k \left\langle \left| \frac{\delta S}{S} \right| \right\rangle \; = \; 0$$

Assuming a lognormal dynamics for $S$ with realized volatility $\widehat{\sigma}$ and $\delta t$ small, $\frac{\delta S}{S} = \widehat{\sigma}\sqrt{\delta t}Z$ where $Z$ is a standard normal variable, thus $\left\langle \frac{\delta S}{S^2}^2 \right\rangle = \widehat{\sigma}^2\delta t$ and $\left\langle \left| \frac{\delta S}{S} \right| \right\rangle = \gamma\widehat{\sigma}\sqrt{\delta t}$ with $\gamma = \sqrt{\frac{2}{\pi}}$. $\widehat{\sigma}^*$ is given by:

$$\widehat{\sigma}^* \; = \; \sqrt{\widehat{\sigma}^2 + \varepsilon_\Gamma k\gamma\frac{\widehat{\sigma}}{\sqrt{\delta t}}}$$

Depending on the sign of the option's gamma, we need to use either $\widehat{\sigma}^*_{\Gamma+}$ or $\widehat{\sigma}^*_{\Gamma-}$, given by:

$$\widehat{\sigma}^*_{\Gamma+} \; = \; \sqrt{\widehat{\sigma}^2 + k\gamma\frac{\widehat{\sigma}}{\sqrt{\delta t}}} \qquad \widehat{\sigma}^*_{\Gamma-} \; = \; \sqrt{\widehat{\sigma}^2 - k\gamma\frac{\widehat{\sigma}}{\sqrt{\delta t}}} \tag{2.139}$$

with $\widehat{\sigma}^*_{\Gamma+} \; \geq \; \widehat{\sigma}^*_{\Gamma-}$.

Expression (2.139) for $\widehat{\sigma}^*_{\Gamma+}/\widehat{\sigma}^*_{\Gamma-}$ was first published by Hayne E. Leland in [68]. When trading a call option, whose gamma is always positive, we use $\widehat{\sigma}^*_{\Gamma+}$ when selling it and $\widehat{\sigma}^*_{\Gamma-}$ when buying it.

What if we trade a call spread, or generally an option payoff whose gamma has varying sign? This is where the UVM is called for. We use the UVM with:

$$\sigma_{\min} \; = \; \widehat{\sigma}^*_{\Gamma-}, \quad \sigma_{\max} \; = \; \widehat{\sigma}^*_{\Gamma+}$$

Two final observations are in order:

- $\gamma$ depends on the distribution we assume for daily returns. It equals $\sqrt{\frac{2}{\pi}}$ for Gaussian returns. In practice, daily returns of equities are better modeled with a Student distribution – see Chapter 10 for examples of actual distributions of index returns. In this respect, $\gamma = \sqrt{\frac{2}{\pi}}$ is an over-estimation.

- The period $\delta t$ of our delta rehedging schedule appears explicitly in (2.139). As $\delta t \to 0$, rehedging costs become prohibitive, to the point where $\widehat{\sigma}^*_{\Gamma-}$ does not exist anymore. When rehedging is frequent, or bid/offer spreads are large, rather than use the Black-Scholes delta and charge for the costs this incurs, it is preferable to go back to square one and cast the delta-hedging strategy as a stochastic control problem that maximizes a utility function which balances the costs generated by hedging with the benefit of a reduced uncertainty of our final P&L. This was done by Mark Davis, Vassilios Panas and Thaleia Zhariphopoulou in [35]. In practice, an exponential utility function $e^{-\lambda P\&L}$ is well suited as the ensuing delta strategy is independent on the initial wealth.[31]

[31] Solving this stochastic control problem numerically is tricky, as its solution is of the bang-bang type. The optimal hedging strategy is a function of $t$, $S$ and the current delta: $\Delta$. The $(S, \Delta)$ plane splits into three zones separated by two lines $\Delta_{\pm}(t, S)$. For $\Delta_{-}(t, S) < \Delta < \Delta_{+}(t, S)$, no adjustment is needed, if $\Delta < \Delta_{-}(t, S)$ (resp. $\Delta > \Delta_{+}(t, S)$) we need to (instantaneously) increase (resp. decrease) our delta so that it equals $\Delta_{-}(t, S)$ (resp. $\Delta_{+}(t, S)$). In the limit $k \to 0$ both $\Delta_{-}$ and $\Delta_{+}$ tend to the Black-Scholes delta $\Delta_{BS}$. For small values of $k$, Elizabeth Walley and Paul Wilmott show in [84] that $\Delta_{\pm} = \Delta_{BS} \pm \frac{1}{S}\left(\frac{3}{4}e^{-r(T-t)}\frac{k}{\lambda}\Gamma_{\$}^{2}\right)^{\frac{1}{3}}$. $\Gamma_{\$} = S^2\frac{d^2 P_{BS}}{dS^2}$ is the dollar gamma – remember $k$ is the total bid/offer spread.

## Chapter's digest

### 2.2 From prices to local volatilities

▶ Given a full volatility surface $\widehat{\sigma}_{KT}$ that complies with the following no-arbitrage conditions:

$$\begin{aligned} e^{qT_1} C\left(\alpha F_{T_1}, T_1\right) &\leq e^{qT_2} C\left(\alpha F_{T_2}, T_2\right) \\ \frac{d^2 C\left(K, T\right)}{dK^2} &\geq 0 \end{aligned}$$

there exists one volatility function $\sigma(t, S)$ given by the Dupire formula (2.3):

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

More generally, for any stochastic volatility model that recovers the market smile, whose instantaneous volatility is $\sigma_t$, the following condition holds:

$$E[\sigma_t^2 | S_t = S] = \sigma\left(t, S\right)^2$$

▶ In the absence of arbitrage, implied volatilities of vanilla options obey the convex order condition:

$$T_2 \widehat{\sigma}^2_{\alpha F_{T_2}, T_2} \geq T_1 \widehat{\sigma}^2_{\alpha F_{T_1}, T_1}$$

This condition is shared by any family of convex payoffs $f(S_T) = h(\frac{S_T}{F_T})$.

❧❧❧❧❧

### 2.3 From implied volatilities to local volatilities

▶ When there are no cash-amount dividends, local volatilities are obtained directly as a function of implied volatilities through (2.19):

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

▶ When cash-amount dividends are present, there exists (a) an exact solution based on a mapping from $S_t$ to an asset that does not jump across dividend dates, (b) an approximate solution that allows one to use the same formula as in the no-dividend case, except the definition of $y$ changes so that, in particular, it complies with the matching conditions for implied volatilities across dividend dates.

❧❧❧❧❧

## 2.4 From local volatilities to implied volatilities

▶ Given a local volatility function $\sigma(t,S)$, implied volatilities satisfy the following condition: (2.32):

$$\widehat{\sigma}_{KT}^2 = \frac{E_{\sigma(S,t)}\left[\int_0^T e^{-rt}\,\sigma\left(t,S\right)^2\; S^2\frac{d^2P_{\widehat{\sigma}_{KT}}}{dS^2}dt\right]}{E_{\sigma(S,t)}\left[\int_0^T e^{-rt}\;S^2\frac{d^2P_{\widehat{\sigma}_{KT}}}{dS^2}dt\right]}$$

▶ For weakly local volatilities $\widehat{\sigma}_{KT}$ is given, at order one in the perturbation with respect to a time-dependent volatility $\sigma_0(t)$ by (2.40):

$$\widehat{\sigma}_{KT}^2 = \frac{1}{T}\int_0^T dt\int_{-\infty}^{+\infty}dy\,\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\,u\left(t,F_te^{\frac{\omega_t}{\omega_T}x_K+\frac{\sqrt{(\omega_T-\omega_t)\omega_t}}{\sqrt{\omega_T}}y}\right)$$

where $\omega_t = \int_0^t\sigma_0^2(u)du$. When expanding around a constant volatility $\sigma_0$, this simplifies to (2.42):

$$\widehat{\sigma}_{KT} = \frac{1}{T}\int_0^T dt\int_{-\infty}^{+\infty}dy\,\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\,\sigma\left(t,F_te^{\frac{t}{T}x_K+\frac{\sqrt{(T-t)t}}{\sqrt{T}}y}\right)$$

▶ For a local volatility function of the form

$$\sigma\left(t,S\right) = \overline{\sigma}\left(t\right)+\alpha\left(t\right)x+\frac{\beta(t)}{2}x^2,\; x=\ln\frac{S}{F_t}$$

the ATMF skew and curvature are given, at order one in $\alpha$ and $\beta$ by (2.48) and (2.49):

$$\mathcal{S}_T = \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{F_T} = \frac{1}{T}\int_0^T\frac{t}{T}\alpha(t)dt$$

$$\left.\frac{d^2\widehat{\sigma}_{KT}}{d\ln K^2}\right|_{F_T} = \frac{1}{T}\int_0^T\left(\frac{t}{T}\right)^2\beta(t)dt$$

If $\alpha(t)$ decays as a power law, so does $\mathcal{S}_T$ for large $T$, with the same exponent: the exponent of the decay of $\alpha(t)$ can be read off the market smile.

▶ For $T\to 0$ we have the exact result:

$$\frac{1}{\widehat{\sigma}\left(T=0,K\right)} = \frac{1}{\ln\frac{K}{S}}\int_S^K\frac{1}{\sigma\left(T=0,\;\mathcal{S}\right)}\frac{d\mathcal{S}}{\mathcal{S}}$$

❧❧❧❧❧

## 2.5 The dynamics of the local volatility model

▶ Given a fixed local volatility function of the above form, as the spot moves, implied volatilities move. At first order in $\alpha(t)$ the sensitivity of the ATMF volatility to a spot move is given by

$$\frac{d\widehat{\sigma}_{F_T T}}{d \ln S_0} \;=\; \frac{1}{T}\int_0^T \alpha(t)dt$$

which can be expressed using the term-structure of the ATMF skew:

$$\frac{d\widehat{\sigma}_{F_T T}}{d \ln S_0} \;=\; \mathcal{S}_T + \frac{1}{T}\int_0^T \mathcal{S}_t dt$$

▶ How the ATMF volatility moves when the spot moves is quantified by the Skew Stickiness Ratio (SSR) – a dimensionless number defined as:

$$\mathcal{R}_T \;=\; \frac{1}{\mathcal{S}_T}\frac{d\widehat{\sigma}_{F_T T}}{d \ln S_0}$$

For a weakly local volatility function:

$$\mathcal{R}_T \;=\; 1 + \frac{1}{T}\int_0^T \frac{\mathcal{S}_t}{\mathcal{S}_T}dt$$

which implies $\lim_{T\to 0}\mathcal{R}_T \;=\; 2$, an exact result shared by all diffusive models.

For an ATMF skew decaying as a power law with exponent $\gamma$, in the local volatility model, for large $T$, for a weakly local volatility function:

$$\mathcal{R}_T \;\to\; \frac{2-\gamma}{1-\gamma}$$

For typical equity smiles $\gamma = \frac{1}{2}$, thus $\mathcal{R}_\infty = 3$.

▶ We have the exact result that $\lim_{T\to 0}\mathcal{R}_T = 2$. In addition, for a local volatility function that is a function of $\frac{S}{F_t}$ only, $\mathcal{R}_T = 2$, $\forall T$.

▶ Expressions above for $\left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{F_T}$, $\left.\frac{d^2\widehat{\sigma}_{KT}}{d\ln K^2}\right|_{F_T}$, $\mathcal{R}_T$ are obtained in an order-one expansion around a constant volatility $\sigma_0$. See (2.60) and (2.65) for formulas in an expansion around a deterministic volatility $\overline{\sigma}(t)$.

▶ In the local volatility model, implied volatilities are a function of $S$. The instantaneous volatility of the ATMF volatility is thus proportional to the instantaneous volatility of $S$, which is equal to $\widehat{\sigma}_{F_t t}$. vol$(\widehat{\sigma}_{F_T T})$ is given by:

$$\text{vol}(\widehat{\sigma}_{F_T T}) \;=\; \mathcal{R}_T\mathcal{S}_T\frac{\widehat{\sigma}_{F_t 0}}{\widehat{\sigma}_{F_T T}} \;=\; \left(1+\frac{1}{T}\int_0^T \frac{\mathcal{S}_t}{\mathcal{S}_T}dt\right)\mathcal{S}_T\frac{\widehat{\sigma}_{F_t t}}{\widehat{\sigma}_{F_T T}}$$

Thus, for short maturities, $\text{vol}(\widehat{\sigma}_{F_T T}) = 2\mathcal{S}_T$ while for long maturities: $\text{vol}(\widehat{\sigma}_{F_T T}) = \frac{2-\gamma}{1-\gamma}\mathcal{S}_T$, for a flat term structure of ATMF volatilities where $\gamma$ is the characteristic exponent of the decay of the ATMF skew.

▶ At order one in $\alpha(t)$ the ATMF skew is related to the weighted average of the instantaneous covariance of $\ln S$ and the ATMF volatility for the residual maturity:

$$\mathcal{S}_T \;=\; \frac{1}{\widehat{\sigma}_T^2 T}\int_0^T \frac{T-t}{T}\big\langle d\ln S_t \, d\widehat{\sigma}_{F_T T}(t)\big\rangle$$

This relationship is derived more generally in Chapter 8, Section 8.4. One consequence of this formula is that, with respect to time-homogeneous stochastic volatility models, a local volatility model calibrated to the same smile generates larger SSRs and weaker future skews.

❧❧❧❧❧

## 2.6 Future skews and volatilities of volatilities

▶ In the local volatility model, skews observed at future dates for a given residual maturity are typically weaker than spot-starting skews. For a spot-starting skew $\mathcal{S}_\theta(\tau = 0)$ that decays as a function of residual maturity $\theta$ with characteristic exponent $\gamma$, the skew at a future date $\tau$ for the same residual maturity $\theta$ scales like, for small $\theta$:

$$\mathcal{S}_\theta(\tau) \;\propto\; \left(\frac{\theta}{\tau}\right)^\gamma \mathcal{S}_\theta(\tau = 0)$$

▶ Investigating model-generated future skews is useful for assessing local-volatility prices of options that are subject to forward-smile risk. One should bear in mind that these future skews – and future levels of volatility of volatility – cannot be locked and will vary as the model is recalibrated to market smiles. Thus, unpredictable gamma/theta carry P&Ls will impact substantially the P&L of a hedged position, in case residual gammas and cross-gammas are sizeable.

❧❧❧❧❧

## 2.7 Delta and carry P&L

▶ The carry P&L of a delta and vega-hedged option position in the local volatility model can be expressed in the usual gamma-theta form, with payoff-independent break-even levels for spot variance, spot/volatility covariances and volatility/volatility covariances. The local volatility model is a genuine diffusive market model.

▶ The carry P&L of a delta-hedged, vega-hedged option position can be equivalently expressed either in terms of spot and implied volatilities (2.105), or as a function of spot and option prices (2.107). The "real" delta of the local volatility model is given by the derivative of the price with respect to $S$, keeping *vanilla option prices* fixed. The delta computed by keeping *fixed implied volatilities* is called the sticky-strike delta. Regardless of the particular parametrization used for vanilla option prices, once the deltas of the hedging vanilla options are included, the "real" delta is recovered.

▶ The delta of the local volatility model – computed with a fixed local volatility function – has no particular significance or usefulness. Furthermore the delta of a vanilla option in the local volatility model is an irrelevant notion. More generally, the issue of outputting a delta of one asset (a vanilla option) on another (the spot) is irrelevant in a market model.

❧❧❧❧❧

## 2.9 The vega hedge

▶ Which portfolio of vanilla options immunizes our derivative position at order one against all perturbations of the local volatility function? Denote by $\mu(\tau, K)$ the density of vanilla options of maturity $\tau$, strike $K$, in the hedging portfolio. $\mu(\tau, K)$ is given by formula (2.121):

$$\mu(\tau, K) = -\frac{1}{K^2}\mathcal{L}\phi(\tau, K)$$

where $\phi(t, S)$ is the conditional dollar gamma, given by formula (2.118):

$$\phi(t, S) = E_\sigma\left[S^2\frac{d^2P}{dS^2}(t, S, \bullet)\,|S, t\right]$$

and operator $\mathcal{L}$ is defined by:

$$\mathcal{L}f = \frac{df}{dt} + (r-q)\,S\frac{df}{dS} + \frac{1}{2}S^2\frac{d^2}{dS^2}\left(\sigma^2(t, S)\,f\right) - rf$$

In the case of a flat local volatility function, $\phi(t, S)$ is easily computed in a Monte Carlo simulation – see expression (2.123).

❧❧❧❧❧

## 2.10 Markov-functional models

▶ In Markov-functional models, $S_t$ is a function of a process $W_t$: $S_t = f(t, W_t)$, where $W_t$ is typically a Brownian motion or an Ornstein-Ühlenbeck process. Markov-functional models are special instances of local volatility and can be calibrated at most to the smile of a single maturity. Smiles for intermediate maturities depend on the choice for the underlying process $W_t$.

▶ Pricing a multi-asset European derivative using a Gaussian copula together with marginals supplied by each asset's respective vanilla smile is equivalent to pricing with a multi-asset local volatility model, with the correlation matrix equal to the Gaussian copula's correlation matrix.

❧❧❧❧❧

**Appendix A – the Uncertain Volatility Model**

▶ In the UVM, minimum and maximum volatility levels $\sigma_{\min}$, $\sigma_{\max}$ are specified. The UVM ensures that no money is lost as long as the realized volatility lies in the interval $[\sigma_{\min}, \sigma_{\max}]$. The seller's price of a derivative with payoff $f(S_T)$ in the UVM solves PDE (2.130) – it is also characterized by:

$$P = \max_{\sigma_t \in [\sigma_{\min}, \sigma_{\max}]} E[f(S_T)]$$

▶ Rather than pricing a derivative fully in the UVM, one can avail oneself of market-traded options to lower as much as possible the sensitivity of the hedged position to realized volatility, and instead use the UVM for the hedged position: this is the idea in the Lagrangian UVM. The seller's price for the derivative can be defined as:

$$\overline{\mathcal{P}}(F) = \max_{\substack{\sigma_t \in [\sigma_{\min}, \sigma_{\max}] \\ E_\sigma[O_i] = \mathcal{P}_{\text{Market}}(O_i)}} E_\sigma[F]$$

where $O_i$ are the market-traded options. Likewise the buyer's price is defined by:

$$\underline{\mathcal{P}}(F) = \min_{\substack{\sigma_t \in [\sigma_{\min}, \sigma_{\max}] \\ E_\sigma[O_i] = \mathcal{P}_{\text{Market}}(O_i)}} E_\sigma[F]$$

Options that can be exactly replicated by means of a static position in market-traded options are such that $\overline{\mathcal{P}}(F) = \underline{\mathcal{P}}(F)$.

$\overline{\mathcal{P}}(F) - \underline{\mathcal{P}}(F)$ otherwise quantifies the non-vanilla character of the payoff at hand.

▶ The UVM can be used to price-in transaction costs on the delta-hedge: $\sigma_{\min}$, $\sigma_{\max}$ are given by Leland's formula.

# Chapter 3

## Forward-start options

Before embarking on stochastic volatility, we pause to study the case of forward-start options – also called cliquets, which we briefly touched upon in Section 1.3. Characterizing the risks of cliquets and how their pricing should be approached provides precious clues as to which aspects of the dynamics of implied volatilities are relevant for pricing these (popular) options, and which features of the vanilla smile a model should be calibrated to.

That cliquet prices are in fact loosely constrained by vanilla smiles is made plain in Section 3.1.7 where we compute lower and upper bounds on the price of a forward-start call option, given vanilla smiles.

We then assess how forward-smile risk is handled in the local volatility model. We work out the example of a forward-start call option in detail; this sheds light on the suitability of local volatility with regard to forward-start options.

### 3.1 Pricing and hedging forward-start options

Many exotic options are sensitive to forward-smile risk, that is risk associated with the uncertainty about market implied volatilities observed in the future, or future smiles – the case of barrier options was briefly examined in the introduction. Among exotics, forward-start options, or cliquets, form a very popular class of derivatives whose prices are purely determined by the distribution of *forward* returns in the pricing model.

The payoffs of cliquets involve the ratio of a security's price observed at two different dates $T_1, T_2$: the payoff at time $T_2$ is $g\left(\frac{S_{T_2}}{S_{T_1}}\right)$. It is a function of the *forward* return $\frac{S_{T_2}}{S_{T_1}} - 1$. As shown below in Section 3.1.3, any European payoff can be expressed as a linear combination of call and put option payoffs. Market smiles for maturities $T_1$ (resp. $T_2$) determine prices of payoffs of the form $f(S_{T_1})$, (resp. $f(S_{T_2})$), but payoffs of the form $g\left(\frac{S_{T_2}}{S_{T_1}}\right)$ require modeling assumptions. It is not clear at this stage what, if any, implied volatility data of maturities $T_1$ and $T_2$ are relevant for pricing payoff $g$.

The following analysis of the risks of cliquets is typical of how one approaches the problem of pricing and hedging an exotic option, by:

- first finding a pricing model *and* a hedge portfolio that, within the chosen model, hedges the gamma/theta and vega risks,
- then estimating the costs of rebalancing the vanilla hedge.

**The forward smile**

Throughout this book, we use the expression *forward-smile risk* to designate the risk associated with the realization of *future smiles*. We do not use the notion of *forward smile*. The forward smile $\widehat{\sigma}_k^{T_1 T_2}$ for two dates $T_1, T_2 > T_1$ is obtained, in a given model, by:

- pricing a forward-start call for different values of moneyness $k$, whose payoff is: $\left(\frac{S_{T_2}/F_{T_2}}{S_{T_1}/F_{T_1}} - k\right)^+$, whose undiscounted price is $P(k)$, where $F_T$ is the forward for maturity $T$.
- implying a Black-Scholes volatility $\widehat{\sigma}_k^{T_1 T_2}$ through:

$$P(k) \ = \ P_{BS}\left(S = 1, K = k, T = T_2 - T_1, r = 0, q = 0;\ \widehat{\sigma}_k^{T_1 T_2}\right)$$

$\widehat{\sigma}_k^{T_1 T_2}$ is a well-defined function of $k$, but is an impractical object that has no historical counterpart. Indeed, $\widehat{\sigma}_k^{T_1 T_2}$ is the future implied volatility for moneyness $k$, averaged over all realizations of future smiles in the model: it is an aggregate of future smile risk and volatility-of-volatility risk.

Therefore, it does not make sense to assess the suitability of a model by comparing $\widehat{\sigma}_k^{T_1 T_2}$ with typical market smiles of maturity $T_2 - T_1$, if anything because the forward smile is invariably more convex, due to its being an average and the connection between price and implied volatility being nonlinear.[1]

The forward smile is thus a notion of limited usefulness – we do not use in the sequel.

### 3.1.1 A Black-Scholes setting

In the standard Black-Scholes model, implied volatilities have no term structure. Since a cliquet involves $S_{T_1}$ and $S_{T_2}$, it is natural to use a Black-Scholes model with time-dependent instantaneous volatility $\sigma(t)$: in such a model, implied volatilities are maturity-dependent but the smile for any given maturity is flat and the implied volatility for maturity $T$, $\widehat{\sigma}_T$, is given by:

$$\widehat{\sigma}_T^2 \ = \ \frac{1}{T-t} \int_t^T \sigma(u)^2 \, du$$

Because of homogeneity, the price of a cliquet in this model does not depend on $S$ and, besides interest and repo rates, only depends on the integrated variance over the interval $[T_1, T_2]$: it is a function of the forward volatility $\widehat{\sigma}_{T_1 T_2}$:

$$P \ = \ e^{-r(T_1 - t)} G\left(\widehat{\sigma}_{T_1 T_2}\right) \tag{3.1}$$

[1]See the discussion of the forward smile of the Heston model in [8].

defined by:

$$\widehat{\sigma}^2_{T_1T_2} = \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\sigma(t)^2\,dt = \frac{(T_2-t)\,\widehat{\sigma}^2_{T_2}-(T_1-t)\,\widehat{\sigma}^2_{T_1}}{T_2-T_1} \tag{3.2}$$

We need to answer two questions:

- How can this forward-start option be hedged?
- Which payoffs of maturities $T_1$ and $T_2$ should $\widehat{\sigma}_{T_1}$ and $\widehat{\sigma}_{T_2}$ be calibrated to? Note that implied volatilities can be defined for any payoff that is convex. These payoffs need to be such that their implied volatilities satisfy the convex order condition:

$$(T_2-t)\widehat{\sigma}^2_{T_2} \geq (T_1-t)\widehat{\sigma}^2_{T_1} \tag{3.3}$$

  so that $\widehat{\sigma}_{T_1T_2}$ in (3.2) is well-defined.

Prior to $T_1$, the cliquet's delta and gamma vanish and the cliquet is only sensitive to $\widehat{\sigma}_{T_1T_2}$. At time $t=T_1$, $S_{T_1}$ is known: the cliquet becomes a standard European option of maturity $T_2$.

The above formula shows that $\widehat{\sigma}_{T_1T_2}$ is a function of implied volatilities $\widehat{\sigma}_{T_1}$, $\widehat{\sigma}_{T_2}$. Trading single vanilla options of maturities $T_1$, $T_2$ is inappropriate as, unlike the forward-start option, their sensitivities to $\widehat{\sigma}_{T_1}$, $\widehat{\sigma}_{T_2}$ will vary as $S$ moves and their combined gamma will likely not vanish. Does there exist a portfolio of vanilla options of maturities $T_1$, $T_2$ whose vega is spot-independent?

### 3.1.2 A vanilla portfolio whose vega is independent of $S$

Let $\rho(K)$ be the density of vanilla options of strike $K$ in the portfolio. We make no distinction between call or put options struck at the same strike as they have the same vega. Because of the homogeneity of degree 1 in $S$ and $K$ of the Black-Scholes formula, the price of a vanilla option of strike $K$ can be written as: $P=Kf\left(\frac{S}{K}\right)$ and likewise for the vega:

$$\mathrm{Vega}_K(S) = K\varphi\left(\frac{S}{K}\right)$$

The vega of our portfolio thus reads:

$$\mathrm{Vega}_\Pi(S) = \int dK\rho(K)K\varphi\left(\frac{S}{K}\right)$$

After switching to variable $u=\frac{S}{K}$ this becomes:

$$\mathrm{Vega}_\Pi(S) = \int \frac{du}{u^3}\,S^2\rho\left(\frac{S}{u}\right)\varphi(u) \tag{3.4}$$

$\text{Vega}_\Pi(S)$ is independent of $S$ only if we choose:

$$\rho(K) \;\propto\; \frac{1}{K^2} \tag{3.5}$$

A portfolio of European options with a density proportional to $\frac{1}{K^2}$ is such that its vega in the Black-Scholes model does not depend on $S$. Up to an affine function of $S$ the corresponding payoff is in fact $\ln S$ – it is called the *log contract* and was first proposed by Anthony Neuberger in [75].

We now verify this by recalling a representation of an arbitrary European payoff in terms of cash, forwards and a portfolio of vanilla options.

### 3.1.3 Digression: replication of European payoffs

The derivation of the well-known formula expressing this decomposition[2] starts from the identity:

$$\begin{aligned} f(S) &= f(K_0) + \int_{K_0}^{S} \frac{df}{dK} dK \\ &= f(K_0) + \int_{K_0}^{\infty} \theta(S-K)\frac{df}{dK}dK \;-\; \int_0^{K_0} \theta(K-S)\frac{df}{dK}dK \end{aligned}$$

where $\theta(x)$ is the Heaviside function ($\theta(x) = 1$ if $x \geq 0$, $\theta(x) = 0$ otherwise) and where the second line can be checked by taking either the case $S > K_0$ or the case $S < K_0$. Integrate now by parts using $(K-S)^+$ as primitive of $\theta(K-S)$ and $-(S-K)^+$ as primitive of $\theta(S-K)$. We get:

$$\begin{aligned} f(S) &= f(K_0) + \int_{K_0}^{\infty} \frac{d^2 f}{dK^2}(S-K)^+ dK + \int_0^{K_0} \frac{d^2 f}{dK^2}(K-S)^+ dK \\ &\qquad + [-\frac{df}{dK}(S-K)^+]_{K_0}^{\infty} - [\frac{df}{dK}(K-S)^+]_0^{K_0} \end{aligned}$$

which after simplification gives:

$$\begin{aligned} f(S) &= f(K_0) \;+\; \left.\frac{df}{dK}\right|_{K_0}(S-K_0) \\ &\qquad + \int_0^{K_0} \frac{d^2 f}{dK^2}(K-S)^+ dK + \int_{K_0}^{\infty} \frac{d^2 f}{dK^2}(S-K)^+ dK \end{aligned} \tag{3.6}$$

This expresses $f$ as a linear combination of an affine function and a continuous density of calls struck above $K_0$ and puts struck below $K_0$. Let $P_f$ be the price of the

[2] See the article by Peter Carr and Dilip Madan ([25]) who trace this result back to the work of Breeden & Litzenberger ([18]), Green & Jarrow ([50]) and Nachman ([74]).

options that pays $f(S_T)$. $P_f$ is the sum of the prices of the different contributions to $f$ :

$$P_f = f(K_0)\, e^{-r(T-t)} + \left.\frac{df}{dK}\right|_{K_0} \left(Se^{-q(T-t)} - K_0 e^{-r(T-t)}\right) + \int_0^{K_0} \frac{d^2 f}{dK^2} P_K dK + \int_{K_0}^{\infty} \frac{d^2 f}{dK^2} C_K dK$$

where $C_K, P_K$ denote, respectively, the prices of a call and a put struck at $K$. Choosing for $K_0$ the forward for maturity $T$, $F_T$, yields:

$$P_f = f(F_T)\, e^{-r(T-t)} + \int_0^{F_T} \frac{d^2 f}{dK^2} P_K dK + \int_{F_T}^{\infty} \frac{d^2 f}{dK^2} C_K dK \qquad (3.7)$$

This is an equality of prices – one should not forget the forwards in the replication of $f$.

Conversely, consider a portfolio consisting of a density $\rho(K)$ of vanilla options of strike $K$. Identity (3.6) shows that the resulting payoff – $\int_0^{\infty} \rho(K)(S-K)^+ dK$, or $\int_0^{\infty} \rho(K)(K-S)^+ dK$ if we use put options – is obtained, up to an affine function of $S$, by simply integrating $\rho(K)$ twice.

### 3.1.4 A vanilla hedge

Starting with the density $\frac{1}{K^2}$ derived above and integrating twice, we recover the payoff of the log contract: $-\ln S$.

For reasons that will become clear when we discuss variance swaps, we prefer to work with payoff $-2\ln S$ rather than $\ln S$. We will henceforth denote by "log contract" the payoff $-2\ln S$. It is replicated with a density of vanilla options equal to $\frac{2}{K^2}$.

The price $Q^T(t,S)$ of a log contract of maturity $T$ in the Black-Scholes model is given by:

$$Q^T(t,S) = -2e^{-r(T-t)}\left(\ln S + (r-q)(T-t) - \frac{\widehat{\sigma}_T^2}{2}(T-t)\right) \qquad (3.8)$$

and its sensitivity to $\widehat{\sigma}_T$ is:

$$\frac{dQ^T}{d\widehat{\sigma}_T} = 2e^{-r(T-t)}(T-t)\,\widehat{\sigma}_T$$

We now resume our discussion of the vega hedge of our forward-start option $P$. Using equations (3.1) and (3.2) we get the sensitivities of $P$ to $\widehat{\sigma}_{T_1}$ and $\widehat{\sigma}_{T2}$:

$$\frac{dP}{d\widehat{\sigma}_{T_2}} = e^{-r(T_1-t)}(T_2-t)\,\widehat{\sigma}_{T_2}\mathcal{N}$$
$$\frac{dP}{d\widehat{\sigma}_{T_1}} = -e^{-r(T_1-t)}(T_1-t)\,\widehat{\sigma}_{T_1}\mathcal{N}$$

where prefactor $\mathcal{N}$ is:

$$\mathcal{N} = \frac{1}{(T_2 - T_1)\,\widehat{\sigma}_{T_1T_2}} \frac{dG}{d\widehat{\sigma}_{T_1T_2}}$$

Thus the portfolio

$$\Pi = -P + \frac{\mathcal{N}}{2}\left(e^{r(T_2-T_1)}Q^{T_2} - Q^{T_1}\right) \tag{3.9}$$

has zero sensitivity to both $\widehat{\sigma}_{T_1}$ and $\widehat{\sigma}_{T_2}$. Notice that the hedge ratios for $Q^{T_1}$ and $Q^{T_2}$ depend neither on $t$, nor on $S$: the hedge is stable as time elapses and $S$ moves, and will only need to be readjusted whenever the forward volatility $\widehat{\sigma}_{T_1T_2}$ varies.

What about the gamma of the vanilla hedge? Taking twice the derivative of equation (3.8) we get:

$$S^2\frac{d^2Q^T}{dS^2} = 2e^{-r(T-t)} \tag{3.10}$$

Thus the dollar gamma of the log contract – up to the usual discounting factor – is equal to 2. The combination $e^{r(T_2-T_1)}Q^{T_2} - Q^{T_1}$ has thus vanishing gamma and consequently, in our deterministic volatility model, vanishing theta as well.

We have been able to choose a hedging model and have assembled a static[3] vanilla portfolio that perfectly hedges at order one our cliquet against variations of $\widehat{\sigma}_{T_1T_2}$.

**Checking that $\widehat{\sigma}_{T_1T_2}$ is well-defined**

Implied volatilities of log contracts of maturities $T_1$ and $T_2$ satisfy the convex order condition (3.3) because of the convexity of the log contract. The log-contract falls in the class of payoffs considered in Section 2.2.2.2, page 32, for which (a) an implied volatility can be defined, (b) the convex order condition (3.3) holds.

When there are no cash-amount dividends, log contract implied volatilities $\widehat{\sigma}_T$ are expressed directly as an average of implied volatilities of vanilla options – see formula (4.21), page 142.

### 3.1.5 Using the hedge in practice – additional P&Ls

#### 3.1.5.1 Before $T_1$– volatility-of-volatility risk

Market implied volatilities are not only maturity- but also strike-dependent: how should we define $\sigma(t)$ or $\widehat{\sigma}_T$?

We *decide* to define $\widehat{\sigma}_T$ as the implied volatility of the log contract of maturity $T$. $\widehat{\sigma}_T$ is well-defined as the function $\ln S$ is concave: we only need to invert equation (3.8). Using this definition for $\widehat{\sigma}_T$ ensures that our hedge instruments have their

[3] In the sense that it does not need to be readjusted as $t$ advances and $S$ moves.

right market prices. The P&L over $[t,\ t+\delta t]$ of portfolio $\Pi$ reads:

$$\begin{aligned} P\&L_\Pi \;=\; & -\Big(e^{-r(T_1-t-\delta t)}G\left(\widehat{\sigma}_{T_1T_2}+\delta\widehat{\sigma}_{T_1T_2}\right)-(1+r\delta t)\,e^{-r(T_1-t)}G\left(\widehat{\sigma}_{T_1T_2}\right)\Big) \\ & +\frac{\mathcal{N}}{2}e^{r(T_2-T_1)}\Big(Q^{T_2}_{t+\delta t}-(1+r\delta t)\,Q^{T_2}_t+\frac{dQ^{T_2}}{dS}\left(\delta S-rS\delta t\right)\Big) \\ & -\frac{\mathcal{N}}{2}\Big(Q^{T_1}_{t+\delta t}-(1+r\delta t)\,Q^{T_1}_t+\frac{dQ^{T_1}}{dS}\left(\delta S-rS\delta t\right)\Big) \end{aligned}$$

where $Q^T_t$ (resp. $Q^T_{t+\delta t}$) is a shorthand notation for $Q^T\left(t,S,\widehat{\sigma}_T\right)$ (resp. $Q^T\left(t+\delta t,S+\delta S,\widehat{\sigma}_T+\delta\widehat{\sigma}_T\right)$). Let us expand this P&L at order 1 in $\delta t$ and order 2 in $\delta S$.

If $\delta\widehat{\sigma}_{T_2}=\delta\widehat{\sigma}_{T_1}=0$, since the cliquet has vanishing gamma/theta and so does our log contract hedge, we get zero P&L.

The only contribution to the P&L is then generated by $\delta\widehat{\sigma}_{T_1},\delta\widehat{\sigma}_{T_2}$. Using expression (3.8) for $Q^T$, we get:

$$\begin{aligned} P\&L_\Pi \;=\; & -e^{-r(T_1-t)}\Big(G\left(\widehat{\sigma}_{T_1T_2}+\delta\widehat{\sigma}_{T_1T_2}\right)-G\left(\widehat{\sigma}_{T_1T_2}\right)\Big) \\ & +\frac{\mathcal{N}}{2}e^{-r(T_1-t)}\left(T_2-t\right)\Big(\left(\widehat{\sigma}_{T_2}+\delta\widehat{\sigma}_{T_2}\right)^2-\widehat{\sigma}^2_{T_2}\Big) \\ & -\frac{\mathcal{N}}{2}e^{-r(T_1-t)}\left(T_1-t\right)\Big(\left(\widehat{\sigma}_{T_1}+\delta\widehat{\sigma}_{T_1}\right)^2-\widehat{\sigma}^2_{T_1}\Big) \end{aligned}$$

which yields:

$$\begin{aligned} P\&L_\Pi \;=\; & -e^{-r(T_1-t)}\Big(G\left(\widehat{\sigma}_{T_1T_2}+\delta\widehat{\sigma}_{T_1T_2}\right)-G\left(\widehat{\sigma}_{T_1T_2}\right)\Big) \\ & +\frac{1}{2\widehat{\sigma}_{T_1T_2}}\frac{dG}{d\widehat{\sigma}_{T_1T_2}}e^{-r(T_1-t)}\ \Big(\left(\widehat{\sigma}_{T_1T_2}+\delta\widehat{\sigma}_{T_1T_2}\right)^2-\widehat{\sigma}^2_{T_1T_2}\Big) \\ =\; & e^{-r(T_1-t)}\left[-\left(G\left(\widehat{\sigma}_{T_1T_2}+\delta\widehat{\sigma}_{T_1T_2}\right)-G\left(\widehat{\sigma}_{T_1T_2}\right)\right)+\frac{dG}{d(\widehat{\sigma}^2_{T_1T_2})}\delta(\widehat{\sigma}^2_{T_1T_2})\right] \end{aligned} \tag{3.11}$$

In the derivation of (3.11) we have used variances $\widehat{\sigma}^2_{T_1},\widehat{\sigma}^2_{T_2},\widehat{\sigma}^2_{T_1T_2}$ rather than volatilities $\widehat{\sigma}_{T_1},\widehat{\sigma}_{T_2},\widehat{\sigma}_{T_1T_2}$. The natural reason for analyzing our P&L using variances rather than volatilities is that the price of the hedge instruments – log-contracts – is affine in $\widehat{\sigma}^2_{T_1},\widehat{\sigma}^2_{T_2}$, rather than $\widehat{\sigma}_{T_1},\widehat{\sigma}_{T_2}$; just as we use $S$ rather than $\sqrt{S}$ for the sake of writing out the P&L of a delta-hedged option.

Expression (3.11) shows that $P\&L_\Pi$ is a function of $\delta(\widehat{\sigma}^2_{T_1T_2})$, not $\delta(\widehat{\sigma}^2_{T_1})$ and $\delta(\widehat{\sigma}^2_{T_2})$ separately.

(3.11) is *not* an expansion of $P\&L_\Pi$ at order two in $\delta\widehat{\sigma}_{T_1T_2}$: it is the actual P&L generated by a change of $\widehat{\sigma}_{T_1T_2}$. One can check on equation (3.11) that the order-one contribution from $\delta(\widehat{\sigma}^2_{T_1T_2})$ vanishes – as it should. This expression highlights the fact that the value of the log contract hedge – the second piece inside the brackets – is exactly quadratic in $\widehat{\sigma}_{T_1T_2}$: this is also apparent in (3.8).

If $G$ is an affine function of $\widehat{\sigma}^2_{T_1T_2}$, the hedge is perfect and our P&L is exactly zero until we reach $T_1$. This is the case for two particular payoffs: one is given by

$g\left(\frac{S_{T_2}}{S_{T_1}}\right) = \ln\left(\frac{S_{T_2}}{S_{T_1}}\right)$, which is a linear combination of our two log contracts and has no commercial interest whatsoever. The other one is the same payoff, but delta-hedged on a daily basis: it is called a forward variance swap and pays the realized quadratic variation over the interval $[T_1, T_2]$.[4]

Usually $G$ will *not* be an affine function of $\widehat{\sigma}^2_{T_1T_2}$. P&L (3.11) thus starts with a term of order two in $\delta(\widehat{\sigma}^2_{T_1T_2})$. Keeping this term only yields:

$$P\&L_{\Pi} \;=\; -\frac{e^{-r(T_1-t)}}{2}\frac{d^2G}{d(\widehat{\sigma}^2_{T_1T_2})^2}\left(\delta(\widehat{\sigma}^2_{T_1T_2})\right)^2 \tag{3.12}$$

For an at-the-money forward call, $g$ is given by:

$$g\left(\frac{S_{T_2}}{S_{T_1}}\right) \;=\; \left(\frac{S_{T_2}}{S_{T_1}} - 1\right)^+ \tag{3.13}$$

If $T_2 - T_1$ is small, for vanishing interest rate and repo, $G$ is approximately linear in $\widehat{\sigma}$:

$$G\left(\widehat{\sigma}_{T_1T_2}\right) \;\simeq\; \frac{\widehat{\sigma}_{T_1T_2}}{\sqrt{2\pi}}\sqrt{T_2 - T_1}$$

Using (3.12), at order 2 in $\delta(\widehat{\sigma}^2_{T_1T_2})$, $P\&L_{\Pi}$ reads:

$$P\&L_{\Pi} \;\simeq\; e^{-r(T_1-t)}\frac{\sqrt{T_2-T_1}}{\sqrt{2\pi}}\frac{1}{2\widehat{\sigma}_{T_1T_2}}\left(\delta\widehat{\sigma}_{T_1T_2}\right)^2 \tag{3.14}$$

We thus make money every time $\widehat{\sigma}_{T_1T_2}$ moves – this will occur generally for all payoffs whose value is a *concave* function of $\widehat{\sigma}^2_{T_1T_2}$.

In the general case, $P\&L_{\Pi}$ will then not vanish – it is generated by the dynamics of $\widehat{\sigma}_{T_1T_2}$: we call this volatility-of-volatility risk. The estimation of $P\&L_{\Pi}$ over $[0, T_1]$ entails making an assumption for the volatility of $\widehat{\sigma}_{T_1T_2}$. The resulting extra charge – or gain – has to be added to the price $P = e^{-rT_1}G\left(\widehat{\sigma}_{T_1T_2}\left(t=0\right)\right)$ quoted at time $t = 0$. In the case of the at-the-money forward option, this charge will be negative, as we reduce the price charged to the client by an estimation of the positive P&Ls (3.14) pocketed every time $\widehat{\sigma}_{T_1T_2}$ moves.[5]

Up to $t = T_1$, our pricing and hedging scheme works, provided we have included this extra charge – or gain – in our price.

#### 3.1.5.2 At $T_1$ – forward-smile risk

At $T_1$, the cliquet turns into a European option of maturity $T_2$. As seen in Section 3.1.3, it can be replicated with vanilla options of maturity $T_2$, hence its value is a function of implied volatilities for maturity $T_2$ observed at $T_1$: $\widehat{\sigma}_{KT_2}\left(T_1\right)$.

---

[4] Variance swaps are abundantly discussed in Chapter 5.

[5] Making random positive P&L may seem less serious than randomly losing money. However, not adjusting the price for this random gain will result in the loss of the trade – this is similar to trying to buy a vanilla option for its intrinsic value.

Imagine that the cliquet is an at-the-money forward call, whose payoff is given in equation (3.13). At $T_1$ the *market* price of the cliquet is then:

$$P_{BS}\left(S_{T_1}, K = S_{T_1}, T_2; \widehat{\sigma}_{K=S_{T_1}T_2}(T_1)\right) \tag{3.15}$$

In our hedging scheme both the cliquet and its hedge are risk-managed in a Black-Scholes model with volatilities $\widehat{\sigma}(T)$ defined as implied volatilities of log contracts. Equation (3.2) shows that, at $t = T_1$, $\widehat{\sigma}_{T_1T_2} = \widehat{\sigma}_{T_2}$, thus the *model* price of the cliquet at $T_1$ is:

$$P_{BS}\left(S_{T_1}, K = S_{T_1}, T_2; \widehat{\sigma}_{T_2}(T_1)\right) \tag{3.16}$$

Compare this formula with expression (3.15). They are identical, except in our pricing and hedging scheme the at-the-money call is valued using the implied volatility of the log contract of the same maturity, $T_2$, rather than the market implied volatility of the at-the-money call. The price we quote at $t = 0$ for our forward-start at-the-money option has then to include a provision to cover for the difference between (3.15) and (3.16).

The risk created by the uncertainty as to the smile prevailing at $T_1$ for maturity $T_2$ – *given a known level of the log-contract implied volatility* $\widehat{\sigma}_{T_2}(T_1)$ – is called forward-smile risk. For a general cliquet payoff $g\left(S_{T_2}/S_{T_1}\right)$, we need to supplement the initial price $P = e^{-rT_1}G\left(\widehat{\sigma}_{T_1T_2}(t=0)\right)$ with an extra charge to cover for the difference between the market price of the payoff at $T_1$ and the Black-Scholes price computed with volatility $\widehat{\sigma}_{T_2}(T_1)$.

What if our forward-start option has no – or hardly any – sensitivity to $\widehat{\sigma}_{T_2}(T_1)$? This is the case for a digital payoff:

$$g\left(\frac{S_{T_2}}{S_{T_1}}\right) = \mathbf{1}_{\frac{S_{T_2}}{S_{T_1}}>1}$$

or a narrow call spread or put spread struck around $S_{T_1}$. In this case $G$ has negligible or vanishing sensitivity to $\widehat{\sigma}_{T_1T_2}$. Using the term structure of log-contracts as hedge instruments does not make sense anymore, and the adjustment for forward-smile risk represents in fact the bulk of the forward-start option price: our cliquet is a pure forward-smile instrument.

### 3.1.6 Cliquet risks and their pricing: conclusion

- We *choose* to price and risk-manage our cliquet in a Black-Scholes model with time-dependent volatility, recalibrated every day on market implied volatilities $\widehat{\sigma}(T)$ of log contracts. We are then able to exactly gamma-hedge and vega-hedge our cliquet, and have to make two adjustments to our price: $\delta P_1$ to cover for volatility-of-volatility risk over $[0, T_1]$, that is P&Ls (3.12), and $\delta P_2$ to cover for forward-smile risk at $T_1$, that is the difference between (3.15) and (3.16). The price we quote for this cliquet at $t = 0$ is then:

$$P = e^{-rT_1}G\left(\widehat{\sigma}_{T_1T_2}(t=0)\right) + \left(\delta P_1 + \delta P_2\right) \tag{3.17}$$

How should we estimate $\delta P_1$ and $\delta P_2$? In case there is no market for volatility of volatility and forward smile we can do the following:

- $\delta P_1$: use historical data of log contract implied volatilities to estimate conservatively the expected realized volatility of $\widehat{\sigma}_{T_1 T_2}$ over $[0, T_1]$.
- $\delta P_2$: use historical data of market smiles of residual maturity $T_2 - T_1$ to estimate conservatively the difference between implied volatilities of our payoff and that of the log contract.

If instead we are able to offset these risks on the market – for example by trading other cliquets – $\delta P_1$ and $\delta P_2$ are computed using *implied* values in place of *historical* values. Practically, whether we decide to use calibrated or chosen values for volatility of volatility or the future smile, we will use a stochastic volatility model to estimate $\delta P_1 + \delta P_2$. The model generates a global price that aggregates all risks – as priced by the model. $\delta P_1$ and $\delta P_2$ are evaluated initially at $t = 0$.

- As we risk-manage the cliquet from $t = 0$ to $t = T_1$, we need to ensure that (a) $\delta P_1$ converges to zero at $T_1$, (b) $\delta P_2$ converges to the adjustment corresponding to the exact difference between the implied volatilities of our forward-start payoff and that of the log contract of maturity $T_2$, as observed at $T_1$ – whether $\delta P_1$, $\delta P_2$ have been calculated by hand or evaluated in a model.

- In a model, $\delta P_1$ will automatically converge to zero as $t \to T_1$. For $\delta P_2$ to converge to the right value, though, the model has to be such that it recovers at $T_1$ the smile of maturity $T_2$.

$\delta P_2$ makes up the bulk of volatility risk for options that have mostly forward-smile risk – forward ATM call spreads or digital payoffs: $G$ has hardly any sensitivity to $\widehat{\sigma}_{T_1 T_2}$ and $\delta P_1 \simeq 0$.

**Pricing forward-smile risk in a model**

- In the *continuous* forward variance models of Chapter 7, there are no separate handles on (a) the spot-starting vanilla smile and (b) future smiles. Thus, at inception, the model should be parametrized so that the desired future smile at $T_1$ is obtained, while as $t \to T_1$, the model is calibrated to the vanilla smile.

- The local-stochastic volatility models of Chapter 12, on the other hand, are calibrated to the vanilla smile, thus $\delta P_2 \to 0$ as $t \to T_1$. The price we pay for this is less control on future smiles generated by the model – that is the value of $\delta P_2$ at $t = 0$.

  We refer the reader to Section 12.6.1 of Chapter 12 for a comparison of prices of forward-start options in different models calibrated to the same vanilla smile.

- The *discrete* forward variance models discussed in Section 7.8 of Chapter 7 offer maximum flexibility: we have a handle on the term-structure of forward skew, while still retaining the capability of matching the short market spot-starting smile. In addition they also afford separate control of $\delta P_1$ (volatility-of-volatility risk) and $\delta P_2$ (forward-smile risk).

Our analysis has shown that the risk that can be hedged using vanilla options is the forward volatility $\widehat{\sigma}_{T_1 T_2}$. Other risks – forward smile and volatility-of-volatility risk – cannot be hedged with vanilla options and have to be priced-in using exogenous parameters.

Contrary to what is sometimes heard on trading desks, an at-the-money forward call option *does* have forward smile sensitivity. The fact that it is at-the-money has no special significance, as forward implied volatilities of vanilla options – be they at-the-money or not – cannot be locked, unlike forward log-contract implied volatilities.[6] Consequently, for cliquets with payoff $g\left(\frac{S_{T_2}}{S_{T_1}}\right)$, the only information available in vanilla option prices to be used in the calibration of our model is the term structure of implied volatilities $\widehat{\sigma}_T$ of log contracts, as these are our hedge instruments.

We have no reason to use other vanilla option data: market prices of other instruments should be included in the calibration only if one is able to exactly pinpoint which instruments to use and which risk they offset.

We study in Section 3.1.9 further below the interesting example of payoffs $S_{T_1} g\left(\frac{S_{T_2}}{S_{T_1}}\right)$ which call for different hedging instruments, thus requiring calibration on a different set of European payoffs, leading to the definition of yet another class of forward volatilities.

The notion that a pricing model should generally be calibrated to the vanilla smile on the grounds that, hopefully, vanilla option prices provide information that should, somehow, be used, is fallacious at best and dangerous at worst. Going back to our forward-start option, this amounts to letting the model arbitrarily link the values of $\delta P_1$, $\delta P_2$ to the vanilla smile, when in practice no trading strategy is able to lock this dependency.

We now provide an illustration of the fact that cliquet prices are in fact rather loosely confined by vanilla smiles.

### 3.1.7 Lower/upper bounds on cliquet prices from the vanilla smile

As a simple example of a cliquet consider a forward ATM call that pays $\left(\frac{S_{T_2}}{S_{T_1}} - 1\right)^+$ at $T_2$. Let the hedging instruments be the underlying itself and call options of maturities $T_1$ and $T_2$. Assume zero interest rate and repo for simplicity.

[6] As will be discussed further on, log contracts are not traded, but variance swaps – which for practical purposes can be considered delta-hedged log contracts – are. Their gamma and vega have the same properties as those of the log contract.

Can the information in market smiles of maturities $T_1$ and $T_2$ be used to bound the price of our cliquet? We would like to derive model-independent lower and upper bounds such that if we were quoted a price that lay outside these bounds – say a lower price than the lower bound – a strategy consisting of (a) a long position in the cliquet entered at that price, (b) a static position in vanilla options of maturities $T_1$ and $T_2$ entered at market price, (c) a delta strategy over $[T_1, T_2]$, would generate positive P&L at $T_2$, for all $(S_{T_1}, S_{T_2})$ configurations.

The spread between lower and upper bound quantifies how much the cliquet differs from a statically replicable payoff and provides a measure of how vanilla-like – or unlike – the cliquet is. In case our cliquet payoff could be in fact synthesized by a combination of: (a) a fixed cash amount, (b) a static position in vanilla options of maturities $T_1$ and $T_2$, (c) a delta strategy set up at $T_1$ and unwound at $T_2$, then lower and upper bounds would coincide.

Consider thus a trading strategy that consists of (a) a cash amount $c$, (b) a static position in $\lambda_i$ call options of strikes $K_i$ / maturity $T_1$ and $\mu_j$ call options of strikes $K_j$ / maturity $T_2$, (c) a delta position $\Delta(S_{T_1})$ starting at $T_1$ and unwound at $T_2$, such that it super-replicates the cliquet's payoff.

Since entering the delta position $\Delta(S_{T_1})$ at $T_1$ does not require any cash outlay, the initial amount of cash needed for setting up this strategy is:

$$c + \Sigma_i \lambda_i C_{K_i T_1} + \Sigma_j \mu_j C_{K_j T_2}$$

Among all super-replicating strategies, that with the lowest initial cost provides the upper model-independent bound $UB$ for the cliquet's price that is compatible with market prices of hedging instruments. Mathematically:

$$UB = \min_{\lambda,\mu,\Delta(S_1),c} \left( c + \sum_i \lambda_i C_{K_i T_1} + \sum_j \mu_j C_{K_j T_2} \right)$$

such that:

$$c + \sum_i \lambda_i \left(S_1 - K_i\right)^+ + \sum_j \mu_j \left(S_2 - K_j\right)^+ + \Delta(S_1)\left(S_2 - S_1\right) \geq \left(\frac{S_2}{S_1} - 1\right)^+ \qquad \forall\,(S_1, S_2)$$

The model-independent lower bound can be defined analogously.[7]

We choose a discrete set of strikes $K_i$ for options of maturity $T_1$, a discrete set of strikes $K_j$ for options of maturity $T_2$ and a finite-dimensional basis for $\Delta\left(S_1\right)$. We

[7] The dual formulation of this problem corresponds to finding a joint distribution $\rho(S_{T_1}, S_{T_2})$ such that (a) the marginal distributions of $S_{T_1}$ and $S_{T_2}$ match the respective market smiles, (b) the martingality condition $E[S_{T_2}|S_{T_1}] = S_{T_1}$ holds and (c) the cliquet price is maximal.

See also the related discussion in the context of the $\lambda$-UVM in Appendix A of Chapter 2, page 90. In case prices of market instruments used for constraining the joint distribution are arbitrageable, this will manifest itself in the result that bounds are infinite.

then need to minimize an affine function of the vectors $\lambda$, $\mu$ and the components of $\Delta(S_1)$ on the basis we have chosen, subject to a set of constraints that are linear as well in $\lambda$, $\mu$, $\Delta(S_1)$: one constraint for each couple $(S_1, S_2)$. Practically we choose a (large) discrete set of such couples.

Numerically, this problem is solved with the simplex algorithm – we refer the reader to [57] for a more general account of sub and super-replicating strategies.

- Let us take $T_1 = 1$ year, $T_2 = 2$ years, vanishing interest rate and repo, and assume that the vanilla smiles for maturities $T_1$ and $T_2$ are flat with implied volatilities all equal to $20\%$. The model-independent lower and upper bounds for the implied volatility of the forward ATM call are: $\widehat{\sigma}_{\min} \simeq 9\%$, $\widehat{\sigma}_{\max} \simeq 25\%$.[8]

- Now, keeping the same smiles at $T_1$ and $T_2$ let us compute lower ($CS_{\min}$) and upper bounds ($CS_{\max}$) for the price of a $95\%/105\%$ forward call spread: $\left(\frac{S_{T_2}}{S_{T_1}} - 95\%\right)^+ - \left(\frac{S_{T_2}}{S_{T_1}} - 105\%\right)^+$.

  We get $CS_{\min} \simeq 1.6\%$, $CS_{\max} \simeq 7.7\%$. These prices correspond approximately to skews $(\widehat{\sigma}_{95\%} - \widehat{\sigma}_{105\%})_{\min} = -8\%$ and $(\widehat{\sigma}_{95\%} - \widehat{\sigma}_{105\%})_{\max} \simeq 8\%$. Notice how wide this range is – the typical order of magnitude of an index skew for a one-year maturity is $\widehat{\sigma}_{95\%} - \widehat{\sigma}_{105\%} = 3\%$.

- Let us compute again lower and upper bounds for our call spread, this time adding as an extra constraint the price of the forward ATM call, computed with an implied volatility of $20\%$: $7.96\%$.

  We now get: $CS_{\min} \simeq 1.8\%$, $CS_{\max} \simeq 7.2\%$. The additional information on the joint distribution of $S_{T_1}, S_{T_2}$ supplied by the forward ATM call has narrowed the price range only slightly.

- This time let us use as extra constraint the price of the forward call spread $\left(\frac{S_{T_2}}{S_{T_1}} - 90\%\right)^+ - \left(\frac{S_{T_2}}{S_{T_1}} - 110\%\right)^+$ computed with implied volatilities $\widehat{\sigma}_{90\%} = 22\%$, $\widehat{\sigma}_{110\%} = 18\%$. The price of this call spread is $10.7\%$.

  We now get: $CS_{\min} \simeq 4.25\%$, $CS_{\max} \simeq 6.7\%$, which corresponds approximately to $(\widehat{\sigma}_{95\%} - \widehat{\sigma}_{105\%})_{\min} = -1\%$ and $(\widehat{\sigma}_{95\%} - \widehat{\sigma}_{105\%})_{\max} \simeq 5.5\%$. Observe how inclusion of the price of a payoff whose risk is congruent with that of our $95\%/105\%$ call spread has tightened significantly the price range of the latter.

[8] I thank Pierre Henry-Labordère for sharing these results, which are obtained numerically – hence the symbol $\simeq$. While results depend somewhat on the discretization chosen for the $(S_1, S_2)$ couples, the same parameters have been used throughout. Note that the simplex algorithm is able to deal with a very large number of constraints.

### 3.1.8 Calibration on the vanilla smile – conclusion

The conclusion from the above results is that vanilla smiles hardly constrain cliquet prices. Only market prices of payoffs whose risks are congruent with those of the exotic at hand are able to narrow down the price range of the latter.

Our example illustrates the fact that an exotics business is not a brokerage business – otherwise exotics could be statically hedged by vanillas and hence simply priced with models whose parameters are fully determined by the vanilla smile. Rather, running a book of exotics entails taking controlled risks on exotic parameters such as: forward smile, volatility of volatility, smile of volatility of volatility, spot/volatility covariance. A suitable model should offer the capability of specifying these parameter levels exogenously.

Ideally the model should also be able to match market prices of vanilla options (really) used as hedges.

When this proves impossible, it is more reasonable to adjust the exotic's price to cover for the difference between model and market prices of hedging instruments, than to corrupt the dynamics in the model in order to calibrate vanilla option prices, with the prospect of mispricing future carry P&Ls.

### 3.1.9 Forward volatility agreements

So far we have defined cliquets as payoffs involving ratios of a spot price observed at dates $T_1$ and $T_2$. As a consequence their price in the Black-Scholes model does not depend on $S$ and, besides interest and repo rates, is a function of forward volatility only.

One may also be interested in payoffs involving *absolute*, rather than *relative* performances of an asset – typically for assets that move in narrow ranges, such as FX exchange rates.

Consider for example the payoff $(S_{T_2} - kS_{T_1})^+$: this is known in FX as a *forward volatility agreement* (FVA).[9] More generally, consider payoffs of type $S_{T_1} g\left(\frac{S_{T_2}}{S_{T_1}}\right)$. In a Black-Scholes framework with deterministic time-dependent volatility the value at $T_1$ of this option is $S_{T_1} G(\widehat{\sigma}_{T_1T_2})$ where $\widehat{\sigma}_{T_1T_2}$ is defined in (3.2). $G$ also involves interest and repo rates over $[T_1, T_2]$.The value of our forward-start option at time $t < T_1$ is:

$$P(t, S) = Se^{-q(T_1-t)} G(\widehat{\sigma}_{T_1T_2}) \tag{3.18}$$

Just as in Section (3.1.1) we need to answer the following questions:

- Which vanilla payoffs should be used as hedges and whose implied volatilities should be used for defining $\widehat{\sigma}_{T_1T_2}$?
- Which additional P&Ls do we need to estimate?

[9] This is one type of FVA – other types of FVAs include the regular forward-start vanilla option on the relative performance of the spot.

#### 3.1.9.1 A vanilla portfolio whose vega is linear in $S$

In comparison with (3.1), (3.18) involves $S$ as a prefactor. While in Section 3.1.2 we looked for European hedges of maturities $T_1$, $T_2$ whose vegas do not depend on $S$, we now look for payoffs whose vegas are proportional to $S$. From expression (3.4) for the vega of a portfolio of vanilla options in the Black-Scholes model, $\text{Vega}_\Pi(S)$ is proportional to $S$ only if $\rho(K)$ has the form:

$$\rho(K) \propto \frac{1}{K}$$

Integrating twice with respect to $K$ yields the payoff:

$$f(S) = S\ln S$$

which we call the $S\ln S$ contract. In the Black-Scholes model the price $R^T(t,S)$ of the $S\ln S$ contract of maturity $T$ is:

$$R^T(t,S) = Se^{-q(T-t)}\left(\ln S + (r-q)(T-t) + \frac{(T-t)\widehat{\sigma}_T^2}{2}\right) \tag{3.19}$$

and its sensitivity to $\widehat{\sigma}_T$ is given by:

$$\frac{dR^T}{d\widehat{\sigma}_T} = Se^{-q(T-t)}(T-t)\,\widehat{\sigma}_T \tag{3.20}$$

The sensitivities of $P$ in (3.18) to $\widehat{\sigma}_{T_1}, \widehat{\sigma}_{T_2}$ are given by:

$$\frac{dP}{d\widehat{\sigma}_{T_2}} = Se^{-q(T_1-t)}(T_2-t)\,\widehat{\sigma}_{T_2}\mathcal{N} \tag{3.21}$$

$$\frac{dP}{d\widehat{\sigma}_{T_1}} = -Se^{-q(T_1-t)}(T_1-t)\,\widehat{\sigma}_{T_1}\mathcal{N} \tag{3.22}$$

where $\mathcal{N}$ is given by:

$$\mathcal{N} = \frac{1}{(T_2-T_1)\,\widehat{\sigma}_{T_1T_2}}\frac{dG}{d\widehat{\sigma}_{T_1T_2}}$$

Observe how the vega of the $S\ln S$ contract in (3.20) exactly matches the $\widehat{\sigma}_{T_1}$ and $\widehat{\sigma}_{T_2}$ vegas of $P$, both in its dependence on $S$ and the repo rate. The portfolio

$$\Pi = -P + \mathcal{N}\left(e^{q(T_2-T_1)}R^{T_2} - R^{T_1}\right) \tag{3.23}$$

has vanishing sensitivity to $\widehat{\sigma}_{T_1}$ and $\widehat{\sigma}_{T_2}$. As is apparent in (3.23) the hedge ratios for $S\ln S$ payoffs of maturities $T_1$ and $T_2$ do not depend on $S$ or $t$: our hedge remains stable if $S$ moves or time advances. Only when $\widehat{\sigma}_{T_1T_2}$ moves will it need to be readjusted.

Taking the second derivatives of (3.18) with respect to $S$ shows that the gamma of our forward-start option vanishes. What about the gamma of our hedge? From expression (3.19) of $R^T$ we can see that the coefficient of $S\ln S$ in the portfolio $e^{q(T_2-T_1)}R^{T_2} - R^{T_1}$ vanishes: our hedge has vanishing gamma and vanishing theta as well.

**Is $\widehat{\sigma}_{T_1T_2}$ well-defined?**

The answer is yes. Implied volatilities of $S\ln S$ payoffs of maturities $T_1$ and $T_2$ satisfy the convex order condition (3.3) because of the convexity of the $S\ln S$ payoff. The $S\ln S$ payoff falls in the class of payoffs considered in Section 2.2.2.2, page 32, for which (a) an implied volatility can be defined, (b) the convex order condition (3.3) holds.

When there are no cash-amount dividends, implied volatilities $\widehat{\sigma}_T$ of $S\ln S$ payoffs are expressed directly as an average of implied volatilities of vanilla options – see formula (4.22), page 143.

### 3.1.9.2 Additional P&Ls and conclusion

Consider portfolio $\Pi$ and let us write the P&L during $\delta t$, at order two in $\delta S$, $\widehat{\sigma}_{T_1}, \widehat{\sigma}_{T_2}$. In contrast with the forward-start option in Section 3.1.5, $S$ appears explicitly in $P$ and $R^{T_1}, R^{T_2}$. In addition to the forward-start option and the $S\ln S$ contracts of maturities $T_1, T_2$, our hedge portfolio also includes a delta position.

Our P&L comprises:

- no order-one contributions from $\delta S$, $\delta\widehat{\sigma}_{T_1}, \delta\widehat{\sigma}_{T_2}$ as, by construction, the sensitivities of $\Pi$ to $\widehat{\sigma}_{T_1}, \widehat{\sigma}_{T_2}$ vanish and $\Pi$ is delta-hedged.
- no $\delta t$ and $\delta S^2$ terms as the portfolio's theta and gamma vanish.
- no order-two $\delta S\delta\widehat{\sigma}_{T_1}$ and $\delta S\delta\widehat{\sigma}_{T_2}$ contributions. Indeed, the sensitivities of the forward-start option's price $P$ in (3.21), (3.22) and of the $S\ln S$ contract's price $R^T$ in (3.20) to $\widehat{\sigma}_{T_1}, \widehat{\sigma}_{T_2}$ are proportional to $S$. If $\frac{d\Pi}{d\widehat{\sigma}_{T_1}}, \frac{d\Pi}{d\widehat{\sigma}_{T_2}}$ vanish for a particular value of $S$, they do so for all values of $S$: $\frac{d^2\Pi}{dSd\widehat{\sigma}_{T_1}} = \frac{d^2\Pi}{dSd\widehat{\sigma}_{T_2}} = 0$.

At order two our P&L thus only comprises terms in $\delta\widehat{\sigma}^2_{T_1}, \delta\widehat{\sigma}^2_{T_2}, \delta\widehat{\sigma}_{T_1}\delta\widehat{\sigma}_{T_2}$. Moreover, from (3.19), it is apparent that $e^{q(T_2-T_1)}R^{T_2} - R^{T_1}$ is a function of $\widehat{\sigma}_{T_1T_2}$, rather than a separate function of $\widehat{\sigma}_{T_1}, \widehat{\sigma}_{T_2}$. Note that it is simply an affine function of $\widehat{\sigma}^2_{T_1T_2}$. Using this variable, rather than $\widehat{\sigma}_{T_1T_2}$, we can write down our P&L at order two in $\delta(\widehat{\sigma}^2_{T_1T_2})$ directly – note the similarity with (3.12).

$$P\&L_\Pi = -\frac{Se^{-q(T_1-t)}}{2}\frac{d^2G}{d(\widehat{\sigma}^2_{T_1T_2})^2}\left(\delta(\widehat{\sigma}^2_{T_1T_2})\right)^2 \tag{3.24}$$

**Conclusion**

Again we have been able to find a portfolio of European payoffs that provides a suitable vega, gamma and theta-hedge for our forward-start option. The implied volatilities we use are those of $S\ln S$ contracts.

The price we quote consists of the three pieces in (3.17). $\delta P_1$ is an estimation of P&Ls (3.24) over $[0, T_1]$. Because $S$ appears as a prefactor, the estimation of $\delta P_1$ may also involve an assumption about the correlation of $S$ and the realized volatility of the forward variance $\widehat{\sigma}^2_{T_1T_2}$. $\delta P_2$ represents an estimation of the difference between the market price at $T_1$ of a vanilla option of strike $kS_{T_1}$, maturity $T_2$, and its price computed with the implied volatility at $T_1$ of the $S\ln S$ contract of maturity $T_2$.

A hasty assessment of our forward-start option starting with expression (3.18) in the Black-Scholes model would have singled out the cross spot/forward volatility covariance as one of the main risks in our product. Our analysis demonstrates, however, that spot/volatility covariance risk is not relevant; it is offset by utilizing the right hedge instruments – $S\ln S$ payoffs.

---

## 3.2 Forward-start options in the local volatility model

We now assess how the local volatility model prices forward-smile risk.

Once the local volatility model is calibrated on a given smile, log contract volatilities $\widehat{\sigma}_T$ will be exactly calibrated, as the log contract is a European payoff. However, calibration to the market smile also determines the dynamics of $\widehat{\sigma}_T$ – and especially that of $\widehat{\sigma}_{T_1T_2}$. The local volatility model will price volatility-of-volatility and forward-smile risks – that is adjustments $\delta P_1, \delta P_2$ – according to its own dynamics, which we now characterize assuming that the local volatility function is weakly local.

### 3.2.1 Approximation for $\widehat{\sigma}_T$

Let us approximate the log contract implied volatility $\widehat{\sigma}_T$ by starting from expression (2.35). The reader can check that the derivation leading to equation (2.35) applies generally to any European option for which an implied volatility can be defined, i.e. applies to any payoff that is convex or concave, in particular the log contract.

Calculation is simpler now than it was in Section 2.4.2 for vanilla options as the dollar gamma of the log contract does not depend on $S$. The numerator in (2.35) reads:

$$\begin{aligned}
&\int_0^T dt \int_0^\infty dS\ \rho_{\sigma_0}(t,S)\, e^{-rt}\ \delta u\,(t,S)\ \ S^2 \frac{d^2 P_{\sigma_0}}{dS^2} \\
&= \ e^{-rT} \int_0^T dt \int_0^\infty dS\ \rho_{\sigma_0}(t,S)\, \delta u\,(t,S)
\end{aligned}$$

where we have used (3.10). Using now expression (2.38) for $\rho_{\sigma_0}(t,S)$ we get for the numerator:

$$e^{-rT}\int_0^T dt\int_{-\infty}^{\infty}dy\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,\delta u\left(t,\;F_te^{-\frac{\sigma_0^2t}{2}+\sigma_0\sqrt{t}y}\right)$$

where $F_t$ is the forward for maturity $t$: $F_t = S_0e^{(r-q)t}$. Dividing now by the denominator in (2.35), which is simply equal to $Te^{-rT}$, yields:

$$\widehat{\sigma}_T^2 = \frac{1}{T}\int_0^T dt\int_{-\infty}^{\infty}dy\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,u\left(t,\;F_te^{-\frac{\sigma_0^2t}{2}+\sigma_0\sqrt{t}y}\right) \tag{3.25}$$

(3.25) should be compared to formula (2.40), page 44, previously derived for vanilla options. While (3.25) could be used directly, for consistency with Section 2.5 we derive an equation for volatilities rather than variances. Let us assume that $\sigma(t,S) = \sigma_0 + \delta\sigma(t,S)$. At order one in $\delta\sigma$, $\delta u = 2\sigma_0\delta\sigma$. (3.25) together with $\widehat{\sigma}_T^2 = \sigma_0^2 + 2\sigma_0\delta\widehat{\sigma}_T$ yields:

$$\delta\widehat{\sigma}_T = \frac{1}{T}\int_0^T dt\int_{-\infty}^{\infty}dy\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,\delta\sigma\left(t,\;F_te^{-\frac{\sigma_0^2t}{2}+\sigma_0\sqrt{t}y}\right)$$

Adding $\sigma_0$ to both sides, we get:

$$\widehat{\sigma}_T = \frac{1}{T}\int_0^T dt\int_{-\infty}^{\infty}dy\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,\sigma\left(t,\;F_te^{-\frac{\sigma_0^2t}{2}+\sigma_0\sqrt{t}y}\right) \tag{3.26}$$

Assume that $\sigma$ is given by the smooth form of the same type as in (2.44):

$$\sigma(t,S) = \overline{\sigma}(t) + \alpha(t)\,x + \frac{\beta(t)}{2}x^2 \tag{3.27}$$

where $x = \ln(S/F_t)$. We get:

$$\begin{aligned}
\widehat{\sigma}_T &= \frac{1}{T}\int_0^T\overline{\sigma}(t)\,dt - \frac{1}{T}\int_0^T\alpha(t)\,\frac{\sigma_0^2t}{2}dt + \frac{1}{T}\int_0^T\frac{\beta(t)}{2}\left(\sigma_0^2t + \frac{\sigma_0^4t^2}{4}\right)dt \\
&= \frac{1}{T}\int_0^T\overline{\sigma}(t)\,dt - \frac{\sigma_0^2T}{2}\mathcal{S}_T + \frac{1}{T}\int_0^T\frac{\beta(t)}{2}\left(\sigma_0^2t + \frac{\sigma_0^4t^2}{4}\right)dt
\end{aligned}$$

where we have used formula (2.48) for the ATMF skew $\mathcal{S}_T$.

We also need an expression for how $\widehat{\sigma}_T$ moves when $S_0$ moves; for this we go back to the general expression of $\widehat{\sigma}_T$ in (3.26). Let us use the notation $S(t,y) = F_t\exp\left(-\frac{\sigma_0^2t}{2} + \sigma_0\sqrt{t}y\right)$ and take the derivative of (3.26) with respect to $\ln S_0$.

$$\begin{aligned}
\frac{d\widehat{\sigma}_T}{d\ln S_0} &= \frac{1}{T}\int_0^T dt \int_{-\infty}^{\infty} dy \, \frac{1}{\sqrt{2\pi}} \, e^{-\frac{y^2}{2}} \, \frac{d\sigma\left(t, S\left(t, y\right)\right)}{d\ln S_0} \\
&= \frac{1}{T}\int_0^T dt \int_{-\infty}^{\infty} dy \, \frac{1}{\sqrt{2\pi}} \, e^{-\frac{y^2}{2}} \, \frac{d\sigma}{d\ln S}\left(t, S\left(t, y\right)\right) \frac{d\ln S\left(t, y\right)}{d\ln S_0} \\
&= \frac{1}{T}\int_0^T dt \int_{-\infty}^{\infty} dy \, \frac{1}{\sqrt{2\pi}} \, e^{-\frac{y^2}{2}} \, \frac{d\sigma}{d\ln S}\left(t, S\left(t, y\right)\right)
\end{aligned}$$

Substituting now in this equation the smooth form (3.27) for $\sigma\left(t, S\right)$ yields:

$$\begin{aligned}
\frac{d\widehat{\sigma}_T}{d\ln S_0} &= \frac{1}{T}\int_0^T dt \int_{-\infty}^{\infty} dy \frac{1}{\sqrt{2\pi}} \, e^{-\frac{y^2}{2}} \left(\alpha\left(t\right) + \beta(t)\Big(-\frac{\sigma_0^2 t}{2} + \sigma_0\sqrt{t}y\Big)\right) \\
&= \frac{1}{T}\int_0^T \alpha\left(t\right) dt \; - \; \frac{\sigma_0^2 T}{2}\frac{1}{T}\int_0^T \frac{t}{T}\beta(t) dt
\end{aligned}$$

We now drop the contribution from the curvature term $\beta$ as we will not use it. Setting $\beta\left(t\right) = 0$ in the expressions for $\widehat{\sigma}_T$ and $\frac{d\widehat{\sigma}_T}{d\ln S_0}$ above yields the following simpler expressions:

$$\widehat{\sigma}_T = \frac{1}{T}\int_0^T \overline{\sigma}\left(t\right) dt - \frac{\sigma_0^2 T}{2}\mathcal{S}_T \tag{3.28}$$

$$\frac{d\widehat{\sigma}_T}{d\ln S_0} = \frac{1}{T}\int_0^T \alpha\left(t\right) dt \tag{3.29}$$

Taking $\beta\left(t\right) = 0$ in expression (2.47) for implied volatilities of vanilla options yields the following simple result for the ATMF implied volatility: $\widehat{\sigma}_{F_T T} = \frac{1}{T}\int_0^T \overline{\sigma}\left(t\right) dt$. Let us choose $\sigma_0$, the constant volatility level around which the order-one expansion in (2.35) is performed, as $\sigma_0 = \widehat{\sigma}_{F_T T}$. Equation (3.28) now reads:

$$\widehat{\sigma}_T = \widehat{\sigma}_{F_T T} - \frac{\widehat{\sigma}_{F_T T}^2 T}{2}\mathcal{S}_T \tag{3.30}$$

The integral in the right-hand side in (3.29) also appears in (2.59c). (3.29) can be rewritten, with no explicit reference to local volatility $\sigma$ anymore, as:

$$\frac{d\widehat{\sigma}_T}{d\ln S_0} = \frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0} \tag{3.31}$$

While we have derived equation (3.30) using local volatilities, it is a relationship involving implied volatilities only: it relates the implied volatility of the log contract, a European payoff, to the smile of vanilla options for the same maturity at order one in the perturbation around a flat local volatility function.[10] We could have derived it in a model-independent fashion, by assuming that the smile for maturity $T$ is given by:

$$\widehat{\sigma}_{KT} \;=\; \widehat{\sigma}_{F_T T} + \mathcal{S}_T \ln\left(\frac{K}{F_T}\right)$$

At order one in $\mathcal{S}_T$ we would have recovered (3.30).

Equation (3.30) implies that for downward sloping smiles ($\mathcal{S}_T < 0$), which are typical for equities, $\widehat{\sigma}_T > \widehat{\sigma}_{F_T T}$. Equation (3.31) states that the rate at which the log contract implied volatility moves as $S$ moves matches that of the ATMF volatility for the same maturity. Equations (3.30) and (3.31) have been obtained for a weakly local volatility of the form

$$\sigma(t, S) \;=\; \overline{\sigma}(t) + \alpha(t) \ln\left(\frac{S}{F_t}\right) \tag{3.32}$$

How accurate are they for real smiles?

Figure 3.1 shows $\widehat{\sigma}_T$, $\widehat{\sigma}_{F_T T}$, as well as $\widehat{\sigma}_T$ computed with formula (3.30) for two Euro Stoxx 50 smiles – with zero rates and repos. Figure 3.2 shows the ratio $\frac{d\widehat{\sigma}_T}{d\ln S_0} / \frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0}$ for the same smiles.

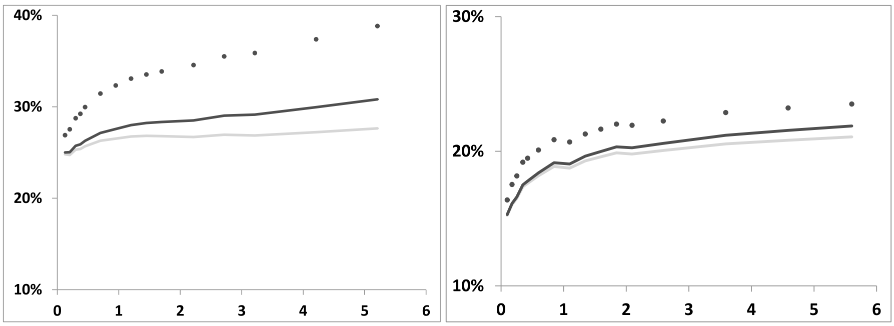


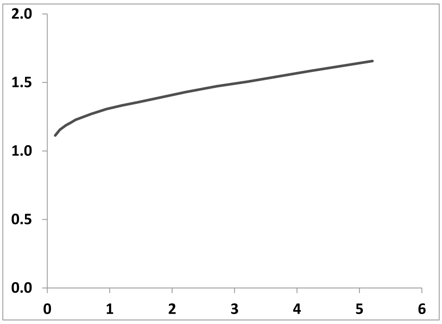

**Figure 3.1**: Implied volatilities as a function of maturity – in years – for the Euro Stoxx 50 smiles of October 4, 2010 (left) and May 16, 2013 (right); dots: $\widehat{\sigma}_T$, light line: $\widehat{\sigma}_{F_T T}$, dark line: approximation (3.30) for $\widehat{\sigma}_T$.

The 1-year smiles of the Euro Stoxx 50 index appear in Figure 2.3, page 58.

It is apparent that formula (3.30) is inaccurate. This is not surprising: the local volatility of the Euro Stoxx 50 market smile is not of type (3.32). In particular the difference between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{F_T T}$ is not entirely determined by the ATMF skew $\mathcal{S}_T$, a very local feature of the smile.

Then note that an expansion of implied volatilities at order one in the local volatility is expected to be less accurate for log contracts than for calls and puts. We

[10]Equivalently, at order one in the perturbation around a flat *implied* volatility surface.



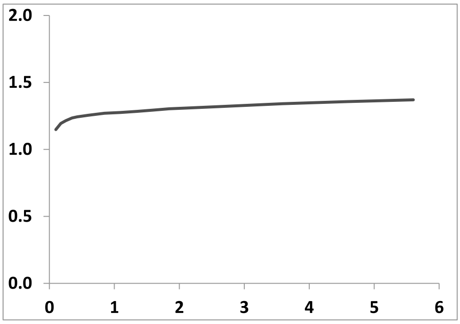




**Figure 3.2**: $\frac{d\widehat{\sigma}_T}{d\ln S_0} / \frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0}$ ratio in the local volatility model, as a function of maturity – in years – for the Euro Stoxx 50 index smiles of October 4, 2010 (left) and May 16, 2013 (right).

already mentioned that, because of the steepness of market smiles, unless one takes into account – at least at order one – the correction to the density in equation (2.32), the approximation is inaccurate.

This issue is magnified in the case of the log contract: for vanilla options, the dollar gamma restricts integration in (2.32) to a region of spot prices around $K$, thus reducing the dependence of implied volatilities of vanilla options to the density for spot values far away from $K$. In contrast, the dollar gamma of the log contract is constant: its implied volatility depends on the density for values of $S$ far away from the initial spot price. Typically densities implied from market smiles are skewed: for low values of $S$, for which $\sigma(t, S)$ is large, the density is larger than the lognormal density used in the approximation.

Formula (3.30) for $\widehat{\sigma}_T$ will then only be acceptable for smooth smiles that are not too steep.

Figure 3.2 shows that the ratio $\frac{d\widehat{\sigma}_T}{d\ln S_0} / \frac{d\widehat{\sigma}_{F_T T}}{d\ln S_0}$ is, in our example, larger than its value of 1 predicted by approximation (3.32), especially for the October 4, 2010 smile, which is particularly steep, presumably because approximation (3.30) underestimates $\widehat{\sigma}_T$ in the first place. While numerically inaccurate, approximation (3.31) does give however the right order of magnitude for $\frac{d\widehat{\sigma}_T}{d\ln S_0}$.

We now use it to estimate the volatility of $\widehat{\sigma}_{T_1T_2}$, assuming for simplicity that the term structure of $\widehat{\sigma}_T$ is flat: $\widehat{\sigma}_{T_1} = \widehat{\sigma}_{T_2}$ . From the definition (3.2) and using approximation (3.29) we have:

$$\begin{aligned} d\widehat{\sigma}_{T_1T_2} &= \frac{T_2 - t}{T_2 - T_1}\frac{\widehat{\sigma}_{T_2}}{\widehat{\sigma}_{T_1T_2}}d\widehat{\sigma}_{T_2} - \frac{T_1 - t}{T_2 - T_1}\frac{\widehat{\sigma}_{T_1}}{\widehat{\sigma}_{T_1T_2}}d\widehat{\sigma}_{T_1} \\ &\simeq \left(\frac{1}{T_2 - T_1}\int_t^{T_2}\alpha(u)\,du - \frac{1}{T_2 - T_1}\int_t^{T_1}\alpha(u)\,du\right)d\ln S_t \\ &\simeq \left(\frac{1}{T_2 - T_1}\int_{T_1}^{T_2}\alpha(u)\,du\right)d\ln S_t \end{aligned} \tag{3.33}$$

The volatility of $\widehat{\sigma}_{T_1T_2}$ over $[0, T_1]$ is then entirely set by the skew of the local volatility function for $t \in [T_1, T_2]$, and presumably will bear no resemblance to historical or implied levels of volatility of volatility. In particular, (3.33) shows that if the local volatility is flat over the interval $[T_1, T_2]$, $\widehat{\sigma}_{T_1T_2}$ is frozen.

Moreover, for $T_2 - T_1$ fixed, the volatility of $\widehat{\sigma}_{T_1T_2}$ is not a function of $T_1 - t$: the local volatility model is not time-homogeneous, in contrast to the stochastic volatility models of Chapter 7.

### 3.2.2 Future skews in the local volatility model

We now turn to the forward skew in the local volatility model. The smile prevailing at $T_1$ for maturity $T_2$ is fully determined by $\sigma(t, S)$ for $t \in [T_1, T_2]$. The local volatility function in (3.32) falls in the class studied in Section 2.5, with $\beta(t) \equiv 0$.

Let us use approximation (3.30) for the difference between the log contract and ATMF implied volatilities. We have:

$$\widehat{\sigma}_{T_2}(S_{T_1}, T_1) - \widehat{\sigma}_{F_{T_2}(S_{T_1})T_2}(S_{T_1}, T_1) = \frac{(T_2 - T_1)\,\widehat{\sigma}^2_{F_{T_2}(S_{T_1})T_2}}{2}\mathcal{S}_{T_2-T_1}(S_{T_1}, T_1)$$

where $\mathcal{S}_{T_2-T_1}(S_{T_1}, T_1)$ is the ATMF skew for the residual maturity $T_2 - T_1$, at time $T_1$, as a function of $S_{T_1}$.

Let us denote the residual maturity by: $\theta = T_2 - T_1$. $\mathcal{S}_\theta(S_{T_1}, T_1)$ is given by expression (2.90), page 63:

$$\mathcal{S}_\theta(S_{T_1}, T_1) = \int_0^1 \alpha(T_1 + u\theta)u\,du \tag{3.34}$$

which can be contrasted with the spot-starting ATMF skew for the same residual maturity:

$$\mathcal{S}_\theta(S_0, 0) = \int_0^1 \alpha(u\theta)u\,du$$

For typical equity smiles, $\alpha(t)$ is of the form (2.51) and decays with an exponent $\gamma \simeq \frac{1}{2}$. For short residual maturities ($\theta \ll T_1$), result (2.92), page 64, implies that:

$$\mathcal{S}_\theta(S_{T_1}, T_1) \propto \frac{1}{T_1^\gamma}$$

This formula shows that the ATMF skew generated by the local volatility model at time $T_1$ for residual maturity $\theta$ decays like $1/T_1^\gamma$, thus will be much lower than the ATMF skew observed at $t = 0$ for the same residual maturity:

$$\mathcal{S}_\theta(S_{T_1}, T_1) \ll \mathcal{S}_\theta(S_0, 0)$$

Imagine that our cliquet is an at-the-money forward call whose payoff is (3.13). At $T_1$ we are exposed to the difference between the implied volatilities of the log contract and the at-the-money call for maturity $T_2$. If $T_1$ is far into the future, the local volatility model generates a smile at $T_1$ for maturity $T_2$ which is very weak compared to the smile we are likely to witness on the market.

The local volatility model will underestimate the difference between the log contract and at-the-money implied volatilities, and more generally the differences between implied volatilities of vanilla options with different strikes: it misprices forward-smile risk.

Remember that in the local volatility model, instantaneous volatilities of volatilities are also determined by the ATM skew. From equation (2.83), page 57,

$$\mathrm{vol}(\widehat{\sigma}_{F_T T}) \;=\; \left(\mathcal{S}_T + \frac{1}{T}\int_0^T \mathcal{S}_\tau d\tau\right)\frac{\widehat{\sigma}_{F_0 0}}{\widehat{\sigma}_{F_T T}}$$

Thus future volatilities of volatilities are smaller than their spot-starting values: the local volatility model also misprices volatility-of-volatility risk.

### 3.2.3 Conclusion

The conclusion is that the local volatility model is not suitable for pricing forward-start options, or more generally options that involve volatility-of-volatility and forward-smile risks.

On one hand, volatilities of forward volatilities generated by the model will depend exclusively on the steepness of the skews prevailing at the time of calibration and may lie arbitrarily above or below historical or implied volatilities of volatilities – thus $\delta P_1$ in (3.17), page 111, is mispriced.

On the other hand, forward skews generated in the model will invariably be too low with respect to both market implied forward skews and historical skews – thus $\delta P_2$ in (3.17) is mispriced as well.

To further understand how the local volatility model uses the information in vanilla smiles to price a cliquet, we now consider a forward-start call and apply the technique of Section 2.9 to explicitly derive the vega hedge – as generated by the local volatility model.

### 3.2.4 Vega hedge of a forward-start call in the local volatility model

Consider the following payoff:

$$\left(\frac{S_{T_2}}{S_{T_1}} - k\right)^+ \;=\; \frac{1}{S_{T_1}}\left(S_{T_2} - kS_{T_1}\right)^+ \tag{3.35}$$

To compute the vega hedge in the local volatility model, all we need is the conditional gamma notional, defined in (2.118), page 82:

$$\phi\left(t,S\right) \;=\; E_\sigma\left[S^2\frac{d^2P}{dS^2}\left(t,S,\;\bullet\right)|S,t\right]$$

We consider for simplicity the case of flat local volatility function. Calculations can then be carried out analytically, besides, we do not expect the structure of the hedge portfolio to depend much on this assumption.

We also use vanishing interest and repo rates.

Let us call $\sigma_0$ the constant level of our flat local volatility function. Because of homogeneity, for $t < T_1$ the Black-Scholes price does not depend on $S$, hence $\phi = 0$. At $t = T_1^+$, our option becomes $\frac{1}{S_{T_1}}$ times a standard call of maturity $T_2$ whose strike is $kS_{T_1}$.

**Hedge portfolio for maturity $T_1$**

At $t = T_1^+$, $S_{T_1} = S$ and the dollar gamma is given by:

$$\phi\left(T_1^+, S\right) = \frac{1}{S} S^2 \left.\frac{d^2 P_{BS}\left(T_1 S; K T_2; \sigma_0\right)}{dS^2}\right|_{K=kS}$$

where the $\frac{1}{S}$ prefactor is the $\frac{1}{S_{T_1}}$ in (3.35). We now use the relationship connecting the vega and dollar gamma of a European option in the Black-Scholes model: $\frac{dP}{d\sigma_0} = S^2 \frac{d^2 P}{dS^2} \sigma_0 T$ to rewrite this as:

$$\phi\left(T_1^+, S\right) = \frac{1}{\sigma_0\left(T_2 - T_1\right)} \frac{1}{S} \left.\frac{dP_{BS}\left(T_1 S; K T_2; \sigma_0\right)}{d\sigma_0}\right|_{K=kS}$$

Because $P_{BS}\left(T_1 S; kS\, T_2; \sigma_0\right)$ is homogeneous in $S$, $\phi\left(T_1^+, S\right)$ does not depend on $S$. $\phi$ thus has a discontinuity at $t = T_1$ that does not depend on $S$, which generates a discrete portfolio of vanilla options of maturity $T_1$. Applying operator $\mathcal{L}$ defined in (2.120) on $\phi$ – only $\frac{d}{dt}\phi$ contributes – we get the following expression for the (discrete) density of vanilla options of maturity $T_1$, struck at $K$:

$$\begin{aligned} \Psi_1(K) &= -\frac{1}{K^2} \phi\left(T_1^+, S\right) \\ &= -\frac{1}{K^2} \frac{1}{\sqrt{2\pi\sigma_0^2\left(T_2 - T_1\right)}} e^{-\frac{\left(-\ln k + \frac{\sigma_0^2\left(T_2 - T_1\right)}{2}\right)^2}{2\sigma_0^2\left(T_2 - T_1\right)}} \end{aligned} \tag{3.36}$$

This is an interesting result: $\mu$ is proportional to $\frac{1}{K^2}$. Compare (3.36) with expression (3.9) that specifies the number of log contracts of maturity $T_1$ and $T_2$ needed to hedge a forward-start option, in the framework of Section 3.1.4. Once we recall that one log contract can be synthesized with a continuous density $\frac{2}{K^2}$ of vanilla options, we realize that result (3.36) is exactly identical: perturbation around a flat volatility yields a vega hedge for maturity $T_1$ which is exactly what we used in Section 3.1.4.

Notice that the vanilla portfolio struck at $T_1$ is static, in that it depends only on $T_2 - T_1$: it depends neither on $T_1$, nor on the initial spot level $S_0$.

**Hedge portfolio for maturity $T_2$**

Let us compute $\phi$ for $t \in ]T_1, T_2[$. Using formula (2.39) for the dollar gamma of a call option in the Black-Scholes model, we get the dollar gamma for payoff (3.35):

$$S^2 \frac{d^2 P}{dS^2} = \frac{k}{\sqrt{2\pi\sigma_0^2 (T_2 - t)}} e^{-\frac{\left(\ln\left(\frac{S}{kS_{T_1}}\right) - \frac{\sigma_0^2 (T_2 - t)}{2}\right)^2}{2\sigma_0^2 (T_2 - t)}} \tag{3.37}$$

which we need to average, conditional on the underlying's value being $S$ at time $t$. The Brownian motion $W_{T_1}$ can be written as a function of $W_t$ and an independent Gaussian random variable $Z$:

$$W_{T_1} = \frac{T_1}{t} W_t + \sqrt{T_1 \left(1 - \frac{T_1}{t}\right)} Z$$

which gives:

$$\ln \frac{S_{T_1}}{S_0} = \frac{T_1}{t} \ln \frac{S_t}{S_0} + \sigma_0 \sqrt{T_1 \left(1 - \frac{T_1}{t}\right)} Z$$

Inserting this expression for $S_{T_1}$ in (3.37) yields the following expression for $\phi$:

$$\phi = \int_{-\infty}^{+\infty} \frac{e^{-\frac{Z^2}{2}}}{\sqrt{2\pi}} \frac{k}{\sqrt{2\pi\sigma_0^2 (T_2 - t)}} e^{-\frac{\left(\left(1 - \frac{T_1}{t}\right) \ln \frac{S}{S_0} - \ln k - \frac{\sigma_0^2 (T_2 - t)}{2} - \sigma_0 \sqrt{T_1 \left(1 - \frac{T_1}{t}\right)} Z\right)^2}{2\sigma_0^2 (T_2 - t)}} dZ$$

Computing the integral over $Z$ yields:

$$\phi = \frac{k}{\sqrt{2\pi\sigma_0^2 \left(T_2 - t + T_1 \left(1 - \frac{T_1}{t}\right)\right)}} e^{-\frac{\left(\left(1 - \frac{T_1}{t}\right) \ln \frac{S}{S_0} - \ln k - \frac{\sigma_0^2 (T_2 - t)}{2}\right)^2}{2\sigma_0^2 \left(T_2 - t + T_1 \left(1 - \frac{T_1}{t}\right)\right)}}$$

Let us compute $\phi$ for $t = T_2^-$:

$$\phi\left(t = T_2^-, S\right) = \frac{k}{\sqrt{2\pi\sigma_0^2 T_1 \left(1 - \frac{T_1}{T_2}\right)}} e^{-\frac{\left(\left(1 - \frac{T_1}{T_2}\right) \ln \frac{S}{S_0} - \ln k\right)^2}{2\sigma_0^2 T_1 \left(1 - \frac{T_1}{T_2}\right)}}$$

For $t > T_2$ $\phi = 0$: the discontinuity of $\phi$ in $T_2$ generates a discrete quantity of vanilla options struck at $T_2$, whose density $\Psi_2(K)$ is $\frac{1}{K^2}\phi\left(T_2^-, K\right)$:

$$\Psi_2(K) = \frac{1}{K^2} \frac{k}{\sqrt{2\pi\sigma_0^2 T_1 \left(1 - \frac{T_1}{T_2}\right)}} e^{-\frac{\left(\left(1 - \frac{T_1}{T_2}\right) \ln \frac{K}{S_0} - \ln k\right)^2}{2\sigma_0^2 T_1 \left(1 - \frac{T_1}{T_2}\right)}} \tag{3.38}$$

First note that, unlike the $T_1$ portfolio, the $T_2$ hedge portfolio is not static: $T_1, T_2$ and $S$ appear explicitly. As time advances, $T_1$ and $T_2$ shrink, and $S$ moves. The number of vanilla options of maturity $T_2$ struck at $K$, as given by (3.38) will need to be readjusted. This does not correspond to the $T_2$ hedge we assembled in Section 3.1.4 which, with zero interest rates, consists of a number of log contracts for maturity $T_2$ that is exactly the opposite of that of maturity $T_1$ – see equation (3.9).

**Intermediate maturities**

Beside the discrete portfolios of vanilla options for maturities $T_1$ and $T_2$ that are generated by the discontinuity of $\phi$ at $T_1$ and $T_2$, application of the operator $\mathcal{L}$ defined in (2.120) on $\phi$ generates a continuous density of options for intermediate maturities that can be easily computed numerically.

Our hedge portfolio in Section 3.1.4, on the contrary, only consists of options of maturity $T_1$ and $T_2$.

### 3.2.5 Discussion and conclusion

The density of vanilla options generated by the log contract hedge in Section 3.1.4 was $|\Psi_1(K)|$ for $T_2$ and $-|\Psi_1(K)|$ for $T_1$.

Rather than work directly with the (discrete) density of options struck at $T_2$ and the (continuous) density for $T \in [T_1, T_2]$, let us normalize the former by $|\Psi_1(K)|$ and the latter by $|\Psi_1(K)| / (T_2 - T_1)$, to highlight the deviation with respect to the log contract hedge. As (3.36) shows, $|\Psi_1(K)|$ is proportional to $1/K^2$.

Consider the following at-the-money forward call: $T_1 = 1$ year, $T_2 = 2$ years, $\sigma_0 = 20\%$, $k = 100\%$, and pick $T = 1.5$ years. Figure 3.3 shows the following quantities:

$$\frac{\Psi_2(K)}{|\Psi_1(K)|}\;,\quad \frac{(T_2 - T_1)\,\mu(T,K)}{|\Psi_1(K)|} \tag{3.39}$$

where we multiply $\mu$ by $(T_2 - T_1)$ since $\mu$, unlike $\Psi_1$ and $\Psi_2$, is a continuous density.

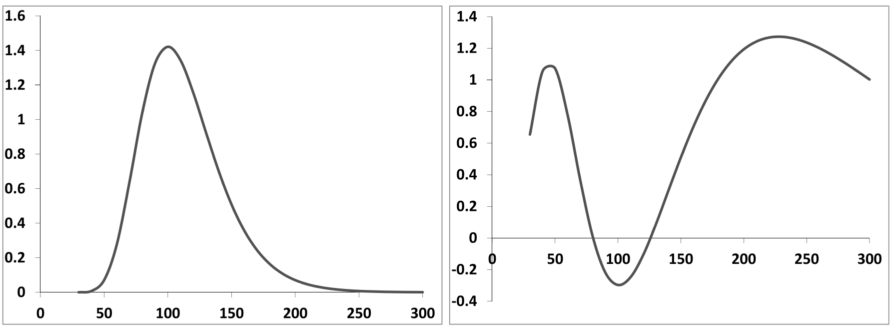



**Figure 3.3**: Densities of vanilla options struck at $T_2$ (left) and at an intermediate maturity $T$, normalized as in (3.39) (right).

Notice how the $T_2$-portfolio deviates from the log contract hedge we used in Section 3.1.4. The local volatility model suggests that we should trade more options struck near the current value of the spot price (100 in our example) and fewer out-of-the-money options. It makes up for this by asking us to trade options for intermediate maturities: a short position for strikes near the current value of the spot price and a long position for far out-of-the-money strikes. Inspection of Figure 3.3 shows that the vanilla hedge for intermediate maturities approximately makes up for the difference between the $T_2$-portfolio and the log contract hedge, which – because of the normalization we use – would correspond in the left-hand graph to a constant value equal to 1.

While the $T_1$-portfolio is static, portfolios for $T_2$ and intermediate maturities in $]T_1, T_2[$ depend explicitly on $T_1, T_2, S_0$: as the spot moves and time advances, they will need to be readjusted. Using this vega hedge in practice would expose us to variations of $S$ and market implied volatilities, thus generating extra P&Ls – of order two in $\delta S$ and $\delta\widehat{\sigma}_{KT}$ – which we cannot expect the local volatility model to have priced in properly.

Carrying out the same analysis for a forward call option struck at $k \neq 100\%$, would have given similar curves, now centered on $K^* = S_0 k^{\frac{T_2}{T_2 - T_1}}$ – as is clear from (3.38).

The conclusion is that it is much more reasonable to use the log contract hedge developed in Section 3.1.4: only when the forward volatility $\widehat{\sigma}_{12}$ moves does the hedge need to be readjusted and the volatility-of-volatility and forward-smile risks can be cleanly isolated and priced.

What about the *price* in the local volatility model? Consider again Figure 3.3 which expresses the dependence of the price of the forward call option to implied volatilities of vanilla options – as seen by the local volatility model. It is difficult to imagine a plausible justification for such peculiar sensitivities, especially as they change when $S$ and $t$ move. They may be more a statement on the local volatility model itself than on the forward call option: besides the *hedge*, the *price* generated by the local volatility model is suspicious as well.

Quite generally, whenever the vega hedge suggested by the local volatility model is *not* static, both the usefulness of the hedge *and* the reliability of the model's price are questionable. We are exposed to the cost of future vega rehedging at then-prevailing market conditions without the ability to gauge the size and sign of these future P&Ls, and cannot trust the local volatility model to have priced them correctly.

## Chapter's digest

### 3.1 Pricing and hedging forward-start options

▶ Cliquets whose payoffs are of the form $g\left(\frac{S_{T_2}}{S_{T_1}}\right)$ are hedged by trading dynamically log-contracts. Implied volatilities of log contracts obey the convex order condition, hence forward volatilities are well-defined, and the sensitivities of log-contracts to forward volatilities are spot-independent. The carry P&L of a vega-hedged cliquet consists of (a) gamma P&L on forward volatility and (b) an adjustment for forward-smile risk, that is uncertainty about the value at $T_1$ of the difference between the implied volatility of the log-contract and that of the payoff at hand.

▶ Calibration on the vanilla smile has little relevance for the pricing of cliquets. Calibrating a model on the vanilla smile in the hope that information contained therein should somehow be reflected in the cliquet price is unreasonable. Indeed, this may result in a price that is overly dependent on the arbitrariness of the connection that a given model establishes between today's smile and its future dynamics. The local volatility model is a case in point.

▶ That vanilla option smiles do not constrain much cliquet prices is confirmed in Section 3.1.7. Model-independent lower and upper bounds are computed for the price of a forward-start call. The farther apart the lower and upper bounds are, the least vanilla-like the cliquet is.

▶ Forward-start options whose payoffs are of the form $S_{T_1} g\left(\frac{S_{T_2}}{S_{T_1}}\right)$ – which include the case of FVAs in the FX world – have a Vega in the Black-Scholes model that is linear in $S$. This calls for a new family of hedging instruments, European payoffs $S_T \ln S_T$ whose vegas are linear in $S$. The implied volatilities of these payoffs obey the convex order condition, thus allowing the definition of forward volatilities. As with payoffs $g\left(\frac{S_{T_2}}{S_{T_1}}\right)$, the carry P&L includes a gamma P&L on these particular forward volatilities.

❧ ❧ ❧ ❧ ❧

### 3.2 Forward-start options in the local volatility model

▶ For a local volatility function of the form $\sigma(t,S) = \overline{\sigma}(t) + \alpha(t) \ln \frac{S}{F_t}$, the dynamics of a forward VS volatility is given at order one in $\alpha$ by (3.33):

$$d\widehat{\sigma}_{T_1 T_2} = \left( \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \alpha(u)\, du \right) d \ln S_t$$

The instantaneous volatility of $d\widehat{\sigma}_{T_1 T_2}$ is set by the slope of the local volatility function for times $t \in [T_1, T_2]$. The volatility of a very short volatility $(T_2 - T_1 \ll T_1)$

is proportional to $\alpha(T_1)$: the model is not time-homogeneous. For typical equity smiles $\alpha(u)$ is decreasing, thus instantaneous volatilities of VS volatilities at future dates will be systematically smaller than those of spot-starting VS volatilities with the same maturity and will also depend on future spot levels. The local volatility model misprices volatility-of-volatility risk.

▶ The skew at a future date $T_1$ for residual maturity $\theta$ is given by expression (3.34):

$$\mathcal{S}_\theta\left(S_{T_1}, T_1\right) = \int_0^1 \alpha(T_1 + u\theta) u du$$

For typical equity smiles $\alpha(t)$ is decreasing, thus $\mathcal{S}_\theta\left(S_{T_1}, T_1\right) \ll \mathcal{S}_\theta\left(S_0, t=0\right)$: future skews in the local volatility model are weaker than spot-starting skews. The local volatility model misprices forward smile risk.

▶ Deriving the vega-hedge in the local volatility model of a forward-start call option with payoff $(\frac{S_{T_2}}{S_{T_1}} - k)^+$ yields a hedge portfolio consisting of a discrete quantity of options of maturity $T_1$ – which exactly matches the log-contract hedge studied in Section 3.1, a discrete portfolio of vanilla options of maturity $T_2$ and a continuous density of options with intermediate maturities.

In the discussion of Section 3.2.5 we argue why this is not a reasonable hedge. Hoping to immunize ourselves from the local volatility model's idiosyncrasies by hedging against perturbations of the local volatility function is not a viable route.

This page intentionally left blank

# Chapter 4

# Stochastic volatility – introduction

Chapter 2 was devoted to local volatility, a very constrained form of stochastic volatility: implied volatilities have $100\%$ correlation among themselves and with $S$, and their volatilities are determined by the market smile used for calibration of the model. Volatilities of implied volatilities – as generated by the model – will likely bear no resemblance to conservative levels based on historically observed volatilities of implied volatilities, or to implied volatilities of volatilities, whenever they are observable.

Can we design models that let us freely specify the dynamics of implied volatilities? We tackle this issue, first starting from a general point of view, then specializing to a modeling framework that can be used practically, based on forward variances of specific European payoffs.

## 4.1 Modeling vanilla option prices

It is natural to try and model the dynamics of the prices of vanilla options directly. Let $C_{KT}$ be the price of a call option with strike $K$, maturity $T$, and let us assume without loss of generality zero interest rates and repos. Because an option is an asset that can be bought or sold, its pricing drift is $r$ – which we have taken to be zero. The dynamics of $C_{KT}$, along with $S$ can be written as:

$$\begin{cases} dS &= \overline{\sigma} S dW^S \\ dC_{KT} &= \Lambda_{KT}\, dW^{KT} \end{cases} \tag{4.1}$$

with the initial conditions:

$$\begin{cases} S_{t=0} &= S_{t=0}^{\text{market}} \\ C_{KT}\,(t=0) &= C_{KT}^{\text{market}}\,(t=0) \end{cases}$$

and the terminal condition:

$$C_{KT}\,(t=T) \;=\; (S_T - K)^+ \tag{4.2}$$

where $W^S$, $W^{KT}$ are Brownian motions, $\Lambda_{KT}$ is the volatility of $C_{KT}$ and $\overline{\sigma}$ is the instantaneous volatility of $S$. We will omit the $t$ subscripts in the processes whenever possible.

### 4.1.1 Modeling implied volatilities

Let us try to work with implied volatilities $\widehat{\sigma}_{KT}$. In the pricing equation, $\widehat{\sigma}_{KT}$ acquires a drift that represents the cost of financing a delta position on $\widehat{\sigma}_{KT}$. Taking a directional position on $\widehat{\sigma}_{KT}$ entails trading the corresponding delta-hedged vanilla option. Expressing that the pricing drift of $C_{KT}$ is the interest rate – here zero – will then determine the pricing – or "risk-neutral" – drift of $\widehat{\sigma}_{KT}$. The joint SDEs of $S$ and $\widehat{\sigma}_{KT}$ read:

$$\begin{cases} dS &= \overline{\sigma} S dW^S \\ d\widehat{\sigma}_{KT} &= \mu_{KT} dt \,+\, \lambda_{KT}\, dW^{KT} \end{cases} \tag{4.3}$$

where $W^{KT}$ is a Brownian motion. $\widehat{\sigma}_{KT}$ is related to $C_{KT}$ through: $C_{KT} = P_{BS}(t, S, \widehat{\sigma}_{KT}, T)$. The terminal condition (4.2) then becomes:

$$\lim_{t \to T} \, (T-t)\, \widehat{\sigma}^2_{KT,\,t} \;=\; 0 \tag{4.4}$$

and the initial condition is simply:

$$\widehat{\sigma}_{KT,\,t=0} \;=\; \widehat{\sigma}^{\text{market}}_{KT,\,t=0}$$

The SDE for $C_{KT}$ reads:

$$\begin{aligned} dC_{KT} \;=\; & \left( \frac{dP_{BS}}{dt} + \frac{\overline{\sigma}^2}{2} S^2 \frac{d^2 P_{BS}}{dS^2} + \frac{\lambda^2_{KT}}{2} \frac{d^2 P_{BS}}{d\widehat{\sigma}^2_{KT}} + \rho \lambda_{KT} \overline{\sigma} S \frac{d^2 P_{BS}}{dS d\widehat{\sigma}_{KT}} \right. \\ & \left. + \, \frac{dP_{BS}}{d\widehat{\sigma}_{KT}} \mu_{KT} \right) dt \,+\, S \frac{dP_{BS}}{dS} \overline{\sigma} dW^S \,+\, \frac{dP_{BS}}{d\widehat{\sigma}_{KT}} \lambda_{KT}\, dW^{KT} \end{aligned}$$

where $\rho$ is the correlation between $W^S$ and $W^{KT}$. Expressing that the drift of $C^{KT}$ vanishes determines $\mu_{KT}$. The derivatives of $P_{BS}$ are all available analytically – the resulting expression for $\mu^{KT}$ is:

$$\mu_{KT} \;=\; \frac{1}{\widehat{\sigma}_{KT}} \left( \frac{\widehat{\sigma}_{KT} - \overline{\sigma}^2}{2\,(T-t)} - \frac{1}{2} d_1 d_2 \lambda^2_{KT} + \frac{d_2}{\sqrt{T-t}} \rho \overline{\sigma} \lambda_{KT} \right)$$

which we directly quote from [80], where $d_1$, $d_2$ are standard expressions appearing in Black-Scholes formulas:

$$d_1 \;=\; \frac{1}{\widehat{\sigma}_{KT}\sqrt{T-t}} \ln \frac{F_T(S_t)}{K} + \frac{\widehat{\sigma}_{KT}\sqrt{T-t}}{2}, \quad d_2 \;=\; d_1 - \widehat{\sigma}_{KT}\sqrt{T-t}$$

The joint dynamics of $(S, \widehat{\sigma}_{KT})$ in (4.3) would now be completely specified if the process $\overline{\sigma}$ was known. The processes for $\overline{\sigma}$ and $\widehat{\sigma}_{KT}$ cannot be chosen arbitrarily, for a solution to (4.3) to exist. For example it is possible to prove that, for short maturities, the at-the-money implied volatility should tend to the instantaneous volatility – see [43] for a proof:[1]

$$\lim_{T \to t} \widehat{\sigma}_{ST,t} \;=\; \overline{\sigma}_t \tag{4.5}$$

[1] Surprisingly, the proof of this simple and natural result is rather technical. Also note that the non-explosion condition imposed on $\mu_{KT}$ in [80] is not needed: as long as (4.4) holds, how fast or how slowly the left-hand side tends to zero does not matter.

One may think that once some technical conditions on processes $\overline{\sigma}$ and $\widehat{\sigma}_{KT}$ are satisfied, we should be able to explicitly construct – at least numerically – a solution to (4.3) that complies with (4.4). The trouble is that once we have a solution to (4.3) over $[0, T]$, the ensuing dynamics for $S$ determines prices of European options for maturities up to $T$, hence the implied volatilities $\widehat{\sigma}_{k\tau}$ of all vanilla options with $\tau < T$.

This is problematic, as our intention in modeling the $\widehat{\sigma}_{KT}$ was to be able to choose their initial values freely. What this means is that information about the initial smile has to be embedded in the process for $\overline{\sigma}$.[2] Moreover, such embedding is presumably convoluted: an arbitrary configuration of the $\widehat{\sigma}_{KT}$ is consistent with a dynamics of $S$ – hence a process $\overline{\sigma}$ may exist – only if the no-arbitrage conditions highlighted in Section 2.2.2 are satisfied. Whenever these conditions are violated, $\overline{\sigma}$ does not exist and this has to manifest itself in the structural impossibility of constructing a process for $\overline{\sigma}$.

The upshot is that direct modeling of implied volatilities of vanilla options is impractical.

---

## 4.2 Modeling the dynamics of the local volatility function

We try here a different line of approach: given a non-arbitrageable configuration of implied volatilities $\widehat{\sigma}_{KT}$ a local volatility function $\sigma(t, S)$ exists, given by the Dupire formula (2.3). There is a one-to-one mapping between the volatility surface and the local volatility function. We are not using a local volatility *model*; we are using the local volatility *function* to represent at any time $t$ the full set of implied volatilities.

We can then generate a dynamics for implied volatilities by generating a dynamics for this local volatility function which we denote by $\sigma_t$: $\sigma_t(\tau, \mathcal{S})$ is the local volatility function associated with the volatility surface at time $t$. For a fixed couple $(\tau, \mathcal{S})$, $\sigma_t(\tau, \mathcal{S})$ is a *process* that exists for $t \leq \tau$.

No-arbitrage restrictions on $\widehat{\sigma}_{KT}$ translate into the simple condition that $\sigma_t^2(\tau, \mathcal{S})$ be positive for all $\tau, \mathcal{S}$. In what follows we work with variances $\sigma^2$. In the local volatility model, the local volatility function is fixed and the instantaneous volatility $\overline{\sigma}_t$ of $S_t$ is given by:

$$\overline{\sigma}_t = \sigma_{t_0}(t, S_t)$$

where the $t_0$ subscript indicates that $\sigma_{t_0}(t, S_t)$ was obtained from implied volatilities observed at time $t_0$. In a model where the local volatility function itself is dynamic, $\overline{\sigma}$ is given by:

$$\overline{\sigma}_t = \sigma_t(t, S_t) \tag{4.6}$$

[2]This is precisely what occurs in the local volatility model, the simplest of all market models.

This follows from the definition of $\sigma_t$. From equation (2.6) the local volatility is given by:

$$\sigma_t^2(\tau, \mathcal{S}) = \frac{E_t[\sigma_\tau^2 \delta(S_\tau - \mathcal{S})]}{E_t[\delta(S_\tau - \mathcal{S})]} = E_t[\sigma_\tau^2 \mid S_\tau = \mathcal{S}]$$

Setting $\mathcal{S} = S_t$ and taking the limit $\tau \to t$ removes the conditionality in the expectation and yields (4.6). In contrast to the local volatility model, the dynamics of $S_t$ is generated by the short end of the volatility function $\sigma_t$ only. Local volatilities $\sigma_t(\tau, \mathcal{S})$ for $\tau > t$ do not appear explicitly in the SDE for $S_t$. Their role is to encode information on implied volatilities $\widehat{\sigma}_{KT}$: prices of vanilla options for maturities $\tau > t$ are derived from $\sigma_t(\tau, \mathcal{S})$ by solving the forward equation (2.7). In our notation:

$$\frac{dC_t^{\mathcal{S}\tau}}{d\tau} + (r-q)\,\mathcal{S}\frac{dC_t^{\mathcal{S}\tau}}{d\mathcal{S}} - \frac{\sigma_t^2(\tau, \mathcal{S})}{2}\mathcal{S}^2\frac{d^2C_t^{\mathcal{S}\tau}}{d\mathcal{S}^2} = -qC_t^{\mathcal{S}\tau}$$

with the initial condition $C_t^{\mathcal{S}t} = (S_t - \mathcal{S})^+$, where $C_t^{\mathcal{S}\tau}$ is the value at time $t$ of a call option of maturity $\tau$, strike $\mathcal{S}$.

We now use notation $\sigma_{\tau\mathcal{S}}$ for $\sigma_t(\tau, \mathcal{S})$ and omit the $t$-subscript in processes whenever possible. The pricing SDEs for $S$, $\sigma_{\tau\mathcal{S}}^2$ are given by:

$$\begin{cases} dS = (r-q)\,Sdt + \sigma_{tS} S dW^S \\ d\sigma_{\tau\mathcal{S}}^2 = \mu_{\tau\mathcal{S}} dt + \lambda_{\tau\mathcal{S}} dW^{\tau\mathcal{S}} \end{cases} \tag{4.7}$$

where $W^{\tau\mathcal{S}}$ is a Brownian motion. Just as in the previous section, we use zero interest rates and repos without loss of generality and determine $\mu_{\tau\mathcal{S}}$ by imposing that option prices have vanishing drift. $\sigma_{\tau\mathcal{S}}^2$ is given by

$$\sigma_{\tau\mathcal{S}}^2 = 2\left.\frac{\frac{dC_{KT}}{dT}}{K^2\frac{d^2C_{KT}}{dK^2}}\right|_{\substack{K=\mathcal{S}\\T=\tau}} \tag{4.8}$$

Both numerator and denominator are prices of linear combinations of vanilla options – a calendar spread for the numerator, a butterfly spread for the denominator – hence their drifts vanish. Remembering from equation (2.5), page 27, that

$$\frac{d^2C_{\mathcal{S}\tau}}{d\mathcal{S}^2} = e^{-r(\tau-t)}E_t[\delta(S_\tau - \mathcal{S})] = e^{-r(\tau-t)}\rho_{\tau\mathcal{S}}$$

where $\rho_{\tau\mathcal{S}}$ is the probability density that $S_\tau = \mathcal{S}$, yields:

$$\mathcal{S}^2\,\sigma_{\tau\mathcal{S}}^2\rho_{\tau\mathcal{S}} = 2e^{r(\tau-t)}\left.\frac{dC_{KT}}{dT}\right|_{\substack{K=\mathcal{S}\\T=\tau}} \tag{4.9}$$

Since the right-hand side of (4.9) has zero drift, so must the left-hand side: $\sigma_{\tau\mathcal{S}}^2\rho_{\tau\mathcal{S}}$ has vanishing drift.[3]

[3] In the left-hand side of (4.9) $\mathcal{S}$ is not a process – only $\sigma_{\tau\mathcal{S}}$ and $\rho_{\tau\mathcal{S}}$ are processes.

$\rho_{\tau\mathcal{S}}$ is a process: it is the (undiscounted) price of a sharp butterfly spread: this implies that its drift vanishes as well. $\rho_{\tau\mathcal{S}}$ is a *function* of $S$ and $t$ and a *functional* of the local volatility function:

$$\rho_{\tau\mathcal{S}} \equiv \rho_{\tau\mathcal{S}}\left(t, S,\ \sigma^2\right)$$

Its SDE then simply reads:

$$d\rho_{\tau\mathcal{S}} = \frac{d\rho_{\tau\mathcal{S}}}{dS}dS + \iint \frac{\delta\rho_{\tau\mathcal{S}}}{\delta\sigma^2_{ux}}d\sigma^2_{ux}$$

where the double integral stands for $\int_t^\tau du \int_0^\infty dx$.

Using now expression (4.7) for $d\sigma^2_{\tau\mathcal{S}}$ and $d\sigma^2_{ux}$, we get the drift of $\sigma^2_{\tau\mathcal{S}}\rho_{\tau\mathcal{S}}$:

$$\rho_{\tau\mathcal{S}}\mu_{\tau\mathcal{S}}dt + \lambda_{\tau\mathcal{S}}\left(S\frac{d\rho_{\tau\mathcal{S}}}{dS}\sigma_{tS}\left\langle dW^S dW^{\tau\mathcal{S}}\right\rangle + \iint \frac{\delta\rho_{\tau\mathcal{S}}}{\delta\sigma^2_{ux}}\lambda_{ux}\left\langle dW^{ux}dW^{\tau\mathcal{S}}\right\rangle\right)$$

Expressing that it vanishes yields:

$$\mu_{\tau\mathcal{S}} = -\frac{\lambda_{\tau\mathcal{S}}}{\rho_{\tau\mathcal{S}}}\left(S\frac{d\rho_{\tau\mathcal{S}}}{dS}\sigma_{tS}\frac{\left\langle dW^S dW^{\tau\mathcal{S}}\right\rangle}{dt} + \iint \frac{\delta\rho_{\tau\mathcal{S}}}{\delta\sigma^2_{ux}}\lambda_{ux}\frac{\left\langle dW^{ux}dW^{\tau\mathcal{S}}\right\rangle}{dt}\right) \tag{4.10}$$

We now derive a more explicit expression for $\frac{\delta\rho_{\tau\mathcal{S}}}{\delta\sigma^2_{ux}}$, which appears in the right-hand side of (4.10). $\rho_{\tau\mathcal{S}}\left(t, S,\ \sigma^2\right)$ is an expectation: $\rho_{\tau\mathcal{S}} = E_t[\delta\left(S_\tau - \mathcal{S}\right)]$. It solves the usual backward equation:

$$\frac{d\rho_{\tau\mathcal{S}}}{dt} + \frac{\sigma^2_{tS}}{2}S^2\frac{d^2\rho_{\tau\mathcal{S}}}{dS^2} = 0 \tag{4.11}$$

with the terminal condition $\rho\left(t = \tau\ S,\ \tau\mathcal{S}\right) = \delta\left(S - \mathcal{S}\right)$. By definition the functional derivative $\frac{\delta\rho_{\tau\mathcal{S}}}{\delta\sigma^2_{ux}}$ is such that, at order one in a perturbation $\delta\sigma^2$ of the local volatility function,

$$\delta\rho_{\tau\mathcal{S}} = \iint \frac{\delta\rho_{\tau\mathcal{S}}}{\delta\sigma^2_{ux}}\,\delta\sigma^2_{ux}$$

Consider a perturbation $\delta\sigma^2$ of $\sigma^2$. At first order in $\delta\sigma^2$, the perturbation $\delta\rho_{\tau\mathcal{S}}$ solves the following equation:

$$\frac{d\delta\rho_{\tau\mathcal{S}}}{dt} + \frac{\sigma^2_{tS}}{2}S^2\frac{d^2\delta\rho_{\tau\mathcal{S}}}{dS^2} = -\frac{\delta\sigma^2_{tS}}{2}S^2\frac{d^2\rho_{\tau\mathcal{S}}}{dS^2} \tag{4.12}$$

with the terminal condition at $t = \tau$: $\delta\rho_{\tau\mathcal{S}} = 0$. The solution to (4.12) is given by the Feynman-Kac theorem:

$$\delta\rho_{\tau\mathcal{S}} = E\left[\int_t^\tau \frac{\delta\sigma^2_{uS_u}}{2}S_u^2\left.\frac{d^2\rho_{\tau\mathcal{S}}}{dS^2}\right|_{S=S_u}du\right]$$

where the expectation is taken with respect to the dynamics generated by the local volatility function $\sigma_{ux}^2$. Let us rewrite this expectation using the probability density for $S$ at time $u$ – which is $\rho_{ux}$:

$$\delta\rho_{\mathcal{TS}} = \frac{1}{2}\int_t^{\mathcal{T}} du \int_0^{\infty} dx\ \rho_{ux} x^2\ \frac{d^2\rho_{\mathcal{TS}}\left(ux,\sigma^2\right)}{dx^2}\ \delta\sigma_{ux}^2$$

which yields:

$$\frac{\delta\rho_{\mathcal{TS}}}{\delta\sigma_{ux}^2} = \frac{1}{2}\rho_{ux}x^2\frac{d^2\rho_{\mathcal{TS}}\left(ux,\sigma^2\right)}{dx^2} \tag{4.13}$$

The sagacious reader will have noticed that we have already derived this result in Section 2.9, in the context of the vega hedge in the local volatility model. $\rho_{\mathcal{TS}}$ is the price of a European option of maturity $\mathcal{T}$: its sensitivity to $\sigma_{ux}^2$ is then given by equation (2.117) which involves the product of the probability density and the dollar gamma: (4.13) is identical to (2.117), page 82. We have our final expression for $\mu_{\mathcal{TS}}$:

$$\begin{aligned}\mu_{\mathcal{TS}} = -\frac{\lambda_{\mathcal{TS}}}{\rho_{\mathcal{TS}}}&\left(S\frac{d\rho_{\mathcal{TS}}\left(tS,\sigma^2\right)}{dS}\sigma_{tS}\frac{\left\langle dW^S dW^{\mathcal{TS}}\right\rangle}{dt}\right. \\ &\left.+\frac{1}{2}\int_t^{\mathcal{T}} du\int_0^{\infty} dx\ \rho_{ux}x^2\frac{d^2\rho_{\mathcal{TS}}\left(ux,\sigma^2\right)}{dx^2}\ \lambda_{ux}\frac{\left\langle dW^{ux} dW^{\mathcal{TS}}\right\rangle}{dt}\right)\end{aligned} \tag{4.14}$$

This expression for $\mu_{\mathcal{TS}}$ was first published by Iraj Kani and Emanuel Derman in [39] – albeit without the first piece in the right-hand side – and more recently rederived by René Carmona and Sergey Nadtochiy in [21], who also prove that the instantaneous volatility $\overline{\sigma}_t$ of $S_t$ is indeed $\sigma_t\left(t,S_t\right)$: this establishes the connection between the SDEs for $\sigma_{\mathcal{TS}}^2$ and $S$ in (4.7).

Drift $\mu_{\mathcal{TS}}$ in (4.14) is computationally expensive: not only is it non-local in the sense that it depends on the whole local volatility function, but it involves *forward* transition densities $\rho_{\mathcal{TS}}\left(ux,\sigma^2\right)$ for all $u$ in $[t,\mathcal{T}]$. It seems difficult to come up with explicit non-trivial solutions to (4.7) based on the direct modeling of local volatilities.

Inspection of expression (4.14) suggests however two simple solutions for which $\mu_{\mathcal{TS}}$ vanishes:

- $\lambda_{\mathcal{TS}} \equiv 0$: this implies $\mu_{\mathcal{TS}} \equiv 0$. The local volatility function is frozen: this recovers the local volatility model.

- $\left\langle dS dW^{\mathcal{TS}}\right\rangle \equiv 0$ and $\left\langle dW^{ux} dW^{\mathcal{TS}}\right\rangle \equiv 0$: local volatilities have zero correlation among themselves and with $S$. All points of the local volatility function have their own uncorrelated dynamics: such a model amounts to randomly drawing the instantaneous volatility of $S$ as time advances in such a way that the expectation of it square matches the square of the local volatility function calibrated on the initial smile: $E[\overline{\sigma}_t^2] = \sigma_{t_0}^2(t, S_t)$. While this appears to generate a non-trivial dynamics for implied volatilities it recovers in fact the local volatility model.[4]

Indeed, imagine sitting at time $t$ with a spot value $S$ and let us compute prices of a vanilla option of maturity $T$ using expression (2.30), derived in Section 2.4.1, where we use as base model the local volatility model with local volatility function $\sigma_{t_0}$:

$$P_{\overline{\sigma}}(t) \;=\; P_{\sigma_0}(t, S) + E_{\overline{\sigma}}\left[\int_t^T \frac{1}{2} e^{-ru} S_u^2 \frac{d^2 P_{\sigma_0}}{dS^2}\left(\overline{\sigma}_u^2 - \sigma_0(u, S_u)^2\right) du\right]$$

As, by construction, $E_{\overline{\sigma}}[\overline{\sigma}_u^2] = \sigma_0(u, S_u)^2$ we get $P_{\overline{\sigma}}(t) = P_{\sigma_0}(t, S)$. This can also be established by noting that, in our model, because each point of the local volatility function is driven by an independent random variable, $E[\overline{\sigma}_u^2 \mid S_u = S] = \sigma_0^2(u, S)$: the general Dupire equation (2.6) then shows that prices in both models are identical. Over any finite time interval the randomness of $\sigma_t$ averages out and the model behaves as though the local volatility function was fixed, equal to $\sigma_0(t, S)$.

Johannes Wissel proposes in [86] an approach that consists in modeling a discrete set of vanilla options prices: he introduces "local implied volatilities", whose relationship to vanilla option prices is more direct than that of local volatilities. The drifts of these local implied volatilities are non-local as well, except we now have a large but finite number of local implied volatilities. Still, an explicit non-trivial example of a model has not been available yet.

### 4.2.1 Conclusion

Again, we hit a snag in our attempt to model the dynamics of implied volatilities. Implied volatilities of vanilla options are unwieldy objects, as are their associated local volatilities: because they are related indirectly to option prices, their drifts are complex. In addition, we are handling the dynamics of a two-dimensional set of processes – reducing the dimensionality may be the price to pay to gain some tractability.

Observe that implied volatilities can be defined for any European payoff, as long as it is convex or concave. Are there particular payoffs whose implied volatilities are easier to handle?

---

[4] I thank Bruno Dupire for pointing this out to me.

## 4.3 Modeling implied volatilities of power payoffs

Consider payoff $S_T^p$, which we call a *power payoff*, following the terminology of Schweizer and Wissel in [81], and denote its price by $Q^{pT}$.[5]

In the absence of cash-amount dividends, which is the assumption we make throughout this section, in the Black-Scholes model with implied volatility $\widehat{\sigma}$, $Q^{pT}$ is given by:

$$Q^{pT} \;=\; e^{-r(T-t)}F_T^p e^{\frac{p(p-1)}{2}(T-t)\widehat{\sigma}^2} \tag{4.15}$$

where $F_T$ is the forward for maturity $T$.

Given the market price $Q^{pT}$ of a power payoff, its implied volatility $\widehat{\sigma}_{pT}$ is obtained by inverting (4.15). Power payoffs are concave for $p \in ]0,1[$ and convex otherwise: $\widehat{\sigma}_{pT}$ is well-defined as long as $p \neq 0$ and $p \neq 1$.

### 4.3.1 Implied volatilities of power payoffs

A power payoff is a European option: its market price can be calculated through replication on vanilla options, using expression (3.7), page 107. However, given a market smile, existence of arbitrary moments of $S_T$ is not guaranteed.

For $p \in [0,1]$ power payoffs have finite prices, as $S^p$ is bounded above by an affine function of $S$. For values of $p > 1$ or values of $p < 0$, market prices of power payoffs may not be finite.

In [67] Roger Lee relates the existence of moments of $S_T$ to the asymptotic behavior of implied volatilities for large and small strikes. Specifically, he shows that:

- When $K \to 0$ or $K \to \infty$, $T\widehat{\sigma}_{KT}^2$ grows at most linearly in $\ln K$, with a slope that cannot be higher than 2 or lower than $-2$.[6] Moreover, the asymptotic slope of $T\widehat{\sigma}_{KT}^2$ as a function of $\ln K$ is related to the largest/lowest index for which moments of $S_T$ are finite:

- Let $p_+ \;= \sup\{p\colon E[S_T^{1+p}] < \infty\}$. Then
$$\limsup_{K\to\infty} \frac{T\widehat{\sigma}_{KT}^2}{\ln K} \;=\; 2 - 4\left(\sqrt{p_+^2 + p_+} - p_+\right)$$

- Let $p_- \;= \sup\{p\colon E[S_T^{-p}] < \infty\}$. Then
$$\limsup_{K\to 0} \frac{T\widehat{\sigma}_{KT}^2}{\ln K} \;=\; -2 + 4\left(\sqrt{p_-^2 + p_-} - p_-\right)$$

[5] This section is based on joint work with Pierre Henry-Labordère.

[6] These bounds on the slope of the integrated variance $f$ as a function of log-moneyness can be derived by imposing that the denominator in the Dupire formula (2.19), page 33, is positive. Assuming that, asymptotically, $f = ay + b$, positivity of the denominator (a butterfly spread) requires that $|a| < 2$.

In practice, even for indexes, implied volatilities exist only in a limited range of strikes: the asymptotic behavior of the smile is then set by the extrapolation chosen by the trader.

Which power payoffs have finite prices then depends on non-observable implied volatility data. Typically $T\widehat{\sigma}_{KT}^2$ is parametrized as a function of $\ln\left(K/F_T\right)$ as the Dupire equation acquires a simple form (see expression (2.19), page 33), and no-arbitrage conditions are more easily handled. Choosing an affine extrapolation in these units amounts – through its slope – to deciding which power payoffs have finite prices.

**Calculating implied volatilities of power payoffs**

Implied volatilities $\widehat{\sigma}_{pT}$ are obtained from market prices $Q^{pT}$ by inverting (4.15). For $p \in ]0,1[$ prices of power payoffs are always finite; there exists a direct formula of $\widehat{\sigma}_{pT}$ as a weighted average of vanilla implied volatilities, which we now derive.

Undiscounted prices of power payoffs are related to the characteristic function $L(p)$ of $x = \ln\frac{S_T}{F_T}$:

$$e^{r(T-t)}\frac{Q^{pT}}{F_T^p} \;=\; E\left[\left(\frac{S_T}{F_T}\right)^p\right] \;=\; E\left[e^{px}\right] \;=\; L(p) \tag{4.16}$$

$L(p)$ is obtained from the market smile through:

$$L(p) \;=\; \int_0^\infty e^{r(T-t)}\frac{d^2C_{KT}}{dK^2}e^{p\ln\frac{K}{F_T}}dK$$

where we have used the fact that the probability density of $S_T$ is given by: $\rho(S_T) = e^{r(T-t)}\left.\frac{d^2C_{KT}}{dK^2}\right|_{K=S_T}$

Andrew Matytsin – see [73] – introduces a measure of moneyness $z$ defined by:

$$z(K) \;=\; \frac{\ln\left(\frac{F_T}{K}\right)}{\widehat{\sigma}_{KT}\sqrt{T}} - \frac{\widehat{\sigma}_{KT}\sqrt{T}}{2} \tag{4.17}$$

Replacing $C_{KT}$ with its expression as a function of $\widehat{\sigma}_{KT}$, he gives the following formula for $L(p)$ :

$$L(p) \;=\; \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}e^{-p\left(\frac{\omega^2}{2}+z\omega\right)}\left(1+p\frac{d\omega}{dz}\right)dz \tag{4.18}$$

where $\omega(z) \;=\; \widehat{\sigma}_{K(z)T}\sqrt{T}$.

Let us introduce a $p$-dependent measure of moneyness, $y$, defined by $y = z + p\omega(z)$:

$$y(K) \;=\; \frac{\ln\left(\frac{F_T}{K}\right)}{\widehat{\sigma}_{KT}\sqrt{T}} + \left(p-\frac{1}{2}\right)\widehat{\sigma}_{KT}\sqrt{T} \tag{4.19}$$

For $p=0$, $y(K)$ is the Black-Scholes $d_2$, while for $p=1$, it is $d_1$. Mapping $K \to y$ maps $[0,+\infty]$ into $[-\infty,+\infty]$ and, most importantly, is monotonic.[7] Thus, $K(y,p)$ is well-defined.

From (4.18), performing a change of variable from $z$ to $y$ yields:

$$L(p) = \int_{-\infty}^{+\infty} dy \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{\frac{p(p-1)}{2}\widehat{\sigma}^2_{K(y,p)T}}$$

where $\widehat{\sigma}_{K(y,p)T}$ is the implied volatility for "moneyness" $y$, that is for strike $K$ such that $y$ and $K$ are related through (4.19).

Using now (4.16) and (4.15) we get the following direct relationship between vanilla and power-payoff implied volatilities:

$$e^{\frac{p(p-1)}{2}\widehat{\sigma}^2_p T} = \int_{-\infty}^{+\infty} dy \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{\frac{p(p-1)}{2}\widehat{\sigma}^2_{K(y,p)T}} \tag{4.20}$$

which holds for all $p \in [0,1]$. $\widehat{\sigma}_{K(y,p)T}$ is easily obtained numerically.[8]

Taking the limits $p \to 0$ and $p \to 1$ yields the implied volatilities of two practically important payoffs.

**Implied volatility of the log contract**

Let us take the limit $p \to 0$. For small $p$, $S^p = 1 + p \ln S$, thus $\lim_{p\to 0} \widehat{\sigma}_p$ is the implied volatility of the log contract, which is replicated with a density of vanilla options proportional to $\frac{1}{K^2}$ – see Section 3.1.2. It has the property that its Black-Scholes dollar gamma and vega are independent of $S$. It is closely related to the variance swap, extensively studied in Chapter 5.

Expanding each side of (4.20) at order one in $p$ yields:

$$1 - \frac{p}{2}\widehat{\sigma}^2_0 T = \int_{-\infty}^{+\infty} dy \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \left(1 - \frac{p}{2}\widehat{\sigma}^2_{K(y,0)T} T\right)$$

which supplies the following formula for the log-contract implied volatility:

$$\widehat{\sigma}^2_{\ln S} = \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \widehat{\sigma}^2_{K(y)T} \tag{4.21a}$$

$$y(K) = \frac{\ln\left(\frac{F_T}{K}\right)}{\widehat{\sigma}_{KT}\sqrt{T}} - \frac{\widehat{\sigma}_{KT}\sqrt{T}}{2} \tag{4.21b}$$

---

[7] From Roger Lee's work we know that for $K \to 0$ and $K \to \infty$, $\widehat{\sigma}^2_{KT}T$ grows at most linearly in $\ln K$: $\widehat{\sigma}^2_{KT}T \propto a \ln K$ with $|a| < 2$. Using these two properties one easily shows that $\lim_{K\to 0} z(K) = +\infty$ and $\lim_{K\to\infty} z(K) = -\infty$.
In [46], Masaaki Fukasawa shows that the $K \to y$ mapping is monotonically decreasing for both $p=0$ and $p=1$. Because $y(K)$ is affine in $p$ this implies that $y(K)$ is monotonic for all $p \in [0,1]$ – I am indebted to Ling Ling Cao for this observation.

[8] Choose a set of strikes $K_i$. For each $K_i$, calculate the corresponding value $y_i$ of $y$ using (4.19) and record the couple $(y_i, \ \widehat{\sigma}_{K_iT})$. Then build an interpolation, for example a spline, of these couples to generate the function $\widehat{\sigma}_{K(y,p)T}$.

This expression was first published by Neil Chriss and William Morokoff in [32]. In comparison with formula (3.7), which expresses the price of the power payoff as a weighted integral of vanilla option prices, (4.21a) is less sensitive to numerical discretization of the integral.

For example, if $\widehat{\sigma}_{KT}$ is constant, equal to $\widehat{\sigma}_0$, a Gauss-Hermite quadrature yields $\widehat{\sigma}_{\ln S} = \widehat{\sigma}_0$ no matter how few points we use. In practice using about 10 points provides good accuracy.

**Implied volatility of the S ln S contract**

Now set $p = 1 - \varepsilon$ and take the limit $\varepsilon \to 0$. For small $\varepsilon$, $S^p = S - \varepsilon S \ln S$, thus $\lim_{p \to 1} \widehat{\sigma}_p$ is the implied volatility of the $S \ln S$ contract, which is replicated with a density of vanilla options proportional to $\frac{1}{K}$ – see Section 3.1.9.1.

This payoff has the property that its dollar gamma – hence its vega – is proportional to $S$. Proceeding as above, we get:

$$\widehat{\sigma}^2_{S \ln S} = \int_{-\infty}^{+\infty} dy \, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \widehat{\sigma}^2_{K(y)T} \tag{4.22a}$$

$$y(K) = \frac{\ln\left(\frac{F_T}{K}\right)}{\widehat{\sigma}_{KT}\sqrt{T}} + \frac{\widehat{\sigma}_{KT}\sqrt{T}}{2} \tag{4.22b}$$

### 4.3.2 Forward variances of power payoffs

Power payoffs are of the type $h\left(\frac{S_T}{F_T}\right)$ with $h$ a convex (resp. concave) function for $p < 0$ or $p > 1$ (resp. $p \in ]0, 1[$). They fall in the class considered in Section 2.2.2.2 for which (a) an implied volatility can be defined, (b) the convex order condition for implied volatilities (2.17) holds.

We then have:

$$(T_2 - t)\, \widehat{\sigma}^2_{pT_2} \geq (T_1 - t)\, \widehat{\sigma}^2_{pT_1} \tag{4.23}$$

This allows us to define *positive* forward variances $\xi$, either discrete, or continuous.

- Discrete forward variances $\xi^{pT_1T_2}$ are defined by:

$$\xi^{pT_1T_2} = \frac{(T_2 - t)\, \widehat{\sigma}^2_{pT_2} - (T_1 - t)\, \widehat{\sigma}^2_{pT_1}}{T_2 - T_1} \tag{4.24}$$

- Continuous forward variances $\xi^{pT}$ are defined by:

$$\xi^{pT} = \frac{d}{dT}\left((T - t)\, \widehat{\sigma}^2_{pT}\right) \tag{4.25}$$

The set of $\xi^{pT}$ for all $T$ is called the variance curve for index $p$; these are the state variables whose dynamics we will model.

We could have defined as well forward variances for implied volatilities of vanilla options, for a given moneyness, in Section 4.1.1, but their unwieldiness makes it a poitnless exercise.[9]

### 4.3.3 The dynamics of forward variances

In what follows, we work with continuous forward variances – discrete forward variances will reappear in Chapter 7.8. Again, we work with zero interest rates without loss of generality, and omit the $t$ subscripts for processes.

$$d\xi^{pT} \;=\; \lambda^{pT} dW^{pT} + \bullet\, dt$$

We determine the drift of $\xi^{pT}$ so that the drift of $Q^{pT}$ vanishes. The SDEs for $\xi^{pT}$ and $S$ are:

$$\begin{cases} dS \;=\; \overline{\sigma} S dW^S \\ d\xi^{pT} \;=\; \mu^{pT} dt \;+\; \lambda^{pT} dW^{pT} \end{cases} \tag{4.26}$$

with initial conditions $S_{t=0} = S_{t=0}^{\text{market}}$ and $\xi_{t=0}^{pT} = \xi_{t=0}^{pT\ \text{market}}$. With zero interest rates, repos, $Q^{pT}$ is given by:

$$Q^{pT} \;=\; S^p e^{\frac{p(p-1)}{2}\int_t^T \xi^{p\tau} d\tau} \tag{4.27}$$

In what follows we deal with a single variance curve, for a given index $p$ – we omit it in the notation. The SDE for $Q^T$ is given by:

$$\begin{aligned} \frac{dQ^T}{Q^T} \;=\; & \left[\frac{p\,(p-1)}{2}\left(\overline{\sigma}^2 - \xi^t\right) + \frac{p\,(p-1)}{2}\int_t^T \mu^\tau d\tau \right. \\ & \left. + \frac{p^2\,(p-1)}{2}\overline{\sigma}\int_t^T \lambda^\tau \rho^{S\tau} d\tau \;+\; \frac{p^2\,(p-1)^2}{8}\int_t^T\!\!\int_t^T du dv \rho^{uv}\lambda^u\lambda^v\right] dt \\ & + \left[p\overline{\sigma} dW^S + \frac{p\,(p-1)}{2}\int_t^T \lambda^\tau dW^\tau\, d\tau\right] \end{aligned} \tag{4.28}$$

where

$$\rho^{S\tau} \;=\; \frac{\left\langle dW^S dW^\tau\right\rangle}{dt}, \rho^{uv} \;=\; \frac{\left\langle dW^u dW^v\right\rangle}{dt}$$

[9]Strike $K$ cannot be kept constant as the maturity is varied. Provided $K$ scales linearly with the forward $F_T$: $K_T = \alpha F_T$, the convex order condition for option prices (2.9) holds and translates into the condition (2.15) on implied volatilities:

$$(T_2 - t)\,\widehat{\sigma}^2_{\alpha F_{T_2}, T_2} \;\geq\; (T_1 - t)\,\widehat{\sigma}^2_{\alpha F_{T_1}, T_1}$$

The counterpart of equation (4.25) for vanilla option implied volatilities would thus have read:

$$\xi_t^{\alpha, T} \;=\; \frac{d}{dT}\left((T-t)\,\widehat{\sigma}^2_{\alpha F_T, T}\right)$$

The first two lines in equation (4.28) are the drift of $Q^T$, which has to vanish for all $T$. Taking the derivative of this drift with respect to $T$ yields:

$$\frac{p\,(p-1)}{2}\mu^T + \frac{p^2\,(p-1)}{2}\overline{\sigma}\lambda^T\rho^{ST} + \frac{p^2\,(p-1)^2}{4}\int_t^T \rho^{Tu}\lambda^T\lambda^u du \;=\; 0$$

from which we get the expression of $\mu^T$, already given in [81]:

$$\mu^T \;=\; -\left(p\overline{\sigma}\lambda^T\rho^{ST} \;+\; \frac{p\,(p-1)}{2}\int_t^T \rho^{Tu}\lambda^T\lambda^u du\right) \tag{4.29}$$

Taking $T \to t$ in (4.28) leaves only the first term in the drift of $Q^T$: $\frac{p(p-1)}{2}\left(\overline{\sigma}^2 - \xi^t\right)$. We then get the extra condition:

$$\xi_t^t \;=\; \overline{\sigma}_t^2 \tag{4.30}$$

Thus, the short end of the variance curve is equal to the instantaneous variance of $S$, *for all values of* $p$.

Using now the expression of $\mu^T$ in (4.29) yields the final SDEs for $S$ and $\xi^T$ – reinstating the $p$ and $t$ indices:

$$dS_t \;=\; \sqrt{\xi_t^t}S_t dW_t^S \tag{4.31a}$$

$$d\xi_t^{pT} \;=\; -\left(p\sqrt{\xi_t^t}\lambda_t^T\rho^{ST} \;+\; \frac{p\,(p-1)}{2}\int_t^T \rho^{Tu}\lambda_t^T\lambda_t^u du\right)dt \;+\; \lambda_t^T dW_t^T \tag{4.31b}$$

### 4.3.4 Markov representation of the variance curve

Generally, the solution of (4.31b) requires evolving each forward variance $\xi^T$ individually in a Monte-Carlo simulation – this is impractical. It may be that, for a well chosen covariance structure of forward variances, each $\xi^T$ can be written as a function of a finite number of state variables, i.e. possesses a Markov representation. Then one only needs to evolve a finite set of state variables to generate the full variance curve at time $t$.

The question of building Markov representations has been especially addressed in the context of yield curve modeling. Formula (4.27) for $Q^{pT}$ is similar to the expression of a zero-coupon bond, up to the factor $S^p$, with $\xi^T$ playing the role of the forward rate for date $T$. It is then tempting, following the work of Oren Cheyette in [30] in the context of the HJM framework for the yield curve, to try and derive a class of solutions to (4.31b) that have a Markov representation.

We carry out the typical derivation one would go through in a yield curve context – for the sake of it, since we must warn the reader that the outcome is fruitless in the case of power payoffs.

While we do not carry the $p$ index, we will now carry $t$ subscripts. Let us assume, following [30], that the covariance structure of the $\xi_t^T$ is such that SDE (4.31b) for

$\xi^T$ has the following particular form:

$$dS_t \;=\; \sqrt{\xi_t^t}\, S_t dW_t^S \tag{4.32a}$$

$$d\xi_t^T \;= \mu_t^T dt \;+\; \sum_{i=0}^{n} \alpha_i\left(T\right)\beta_{it} dW_t^i \tag{4.32b}$$

where $W_i$ , $i = 0 \ldots n$ are $n$ correlated Brownian motions, $\beta_{it}$ are processes, and $\alpha_i\left(T\right)$ are functions of $T$. The correlations of $W_t^i$ and $W_t^j$ is $\rho_{ij}$ and the correlation of $W_t^i$ and $W_t^S$ is $\rho_{iS}$.

The volatility structure in the above equations is not inapt. For example, taking $\beta_{it} = e^{k_i t}$, a function, and $\alpha_i\left(T\right) = \alpha_i e^{-k_i T}$ yields:

$$d\xi_t^T \;= \mu^T dt \;+\; \sum_{i=0}^{n} \alpha_i e^{-k_i(T-t)} dW_t^i$$

The dynamics of the $\xi^T$ is time-homogeneous – volatilities and correlations of $\xi^T, \xi^{T'}$ are only a function of $T-t$ and $T-t, T'-t$, respectively – the "volatilities" of $\xi^T$ being expressed as a linear combination of exponentials.

Mirroring the derivation in [30], we integrate (4.32b) and try to express $\xi_t^T$ as a function of as few processes as possible. We have:

$$\begin{aligned}
\xi_t^T \;=&\; \xi_0^T - p\Sigma_i \int_0^t \alpha_i(T)\beta_{i\tau}\rho_{iS}\sqrt{\xi_\tau^\tau}d\tau \\
&- \frac{p\left(p-1\right)}{2}\int_0^t\left(\int_\tau^T \Sigma_{ij}\rho_{ij}\alpha_i(T)\beta_{i\tau}\alpha_j(u)\beta_{j\tau}du\right)d\tau + \Sigma_i\int_0^t \alpha_i(T)\beta_{i\tau}dW_\tau^i \\
=&\; \xi_0^T + \Sigma_i\alpha_i(T)\Bigg[-p\rho_{iS}\int_0^t \beta_{i\tau}\sqrt{\xi_\tau^\tau}d\tau \\
&\qquad -\frac{p\left(p-1\right)}{2}\Sigma_j\int_0^t \rho_{ij}\beta_{i\tau}\beta_{j\tau}\left(A_j(T)-A_j(\tau)\right)d\tau + \int_0^t \beta_{i\tau}dW_\tau^i\Bigg]
\end{aligned}$$

where we introduce $A_j(\tau) = \int_0^\tau \alpha_j(u)du$.

We now define process $B_{ij,t} = \rho_{ij}\int_0^t \beta_{it}\beta_{jt}$. A little manipulation yields:

$$\begin{aligned}
\xi_t^T \;=&\; \xi_0^T + \Sigma_i\alpha_i(T)\left[-\frac{p\left(p-1\right)}{2}\Sigma_j B_{ij,t}\left(A_j^T - A_j^t\right)\right. \\
&\left. + \int_0^t\left(-p\rho_{iS}\beta_{i\tau}\sqrt{\xi_\tau^\tau}d\tau - \frac{p\left(p-1\right)}{2}\Sigma_j B_{ij,t}\alpha_j(\tau)d\tau + \beta_{i\tau}dW_\tau^i\right)\right]
\end{aligned}$$

On top of processes $B_{ij,t}$ we thus need to define $n$ processes $x_{it}$:

$$\begin{aligned}
x_{i0} \;&=\; 0 \\
dx_{it} \;&=\; -p\rho_{iS}\beta_{it}\sqrt{\xi_t^t}dt - \frac{p\left(p-1\right)}{2}\Sigma_j B_{ij,t}\alpha_j(t)dt + \beta_{it}dW_t^i
\end{aligned}$$

The variance curve is given at time $t$ by:

$$\xi_t^T = \xi_0^T + \Sigma_i \alpha_i(T) \left[ -\frac{p\,(p-1)}{2} \Sigma_j B_{ij,t} \left( A_j^T - A_j^t \right) + x_{it} \right] \tag{4.33}$$

Setting $T = t$ gives the expression of the short end of the curve – the instantaneous variance of $S_t$:

$$\xi_t^t = \xi_0^t + \Sigma_i \alpha_i(t) x_{it} \tag{4.34}$$

Thus, in a Monte-Carlo simulation of our model, we only need to evolve processes $x_{it}$, $V_{ij,t}$ and $S_t$, according to the following SDEs:

$$\begin{aligned} dx_{it} &= -p\rho_{iS}\beta_{it}\sqrt{\xi_0^t + \Sigma_i\alpha_i(t)x_{it}}dt - \frac{p\,(p-1)}{2}\Sigma_j B_{ij,t}\alpha_j(t)dt \quad &(4.35a)\\ &\quad + \beta_{it} dW_t^i \\ dB_{ij,t} &= \rho_{ij}\beta_{it}\beta_{jt}\, dt &(4.35b)\\ dS_t &= \sqrt{\xi_0^t + \Sigma_i\alpha_i(t)x_{it}}\; S_t\, dW_t^S &(4.35c) \end{aligned}$$

The variance curve at time $t$ is given by (4.33). Inspection of equations (4.35) shows that processes $\beta_{it}$ can depend arbitrarily on the $x_{it}$ and $V_{ij,t}$, thus $\beta_{it}$ can, for example, be an arbitrary function of the variance curve $\xi_t$ at time $t$. This would generate the equivalent of a "local volatility" model for the $\xi_t^T$.

Unlike forward rates though, the instantaneous variance cannot be negative. Looking at the SDE for processes $x_i$ above it is not clear that it is possible to define processes $\beta_i$ and functions $\alpha_i$ that ensure that $\xi_t^t = \xi_0^t + \Sigma_i\alpha_i(t)x_{it} \geq 0$.[10]

The conclusion is that, unfortunately the ansatz (4.32b) used in [30] cannot be transposed to the framework of forward variances. This does not mean there are no low-dimensional Markov representations of the variance curve.

Indeed, any stochastic volatility model written on the instantaneous variance $V_t = \xi_t^t$ does provide a Markov representation of the variance curve for any value of $p$, though it may not be explicit. Think for example of the Heston model.

We refer the reader to Chapter 7 for examples of models with low-dimensional Markov representations for forward variances associated to $p \to 0$.

### 4.3.5 Dynamics for multiple variance curves

Given an initial variance curve $\xi_{t=0}$ for a particular value $p^*$, SDEs (4.31) generate the joint dynamics of $(S_t, \xi_t^{p^*})$. For a given market smile, we will generally have a

[10] Why not assume a lognormal dynamics for $\xi_t^T$: $d\xi_t^T = \xi_t^T \left(\mu_t^T dt + \Sigma_i\alpha_i\,(T)\,\beta_{it} dW_t^i\right)$ and look for a Markov representation of $\ln \xi_t^T$? We encourage the reader to try for herself; there does not seem to be a solution unless $p = 0$ or $p = 1$.

set of values of $p$ for which prices of power payoffs are finite, hence the $\xi_{t=0}^{pT}$ are well-defined. Is it possible to generate a joint dynamics for $S_t$ and a set of $\xi_t^p$?

A solution to SDEs (4.31) for a given $p^*$ provides the full dynamics of $S_t$, hence prices for power payoffs and a dynamics for variance curves $\xi_t^p$ with $p \neq p^*$. If our objective is to be able to independently set the initial values of variances curves for multiple values of $p$, this implies that information about initial curves $\xi_{t=0}^p$ has to be embedded in the SDE (4.31b) for $\xi_t^{p^*}$. It is not clear how this can be done practically.

The joint dynamics of multiple variance curves has to comply with condition (4.30) which expresses that the short ends of variance curves collapse one onto another at all times, as $\xi_t^{pt}$ is the instantaneous variance of $S_t$. The author does not know of an example of direct modeling of the joint dynamics for multiple curves that is able to calibrate to market prices of power payoffs. Obviously the local volatility model provides a solution – albeit not explicit and not very exciting as it is driven by a single Brownian motion.

Even though we are not able to handle the dynamics of multiple curves, we are able to explicitly construct the joint dynamics of $S_t$ and the variance curve, for a particular value $p^*$. We generate a dynamics for the full volatility surface by modeling the dynamics of one particular variance curve $\xi_t^{p^*}$. This produces a dynamics for other variance curves $\xi_t^p$ with $p \neq p^*$ that obeys (4.31b), except we cannot set their initial condition. We will generally have $\xi_{t=0}^p \neq \xi_{t=0}^{p\text{market}}$: our models will not provide exact calibration to the vanilla smile.[11]

### 4.3.6 The log contract, again

Consider the payoff $\frac{S^p-1}{p}$: it has the same implied volatility as the power payoff of index $p$. Taking the limit $p \to 0$:

$$\lim_{p\to 0} \frac{S^p - 1}{p} \;=\; \ln S$$

We have already made this observation in Section 4.3.1 and have derived an expression for the implied volatility of the log contract, in case there are no cash-amount dividends.

We first encountered the log contract in the discussion of cliquet hedges in Chapter 3. Using the same notation as in Section 3.1.4, we simply denote by $\widehat{\sigma}_T$ the implied volatility of the log contract of maturity $T$ and by $\zeta_t$ the associated variance curve:

$$\zeta_t^T \;\equiv\; \xi^{p=0,T} \;=\; \frac{d}{dT}\left((T-t)\,\widehat{\sigma}_T^2\,(t)\right)$$

Log contracts have finite prices: they do not require anything beyond the non-abitrageability of the smile, thus $\widehat{\sigma}_T$ is always well-defined. Power payoffs satisfy the

[11] At least we are able to calibrate exactly the term structure of implied volatilities $\widehat{\sigma}_{p^*T}$. In the case of implied volatilities of vanilla options, we were not even able to handle the dynamics of *one* $\widehat{\sigma}_{KT}$, let alone a term structure $\widehat{\sigma}_{K_T T}$.

convex order condition, thus $\zeta_t^T$ is positive. Taking the limit $p \to 0$ in (4.31) yields the following joint dynamics for $(S_t, \xi_t)$:

$$\begin{cases} dS_t &= \sqrt{\zeta_t^t} S_t dW_t^S \\ d\zeta_t^T &= \lambda_t^T dW_t^T \end{cases} \tag{4.36}$$

Thus, forward variances associated to log contracts have no drift.

In practice, log contracts themselves are not traded, as vanilla options are not traded over a sufficiently wide range of strikes to allow for exact replication. Moreover, our analysis does not carry over to the case of dividends with fixed cash amounts – which cannot be represented by a proportional yield $q$ – as the property that $E[S_{T_2}|S_{T_1}] = \frac{F_{T_2}}{F_{T_1}}$ needed to prove the convex order condition (4.23) no longer holds.

Luckily, closely related instruments known as variance swaps are traded; (4.36) will still hold, except the $\zeta_t^T$ will be replaced by variance swap forward variances – they are the basic building blocks in the models of Chapter 7.

Before we do this, we pause to study variance swaps in detail.

## Chapter's digest

▶ Hoping to construct a market model for vanilla options by modeling implied volatilities of vanilla options directly is a dead end.

▶ There is a one-to-one mapping between a non-arbitrageable vanilla smile and its corresponding local volatility function. Morover, no-arbitrage conditions simply translate in the requirement that local volatilities be real. It is then tempting to specify a dynamics for the local volatility function. It turns out, however, that the drift of local volatilities is non-local and involves forward transition densities, thus is computationally too expensive. Again, we reach an impasse – rather than trying to model the dynamics of implied volatilities of vanilla payoffs – or their associated local volatilities – are there other types of convex payoffs whose implied volatilities are less unwieldy objects?

▶ One good candidate is the family of power payoffs $\left(\frac{S_T}{F_T}\right)^p$. For each value of $p$ such that the market price of the corresponding power payoff is finite for all $T$, a term structure of forward variances $\xi^{pT}$ can be defined. Setting the volatilities of the $\xi^{pT}$ determines their drifts. For general values of $p$, it is not clear that there exist particular forms of the volatilities of the $\xi^{pT}$ that give rise to a Markov representation of the variance curve. For $p \to 0$, however, the drift of $\xi^{pT}$ vanishes. $\xi^{p=0T}$ are forward variances associated to the term structure of log contracts, which are closely related to variance swaps.

# *Chapter 5*

## *Variance swaps*

This chapter is devoted to variance swaps (VS) and their connection to delta-hedged log contracts – it is a prerequisite for the chapters that follow.

We show how a VS can be synthesized using European payoffs, characterize the impact of large returns on its replication and assess the relevance of pricing a VS in a jump-diffusion model. Finally we analyze the impact of cash-amount dividends on the VS replication, as well as the effect of interest-rate volatility.

We then study the replication of weighted variance swaps.

This is followed by two appendices – Appendix A on timer options and Appendix B on the perturbation of the lognormal density.

### 5.1 Variance swap forward variances

A variance swap (VS) contract pays at maturity the realized variance of a financial underlying, computed as the sum of the squares of daily log-returns. The market convention for the VS payoff is:

$$\frac{252}{N}\sum_{i=0}^{N-1}\ln^2\left(\frac{S_{i+1}}{S_i}\right)-\widehat{\sigma}^2_{\mathrm{VS},T}\left(t\right) \tag{5.1}$$

where $N$ is the number of trading days for the maturity of the variance swap, $S_i$ are the daily closing quotes of the underlying. $\widehat{\sigma}_{\mathrm{VS},T}\left(t\right)$ is set so that the initial value at time $t$ of the VS is zero and is called the VS volatility for maturity $T$.[1] 252 is the typical number of trading days in a year. Because the distribution of trading days is not uniform throughout the year, the ratio $\frac{N}{252}$ is generally not equal to the year fraction for the maturity of the VS and $\widehat{\sigma}_{\mathrm{VS}}$ is a biased estimator of realized volatility, especially for short maturities. We prefer to work with the following convention:

$$\frac{1}{T-t}\sum_{i=0}^{N-1}\ln^2\left(\frac{S_{i+1}}{S_i}\right)-\widehat{\sigma}^2_{\mathrm{VS},T}\left(t\right) \tag{5.2}$$

[1]VS term sheets include a prefactor $1/(2\widehat{\sigma}_{\mathrm{VS}})$ so that, for a small difference between realized volatility $\sigma_r$ and VS volatility $\widehat{\sigma}_{\mathrm{VS}}$, the payout of a VS contract is simply $\sigma_r-\widehat{\sigma}_{\mathrm{VS}}$.

where the $S_i$ are observed at dates $t_i$, such that $t_N = T$, which we rewrite for notational economy as:

$$\frac{1}{T-t}\sum_t^T \ln^2\left(\frac{S_{i+1}}{S_i}\right) - \widehat{\sigma}^2_{\mathrm{VS},T}(t) \tag{5.3}$$

Imagine taking at time $t$ a long position in $(T_2 - t)$ VSs of a maturity $T_2$ and a short position in $(T_1 - t)\, e^{-r(T_2 - T_1)}$ VSs of maturity $T_1$ with $T_2 > T_1$. The market implied volatilities of these two VSs are, at time $t$, $\widehat{\sigma}_{\mathrm{VS},T_2}(t)$ and $\widehat{\sigma}_{\mathrm{VS},T_1}(t)$. From (5.3), the payoff of this position, capitalized at $T_2$ is:

$$\begin{aligned}
&\sum_{T_1}^{T_2} \ln^2\left(\frac{S_{i+1}}{S_i}\right) - \left((T_2 - t)\,\widehat{\sigma}^2_{\mathrm{VS},T_2}(t) - (T_1 - t)\,\widehat{\sigma}^2_{\mathrm{VS},T_1}(t)\right) \\
&= \sum_{T_1}^{T_2} \ln^2\left(\frac{S_{i+1}}{S_i}\right) - (T_2 - T_1)\,\widehat{\sigma}^2_{\mathrm{VS},T_1T_2}(t) \qquad (5.4)
\end{aligned}$$

where we have introduced the discrete forward variance $\widehat{\sigma}_{\mathrm{VS},T_1T_2}$ defined as:

$$\widehat{\sigma}^2_{\mathrm{VS},T_1T_2}(t) = \frac{(T_2 - t)\,\widehat{\sigma}^2_{\mathrm{VS},T_2}(t) - (T_1 - t)\,\widehat{\sigma}^2_{\mathrm{VS},T_1}(t)}{T_2 - T_1}$$

$\widehat{\sigma}^2_{\mathrm{VS},T_1T_2}(t)$ is positive by construction, as the value at time $t$ of the second-hand side of (5.4) vanishes and its first piece is positive by construction. Imagine we unwind our position by entering at a later time $t' < T_1$ the reverse position: selling $(T_2 - t')$ VSs of maturity $T_2$ and buying $(T_1 - t')\, e^{-r(T_2 - T_1)}$ VSs of maturity $T_1$ at market implied volatilities prevailing at $t'$. This cancels the contribution from the realized variance over $[T_1, T_2]$ and the P&L of our strategy capitalized at time $T_2$ is:

$$(T_2 - T_1)\left(\widehat{\sigma}^2_{\mathrm{VS},T_1T_2}(t') - \widehat{\sigma}^2_{\mathrm{VS},T_1T_2}(t)\right) \tag{5.5}$$

It no longer involves the realized variance of $S$ and only depends on the variation of implied VS volatilities over $[t, t']$. (5.5) shows that we are able to generate a P&L that is linear in the variation of $\widehat{\sigma}^2_{\mathrm{VS},T_1T_2}$ over $[t, t']$ at no cost.

To produce a P&L that is linear in the variation of an equity underlying $S$, we borrow money to buy the underlying share and need to pay interest while we hold the share, hence the non-vanishing pricing drift of $S$. In the case of forward VS variances, no money is needed to materialize P&L (5.5): $\widehat{\sigma}^2_{\mathrm{VS},T_1T_2}$ has vanishing pricing drift.

$\widehat{\sigma}^2_{\mathrm{VS},T_1T_2}$ is a discrete forward variance. We can similarly define continuous VS forward variances, which we simply denote by $\xi_t^T$, given by:

$$\xi_t^T = \frac{d}{dT}\left((T - t)\,\widehat{\sigma}^2_{\mathrm{VS},T}(t)\right)$$

The $\xi^T$ are driftless as well; in a diffusive setting:

$$d\xi_t^T = \bullet\, dW_t^T \tag{5.6}$$

## 5.2 Relationship of variance swaps to log contracts

Consider the VS payoff (5.2). As $\widehat{\sigma}^2_{\mathrm{VS},T}$ is a constant, we will focus on the first piece in (5.2) and simply take for the VS payoff:

$$\sum_{i=0}^{N-1} \ln^2\left(\frac{S_{i+1}}{S_i}\right) \tag{5.7}$$

Let us assume for simplicity that dates $t_i$ are equally spaced by $\Delta t$. If $\ln\left(\frac{S_{i+1}}{S_i}\right)$ is small, at order two in $\frac{\delta S_i}{S_i}$ the payoff (5.7), discounted at $t=0$ can be rewritten as:

$$e^{-rT}\sum_{i=0}^{N-1}\left(\frac{\delta S_i}{S_i}\right)^2 = \sum_{i=0}^{N-1} e^{-rt_i}e^{-r(T-t_i)}\left(\frac{\delta S_i}{S_i}\right)^2 \tag{5.8}$$

where $\delta S_i = S_{i+1} - S_i$. We recognize the typical expression of the discounted sum of the gamma portion of the usual daily gamma/theta P&Ls (1.9) of a delta-hedged option risk-managed at zero volatility. The gamma/theta P&L reduces to its gamma part only if we choose a vanishing implied volatility for risk-managing the option.

Is it possible to find a European payoff whose dollar gamma – when risk-managed at vanishing volatility – matches that of the VS? This condition reads:

$$\frac{1}{2}S^2\frac{d^2P_{\widehat{\sigma}=0}}{dS^2} = e^{-r(T-t)} \tag{5.9}$$

The answer is yes – it is, up to a factor, the log contract. The value $Q^T(t,S)$ of payoff $-2\ln S_T$ in the Black-Scholes model with implied volatility $\widehat{\sigma}$, assuming there are no dividends with fixed cash amounts, is given by (3.8):

$$Q^T(t,S) = -2e^{-r(T-t)}\Big(\ln S + (r-q)(T-t) - \frac{\widehat{\sigma}^2}{2}(T-t)\Big) \tag{5.10}$$

Take $\widehat{\sigma}=0$ – the delta and gamma of $Q^T$ do not depend on $\widehat{\sigma}$ and we get:

$$\frac{dQ^T_{\widehat{\sigma}=0}}{dS} = -e^{-r(T-t)}\frac{2}{S}, \quad \frac{1}{2}S^2\frac{d^2Q^T_{\widehat{\sigma}=0}}{dS^2} = e^{-r(T-t)} \tag{5.11}$$

$Q^T_{\widehat{\sigma}=0}$ indeed fulfills condition (5.9) – we could have obtained (5.10) by straight integration of (5.9). Risk-managing the European payoff $-2\ln S_T$ with zero implied volatility exactly produces – at second order in $\delta S_i$ – payoff (5.7). How much should we charge for it – or, equivalently, what is $\widehat{\sigma}^2_{\mathrm{VS},T}$ so that the value at $t=0$ of the VS contract vanishes?

If the market price at $t=0$ of the log contract were $Q^T_{\widehat{\sigma}=0}$ we would not need to charge anything – we would set $\widehat{\sigma}_{\mathrm{VS},T} = 0$. In reality, the log contract has a market price $Q^T_{\text{market}}$: purchasing the log contract generates a mark-to-market P&L $-(Q^T_{\text{market}} - Q^T_{\widehat{\sigma}=0})$ for us, which we charge to the client as the premium of the variance swap. This premium is $T\widehat{\sigma}^2_{\text{VS},T}$ and is paid at maturity. $\widehat{\sigma}_{\text{VS},T}$ is then given by:

$$\widehat{\sigma}^2_{\text{VS},T} = \frac{e^{rT}}{T}\left(Q^T_{\text{market}} - Q^T_{\widehat{\sigma}=0}\right) \tag{5.12}$$

Given the market price $Q^T_{\text{market}}$ we can invert (5.10) to back out the log contract implied volatility $\widehat{\sigma}_T$. Substituting expression (5.10) for $Q^T_{\text{market}}$ in (5.12) then yields:

$$\widehat{\sigma}_{\text{VS},T} = \widehat{\sigma}_T \tag{5.13}$$

$$\xi^T_t = \zeta^T_t \tag{5.14}$$

The log contract is replicated with a vanilla portfolio using a density proportional equal to $\frac{2}{K^2}$ – see Section 3.1.3:

$$\begin{aligned} -2\ln S &= -2\ln S_0 - \frac{2}{S_0}(S - S_0) \\ &\quad + \int_0^{S_0} \frac{2}{K^2}(K-S)^+ dK + \int_{S_0}^{\infty} \frac{2}{K^2}(S-K)^+ dK \end{aligned} \tag{5.15}$$

At order two in $\frac{\delta S}{S}$ the payoff of a VS is then synthesized by delta-hedging this portfolio until maturity with zero implied volatility: $\widehat{\sigma}_T$ is simply computed as the implied volatility of the replicating vanilla portfolio and is model-independent.[2] As a consequence, forward variances of log contracts and VSs are identical objects. (5.12) can be rewritten as:

$$\widehat{\sigma}^2_{\text{VS},T} = \frac{e^{rT}}{T}\int_0^{\infty} \frac{2}{K^2}\left(P^{KT}_{\text{market}} - P^{KT}_{\widehat{\sigma}=0}\right) dK \tag{5.16}$$

where $P^{KT}$ is the price of a vanilla option of strike $K$, maturity $T$. $P^{KT}_{\widehat{\sigma}=0}$ , the price for a vanishing volatility is simply the intrinsic value computed for the forward and discounted to $t = 0$; for a call option:

$$P^{KT}_{\widehat{\sigma}=0} = e^{-rT}\left(Se^{(r-q)T} - K\right)^+$$

$(P^{KT}_{\text{market}} - P^{KT}_{\widehat{\sigma}=0})$ is identical for a call or a put struck at $K$ because of call-put parity – no need to distinguish between both types of vanilla options.[3]

While $\widehat{\sigma}_T$ is well-defined whenever the market smile is non-arbitrageable, it is very sensitive to the extrapolation chosen for implied volatilities outside the range

[2]It is this property that prompted banks to start offering variance swaps in the nineties: at the time, variance swaps were exotic instruments that trading desks hedged with vanilla options. Since then, on indexes, they have become emancipated from their vanilla replication and exist as independent instruments.

[3](5.12) and (5.16) hold as long as there are no dividends with fixed cash amounts; in the general case, expression (5.47) applies.

of strikes traded on the market. In practice, for very liquid securities such as indexes, market-makers do the reverse and infer implied volatilities for low strikes from market quotes of VSs.[4]

The delta and gamma of the log contract in (5.11) do not depend on the implied volatility $\widehat{\sigma}$: we could as well have chosen to risk-manage the log contract with a non-zero implied volatility. The most natural choice is $\widehat{\sigma} = \widehat{\sigma}_T$: delta-hedging the log contract then generates a gamma P&L *and* a theta P&L that exactly match both pieces in (5.2).

The idea of using zero implied volatility proves useful when analyzing more complex payoffs involving realized variance weighted by a function of spot value, such as conditional variance swaps, for which squared daily returns are accumulated only when $S_i$ lies within an interval, typically $[0, L]$, $[L, H]$ or $[H, \infty]$ – and also in the case of fixed amount dividends.

Weighted variance swaps – and in particular conditional VSs – are dealt with in Section 5.9, page 176.

### 5.2.1 A simple formula for $\widehat{\sigma}_{\text{VS},T}$

In (5.16) $\widehat{\sigma}_{\text{VS},T}$ is expressed in terms of market prices of vanilla options. (5.16) holds whenever cash-amount dividends are not present. Otherwise it is replaced with (5.47) – see Section 5.6.2 further below.

In the absence of cash-amount dividends, $\widehat{\sigma}_{\text{VS},T}$ can equivalently be computed as the implied volatility of the log contract – or of a set of European payoffs otherwise; see the derivation in Section 5.3.1 below. This is a more efficient method than using (5.16) as it is less sensitive to the discretization of the replicating portfolio in (5.15) – in particular, for a flat volatility surface, we trivially recover the exact value of $\widehat{\sigma}_{\text{VS},T}$.

Still, with no cash-amount dividends present, there is a direct expression of $\widehat{\sigma}_{\text{VS},T}$ as a weighted average of implied volatilities of vanilla options. This was obtained in the context of power payoffs, in Section 4.3 of Chapter 4.

We simply quote result (4.21) and refer the reader to page 142 for its derivation:

$$\widehat{\sigma}^2_{\text{VS},T} = \int_{-\infty}^{+\infty} dy\, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \widehat{\sigma}^2_{K(y)T} \tag{5.17a}$$

$$y(K) = \frac{\ln\left(\frac{F_T}{K}\right)}{\widehat{\sigma}_{KT}\sqrt{T}} - \frac{\widehat{\sigma}_{KT}\sqrt{T}}{2} \tag{5.17b}$$

[4]For long-dated variance swaps, the identity (5.13) has to be adjusted to take interest-rate volatility into account. Black-Scholes implied volatilities of European payoffs are in fact implied volatilities of the *forward* for the option's maturity. In contrast, a variance swap pays the realized variance of the *spot*. Interest-rate volatility introduces a difference between the realized variances of the spot and the forward, which is material for long-dated variance swaps. See Section 5.8 below for an estimation of this effect.

## 5.3 Impact of large returns

The key property that synthesizing the VS payoff boils down to delta-hedging a log contract – hence that $\widehat{\sigma}_{\mathrm{VS},T} = \widehat{\sigma}_T$ – has been derived in an expansion at order 2 in $\frac{\delta S}{S}$ of the VS payoff and the P&L of a delta-hedged log contract. How robust is it? What if returns are large?

We now consider two canonical examples of dynamics for $S_t$: a diffusive model and a jump-diffusion model. Unlike the former, the latter is able to generate large returns, even at short time scales, with a probability proportional to $\Delta t$. We consider the limit of very frequent observations of $S$ which enables us to explicitly compute all quantities of interest.

The case of real underliers is investigated next.

### 5.3.1 In diffusive models

The price of the log contract of maturity $T$, risk-managed at zero implied volatility, $P^T_{\widehat{\sigma}=0}$, satisfies the following condition:

$$\frac{1}{2} S^2 \frac{d^2 P^T_{\widehat{\sigma}=0}}{dS^2} \;=\; e^{-r(T-t)} \tag{5.18}$$

Assume that $S$ follows a diffusive dynamics:

$$dS_t \;=\; (r-q) S_t dt + \overline{\sigma}_t S_t dW_t \tag{5.19}$$

where instantaneous volatility $\overline{\sigma}$ is an arbitrary process.

Remember expression (2.30) in Section 2.4.1 relating the price of a payoff in an arbitrary *diffusive* model $P_{\overline{\sigma}}$ with instantaneous volatility $\overline{\sigma}$ to the price $P_\sigma$ in a local volatility model whose local volatility function is $\sigma\left(t,S\right)$:

$$P_{\overline{\sigma}}\left(t=0\right) \;=\; P_\sigma\left(0,S_0\right) \;+\; E_{\overline{\sigma}}\left[\int_0^T \frac{1}{2} e^{-rt} S_t^2 \frac{d^2 P_\sigma}{dS^2}\left(\overline{\sigma}_t^2 - \sigma(t,S_t)^2\right) dt\right] \tag{5.20}$$

We now set $\sigma\left(t,S\right) \equiv 0$ and use the notation: $P^T_{\overline{\sigma}}$ for $P_{\overline{\sigma}}$ and $P^T_{\widehat{\sigma}=0}$ for $P_\sigma$. Using identity (5.18) in (5.20) yields:

$$P^T_{\overline{\sigma}} \;=\; P^T_{\widehat{\sigma}=0} + e^{-rT} E_{\overline{\sigma}}\Big[\int_0^T \overline{\sigma}_t^2 dt\Big] \tag{5.21}$$

where all prices are evaluated at $t=0$.

We now turn to the variance swap. We have:

$$d\ln S_t \;=\; (r-q-\frac{1}{2}\overline{\sigma}_t^2)dt \;+\; \overline{\sigma}_t dW_t$$

where the first term in the right-hand side is of order $dt$ and the second term is of order $\sqrt{dt}$. Square this expression and take the limit $dt \to 0$. The only contribution at order $dt$ comes from the square of the second term and we get:

$$\lim_{dt\to 0} \frac{1}{dt} \left(d\ln S_t\right)^2 = \overline{\sigma}_t^2$$

which implies that:

$$\lim_{\Delta t\to 0} \sum_{i=0}^{N-1} \ln^2\left(\frac{S_{i+1}}{S_i}\right) = \int_0^T \overline{\sigma}_t^2 dt$$

Thus, in the limit of frequent observations, the VS implied volatility $\widehat{\sigma}_{\mathrm{VS},T}$, defined by (5.2), is given by:

$$\widehat{\sigma}_{\mathrm{VS},T}^2 = \frac{1}{T} E\Big[\lim_{\Delta t\to 0} \sum_{i=0}^{N-1} \ln^2\left(\frac{S_{i+1}}{S_i}\right)\Big] = E\Big[\frac{1}{T}\int_0^T \overline{\sigma}_t^2 dt\Big] \tag{5.22}$$

How frequent is frequent? The log-return over $[t_i,\ t_{i+1}]$ is given by:

$$\ln\left(\frac{S_{i+1}}{S_i}\right) = \int_{t_i}^{t_{i+1}} \overline{\sigma}_t dW_t + \int_{t_i}^{t_{i+1}} \left(r - q - \frac{\overline{\sigma}_t^2}{2}\right) dt$$

The orders of magnitude of the two pieces in the right-hand side are, respectively, $\overline{\sigma}\sqrt{\Delta t}$ and $\overline{\sigma}^2\Delta t$, with $\Delta t = t_{i+1} - t_i$. The second piece can be safely ignored whenever $\overline{\sigma}^2\Delta t \ll \overline{\sigma}\sqrt{\Delta t}$, that is $\overline{\sigma}\sqrt{\Delta t} \ll 1$, which is typically the case for volatility levels of equity underlyings.

Assume now that our diffusive model is calibrated to the market smile: $P^T_{\overline{\sigma}} = P^T_{\mathrm{Market}}$. From (5.21) and (5.22):

$$\widehat{\sigma}_{\mathrm{VS},T}^2 = E_{\overline{\sigma}}\Big[\frac{1}{T}\int_0^T \overline{\sigma}_t^2 dt\Big] = \frac{e^{rT}}{T}\left(P^T_{\mathrm{Market}} - P^T_{\widehat{\sigma}=0}\right) \tag{5.23}$$

This identity implies that any diffusive model calibrated to the vanilla smile prices VSs identically. Denote by $\widehat{\sigma}_T$ the implied volatility of the log contract; $\widehat{\sigma}_T$ is such that $P^T_{\mathrm{Market}} = P^T_{\widehat{\sigma}_T}$.

Using (5.20) again, still setting $\sigma(t,S) \equiv 0$, but this time choosing for $\overline{\sigma}_t$ the constant volatility $\widehat{\sigma}_T$ yields:

$$P^T_{\mathrm{Market}} = P^T_{\widehat{\sigma}} = P^T_{\widehat{\sigma}=0} + e^{-rT} T \widehat{\sigma}_T^2$$

Inserting this expression of $P^T_{\mathrm{Market}}$ in (5.23) yields our final result:

$$\widehat{\sigma}_{\mathrm{VS},T} = \widehat{\sigma}_T$$

$\widehat{\sigma}_T$ and $\widehat{\sigma}_{\mathrm{VS},T}$ are identical and are simply related to the quadratic variation of $\ln S_t$ over $[0, T]$. VS forward variances $\xi^T$ are then identical to log contract forward variances $\zeta^T$ and the instantaneous variance of $S_t$ is the short end of the variance curve: $\overline{\sigma}_t = \sqrt{\xi_t^t}$. The joint dynamics of $S_t$ and $\xi_t$ is given by equations (4.36):

$$\begin{cases} dS_t &= \sqrt{\xi_t^t} S_t dW_t^S \\ d\xi_t^T &= \lambda_t^T dW_t^T \end{cases}$$

**The case of cash-amount dividends**

When there are cash-amount dividends the log contract no longer has constant dollar gamma. However, we show in Section 5.6.2 below that the log contract can be supplemented with European payoffs of intermediate maturities to generate a portfolio that, risk-managed at zero implied volatility, again satisfies condition (5.18).

In conclusion, in a diffusive setting:

- The VS volatility is the implied volatility of a portfolio of European options. This portfolio reduces to the log contract when there are no cash-amount dividends
- Any diffusive model calibrated to the vanilla smile yields the same value for $\widehat{\sigma}_{\mathrm{VS},T}$

### 5.3.2 In jump-diffusion models

Assume that, in addition to the diffusion in (5.19), $S_t$ is allowed to abruptly jump at times generated by a Poisson process with constant intensity $\lambda$ – we take zero interest rate and repo without loss of generality:

$$dS_t \;=\; \overline{\sigma}_t S_t dW_t^S + S_{t^-}\left(J dN_t - \lambda \overline{J} dt\right) \tag{5.24}$$

where relative magnitudes $J$ of successive jumps are iid random variables, $\overline{J} = E[J]$ and $N_t$ is the counting process of the underlying Poisson process. $J$ and $N_t$ are assumed to be independent and $E[dN_t] = \lambda dt$. The drift in (5.24) is the compensator of the jump process; it ensures that the financing cost of $S$ vanishes: $E[dS_t] = 0$.[5] We have for $S_T$:

$$S_T \;=\; S_0 e^{-\frac{1}{2}\int_0^T \overline{\sigma}_t^2 dt + \int_0^T \overline{\sigma}_t dW_t^S} e^{-\lambda \overline{J} T} \prod_{i=1}^{N_T} (1 + J_i)$$

where $N_T$ is the (random) number of jumps occurring over $[0, T]$, whose probability distribution is $p_n = p\left(N_T = n\right) = e^{-\lambda T} \frac{(\lambda T)^n}{n!}$. In particular, $E[N_T] = \lambda T$. We get

[5] See Appendix A of Chapter 10, page 407, for the interpretation of the pricing equation in jump-diffusion models.

for the price of the log contract:

$$\begin{aligned}
& E[-2\ln S_T] \\
&= -2\ln S_0 + 2\Big(\lambda\overline{J}T + \frac{1}{2}E\Big[\int_0^T \overline{\sigma}_t^2 dt\Big]\Big) - 2\Sigma_{n=0}^{\infty}\, p_n E\big[\ln\left((1+J)^n\right)\big] \\
&= -2\ln S_0 + 2\Big(\lambda\overline{J}T + \frac{1}{2}E\Big[\int_0^T \overline{\sigma}_t^2 dt\Big]\Big) - 2\lambda T\overline{\ln(1+J)}
\end{aligned}$$

where we have used the fact that the $J_i$ are iid and independent of $N_t$ and $\Sigma_{n=0}^{\infty}\, np_n = \lambda T$ . Inverting (5.10) we get the implied volatility of the log contract:

$$\widehat{\sigma}_T^2 \;=\; E\Big[\frac{1}{T}\int_0^T \overline{\sigma}_t^2 dt\Big] \;-2\lambda\,\overline{\ln(1+J) - J} \tag{5.25}$$

Let us now turn to the VS payoff (5.7). In the limit of frequent observations:

$$\widehat{\sigma}_{\mathrm{VS},T}^2 \;=\; \frac{1}{T}E\Big[\lim_{\Delta\to 0}\sum_{i=0}^{N-1}\ln^2\Big(\frac{S_{i+1}}{S_i}\Big)\Big] \;=\; \frac{1}{T}\int_0^T E\Big[\frac{(d\ln S_t)^2}{dt}\Big]dt$$

We have:

$$d\ln S_t \;=\; -\Big(\frac{\overline{\sigma}_t^2}{2} + \lambda\overline{J}\Big)dt + \overline{\sigma}_t dW_t + \ln(1+J)\,dN_t$$

Squaring this expression and taking the expectation yields:

$$\frac{E[(d\ln S_t)^2]}{dt} \;=\; E[\overline{\sigma}_t^2] \;+\; \lambda\,\overline{\ln^2(1+J)} \tag{5.26}$$

which gives, in the limit of frequent observations:

$$\widehat{\sigma}_{\mathrm{VS},T}^2 \;=\; E\Big[\frac{1}{T}\int_0^T \overline{\sigma}_t^2 dt\Big] \;+\; \lambda\,\overline{\ln^2(1+J)} \tag{5.27}$$

In the right-hand side of (5.26) the contribution of jumps to $E[(d\ln S_t)^2]/dt$ is weighted by $\lambda$, as for small $dt$ there can be at most one jump, with probability $\lambda dt$. If spot observations are not sufficiently frequent, more than one jump can occur during the interval $\Delta t$, with the result that the contribution of jumps is no longer linear in $\lambda$: in addition to $\overline{\sigma}\sqrt{\Delta t} \ll 1$, we also need $\lambda\Delta t \ll 1$.

Compare (5.25) and (5.27). $\widehat{\sigma}_T$ and $\widehat{\sigma}_{VS,T}$ are not equal anymore and their difference is given by:[6]

$$\widehat{\sigma}_{\mathrm{VS},T}^2 - \widehat{\sigma}_T^2 \;=\; \lambda\,\overline{\ln^2(1+J) + 2\ln(1+J) - 2J} \tag{5.28}$$

[6]The difference between $\widehat{\sigma}_{\mathrm{VS},T}^2$ and $\widehat{\sigma}_T^2$ does not depend on $T$ because in our example the jump process has constant intensity.

remember that we are using $\xi^T$ to denote VS forward variances and $\zeta^T$ to denote log-contract forward variances. (5.28) implies that $\xi_t$, $\zeta_t$ are related though:

$$\xi_t^T - \zeta_t^T \;=\; \lambda\,\overline{\ln^2(1+J) + 2\ln(1+J) - 2J}$$

$\zeta_t$, $\xi_t$ are both driftless and the joint dynamics of $S_t$, $\zeta_t$, $\xi_t$, is given by:

$$\begin{cases} dS_t &= \overline{\sigma}_t S_t dW_t^S + S_{t^-}\left(JdN_t - \lambda\overline{J}dt\right) \\ d\zeta_t^T &= \lambda_t^T dW_t^T \\ d\xi_t^T &= \lambda_t^T dW_t^T \end{cases}$$

where $\overline{\sigma}_t$ equals neither $\sqrt{\xi_t^t}$ nor $\sqrt{\zeta_t^t}$. We can use (5.27) to express $\overline{\sigma}_t$ as a function of $\xi_t^t$:

$$\overline{\sigma}_t \;=\; \sqrt{\xi_t^t - \lambda\,\overline{\ln^2(1+J)}}$$

The fact that $\overline{\sigma}_t$, $\sqrt{\xi_t^t}$ and $\sqrt{\zeta_t^t}$ are all different is typical of models for $S_t$ that are not pure diffusions. We now estimate the order of magnitude of this difference and how it is related to the market smile used for calibration.

### 5.3.3 Difference of VS and log-contract implied volatilities

In the discussion that follows, we consider "pure" jump-diffusion/Lévy models, that is models with no dynamical variables besides $S_t$. This excludes mixtures of stochastic volatility and jump/Lévy processes, for example models where $\overline{\sigma}_t$ is a process correlated with $S_t$.

We will henceforth use "jump-diffusion" to denote jump/Lévy processes, that is processes with independent stationary increments for $\ln S_t$. In the context of the preceding section, this amounts to setting $\overline{\sigma}_t$ constant.

The dynamics in (5.24) serves as a basic prototype for the class of Lévy processes, but our conclusions have general relevance.

Let us expand the right-hand side of (5.28) in powers of $J$. The first non-vanishing contribution comes from $J^3$ and we get:

$$\widehat{\sigma}^2_{\mathrm{VS},T} - \widehat{\sigma}^2_T \;\simeq\; -\frac{1}{3}\lambda\overline{J^3} \tag{5.29}$$

The fact that the first non-vanishing contribution is of order 3 in $J$ is expected: at order 2 in $J$ the effect of jumps is identical to that of a simple diffusion – see Appendix A of Chapter 10: whether the return was generated by a jump or by Brownian motion is immaterial.

Higher-order terms – $\lambda\overline{J^4}$, $\lambda\overline{J^5}$, etc. – contribute as well, but the order-three term is the leading contribution in the limit of small and frequent jumps. Indeed, when taking the limit $J \to 0$, $\lambda$ should be increased so that the contribution of jumps to the quadratic variation of $\ln S$ – and the quadratic variation itself – stays fixed. Equation (5.26) shows that as $J \to 0$, for $\lambda\overline{J^2}$ to stay constant, $\lambda$ has to scale like $\frac{1}{J^2}$. Terms $\lambda\overline{J^n}$ are then of order $J^{n-2}$: the largest contribution is generated by $n = 3$.

**Calibrating jump parameters to the vanilla smile**

In jump-diffusion models, jumps not only introduce a difference between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\text{VS},T}$; they also generate a smile. In Appendix A of Chapter 10 – see equation (10.26), page 413 – it is shown that in the limit $J \to 0$ the ATMF skew for maturity $T$, $\mathcal{S}_T$, is given, at order one in the skewness of $\ln S_T$, by:

$$\mathcal{S}_T \simeq \frac{\lambda\overline{J^3}}{6\widehat{\sigma}_T^3 T} \tag{5.30}$$

which, using (5.29), gives the following approximation for the difference of $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\text{VS},T}$ as a function of the ATMF skew for maturity $T$:

$$\widehat{\sigma}_{\text{VS},T}^2 - \widehat{\sigma}_T^2 \simeq -2\widehat{\sigma}_T^3 \mathcal{S}_T T \tag{5.31}$$

Assuming the right-hand side is small:

$$\widehat{\sigma}_{\text{VS},T} \simeq \widehat{\sigma}_T \left(1 - \widehat{\sigma}_T \mathcal{S}_T T\right) \tag{5.32}$$

Thus, given a market smile, assuming a diffusion for $S$ leads to:

$$\widehat{\sigma}_{\text{VS},T} = \widehat{\sigma}_T$$

while assuming a jump-diffusion process leads to (5.28), which for weak smiles is approximated by (5.32):

$$\widehat{\sigma}_{\text{VS},T} \simeq \widehat{\sigma}_T \left(1 - \widehat{\sigma}_T \mathcal{S}_T T\right)$$

Consider the case of a one-year maturity VS on an equity index. Typically, the difference of the implied volatilities for strikes $95\%$ and $105\%$ is of the order of 2 points of volatility, which gives $\mathcal{S}_T = -0.02/\ln(105/95) = -0.2$. Taking $\widehat{\sigma}_T = 20\%$, yields $\widehat{\sigma}_T \mathcal{S}_T T = -4\%$. Equation (5.32) then yields a difference of about one point of volatility between $\widehat{\sigma}_{\text{VS},T}$ and $\widehat{\sigma}_T$.

The size of this correction has prompted some to argue that replicating variance swaps with log contracts is inadequate and that variance swaps should be priced with jump/Lévy models calibrated on the market smile. Is this reasonable?

As we now show, the difference between $\widehat{\sigma}_{\text{VS},T}$ and $\widehat{\sigma}_T$ is due to the non-vanishing skewness of returns *at short time scales*. In the case of a jump-diffusion model, this skewness is inferred from the market smile of maturity $T$. Is this model-mediated relationship between *skewness* of short returns and *skew* of the $T$-maturity smile robust?

### 5.3.4 Impact of the skewness of daily returns – model-free

We now look at things in model-free fashion, with zero interest and repo, for simplicity.

Let $r_i$ be the log-return of $S$ over $[t_i, t_{i+1}]$: $r_i = \ln(\frac{S_{i+1}}{S_i})$, and imagine that we have sold a variance swap and are long a delta-hedged log contract. We assume that we are keeping this static position until $t = T$, risk-managing the VS at a fixed implied volatility $\widehat{\sigma}_{\text{VS},T}$ and the log contract at a fixed implied volatility $\widehat{\sigma}_T$.

Our total P&L during $\Delta t$ is the difference between the P&L of the delta-hedged log contract and the payoff of the VS over time interval $[t_i, t_{i+1}]$:

$$\begin{aligned} P\&L &= \left(Q^T(t_{i+1}, S_{i+1}) - Q^T(t_i, S_i)\right) - \frac{dQ^T}{dS}(t_i, S_i)\,(S_{i+1} - S_i) \qquad (5.33)\\ &\quad - (r_i^2 - \widehat{\sigma}_{\text{VS},T}^2 \Delta t) \\ &= \left(2(e^{r_i} - 1) - 2r_i - \widehat{\sigma}_T^2 \Delta t\right) - \left(r_i^2 - \widehat{\sigma}_{\text{VS},T}^2 \Delta t\right) \qquad (5.34)\\ &= \left(2(e^{r_i} - 1) - 2r_i - r_i^2\right) - \left(\widehat{\sigma}_T^2 - \widehat{\sigma}_{\text{VS},T}^2\right)\Delta t \qquad (5.35) \end{aligned}$$

where the first piece in the right-hand side of (5.33) is the P&L of the delta-hedged log contract over $[t_i, t_{i+1}]$. We have used the expressions of $Q^T$ and $\frac{dQ^T}{dS}$ given in (5.10) and (5.11) to get (5.35).

Expand this P&L in powers of $r_i$. Up to order 2 in $r_i$, the payoff of the VS and the P&L of the delta-hedged log contract match. The first non-vanishing contribution comes from the order-three term and we get:

$$P\&L \simeq \frac{r_i^3}{3} - \left(\widehat{\sigma}_T^2 - \widehat{\sigma}_{\text{VS},T}^2\right)\Delta t \qquad (5.36)$$

The combination of a static long position in a variance swap and a short position in a delta-hedged log contract generates at lowest order the daily P&L (5.36) which looks like a gamma/theta P&L except it involves $r_i^3$ rather than $r_i^2$.

Writing that the P&L in (5.36) vanishes on average yields the following relationship between the skewness of daily log-returns and the difference between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\text{VS},T}$:

$$\widehat{\sigma}_{\text{VS},T}^2 - \widehat{\sigma}_T^2 \simeq -\frac{\langle r^3 \rangle}{3\Delta t} \simeq -\frac{s_{\Delta t}}{3}\widehat{\sigma}_T^3\sqrt{\Delta t} \qquad (5.37)$$

where $s_{\Delta t}$ denotes the skewness of daily returns defined by: $s_{\Delta t} = \langle r^3 \rangle / \langle r^2 \rangle^{\frac{3}{2}}$. For the sake of relating $\langle r^3 \rangle$ to $s_{\Delta t}$ we have taken $\langle r^2 \rangle = \widehat{\sigma}_T^2 \Delta t$; at order one in the difference $\widehat{\sigma}_{\text{VS},T} - \widehat{\sigma}_T$ we could have equivalently used $\widehat{\sigma}_{\text{VS},T}^2 \Delta t$.

Assuming that the right-hand side is small, this results in the following adjustment for $\widehat{\sigma}_{\text{VS},T}$, at order one:

$$\frac{\widehat{\sigma}_{\text{VS},T}}{\widehat{\sigma}_T} - 1 \simeq -\frac{s_{\Delta t}}{6}\widehat{\sigma}_T\sqrt{\Delta t} \qquad (5.38)$$

The interpretation of the results in Sections 5.3.1 and 5.3.2 is now clear:

- In diffusive models $\langle r^3 \rangle$ scales like $\Delta t^{3/2}$. As $\Delta t \to 0$ the contribution of $\langle r^3 \rangle$ becomes negligible with respect to that of $\langle r^2 \rangle$, which scales like $\Delta t$: $\widehat{\sigma}_{\text{VS},T} = \widehat{\sigma}_T$.

- In jump-diffusion models, the portion of $\langle r^3 \rangle$ that is generated by jumps is proportional to $\lambda \Delta t \overline{J^3}$, thus scales like $\Delta t$, just as $\langle r^2 \rangle$. Hence, $\widehat{\sigma}_T \neq \widehat{\sigma}_{\text{VS},T}$. The implied value of the cubes of daily log-returns is non-vanishing and depends on jump parameters calibrated on the smile of maturity $T$.

### 5.3.5 Inferring the skewness of daily returns from market smiles?

Using (5.37) together with (5.31) yields:

$$-2\widehat{\sigma}_T^3 \mathcal{S}_T T \ \simeq\ -\frac{\widehat{s}_{\Delta t}}{3}\widehat{\sigma}_T^3\sqrt{\Delta t}$$

which gives:

$$\widehat{s}_{\Delta t} \ \simeq\ 6\frac{\mathcal{S}_T T}{\sqrt{\Delta t}} \tag{5.39}$$

Consider the case of a one-year maturity and let us use the same level of ATMF skew as in the numerical example on page 161. Taking $\mathcal{S}_T = -0.2$ and $\Delta t = \frac{1}{252}$ gives $\widehat{s}_{\Delta t} \simeq -19$. This is a very large value, much larger than its historical average – see below.

This value of $\widehat{s}_{\Delta t}$ is derived from the smile of maturity $T$, the VS maturity, through equation (5.39). How is it that, out of a calibration to the market smile of maturity $T$, the jump-diffusion model is able to predict the value of the skewness of returns at short time scales? Is this prediction robust?

Expression (5.39) shows that, had we used a different value for $T$, we would have obtained a different estimate for $\widehat{s}_{\Delta t}$ – unless $\mathcal{S}_T$ scales like $\frac{1}{T}$.

It is a well-known property of jump-diffusion processes, that, for small jump amplitudes, the ATMF skew they generate scales like $\frac{1}{T}$ – see equation (5.30). Indeed these processes generate independent stationary increments for $\ln S_T$. Cumulants of $\ln S_T$ then scale linearly with $T$, which implies that the skewness $s_T$ of $\ln S_T$ scales like $\frac{1}{\sqrt{T}}$.

As shown in Appendix B, perturbation of the lognormal Black-Scholes density at order one in the third-order cumulant yields identity (5.93) between the ATMF skew $\mathcal{S}_T$ and the skewness $s_T$ for maturity $T$:

$$\mathcal{S}_T \ =\ \frac{s_T}{6\sqrt{T}} \tag{5.40}$$

A scaling of $s_T \propto \frac{1}{\sqrt{T}}$ then translates into the property that the ATMF skew scales like $\frac{1}{T}$, which is what equation (5.30) expresses.

The skew that a jump-diffusion model generates is a direct reflection of the skewness of increments of $\ln S_t$ at short time scales: for weak skews $\mathcal{S}_T \propto \frac{1}{T}$ and equation (5.39) supplies a value for the skewness of daily returns that does not depend on the particular value of $T$ used.

The situation is very different in diffusive stochastic volatility models: the process for $\ln S_t$ is Gaussian at short time scales[7] and develops skewness at longer time scales by the fact that volatility is stochastic and correlated with $S_t$. The smile produced by a stochastic volatility model is *not* a reflection of the non-Gaussian character of returns at short time scales.

Thus, a jump-diffusion model yields a sizeable difference between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\mathrm{VS},T}$ because it assumes a direct relationship between the ATMF skew of maturity $T$ and $\widehat{s}_{\Delta t}$. This is a bold assumption, which, moreover, is not supported by market data as, typically, ATMF skews approximately scale like $\frac{1}{\sqrt{T}}$, rather than $\frac{1}{T}$.[8]

### 5.3.6 Preliminary conclusion

The difference between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\mathrm{VS},T}$ depends on the non-Gaussian character of daily log-returns. Inferring the *implied* skewness of daily log-returns out of vanilla smiles through the filter of a calibrated model leads to a correction to $\widehat{\sigma}_{\mathrm{VS},T}$ which is unreasonably model-dependent.

In fact the *implied* skewness of daily returns could only be accessible if the package consisting of a variance swap and the offsetting log contract was actually traded. While the *realized* skewness of daily returns is typically small – see below – the *implied* skewness that would then be backed out of the difference between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\mathrm{VS},T}$ could be arbitrarily large.[9]

### 5.3.7 In reality

Imagine that there was no active VS market and we were asked to quote a VS – this would have been a typical situation in the late 90s. What would we have done?

---

[7]This statement does not stand in contradiction with the well-known property that the ATMF skew in stochastic volatility models does not vanish in the limit $T \to 0$. We show in Section 8.5 that, at order one in volatility of volatility, the *skewness* of short-maturity log-returns scales like $\sqrt{T}$, hence vanishes at short time scales: returns become Gaussian. The skew is given approximately by (5.40): because the *skewness* scales like $\sqrt{T}$, the *skew* tends to a finite limit. Generally, whether the short-maturity limit of the ATMF *skew* vanishes or not depends on whether the *skewness* vanishes faster or slower than $\sqrt{T}$ as $T \to 0$.

[8]In our discussion we have used the example of a jump-diffusion model with independent increments for $\ln S$: The skew $\mathcal{S}_T$ is then fully generated by the jump process only. In models that are mixtures of jump/Lévy process and stochastic volatility – for example jump-diffusion models where $\overline{\sigma}_t$ is stochastic and correlated with $S_t$, or subordinated Lévy processes where physical time $t$ is replaced with the integral of a random positive diffusive process (see for example [26]) – $\mathcal{S}_T$ is a product of both jump and stochastic volatility portions of the model. Contrary to the jump-diffusion model that we have used so far, in such models the implied skewness of daily returns $\widehat{s}_{\Delta t}$ is not simply a function of the level of the market skew, as part of the latter is generated by stochastic volatility. We are exposed to the additional risk of letting the model determine how much of the vanilla smile is generated by the jump or Lévy component.

[9]So-called daily cliquets – strips of put options on daily index returns with far out-of-the-money strikes, say 80% – are a case in point. Their market prices can be drastically different than prices computed using historical densities of daily returns. See Chapter 10, page 391.

Typically we would have used a conservative estimate of the realized skewness to compute an adjustment to $\widehat{\sigma}_T$ using formula (5.38):[10]

$$\widehat{\sigma}_{\mathrm{VS},T} \simeq \widehat{\sigma}_T \left(1 - \frac{s_{\Delta t}}{6} \widehat{\sigma}_T \sqrt{\Delta t}\right)$$

For indexes, the *realized* skewness $s_{\Delta t}$ of daily log-returns, defined by: $s_{\Delta t} = \left\langle r^3 \right\rangle / \left\langle r^2 \right\rangle^{\frac{3}{2}}$, where we take $\langle r \rangle = 0$, is a (dimensionless) number of order 1.

Perhaps surprisingly, the skewness of daily returns of equity indexes is not always negative. Historical skewness is difficult to measure, as the skewness estimator is very sensitive to large returns. The skewness estimator applied to a historical sample that includes an inordinately large return will be swamped by the contribution of that one return. Whenever one computes realized skewness over sufficiently long periods of time that do not contain large returns, one finds a number with varying sign, of order 1.

Obviously, evaluating $s_{\Delta t}$ for the S&P 500 index over a historical sample that includes October 1987 will yield a large negative number, however its magnitude depends crucially on the size of the window used for its estimation, as this large number is generated by one single return. The notion that evaluating $s_{\Delta t}$ over a historical sample that includes a market crash – say, October 1987 – gives an estimate of $s_{\Delta t}$ that appropriately accounts for the possibility of crashes is thus a misguided idea. Whenever one wishes to adjust the price of a derivative for the impact of large market moves, it is much more reasonable to include an explicit stress-test impact in the derivative's price – see below (5.42) for the particular case of the variance swap.

Taking $s_{\Delta t} = -1$, $\widehat{\sigma}_T = 20\%$, $\Delta t = \frac{1}{252}$, and using (5.38) gives a relative adjustment $-\frac{s_{\Delta t}}{6} \widehat{\sigma}_T \sqrt{\Delta t} = 0.2\%$.

It is in fact possible to assess the magnitude of the contribution of all orders – not just that of $r^3$ – by directly measuring on a historical sample the relative difference of the payoff of the VS and the P&L of the delta-hedged log contract (see the two expressions in (5.34)):

$$\frac{1}{2}\left(\frac{\left\langle r^2 \right\rangle}{\left\langle 2\left(e^r - 1\right) - 2r \right\rangle} - 1\right) \tag{5.41}$$

We have used the factor $1/2$ to convert a relative mismatch of variances into a relative mismatch of volatilities. Figure 5.1 shows the ratio (5.41) computed over 20 years of daily returns of the S&P 500 index, measured with a one-year sliding window.

[10]Because of the unavailability of out-of-the-money options needed in the replication of the log contract, we would also have charged an additional amount to cover for the cost of buying/selling options at then-prevailing market conditions whenever the spot moved.

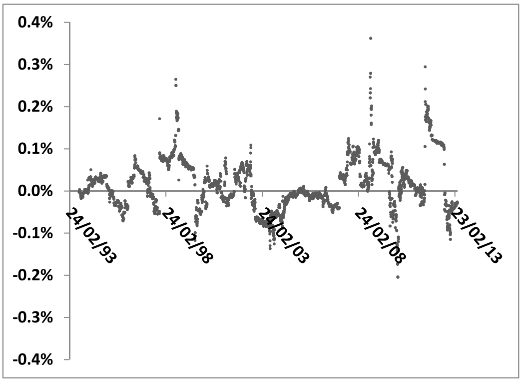



**Figure 5.1**: The ratio (5.41) evaluated for the S&P 500 index over 20 years, with a one-year sliding window.

As is apparent, ratio (5.41) is very noisy, even with a one-year estimator, but $0.2\%$ is the right order of magnitude for the relative adjustment of $\widehat{\sigma}_{\mathrm{VS},T}$ with respect to $\widehat{\sigma}_T$.[11]

This represents a very minute correction: the level of realized skewness of daily returns does not warrant in practice a distinction between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\mathrm{VS},T}$.

On the modeling side, it is possible to build a stochastic volatility model that affords full control of the conditional distribution of daily returns, thereby allowing an assessment of the difference between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\mathrm{VS},T}$. This is done in Chapter 10 and the case of VSs is specifically covered in Section 10.2.4.

In case we would like to adjust $\widehat{\sigma}_{\mathrm{VS},T}$ for the occurrence of large returns, for example to cover the cost of a stress-test P&L on a VS position, it is more reasonable to use formula (5.28). For example, a trading desk that was charged on an annual basis a fraction $\varepsilon$ of its stress-test P&L would quote $\widehat{\sigma}_{\mathrm{VS},T}$ according to:

$$\widehat{\sigma}^2_{\mathrm{VS},T} \;=\; \widehat{\sigma}^2_T + \;\varepsilon\Big(\ln^2(1+J) + 2\ln(1+J) - 2J\Big) \tag{5.42}$$

where $J$ is the amplitude of the stress-test scenario – for equities $J$ would be negative.

In the absence of a VS market we would then have chosen $\widehat{\sigma}_{\mathrm{VS},T} = \widehat{\sigma}_T$ with the possible addition of a reserve policy in the form of an adjustment given by (5.42).

It is important to note that adjustment (5.42) does not quantify the impact of large returns on a VS. It quantifies the impact of large returns on a position consisting of

[11]The largest values of ratio (5.41) are reached during the 1987 crash and lie outside Figure 5.1. The maximum is 2.5% (compare with the scale of the $y$ axis), which is not a meaningful number, as it depends on the width of the window used for the estimator – here one year.

a VS together with the offsetting log contract, i.e. the portion of the impact of large returns on VSs *that is not already accounted for in vanilla option prices.*[12]

Assuming a monthly negative jump of $5\%$ ($\varepsilon = 12$, $J = -5\%$), (5.42) yields for $\widehat{\sigma}_T = 20\%$ an adjustment $\widehat{\sigma}_{\text{VS},T} = \widehat{\sigma}_T + 0.13\%$.

---

## 5.4 Impact of strike discreteness

The previous section has been devoted to the assessment of the difference between the payoff of a VS and the P&L generated from delta-hedging a log contract. In reality, neither is the latter traded, nor can it be perfectly synthesized out of vanilla options, for the simple reason that only discrete strikes trade. The log contract is thus replaced with a piecewise affine profile.

Figure 5.2 shows the relative difference between the VS payoff and the P&L generated by delta-hedging an approximation of the log contract that uses discrete strikes $K_i$, such that the $\ln K_i$ are equally spaced. This difference is expressed as a relative adjustment factor on $\widehat{\sigma}_T$. For $\Delta \ln K \to 0$ this adjustment factor is the ratio (5.41), shown in Figure 5.1 for the case of a 1-year VS contract.

Each point in Figure 5.2 corresponds to the replication of a 1-year VS using S&P 500 daily closing quotes. The log contract is approximated by a strip of vanilla options with $\Delta \ln K = 5\%$ (resp. $1\%$) for the left-hand (resp. right-hand) graph, with $K_{\min} = 10\% S_0$, $K_{\max} = 500\% S_0$, where $S_0$ is the spot value at inception of the 1-year VS. We delta-hedge the vanilla portfolio at constant volatility.[13]

The right-hand graph in Figure 5.2 is similar to Figure 5.1: a strip of vanilla options with strikes spaced $1\%$ apart provides an acceptable replication of the log contract, at least for a 1-year maturity.

The left-hand graph shows that with a coarser discretization, the replication of the VS payoff is much less accurate: a relative adjustment of $\widehat{\sigma}_T$ of $2\%$ translates for $\widehat{\sigma}_T = 20\%$ in an adjustment of about plus or minus half a volatility point. Notice however, that, in contrast with the P&L impact of higher-order returns, the additional P&L generated by the imperfect replication of the log-contract payoff has no reason

---

[12]Interestingly, had the market standard for VS contracts featured standard returns: $\frac{S_{i+1}}{S_i} - 1$, rather than log-returns, the order-three term in expression (5.36) of the P&L would have been equal to $-\frac{2}{3}r^3$, rather than $\frac{1}{3}r^3$, where $r$ is the log-return. At leading order, a short position in a VS contract combined with a long position in its delta-hedged vanilla-option hedge would generate positive, rather than negative P&L, for $J < 0$. Another advantage of using standard returns is that, even in the case of extreme bankruptcy ($S_{i+1} = 0$) the return remains well-defined.

[13]This constant volatility is taken equal to the realized volatility over the 1-year period that follows – a proxy for the actual market implied volatility of the vanilla portfolio, which we would use in reality. In practice the final P&L is not very sensitive to the actual implied volatility used for risk-managing the vanilla portfolio, especially if the log-contract profile is well approximated. Remember that the delta of the log contract in the Black-Scholes model does not depend on the implied volatility.

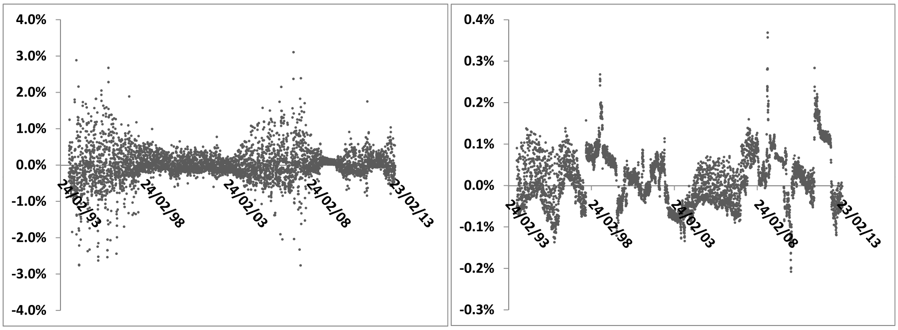



**Figure 5.2**: Mismatch of (a) the payoff of a 1-year VS, (b) the P&L generated by delta-hedging an approximation of the log-contract payoff of maturity 1 year using discrete strikes $K_i$ with $\Delta \ln K = 5\%$ (left) and $1\%$ (right). We have used S&P 500 daily quotes. This difference is expressed as a relative adjustment on the VS volatility $\widehat{\sigma}_T$.

to be biased and should be considered a noise – this is clearly seen in the left-hand graph of Figure 5.2.

---

## 5.5 Conclusion

- If log contracts were traded or, equivalently, if vanilla options of all strikes were available, both VS ($\widehat{\sigma}_{\text{VS},T}$) and log contract ($\widehat{\sigma}_T$) implied volatilities would be market parameters.

  Their difference would supply a measure of the *implied* skewness of daily returns, which could be arbitrarily large, even though its *realized* counterpart is small. In such a situation VS and log-contract forward variances $\xi^T$ and $\zeta^T$ would be different objects and the joint dynamics of $S, \xi^T, \zeta^T$ would read as:

$$\begin{cases} dS_t &= \overline{\sigma}_t S_t dW_t^S \\ d\zeta_t^T &= \lambda_t^T dU_t^T \\ d\xi_t^T &= \psi_t^T dV_t^T \end{cases}$$

  $\overline{\sigma}_t,\ \zeta_t^T, \xi_t^T$ are all different.

- In the absence of liquid log contracts, we have shown that it makes no sense to price VSs using a jump-diffusion or Lévy model calibrated on the vanilla smile. By doing so, we use the model as a tool for inferring the skewness of returns at short time scales out of vanilla smiles; this is incongruous as there is no reason why these quantities should be related – except in the model.

Above all, there is no way of locking in this relationship – that is the difference between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ – by trading vanilla options.

- In practice VSs are much more liquid than far out-of-the-money vanilla options: while $\widehat{\sigma}_{\mathrm{VS},T}$ is a market parameter, $\widehat{\sigma}_T$ usually is not and depends on how one extrapolates implied volatilities for strikes that lie outside the liquid range.

  It is then common practice among index volatility market makers to choose this extrapolation such that replication of the log contract recovers the VS implied volatility – this enforces the equality $\widehat{\sigma}_T = \widehat{\sigma}_{\mathrm{VS},T}$.

- In the following chapters, we will thus make the assumption that the process for $S_t$ is a diffusion, so that $\widehat{\sigma}_T = \widehat{\sigma}_{\mathrm{VS},T}$ and that the instantaneous volatility of $S$ is given by the short end of the variance curve:

$$\begin{cases} dS_t &= \sqrt{\xi_t^t} S_t dW_t^S \\ d\xi_t^T &= \lambda_t^T dW_t^T \end{cases} \tag{5.43}$$

  One exception is Chapter 10 where we examine the impact of the conditional distribution of daily returns on derivative prices. We employ a model that gives us explicit control on the one-day smile and assess the difference $\widehat{\sigma}_{\mathrm{VS},T} - \widehat{\sigma}_T$.

- The two key properties expressed by (5.43) still hold in the presence of cash-amount dividends:

  - Forward VS variances are still driftless as the reasoning used in Section 5.1 still applies.
  - The instantaneous volatility of $S$ is still given by the short end of the variance curve.

### When $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ are different

- For some very liquid indexes such as the S&P 500 or Euro Stoxx 50, far out-of-the-money puts are liquid. The spread between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ that one typically observes, even for short maturities, is not attributable to uncertainty about far out-of-money implied volatilities: the market does make a distinction between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$.

  As discussed in Sections 5.3.7 and 5.4, there are valid reasons for this: (a) terms of order higher than 2 in the P&L of the delta-hedged log contract, whose implied value could be much larger than their realized value in Figure 5.1, (b) the fact that the log contract is in practice approximated by a portfolio of vanilla options with discrete strikes.

- In these cases, one can use expression (5.42) to dissociate $\widehat{\sigma}_{\mathrm{VS},T}$ from $\widehat{\sigma}_T$, with $J$ chosen and $\varepsilon$ – possibly time-dependent – calibrated so that the market VS volatility $\widehat{\sigma}_{\mathrm{VS},T}$ is recovered.

  This is not the same as using a jump model. Equation (5.42) expresses the spread between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ as the *difference* in how a jump of magnitude $J$ impacts a VS relative to a log contract. Vanilla options would still be priced with a *diffusive* model – say a local volatility or stochastic volatility model of type (5.43) calibrated to the vanilla smile – but whenever a spot-starting or forward-starting VS was priced, the adjustment specified in (5.42) would be used to generate $\widehat{\sigma}_{\mathrm{VS},T}$ from $\widehat{\sigma}_T$.

- In a pricing library, realized variance – in the form of squared log-returns – would need to be identified as such so that it can be adjusted automatically. We still simulate SDE (5.43), except $\xi_t^T$ has the status of a log-contract forward variance. Whenever the payoff at hand calls for observation of realized variance, adjustment (5.42) is applied automatically:

$$\ln^2\left(\frac{S_{i+1}}{S_i}\right) \;\rightarrow\; \ln^2\left(\frac{S_{i+1}}{S_i}\right) \;+\; (\lambda\Delta)\,\overline{\ln^2\left(1+J\right)+2\ln\left(1+J\right)-2J}$$

  It is also applied when the payoff calls for observation of implied realized variance – that is a forward VS variance $\zeta_t^T$ – for example in the case of a variance swaption:[14]

$$\zeta_t^T \;=\; \xi_t^T \;+\; \lambda\,\overline{\ln^2\left(1+J\right)+2\ln\left(1+J\right)-2J}$$

  This additive adjustment preserves the martingality of $\zeta_t^T$. $\lambda, J$ are chosen so as to match market values of $\widehat{\sigma}_{\mathrm{VS},T}-\widehat{\sigma}_T$. $\lambda$ is posssibly time-dependent, in which case it is replaced by $\frac{1}{T}\int_0^T\lambda_t dt$ in the formulas above. The exposure to the spread between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ is easily monitored at the book level.

- For longer-dated VSs a large mismatch between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ is more likely caused by a mispricing of the effect of interest-rate volatility – see Section 5.8 below – or an inappropriate dividend model.[15]

[14] In essence, this boils down to considering $\ln\left(S_{i+1}/S_i\right)^2$ as a short-period cliquet, and performing an ad-hoc adjustment for the forward-smile risk over interval $[t_i, t_{i+1}]$.

[15] Given market prices of vanilla options and the forward, one obtains different values for $\widehat{\sigma}_T$ depending upon whether one models dividends as proportional to the spot or as fixed cash amounts. $\widehat{\sigma}_T$ is higher when proportional dividends are used. There is an additional impact for index variance swaps, as the market convention mandates that daily returns be not stripped of the contribution of dividends. As shown in the next section this contribution to $\widehat{\sigma}_T$ is small, but can become unreasonably large if evaluated with an inappropriate dividend model – see Section 5.7 below.

## 5.6 Dividends

Dividends have two effects on the pricing of VSs: on the payoff itself and on its replication with vanilla options.

### 5.6.1 Impact on the VS payoff

Imagine a dividend $d$ falls just after the close of day $t_i$: $S_{t_i^+} = S_{t_i^-} - d$.[16] The log-return over $[t_i, t_{i+1}]$ can then be written as:

$$\ln\left(\frac{S_{i+1}}{S_i}\right) = \ln\left(\frac{S_{i+1}}{S_i - d}\right) + \ln\left(\frac{S_i - d}{S_i}\right)$$

Let us assume that $S$ follows a diffusion with instantaneous volatility $\overline{\sigma}$. Squaring this expression, taking its expectation over $S_{i+1}$ and keeping terms up to order 2 in $\Delta t$ yields:

$$E\left[\ln^2\left(\frac{S_{i+1}}{S_i}\right)\right] = \overline{\sigma}^2\Delta t + \ln^2\left(\frac{S_i - d}{S_i}\right) + 2\left(r - q - \frac{\overline{\sigma}^2}{2}\right)\Delta t \ln\left(\frac{S_i - d}{S_i}\right) \quad (5.44)$$

Let us take the following typical values for a stock: $\overline{\sigma} = 30\%$, $\Delta t = \frac{1}{252}$, $\frac{d}{S_i} = 3\%$, $r - q = 3\%$: the order of magnitudes of the three terms above is, respectively: $3.10^{-4}$, $10^{-3}$, $2.10^{-6}$: the drift of $S$ over $\Delta t$ is so small that the last term can be safely ignored. It turns out that the second term in the right-hand side of (5.44) can be discarded as well for the following reasons:

- For stocks, the usual convention of VS term sheets is to adjust the return over $[t_i, t_{i+1}]$ by the dividend amount $d$. The return used for the sake of computing realized variance is $\ln(\frac{S_{i+1}}{S_i - d})$.

- For indexes, the value of $S_i$ is not adjusted for $d$, however an index is a basket of stocks: it jumps whenever a dividend is paid on one of its components by an amount equal to the dividend times the relative weight of that particular stock in the index. The yearly dividend yield of an index is the same order of magnitude as the yields of the components, except it is spread out over many dates in the year, corresponding to the dividend payment dates of the components.

  Take the example of the Euro Stoxx 50 index, with $n = 50$ dividend dates per year. Let us use a constant volatility $\overline{\sigma} = 20\%$ and assume that each of the $n$ dividends is proportional to $S$, with a proportionality coefficient equal to $q/n$,

[16] The dividend is not paid to stockholders at that time. Rather, anyone who was not owning the share at $t_i$ loses the right to the dividend payment – which occurs at a later date. The effect on $S$ is however identical to that of a dividend payment occurring between $t_i$ and $t_{i+1}$: $S$ jumps by the value of the right to the dividend payment.

so that the yearly dividend yield $q$ for the index is $3\%$. Keeping the first two terms in (5.44) yields for the realized volatility over one year:

$$\sigma_r = \sqrt{\overline{\sigma}^2 + n\ln^2\left(1-\frac{q}{n}\right)}$$

For large $n$, $n\ln^2\left(1-\frac{q}{n}\right)$ scales like $\frac{1}{n}$ and the contribution of dividends become negligible: using the numerical values above yields $\sigma_r = 20.005\%$.

The conclusion is that the impact of dividends on the VS payoff itself is negligible in practice.

### 5.6.2 Impact on the VS replication

Let us take zero interest rate and repo for simplicity and imagine that fixed cash dividends $d_j$ fall at dates $T_j$. Consider a delta-hedged log contract at zero implied volatility. Its value is now given by

$$Q^T(t,S) = -2\ln\left(S - \sum_{t<T_j<T} d_j\right)$$

The dollar gamma is given by:

$$\frac{S^2}{2}\frac{d^2Q^T}{dS^2} = \frac{S^2}{\left(S - \sum_{t<T_j<T} d_j\right)^2} \tag{5.45}$$

Because of the presence of dividends with fixed cash amounts, it is no longer equal to one. It is however possible to assemble a portfolio that has constant dollar gamma by supplementing the log contract of maturity $T$ with a set of European payoffs $E^j$ of maturities $T_j^-$. Denote by $T_N$ the last dividend date before $T$. At time $t = T_N^-$, the value of the log contract is $-2\ln(S-d_N)$. To have a portfolio whose value is $-2\ln S$, we need to go long an additional European payoff $E^N(S)$ of maturity $T_N^-$ such that

$$-2\ln(S-d_N) + E^N(S) = -2\ln S$$

which yields:

$$E^N(S) = 2\ln\left(\frac{S-d_N}{S}\right) \tag{5.46}$$

Working backward in time we can check that European payoffs $E^j$ maturing at previous dates $T_j^-$ have the same form as $E^N$: $E^j(S) = 2\ln\left(\frac{S-d_j}{S}\right)$. In case interest rates are non-vanishing, the quantities of these intermediate payoffs are changed slightly, but keep the form (5.46).[17]

[17] Obviously, payoff $E^j(S)$ is only defined for $S > d_j$. The implied volatility surface used for pricing payoff $E^j$ must ensure that this condition holds. In other words, the smile used as input must be such that the density of $S_{T_j^-}$ vanishes for $S_{T_j^-} < d_j$.

The conclusion is that $\widehat{\sigma}_{\text{VS}}$ is no longer equal to $\widehat{\sigma}_T$ but depends on smiles of intermediate maturities. The VS is replicated statically at second order in $\frac{\delta S}{S}$ by trading a log contract plus a series of European payoffs with maturities corresponding to dividend dates. Using (5.23), we have:

$$\widehat{\sigma}_{\text{VS},T}^2 \;=\; \frac{e^{rT}}{T}\left(Q^T_{\text{market}} - Q^T_{\widehat{\sigma}=0} \;+\; \sum_{T_j<T}\left(E^j_{\text{market}} - E^j_{\widehat{\sigma}=0}\right)\right) \tag{5.47}$$

For index VSs the contribution of dividends to the realized variance – examined in the previous section – has to be added by hand, in the form of additional $\ln^2\left(1-\frac{d_j}{S}\right)$ payoffs – see section below.

---

## 5.7 Pricing variance swaps with a PDE

Indexes are large baskets; computing $\widehat{\sigma}_{\text{VS},T}$ with (5.47) entails replicating payoffs $E_j$ for many maturities, corresponding to the (numerous) dividend dates of the components. It is more convenient to compute $\widehat{\sigma}_{\text{VS},T}$ by using a diffusive model calibrated on the market smile – we know from Section 5.3.1 and equation (5.23) that this yields the same value for $\widehat{\sigma}_{\text{VS},T}$ as (5.47).

As any diffusive model calibrated to the market smile will do, we can simply use a local volatility model. We have:

$$\widehat{\sigma}_{\text{VS},T}^2 \;=\; \frac{1}{T}\int_0^T E[\sigma^2(t,S_t)]dt \tag{5.48}$$

where $\sigma\left(t,S\right)$ is the local volatility function calibrated to the market smile, given by Dupire formula (2.3).[18] The expectation in the right-hand side of (5.48) is evaluated by solving the following backward PDE:

$$\frac{dU}{dt} + \left(r-q\right)S\frac{dU}{dS} + \frac{\sigma^2\left(t,S\right)}{2}S^2\frac{d^2U}{dS^2} \;=\; -\sigma^2\left(t,S\right) \tag{5.49}$$

with terminal condition $U\left(t=T,S\right)=0$ and the following matching condition at each dividend date: $U\left(T_j^-,S\right) \;=\; U\left(T_j^+,S-d_j\right)$. $\widehat{\sigma}_{\text{VS},T}$ is given by:

$$\widehat{\sigma}_{\text{VS},T} \;=\; \sqrt{\frac{U\left(0,S_0\right)}{T}} \tag{5.50}$$

where $S_0$ is the spot value at $t=0$.

[18] The reader can check that the derivation of (2.6) is still valid when there are dividends, however one has to ensure that prices of vanilla options maturing immediately before and after dividend dates obey appropriate matching conditions. For a fixed amount dividend $d_j$ falling at $T_j$: $C(K,T_j^+) = C(K+d_j,T_j^-)$. If one prefers to work with implied volatilities directly, one has to make sure that the equivalent matching condition for implied volatilities holds. These are discussed in Section 2.3.1, page 34.

For indexes, the returns used for calculating the VS payout are not stripped of dividends. This extra contribution materializes as a discontinuity of $U$ at dividend dates. For a dividend $d_j$ falling at $T_j$ this matching condition is – see (5.44):

$$U(T_j^-, S) \;=\; U(T_j^+, S - d_j(S)) \;+\; \ln^2\left(1 - \frac{d_j(S)}{S}\right) \tag{5.51}$$

where $d_j(S)$ expresses that the dividend generally depends on the spot level. As argued in Section 5.6.1, this additional contribution of dividends to index VS levels should be small.

Using a pure cash amount dividend model may however result in a blatant overestimation. In fact, expression (5.51) requires $d_j(S) < S$, so $d_j(S)$ should be replaced with $\max\left(d_j(S), y_j^{\max} S\right)$ where $y_j^{\max}$ is the maximum yield allowed for dividend $d_j$, with $y_j^{\max} < 1$. If $y_j^{\max}$ is too large, the steep smiles of equity indexes may cause the contribution of payoff

$$\ln^2\left(1 - \max\left(\frac{d_j(S)}{S}, y_j^{\max}\right)\right)$$

to become unreasonably large. It is thus important to cap the effective yield of dividends when pricing index VSs.[19]

**Adjustment for large returns**

Imagine now that we would like to adjust the VS volatility for the impact of large returns – as in (5.42). This adjustment expresses the fact that while the VS can be perfectly hedged with vanilla options up to second order in $(S_{i+1} - S_i)$, higher-order terms impact the VS and the replicating vanilla portfolio differently. Expanding equation (5.42) in powers of the jump magnitude $J$ and stopping at the lowest non-trivial order gives:[20]

$$\begin{aligned} \widehat{\sigma}^2_{\mathrm{VS},T} \;&=\; \widehat{\sigma}^2_T + \;\varepsilon\left(\ln^2(1+J) + 2\ln(1+J) - 2J\right) \\ &\simeq\; \widehat{\sigma}^2_T - \frac{1}{3}\varepsilon J^3 \end{aligned} \tag{5.52}$$

where $\varepsilon$ is the annualized probability of a jump. $\widehat{\sigma}_{\mathrm{VS},T}$ is still given by (5.50), except the PDE for $U$ now reads:

$$\frac{dU}{dt} + (r-q)\,S\frac{dU}{dS} + \frac{\sigma^2(t,S)}{2}S^2\frac{d^2U}{dS^2} \;=\; -\left(\sigma^2(t,S) - \frac{1}{3}\varepsilon J^3\right) \tag{5.53}$$

For constant $J$ and $\varepsilon$ the solution of (5.53) is simply $U(0,S_0) - \frac{1}{3}(\varepsilon T)J^3$, where $U(0,S_0)$ is the solution of (5.49).

[19] Capping the yield of dividends opens up a (small) can of worms: with dividends no longer an affine function of $S$, forwards become sensitive to volatility and the technique of Section 2.3.1 for calibrating the local volatility function no longer works exactly.

[20] Note the similarity with (5.36).

PDE (5.53) proves useful in situations when one needs to price weighted VSs consistently with VSs, in circumstances when VS market volatilities do not match the vanilla replication – presumably because VS market prices include an adjustment of type (5.52).

Weighted VSs are covered in Section 5.9 below. In weighted VSs, the realized variance is weighted by a function of the spot, $w(S)$.

Because weighted VSs can be replicated with vanilla options, they can be priced with PDE (5.49) where $\sigma^2(t,S)$ is simply replaced with $w(S)\sigma^2(t,S)$. What about adjustment (5.52) for higher-order terms?

Rather than assuming that $J$ is constant, it is more reasonable to assume that the scale of a return occurring for a spot level $S$ at time $t$ is set by the local volatility $\sigma(t,S)$ – regardless of which component, be it Brownian motion or jump, generated that one return. We thus write:

$$\varepsilon J^3 \;\equiv\; -\mu\sigma^3(t,S)$$

where $\mu$ is a constant chosen so that market prices of VSs are recovered. The PDE for the weighted VS then reads:

$$\frac{dU}{dt} + (r-q)\,S\frac{dU}{dS} + \frac{\sigma^2(t,S)}{2}S^2\frac{d^2U}{dS^2} \;=\; -w(S)\left(\sigma^2(t,S) + \frac{1}{3}\mu\sigma^3(t,S)\right)$$

where $\mu$ is calibrated to market VS quotes.

---

## 5.8 Interest-rate volatility

Assume that there are no cash-amount dividends. In a diffusive setting $\widehat{\sigma}_{\mathrm{VS},T}$ is then equal to the implied volatility of the log contract.

As with any European payoff, the Black-Scholes implied volatility one backs out of a market price is in fact the integrated volatility of the forward $F_t^T = S_t e^{(r_t-q_t)(T-t)}$ rather than the integrated volatility of $S_t$, where $r_t, q_t$ are the interest rate and repo prevailing at $t$ for maturity $T$.

When interest rates are deterministic, the (lognormal) volatilities of $F_t^T$ and $S_t$ are identical and the Black-Scholes implied volatility is that of $S_t$. In the case of stochastic interest rates they are different – note that what matters is the volatility of the interest rate for the residual maturity.

We could still use $\widehat{\sigma}_{\mathrm{VS},T} = \widehat{\sigma}_T$ if the VS contract paid the realized variance of the forward, i.e. the sum of $\ln^2\left(\frac{F_{i+1}^T}{F_i^T}\right)$, however this is not the case.

In what follows, we still use the notation $\widehat{\sigma}_T$ for the implied (Black-Scholes) volatility of the log contract, i.e. the VS volatility of the *forward*, and $\widehat{\sigma}_{\mathrm{VS},T}$ for the VS volatility of the *spot*.

Let us assume that the instantaneous normal volatility at time $t$ of the interest rate $r_t$ of maturity $T$ is constant, equal to $\sigma_r$. This dynamics is generated by the Ho&Lee model, which is a short rate model such that rates for all maturities have the same (normal) volatility, equal to that of the short rate:

$$dr_t = \left(\frac{df_{0,t}}{dt} + \sigma_r^2 t\right) dt + \sigma_r dW_t^r$$

where $f_{0,t}$ is the initial term structure of forward rates.

Denote by $\sigma$ the (lognormal) volatility of $S_t$ and $\rho$ the correlation between $S_t$ and $r_t$. We assume for simplicity zero repo. The instantaneous variance of $F_t^T$ is given by:

$$\begin{aligned} E[\left(d\ln F_t^T\right)^2] &= E[(d\ln S_t + (T-t)\, dr_t)^2] \\ &= \left(\sigma^2 + 2\rho\,(T-t)\,\sigma\sigma_r + (T-t)^2\,\sigma_r^2\right) dt \end{aligned}$$

The integrated variance of the forward reads:

$$\frac{1}{T}\int_0^T E[\left(d\ln F_t^T\right)^2] = \frac{1}{T}\int_0^T \left(\widehat{\sigma}_{\mathrm{VS},T}^2 + 2\rho\,(T-t)\,\widehat{\sigma}_{\mathrm{VS},T}\sigma_r + (T-t)^2\,\sigma_r^2\right) dt$$

We thus get the following relationship between $\widehat{\sigma}_T$ and $\widehat{\sigma}_{\mathrm{VS},T}$ :

$$\widehat{\sigma}_T^2 = \widehat{\sigma}_{\mathrm{VS},T}^2 + \rho\widehat{\sigma}_{\mathrm{VS},T}\sigma_r T + \frac{\sigma_r^2 T^2}{3} \tag{5.54}$$

We leave it to the reader to invert (5.54) to get $\widehat{\sigma}_{\mathrm{VS},T}$ as a function of $\widehat{\sigma}_T$. At order one in $\sigma_r$ the adjustment reads:

$$\widehat{\sigma}_{\mathrm{VS},T} = \widehat{\sigma}_T - \frac{\rho}{2}\sigma_r T \tag{5.55}$$

Interest-rate volatility mostly affects long-maturity VSs. As an example take $\widehat{\sigma}_T = 25\%$, an interest-rate volatility of 5 bps/day, a correlation equal to $50\%$, and a maturity of 5 years. (5.55) yields $\widehat{\sigma}_{\mathrm{VS},T} = 24\%$. The correction is about one point of volatility – this is not a small effect.

In the Ho&Lee model we have a constant $\sigma_r$, independent of $T-t$. We can of course use a time-dependent $\sigma_r$, calibrated on co-terminal swaptions, and the full term structure of VS volatilities rather than assuming a constant spot volatility.

---

## 5.9 Weighted variance swaps

We have so far concentrated on the standard VS, by far the most common variance instrument. Other variance payoffs exist, corresponding to other ways of weighting realized variance, as a function of the underlying. Their payoffs read:

$$\begin{aligned}
&\frac{1}{T-t}\sum_t^T w(S_i)\left(\ln^2\left(\frac{S_{i+1}}{S_i}\right) - \Delta t\,\widehat{\sigma}^2\right) \\
&= \frac{1}{T-t}\sum_t^T w(S_i)\ln^2\left(\frac{S_{i+1}}{S_i}\right) - \frac{\Delta t}{T-t}\widehat{\sigma}^2\sum_t^T w(S_i)
\end{aligned} \tag{5.56}$$

where $\Delta t$ is 1 day and $\widehat{\sigma}$ is the strike of the weighted VS, such that the latter is worth zero at inception.

As the second term in (5.56) is simply a string of European payoffs, we concentrate on the first one.

We now discuss the replication of these payoffs – we refer the reader to original work in [24], [25] and [65].

The standard VS is sensitive to contributions of order 3 in $\delta S$, and to the discretization of the replicating European profile, effects dicussed above, which impact weighted VSs in equal measure.

We focus here on the replication at order two in $\delta S$, paralleling the analysis in Section 5.2.

Can we find a European payoff $f\,(S)$ of maturity $T$ such that, delta-hedging it at zero implied volatility generates as gamma P&L the desired variance payoff? Using the same notation as in (5.9), this condition reads:

$$\frac{1}{2}S^2\frac{d^2P_{\widehat{\sigma}=0}}{dS^2} \;=\; w(S)e^{-r(T-t)} \tag{5.57}$$

$P_{\widehat{\sigma}=0}$ is the price of European payoff $f$ of maturity $T$, at zero implied volatility:

$$P_{\widehat{\sigma}=0}(t,S) \;=\; e^{-r(T-t)}f\big(Se^{\mu(T-t)}\big)$$

where $\mu = r - q$. (5.57) translates into:

$$\frac{1}{2}S^2f''\big(Se^{\mu(T-t)}\big)e^{2\mu(T-t)} \;=\; w\,(S)$$

or equivalently:

$$\frac{1}{2}S^2f''(S) \;=\; w\big(Se^{-\mu(T-t)}\big) \tag{5.58}$$

which must be obeyed $\forall t, \forall S$.

The standard VS corresponds to $w \equiv 1$. Taking $f\,(S) = -2\ln S$ indeed takes care of (5.58).

For other weighting schemes, (5.58) cannot hold, owing to the dependence of the right-hand side on $t$.

Let us then *decide* that $\mu = 0$: we risk-manage a long position in European payoff $f$ (a) at zero implied volatility, (b) with $q = r$. $P_0 \equiv P^{\mu=0}_{\widehat{\sigma}=0}$ solves the Black-Scholes PDE with $q = r,\ \sigma = 0$:

$$\frac{dP_0}{dt} = rP^{\mu=0}_{\widehat{\sigma}=0}, \quad P_0(t = T, S) = f(S) \tag{5.59}$$

The P&L during $\delta t$ of a long position in payoff $f$, delta-hedged, reads:

$$\begin{aligned} P\&L &= \big(P_0\left(t+\delta t, S+\delta S\right) - P_0\left(t, S\right)\big) - \frac{dP_0}{dS}\big(\delta S - (r-q)\, S\delta t\big) \; - \; rP_0\delta t \\ &= e^{-r(T-t)} w(S) \left(\frac{\delta S}{S}\right)^2 \; + \; S\frac{dP_0}{dS}(r-q)\delta t \end{aligned} \tag{5.60}$$

where we have expanded the P&L at order two in $\delta S$, using (5.59) and the property that $\frac{1}{2}S^2\frac{d^2P_0}{dS^2} \;=\; w(S)e^{-r(T-t)}$.

The first portion in (5.60) is exactly what we need to replicate our weighted VS at order two in $\delta S$.[21]

The second portion can be canceled by selling at inception a quantity $(r-q)\delta t$ of European options of maturity $t$ with payoff $S\frac{dP_0}{dS}(t, S) = e^{-r(T-t)}S\frac{df}{dS}$.

In conclusion, a weighted VS of maturity $T$ with weight $w(S)$ is replicated at order two in $\delta S$ with:

- a European payoff $f(S)$ of maturity $T$ with $f$ such that $\frac{d^2f}{dS^2} = 2\frac{w(S)}{S^2}$, delta-hedged at zero implied volatility, and with $q = r$
- a continuous density $(r-q)$ of (unhedged) European options of maturities $\tau$ spanning $[0,\ T]$ whose payoffs are $e^{-r(T-\tau)}S\frac{df}{dS}$

We now review some examples of weighted VSs.

**Gamma swap:** $w(S) = S$

The replicating European payoff of maturity $T$ is $f(S) = 2S\ln S$ and the intermediate European payoffs are log contracts.

In case $r - q = 0$, for example if the underlying is a future, the replicating portfolio only consists of the final payoff $2S\ln S$. Strike $\widehat{\sigma}$ of the gamma swap then equals the implied volatility of this payoff, $\widehat{\sigma}_{S\ln S}$, which is given explicitly as a function of vanilla implied volatilities by:

$$\begin{aligned} \widehat{\sigma}^2_{S\ln S} &= \int_{-\infty}^{+\infty} dy\, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \widehat{\sigma}^2_{K(y)T} \\ y(K) &= \frac{\ln\left(\frac{F_T}{K}\right)}{\widehat{\sigma}_{KT}\sqrt{T}} + \frac{\widehat{\sigma}_{KT}\sqrt{T}}{2} \end{aligned}$$

[21] Factor $e^{-r(T-t)}$ is appropriate, as the gamma P&L is generated at time $t$ while the VS payoff is delivered at $T$.

This formula is derived on page 143, in Section 4.3.1 of Chapter 4.

The $S\ln S$ contract is replicated with a density $\frac{2}{K}$ of vanilla options, as opposed to a density $\frac{2}{K^2}$ for the standard VS, thus lessening the contribution of expensive low-strike vanilla options, in case of a strong skew.

In exchange for relinquishing part of the realized variance for low spot values, we get a lower strike: $\widehat{\sigma}_{S\ln S} \le \widehat{\sigma}_{\mathrm{VS}}$.

**Arithmetic variance swap** $w(S) = S^2$

The payoff of the arithmetic VS is: $\Sigma_i(S_{i+1} - S_i)^2$.

It is not traded, though it enjoys a unique property among weighted VSs: it is exactly replicable, even in the presence of large returns.

Indeed, the replicating European payoff of maturity $T$ is a parabola. The expansion of the carry P&L at order two in $\delta S$ in (5.60) is exact, as there are no higher-order terms.

**Corridor variance swap:** $w(S) = 1_{S\in[L,H]}$

In corridor variance swaps $w(S)$ is an indicator function. They are very popular, either as corridors, or as down-VSs ($w(S) = 1_{S\in[0,H]}$), or up-VSs ($w(S) = 1_{S\in[L,\infty]}$). Their replication has been first studied by Keith Lewis and Peter Carr in [24].

The terminal replicating European payoff is a truncated log contract, synthesized with a density $\frac{2}{K^2}$ of vanilla options of strikes $K \in [L, H]$. $f(S)$ is given by:

$$\begin{array}{ll} -2\ln S & S \in [L,\ H] \\ -2\ln H - \frac{2}{H}(S-H) & S \ge H \\ -2\ln L - \frac{2}{L}(S-L) & S \le L \end{array}$$

The intermediate European payoffs are simple combinations of zero-coupon bonds and vanilla options struck at $L$ and $H$.

How should returns that cross a barrier be treated? Consider for example a situation with $S_i < H$ and $S_{i+1} > H$.

For this particular return, our long, delta-hedged, position in payoff $f$ generates the following P&L – taking zero interest rate for simplicity:

$$\begin{aligned} P\&L &= \big(f(S_{i+1}) - f(S_i)\big) - \frac{df}{dS}(S_i)(S_{i+1} - S_i) \\ &= -2\ln H - \frac{2}{H}(S_{i+1} - H) + 2\ln S_i + \frac{2}{S_i}(S_{i+1} - S_i) \end{aligned}$$

This is not equal to $\left(\frac{S_{i+1}-S_i}{S_i}\right)^2$, which is what the corridor VS payoff would prescribe, at order two in $(S_{i+1} - S_i)$. Expanding our P&L at order two in $(S_i - H)$ and $(S_{i+1} - H)$ yields:

$$P\&L = \frac{(S_{i+1} - S_i)^2}{H^2} - \frac{(S_{i+1} - H)^2}{H^2}$$

We leave it to the reader to work out the other three cases. Unfortunately, none of the provisions typically found in term sheets of corridor VSs matches the P&L that our replication strategy generates.

This mismatch has to be estimated separately and factored in the strike of the corridor VS.

---

## Appendix A – timer options

Let us take zero repo and interest rate and consider a short position in a European option of maturity $T$, delta-hedged in the Black-Scholes model with a fixed implied volatility $\widehat{\sigma}$.

We assume for now that our option's payoff is convex so that $\frac{d^2P_{\widehat{\sigma}}}{dS^2} \geq 0$. Our final P&L at order one in $\delta t$ and two in $\delta S$ is given by expression (1.9), page 9. In the continuous limit:

$$P\&L \;=\; -\int_0^T \frac{S_t^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dS^2}(t,S_t)(\sigma_t^2-\widehat{\sigma}^2)dt \tag{5.61}$$

where $\sigma_t$ is the instantaneous realized volatility and $P_{\widehat{\sigma}}$ is the Black-Scholes expression for the option's price.

In case $\sigma_t$ lies consistently above/below $\widehat{\sigma}$ we will lose/make money. It can happen though that while the realized volatility over the option's maturity, $\sqrt{\frac{1}{T}\int_0^T \sigma_t^2 dt}$, is lower than $\widehat{\sigma}$, $P\&L < 0$ as in expression (5.61) the difference $(\sigma_t^2 - \widehat{\sigma}^2)$ is weighted by the option's dollar gamma, which varies with $t$ and $S$.

Only for VSs, whose dollar gamma is constant, does the final P&L only depend on the difference between realized volatility over $[0,T]$ and implied volatility.

Rather than keeping $\widehat{\sigma}$ fixed, consider adjusting it in real time so as to absorb, over each interval between two delta rehedges, the gamma/theta P&L.[22] Denote by $\widehat{\sigma}_t$ the implied volatility we use at time $t$. During $\delta t$, our total P&L, including the mark-to-market P&L from remarking our option position at $t+\delta t$ with the implied volatility $\widehat{\sigma}_t + \delta\widehat{\sigma}$, is:

$$P\&L \;=\; -\frac{S_t^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dS^2}(\sigma_t^2-\widehat{\sigma}_t^2)\delta t \;-\; \frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}}\delta\widehat{\sigma} \tag{5.62}$$

with $\delta\widehat{\sigma}$ chosen so that $P\&L = 0$.

We now pause to derive an ancillary result relating vega to the dollar gamma in the Black-Scholes model.

[22] This introduction draws from a presentation given by Bruno Dupire at the 2007 Global Derivatives conference.

## A.1 Vega/gamma relationship in the Black-Scholes model

Denote by $V$ the vega: $V = \frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}}$, where $P_{\widehat{\sigma}}$ is the Black-Scholes price with implied volatility $\widehat{\sigma}$. Taking the derivative of the Black-Scholes equation (1.4) with respect to $\widehat{\sigma}$ yields:

$$\frac{dV}{dt} + (r-q)S\frac{dV}{dS} + \frac{\widehat{\sigma}^2}{2}S^2\frac{d^2V}{dS^2} - rV \; = \; -\widehat{\sigma}S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2} \tag{5.63}$$

At maturity, $V(t=T,S)=0, \; \forall S$. $V$ is thus only generated by the source term in (5.63):

$$V(t,S) \; = \; \widehat{\sigma}\int_t^T E_{t,S}\left[e^{-r(\tau-t)}S_\tau^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}(\tau,S_\tau)\right]d\tau \tag{5.64}$$

Setting $x = \ln S$, the Black-Scholes equation reads:

$$\frac{dP_{\widehat{\sigma}}}{dt} + \big(r - q - \frac{\widehat{\sigma}^2}{2}\big)\frac{dP_{\widehat{\sigma}}}{dx} + \frac{\widehat{\sigma}^2}{2}\frac{d^2P_{\widehat{\sigma}}}{dx^2} \; = \; rP_{\widehat{\sigma}} \tag{5.65}$$

Take the derivative of (5.65) $n$ times with respect to $x$. Since neither $r$, nor $q$, nor $\widehat{\sigma}$ depend on $x$, $\frac{d^nP_{\widehat{\sigma}}}{dx^n}$ solves the same PDE as $P_{\widehat{\sigma}}$. We thus have:

$$\frac{d^nP_{\widehat{\sigma}}}{dx^n}(t,x) \; = \; E_{t,x}\left[e^{-r(T-t)}\frac{d^nP_{\widehat{\sigma}}}{dx^n}(T,x_T)\right]$$

$e^{-r(T-t)}\frac{d^nP_{\widehat{\sigma}}}{d\ln S^n}$ is thus a martingale.

Set $n=1$: $\frac{dP_{\widehat{\sigma}}}{d\ln S} = S\frac{dP_{\widehat{\sigma}}}{dS}$ – we get the result that the discounted dollar delta is a martingale.

Now set $n=2$: $\frac{d^2P_{\widehat{\sigma}}}{d\ln S^2} = S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2} + S\frac{dP_{\widehat{\sigma}}}{dS}$. Combining this with the result for $n=1$ implies that the discounted dollar gamma is a martingale as well.

More generally $e^{-r(T-t)}S^n\frac{d^nP_{\widehat{\sigma}}}{dS^n}$ is a martingale for all $n$; this is also true in the Black-Scholes model with deterministic time-dependent volatility.

Thus, the expectation in (5.64) is simply equal to the dollar gamma evaluated at time $t$ for spot $S$. (5.64) then becomes:

$$V(t,S) \; = \; \widehat{\sigma}S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}\int_t^T d\tau$$

Thus

$$\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}} \; = \; S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}\widehat{\sigma}(T-t) \tag{5.66}$$

Going back to (5.62), expressing now $\frac{dP_{\widehat{\sigma}}}{d\widehat{\sigma}}$ in terms of $S^2\frac{d^2P_{\widehat{\sigma}}}{dS^2}$, the condition $P\&L = 0$ translates into:

$$\frac{1}{2}(\sigma_t^2 - \widehat{\sigma}_t^2)\delta t \; + \; (T-t)\widehat{\sigma}_t\delta\widehat{\sigma} \; = \; 0$$

Denote by $Q_t$ the quadratic variation of $\ln S$ realized since $t = 0$: $Q_t = \int_0^t \sigma_\tau^2 d\tau$. Replacing $\sigma_t^2 \delta t$ with $\delta Q_t$, we get, at order one in $\delta t$:[23]

$$\delta\big(\widehat{\sigma}_t^2 (T-t)\big) \;=\; -\delta Q_t$$

$\widehat{\sigma}_t$ is thus given by:

$$\widehat{\sigma}_t^2 \;=\; \frac{\widehat{\sigma}_{t=0}^2 T - Q_t}{T-t} \tag{5.67}$$

Imagine that, over $[0,t]$, the integrated realized volatility matches the initial implied volatility, thus $Q_t = \widehat{\sigma}_{t=0}^2 t$. If we had delta-hedged our option in standard fashion, our P&L at time $t$ would be given by (5.61) with $T \equiv t$. Even though the average realized volatility matches the implied volatility, there is no reason why our theta/gamma P&L would vanish.

In our situation (5.67) shows instead that, if $Q_t = \widehat{\sigma}_{t=0}^2 t$ then $\widehat{\sigma}_t = \widehat{\sigma}_{t=0}$: our option is valued at $t$ with an implied volatility that is equal to its initial value, and we have generated exactly zero carry P&L.

In case $Q_t > \widehat{\sigma}_{t=0}^2 t$ then $\widehat{\sigma}_t < \widehat{\sigma}_{t=0}$: the negative gamma/theta P&L is offset by a positive mark-to-market vega P&L.

We then risk-manage our option by adjusting $\widehat{\sigma}_t$ according to (5.67), generating zero P&L. Two things can happen, according to whether the realized volatility over $[0,T]$ exceeds $\widehat{\sigma}_{t=0}$ or not:

- $Q_T < \widehat{\sigma}_{t=0}^2 T$: from (5.67), $\widehat{\sigma}_T$ is infinite. In the Black-Scholes model, however, the price of a European option is not a separate function of $\widehat{\sigma}$ and $(T-t)$, but a function of $Q = \widehat{\sigma}^2(T-t)$. We will thus use the notation $P_Q$ rather than $P_{\widehat{\sigma}}$. From (5.67), $Q = \widehat{\sigma}_{t=0}^2 T - Q_T$: there is some time value left in our option. Since at $T$ we pay the intrinsic value to the client – given by $P_{Q=0}(T, S_T)$ – we make a net positive P&L given by:

  $$P\&L \;=\; P_{Q=(\widehat{\sigma}_{t=0}^2 T - Q_T)}(T,S_T) - P_{Q=0}(T,S_T)$$

  We have assumed that our option has convex payoff, thus $P\&L \geq 0$.

- $Q_T > \widehat{\sigma}_{t=0}^2 T$: $Q_t$ is a process that starts from zero at $t = 0$ and increases: there exists an intermediate time $\tau$ such that $Q_\tau = \widehat{\sigma}_{t=0}^2 T$. From (5.67), at time $\tau$, $\widehat{\sigma}_\tau = 0$. Moreover, for $t > \tau$, if we kept using (5.67) we would have $\widehat{\sigma}_t^2 < 0$. At $t = \tau$ we thus stop adjusting $\widehat{\sigma}$ and delta-hedge our option over $[\tau, T]$ using $\widehat{\sigma}_\tau$, i.e. zero implied volatility. Our net P&L is then given by an expression similar to (5.61), except it is restricted to the interval $[\tau, T]$ and there is no theta contribution:

  $$P\&L \;=\; -\int_\tau^T \frac{S_u^2}{2}\frac{d^2 P_{\widehat{\sigma}=0}}{dS^2}(u,S_u)\sigma_u^2 dt$$

  Since our option's payoff is convex, this P&L is negative.

[23]In our context $\widehat{\sigma}_t$ is a stochastic process that does not have a diffusive term – it only has a drift. In an expansion at order one in $\delta t$, we only keep terms of order one in $\delta\widehat{\sigma}_t$.

The conclusion is that, unlike what happens with the standard delta-hedging process, with our hedging scheme, the sign of our final P&L is exactly that of the difference between realized and implied volatility or, equivalently, between the realized quadratic variation $Q_T$ and $\widehat{\sigma}^2_{t=0}T$. The magnitude of the P&L is random.

## A.2 Model-independent payoffs based on quadratic variation

With the hedging scheme of the previous section, no P&L is generated until we reach maturity unless $Q_t$ reaches our quadratic variation "budget" $\widehat{\sigma}^2_{t=0}T$. Our P&L at $T$ depends on the remaining quadratic variation "budget" $\widehat{\sigma}^2_{t=0}T - Q_T$. This suggests we could create an exactly hedgeable claim by delivering the payout at a random maturity defined as the time $\tau$ when $Q_\tau$ reaches a pre-specified value $\mathcal{Q}$.

A *timer* option is such an option; it pays a payoff $f(S_\tau)$ at time $\tau$ such that $Q_\tau = \mathcal{Q}$, where $\mathcal{Q}$, called the quadratic variation budget, is specified in the term sheet, in lieu of maturity.[24] From our analysis above, the price of this option – for zero interest rate and repo – is thus simply $P(S, Q, \mathcal{Q}) \equiv \mathcal{P}_{BS}(S, Q; \mathcal{Q})$ where $P_{BS}$ is given by:

$$\mathcal{P}_{BS}(S, Q; \mathcal{Q}) = E\left[f\left(Se^{-\frac{\mathcal{Q}-Q}{2}+\sqrt{\mathcal{Q}-Q}Z}\right)\right] \tag{5.68}$$

where $Z$ is a standard normal variable. It is the Black-Scholes price, calculated with an effective volatility equal to 1 and an effective maturity equal to $\mathcal{Q} - Q$.

The price of a timer option is thus model-independent, as by following the delta-hedging outlined above, we replicate exactly the option's payoff, no matter what the realized volatility is. Physical time does not enter the pricing function.

Timer option prices do not depend on market implied volatilities – this makes timer options attractive for underlyings that lack a liquid options market. Carrying a naked gamma/theta position, in the case of a standard option, is too risky, even though we may have selected a conservative implied volatility. In the timer version, the carry P&L – see below – is smaller. This enables trades that would not be contemplated otherwise.

Timer options effectively started trading around 2007, but the idea of replacing physical time with quadratic variation and designing payoffs that are not sensitive to volatility assumptions well predates timer options – see Avi Bick's 1995 article [14].

In practice, quadratic variation is measured using log returns of daily closes $S_i$:

$$Q_{i+1} = Q_i + \ln^2\left(\frac{S_{i+1}}{S_i}\right)$$

[24] While Société Générale has marketed these options under the name of *timer* options they have also been traded under the name of *mileage* options.

Are there other model-independent payoffs involving the quadratic variation? Write the price of such an option as $P(t,S,Q)$. The P&L during $\delta t$ of a short, delta-hedged position reads, at order one in $\delta t$ and two in $\delta S$:

$$P\&L = -\frac{dP}{dt}\delta t - \frac{S^2}{2}\frac{d^2P}{dS^2}\frac{\delta S^2}{S^2} - \frac{dP}{dQ}\delta Q \tag{5.69a}$$

$$= -\frac{dP}{dt}\delta t - \left[\frac{S^2}{2}\frac{d^2P}{dS^2} + \frac{dP}{dQ}\right]\frac{\delta S^2}{S^2} \tag{5.69b}$$

where, by definition, $\delta Q = \frac{\delta S^2}{S^2}$. This P&L vanishes if the following conditions hold:

$$\frac{dP}{dt} = 0 \tag{5.70a}$$

$$\frac{S^2}{2}\frac{d^2P}{dS^2} + \frac{dP}{dQ} = 0 \tag{5.70b}$$

(5.70a) expresses that $P$ does not depend on physical time. (5.70b) is identical to the Black-Scholes equation, except time is replaced with quadratic variation; we only need to supplement (5.70b) with the terminal profile $P(S,\mathcal{Q})$, where $\mathcal{Q}$ is the quadratic variation budget.

Thus, in addition to the "European" timer option discussed above whose price is given by (5.68), many familiar payoffs have a *timer* counterpart.

We can define a timer barrier option that pays $f(S_\tau)$ when $Q_\tau$ reaches $\mathcal{Q}$ unless $S$ hits a barrier $B$ in which case the option pays a rebate $R(Q_\tau)$. This type of payoff will come in handy when we construct an upper bound for prices of options on realized variance – see Section 7.6.10. The barrier can also be made a function of $Q$.

Path-dependent variables can be used except they are not allowed to involve physical time.[25]

Quadratic variation does not need to accrue uniformly; we can choose to weight realized variance by a function of $S$: $\delta Q_t = \mu(S)\sigma_t^2\delta t$; for example realized variance is not counted whenever $S$ lies above or below a given threshold. (5.70b) is then replaced with:

$$\frac{S^2}{2}\frac{d^2P}{dS^2} + \mu(S)\frac{dP}{dQ} = 0$$

Selecting $\mu(S) = S^2$ corresponds to accruing the quadratic variation of $S$, rather than $\ln S$.

[25]In timer options we make a stochastic time change from physical time to quadratic variation: $t \to Q_t = \int_0^t \sigma_u^2 du$ and from $S_t$ to $S_Q^\star$ defined by: $S_{Q_t}^\star = S_t$; $S_Q^\star$ is lognormally distributed with a quadratic variation of $\ln S^\star$ equal to $Q$.

Consider a path-dependent variable $f$ that is a function of the path of $S_t$ over $[0,T]$, which we denote by $[S_t]_0^T$. A timer option whose payoff involves $f$ remains model-independent only if the condition: $f([S_t]_0^T) = f([S_Q^\star]_0^{Q_T})$ holds.

While the (continuously sampled) $\min_t S_t$ and $\max_t S_t$ satisfy this condition, the Asian average $M_t = \frac{1}{t}\int_0^t S_u du$ does not.

We can also define multi-asset timer options. For example define $Q$ as the quadratic variation of $\ln(\frac{S_2}{S_1})$. An option paying $S_1 f\left(\frac{S_2}{S_1}\right)$ when $Q$ hits $\mathcal{Q}$ is model-independent.

This is easily shown by using $S_1$ as numeraire: the value of all assets, including the value of our timer option as well as $S_2$ are expressed in units of $S_1$. $Q$ is then simply the quadratic variation of $S_2$ expressed in units of $S_1$ – in these new units this option becomes a standard timer option on $S_2$.

It can be shown that these are the only model-independent options that are functions of $S_1, S_2, Q$.

Finally, model-independent payoffs whose prices do not depend on physical time can also be created using $\min_\tau S_\tau$ and $\max_\tau S_\tau$ rather than quadratic variation – see [31].

## A.3 How model-independent are timer options?

Our analysis above applies to the ideal situation of real-time delta-hedging, and for a continuous process for $S$ – hence the expansion at order two in $\delta S$ and order one in $\delta t$ for $\delta Q_t$. What about the real case?

In standard options delta-hedging offsets the directional position on $S$; our P&L starts with terms of order two in $\delta S$ whose contribution, as illustrated in Section 1.2, is not small.

One order is gained with timer options: condition (5.70b) ensures that the $\delta S^2$ term in the gamma P&L is offset by a corresponding change of the quadratic variation. Our P&L now starts with terms of order three.

Consider a "European" timer option, that is with no path-dependence. $P(Q, S)$ solves equation (5.70b). In typical term sheets of timer options, quadratic variation is defined as the sum of squared *log-returns:* $\delta Q = \ln\left(1 + \frac{\delta S}{S}\right)^2$. Expanding $P(S, Q)$ at order 3 in $\delta S$, setting $\delta Q = \frac{\delta S^2}{S^2} - \frac{\delta S^3}{S^3}$ and using (5.70b):

$$\begin{aligned} P\&L &= -\frac{S^2}{2}\frac{d^2P}{dS^2}\frac{\delta S^2}{S^2} - \frac{dP}{dQ}\delta Q - \frac{S^3}{6}\frac{d^3P}{dS^3}\frac{\delta S^3}{S^3} - \frac{d^2P}{dQdS}\delta Q\delta S \\ &= \left(\frac{S^2}{2}\frac{d^2P}{dS^2} + \frac{S^3}{3}\frac{d^3P}{dS^3}\right)\frac{\delta S^3}{S^3} \end{aligned} \tag{5.71}$$

Consider a timer call option. For spot values near the strike, the prefactor in (5.71) is dominated by $\frac{S^2}{2}\frac{d^2P}{dS^2}$: $P\&L \simeq \frac{S^2}{2}\frac{d^2P}{dS^2}\frac{\delta S^3}{S^3}$, to be compared with the gamma P&L of the standard, non-timer, option: $-\frac{S^2}{2}\frac{d^2P}{dS^2}\frac{\delta S^2}{S^2}$. The prefactors are identical; the P&L in the timer version is smaller by a factor $\frac{\delta S}{S}$.

Timer options are thus less risky than their non-timer counterparts but are in practice not fully model-independent.[26] Mathematically, "model independent" means

[26] Timer options on single stocks – which can experience large overnight drawdowns – are particularly risky.

model-independent as long as the process for $S_t$ is a continuous semimartingale. Practically, the meaning of "model-independent" is that the P&L vanishes up to order two in $\delta S$.

We have assumed so far vanishing interest rate and repo. The essence of timer options is that quadratic variation replaces physical time. Financing costs/benefits, however, are paid/received *prorata temporis*: physical time re-enters the picture. Taking into account interest rate and repo, the P&L in (5.69) now reads:

$$\begin{aligned} P\&L &= -\frac{dP}{dt}\delta t - \frac{S^2}{2}\frac{d^2P}{dS^2}\frac{\delta S^2}{S^2} - \frac{dP}{dQ}\delta Q + rP\delta t - (r-q)S\frac{dP}{dS}\delta t \\ &= -\left[\frac{dP}{dt} - rP + (r-q)S\frac{dP}{dS}\right]\delta t - \left[\frac{S^2}{2}\frac{d^2P}{dS^2} + \frac{dP}{dQ}\right]\frac{\delta S^2}{S^2} \end{aligned}$$

The conditions ensuring "model-independence" are now:

$$\frac{S^2}{2}\frac{d^2P}{dS^2} + \frac{dP}{dQ} = 0 \quad (5.72a)$$

$$\frac{dP}{dt} - rP + (r-q)S\frac{dP}{dS} = 0 \quad (5.72b)$$

(5.72b) implies that $P(t, S, Q)$ has the following form:

$$P(t, S, Q) = e^{rt} p\big(Se^{-(r-q)t}, Q\big)$$

(5.72a) yields the following condition for $p(x, Q)$:

$$\frac{x^2}{2}\frac{d^2p}{dx^2} + \frac{dp}{dQ} = 0$$

Setting the terminal condition $p(x, Q = \mathcal{Q}) = f(x)$ then fully determines $p$. At time $\tau$, when $Q_\tau$ reaches $\mathcal{Q}$, we pay the amount $e^{r\tau} f(S_\tau e^{-(r-q)\tau})$ to the client. Commercially, this is less attractive than simply paying $f(S_\tau)$.

Select a (distant) maturity $T$ and redefine $f$ as: $f(x) \to e^{-rT} f(xe^{(r-q)T})$. The payoff at $\tau$ is then $e^{-r(T-\tau)} f(S_\tau e^{(r-q)(T-\tau)})$, which is equivalent to settling the payoff $f(F_\tau^T)$ at $T$, where $F_\tau^T$ is the forward at time $\tau$, spot $S_\tau$, for maturity $T$. $P(t, S, Q)$ is given by:

$$P(t, S, Q) = e^{-r(T-t)} \mathcal{P}_{BS}(Se^{(r-q)(T-t)}, Q; \mathcal{Q}) \quad (5.73)$$

where $P_{BS}$ is defined in (5.68). This specification of a timer option remains in fact model-independent when interest rates are stochastic.

Thus, for non-vanishing interest rate and repo, model-independent payoffs still exist, but become somewhat convoluted.

Equivalently, a standard timer option now acquires a spurious sensitivity to realized – or implied – volatility, as this volatility determines the duration over which financing costs are paid/received. For example, in the Black-Scholes model, with an implied volatility $\widehat{\sigma}$, the price of a standard vanilla timer option becomes:

$$P(t,S,Q) \;=\; e^{-r\frac{\mathcal{Q}-Q}{\widehat{\sigma}^2}}\mathcal{P}_{BS}(Se^{(r-q)\frac{\mathcal{Q}-Q}{\widehat{\sigma}^2}},Q;\mathcal{Q})$$

$P$ explicitly depends on $\widehat{\sigma}$ thus is not model-independent anymore.[27]

Beside a reserve that covers third-order terms in the P&L and an adjustment to account for interest rate, repo and dividends, two additional corrections to the model-independent price are needed.

The final quadratic variation is always larger than the allotted budget $\mathcal{Q}$ as, usually, the expiry of the timer option is defined as the first day when $Q_\tau$ exceeds $\mathcal{Q}$: this overshoot needs to be factored in the price.

Also, term sheets of timer options specify a maximum maturity $T_{\max}$, typically $T_{\max} = 2\frac{\mathcal{Q}}{\widehat{\sigma}^2}$ where $\widehat{\sigma}$ is a reference volatility. The corresponding price adjustment is very model-dependent.

## A.4 Leveraged ETFs

Consider a leveraged ETF (leveraged exchange traded fund [LETF]). The fund's strategy consists in investing in a security $S$, with a fixed leverage $\beta$. Ignoring borrowing costs, we would expect the performance of our investment $I$, over a given time horizon, to be $\beta$ times the performance of $S$: $\ln\left(I_T/I_0\right) \simeq \beta \ln\left(S_T/S_0\right)$.

At time $t$ we need to hold $\beta\frac{I_t}{S_t}$ units of $S$. $\frac{I_t}{S_t}$ will keep changing – except in the uninteresting case $\beta = 1$ . We thus need to rebalance our position – in the case of LETFs on a daily basis. This rebalancing is analogous to the readjustment of the delta of an option position and likely exposes us to realized volatility.

Let $I_t$ be the fund net asset value (NAV) at time $t$ and $S_t$ the security it invests in. As the fund manager, we borrow the amount $\beta I_t$, which we invest in $S$, while accruing interest on the amount $I_t$. Over $[t, T+\delta t]$, our return $r_I = \frac{I_{t+\delta t}}{I_t} - 1$ is given by:

$$r_I \;=\; \beta\left(r_S - (r-q)\delta t\right) + r\delta t$$

where $r_S = \frac{S_{t+\delta t}}{S_t} - 1$. Expanding at order two in $\delta \ln S$ and one in $\delta t$ the log-return $\delta \ln I = \ln(1 + r_I)$ is given by:

$$\begin{aligned}
&\delta \ln I \\
&= \ln(1 + \beta\left(e^{\delta \ln S} - 1 - (r-q)\delta t\right) + r\delta t) \\
&= \beta\delta \ln S + \left(r - \beta(r-q)\right)\delta t - \frac{\beta(\beta-1)}{2}\delta \ln S^2
\end{aligned}$$

[27]Because it is generated by rate and repo sensitivities, the vega of a timer option is very unlike that of its non-timer counterpart.

Keeping terms up to $\delta \ln S^2$ :

$$\delta \ln I \;=\; \beta \delta \ln S + (r - \beta(r-q))\,\delta t - \frac{\beta(\beta-1)}{2}\delta Q$$

where $\delta Q = \frac{\delta S^2}{S^2}$. This yields:

$$I_t \;=\; I(t, S_t, Q_t) \tag{5.74}$$

$$I(t,S,Q) \;=\; I_0 e^{rt}\left(\frac{S}{S_0 e^{(r-q)t}}\right)^{\beta} e^{-\frac{\beta(\beta-1)}{2}Q} \tag{5.75}$$

$I_t$ is the result of a pure delta strategy: (5.74) shows that the fund's NAV replicates (up to second order in $\delta S$) the (exotic) payoff $I(t, S_t, Q_t)$ that involves both $S_t$ and its realized variance – starting from NAV $I_0$ at $t = 0$.[28] The reader can check that $I(t, S, Q)$ indeed satisfies conditions (5.72): the LETF is a perpetual contract, with no quadratic variation budget $\mathcal{Q}$.

Set a maturity $T$: the payoff $I_T$ at $T$ can be generated out of an initial investment $I_0$. Equivalently the payoff $\frac{I_T}{I_0} - e^{rT}$ at $T$ can be synthesized with zero initial cash. Multiply by $e^{-rT}$ and take the limit $\beta \to 0$:

$$\lim_{\beta\to 0}\frac{1}{\beta}\left(e^{-rT}\frac{I_T}{I_0} - 1\right) \;=\; \ln\left(\frac{S_T}{S_0 e^{(r-q)T}}\right) + \frac{Q_T}{2} \tag{5.76}$$

The package in the right-hand side of (5.76) is replicated with zero initial cash. Equivalently, the value at $t = 0$ of a payoff of maturity $T$ that delivers either $-\ln\left(\frac{S_T}{S_0 e^{(r-q)T}}\right)$ or $\frac{Q_T}{2}$ is equal – hence the replication strategy of the VS once again.

Finally, options on LETFs exist as well. Unlike LETFs, LETF options are highly model-dependent.

## Appendix B – perturbation of the lognormal distribution

In the Black-Scholes model implied volatilities are flat. A smile appears whenever the distribution of $S$ is not lognormal, such as in stochastic volatility or jump/Lévy models.

These models collapse onto the Black-Scholes model when a given parameter – say volatility of volatility or jump size – vanishes. It is thus possible to carry out an

[28]For typical values of $\beta$, such as $\beta = 2$ or $\beta = -2$, the contribution from realized variance impacts negatively the performance of the ETF, more so for negative values of $\beta$ – an aspect of LETF trading that some investors seem to have overlooked.

expansion in powers of this parameter, around the Black-Scholes case. In Chapter 8 we go through such an expansion for general forward variance models.

Here we consider the general case of a distribution of $S_T$ which we assume to be slightly non-lognormal. Our aim is to derive the expansion of European option prices at order one in the parameters that quantify the non-lognormality of $S_T$ – the cumulants of $\ln S_T$.

While this idea is not new (see for example [5]), it is important to ensure that, at the chosen order in the expansion, some quantities stay fixed. The constraint that the forward $F_T = E[S_T]$ should be unchanged is typically enforced.

In the context of forward variance models, forward variances $\xi_0^\tau$ are underlyings in their own right whose values should be left unchanged as well.

We thus require that prices of log contracts – hence VS volatilities – be unaffected in the expansion.

Let us start with a lognormal density for $S_T$. The density $\rho_0$ of $z = \ln(\frac{S_T}{F_T})$, where $F_T$ is the forward for maturity $T$, is given by:

$$\rho_0\left(z\right) \;=\; \frac{1}{\sqrt{2\pi\Sigma^2}}e^{-\frac{(z-\mu)^2}{2\Sigma^2}}$$

where $\mu$, $\Sigma$ are the average and standard deviation of the unperturbed lognormal density whose volatility we denote by $\widehat{\sigma}_0$; namely:

$$\begin{aligned}\Sigma &= \widehat{\sigma}_0\sqrt{T}\\ \mu &= -\frac{\widehat{\sigma}_0^2 T}{2} \;=\; -\frac{\Sigma^2}{2}\end{aligned}$$

Let us denote by $L\left(q\right)$ the logarithm of the characteristic function of a given density $\rho$:

$$L\left(q\right) \;=\; \ln\left(\int_{-\infty}^{+\infty} e^{-qz}\rho\left(z\right)dz\right) \tag{5.77}$$

$L\left(q\right)$ is called the cumulant-generating function. For the normal density $\rho_0$, $L_0\left(q\right)$ reads:

$$L_0\left(q\right) \;=\; -\mu q + \frac{\Sigma^2}{2}q^2 \;=\; \frac{\Sigma^2}{2}\left(q+q^2\right) \tag{5.78}$$

$L_0\left(q\right)$ is a polynomial of order 2 – this is a distinguishing feature of normal distributions. For a general density $\rho$, cumulants $\kappa_n$ are defined as the coefficients of the Taylor expansion of $L\left(q\right)$ around $q=0$ – when it exists:

$$L\left(q\right) \;=\; \sum_{n=1}^{\infty}\frac{\left(-1\right)^n}{n!}\kappa_n q^n$$

One can check by taking derivatives of $L(q)$ in (5.77) and evaluating them at $q = 0$ that the first and second cumulants are related respectively to the mean and variance of $\rho$:

$$\kappa_1 = \overline{z}, \;\; \kappa_2 = \overline{(z-\overline{z})^2}$$

where $\overline{f}$ denotes $E[f]$. Likewise, the third and fourth cumulants are related to centered moments of $\rho$:

$$\kappa_3 = \overline{(z-\overline{z})^3}, \quad \kappa_4 = \overline{(z-\overline{z})^4} - 3\overline{(z-\overline{z})^2}^2$$

Since cumulants $\kappa_n$ for $n > 2$ vanish for a normal distribution, it is natural to consider a small perturbation $\delta\kappa_n$ of the cumulants of $L_0(q)$ for $n \geq 3$:

$$L(q) \;=\; L_0(q) + \sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \delta\kappa_n q^n \tag{5.79}$$

We now derive the perturbed density $\rho = \rho_0 + \delta\rho$ at order one in the $\delta\kappa_n$.

## B.1 Perturbing the cumulant-generating function

While we perturb $\rho_0$ to generate a smile for implied volatilities, we wish to keep the forward unchanged. By definition of $z$:

$$E[S_T] \;=\; F_T E[e^z] \;=\; F_T e^{L(-1)}$$

which imposes the constraint $L(-1) = 0$. Inspection of (5.78) shows that $L_0(p)$ obviously obeys this condition.

Imagine that only one $\delta\kappa_n$ – say $\delta\kappa_3$ – is non-vanishing. Then the constraint $L(-1) = 0$ cannot be accommodated unless we shift cumulants of order 1 and 2 by an amount proportional to $\delta\kappa_3$. We then rewrite (5.79) as:

$$L(q) \;=\; L_0(q) \;-\delta\kappa_1 q + \delta\kappa_2 \frac{q^2}{2} \;+\; \sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \delta\kappa_n q^n$$

Several choices are possible:

- Take $\delta\kappa_1 \neq 0$, $\delta\kappa_2 = 0$: this is the choice typically made in the literature, thereby translating the distribution of $z$ by an amount $\delta\kappa_1$ given by $\delta\kappa_1 = -\sum_{n=3}^{\infty} \frac{1}{n!}\delta\kappa_n$ and leaving the standard deviation of $z$ unchanged.[29] This results in the following expression for $L(q)$:

$$L(q) \;=\; L_0(q) + \sum_{n=3}^{\infty} \frac{\delta\kappa_n}{n!} \left((-1)^n q^n + q\right)$$

[29]Choosing not to alter $\kappa_2$ has little financial motivation since, in the presence of a smile, the standard deviation of $\ln S_T$ is not related simply to implied volatilities of vanilla options.

- Take $\delta\kappa_1 = 0$, $\delta\kappa_2 = -2\sum_{n=3}^{\infty}\frac{1}{n!}\delta\kappa_n$. This yields:

$$L(q) \;=\; L_0(q) + \sum_{n=3}^{\infty}\frac{\delta\kappa_n}{n!}\left((-1)^n q^n - q^2\right) \tag{5.80}$$

- Take both $\delta\kappa_1 \neq 0$, $\delta\kappa_2 \neq 0$:

$$L(q) \;=\; L_0(q) + \sum_{n=3}^{\infty}\frac{\delta\kappa_n}{n!}\left((-1)^n q^n - q^2 \;+\; \theta_n\left(q+q^2\right)\right) \tag{5.81}$$

where the $\theta_n$ are arbitrary.

## B.2 Choosing a normalization and generating a density

In diffusive models the VS volatility for maturity $T$ is equal to the implied volatility of the log contract and is given by:

$$\widehat{\sigma}^2_{\mathrm{VS},T}T \;=\; \widehat{\sigma}^2_T T \;=\; E[-2\ln\left(S_T/F_T\right)]$$

From the definition (5.77) of the cumulant-generating function, we get:

$$\widehat{\sigma}^2_T \;=\; \frac{2}{T}\left.\frac{dL}{dq}\right|_{q=0} \;=\; -\frac{2}{T}\kappa_1 \tag{5.82}$$

Thus, keeping $\kappa_1$ unchanged guarantees that the log-contract implied volatility is unchanged. For diffusive models, this guarantees that $\widehat{\sigma}_{\mathrm{VS},T}$ is also unchanged.

This is a very desirable feature for stochastic volatility models: forward VS variances are underlyings whose initial values should be left unchanged in a perturbation of the model's parameters. We then choose expression (5.80) for $L(q)$.

It is a classical result that a general perturbation in the cumulants of $L_0(q)$:

$$L(q) \;=\; L_0(q) + \sum_{n=1}^{\infty}\frac{\delta\kappa_n}{n!}q^n$$

translates at order one in the $\delta\kappa_n$ into the following perturbation of the density:

$$\begin{aligned}\rho(z) \;&=\; \rho_0(z) + \delta\rho(z)\\ \delta\rho(z) \;&=\; \sum_{n=1}^{\infty}\frac{\delta\kappa_n}{\Sigma^n\sqrt{n!}}H_n\Big(\frac{z-\mu}{\Sigma}\Big)\rho_0(z)\end{aligned} \tag{5.83}$$

where the $H_n$ are a family of orthogonal polynomials – the Hermite polynomials – defined by:

$$H_n(z) = \frac{(-1)^n}{\sqrt{n!}}e^{\frac{z^2}{2}}\frac{d^n}{dz^n}\left(e^{-\frac{z^2}{2}}\right) \tag{5.84}$$

with the following properties:

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}H_n(z)\,H_m(z)\,dz &= \delta_{nm}\\ \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}H_n(z)\,e^{-qz}dz &= \frac{(-1)^n q^n}{\sqrt{n!}}e^{\frac{q^2}{2}}\end{aligned}$$

Using (5.84), (5.83) can be rewritten in a simpler form – this is the Gram-Charlier formula:

$$\delta\rho(z) = \sum_{n=1}^{\infty}\frac{\delta\kappa_n}{n!}(-1)^n\frac{d^n\rho_0(z)}{dz^n}$$

Choosing now normalization (5.80) results in the following expression of $\delta\rho$ at order one in $\delta\kappa_n$, $n \geq 3$:

$$\delta\rho(z) = \sum_{n=3}^{\infty}\frac{\delta\kappa_n}{n!}\left((-1)^n\frac{d^n}{dz^n} - \frac{d^2}{dz^2}\right)\rho_0(z) \tag{5.85}$$

Normalization (5.81) would have yielded:

$$\delta\rho(z) = \sum_{n=3}^{\infty}\frac{\delta\kappa_n}{n!}\left((-1)^n\frac{d^n}{dz^n} - \frac{d^2}{dz^2} + \theta_n\left(\frac{d^2}{dz^2}+\frac{d}{dz}\right)\right)\rho_0(z) \tag{5.86}$$

The perturbed density is $\rho = \rho_0 + \delta\rho$.

## B.3 Impact on vanilla option prices and implied volatilities

Denote by $\delta P$ the perturbation in the price of a vanilla option generated by $\delta\rho$. Its payoff – for a call - is given by:

$$(S_T - K)^+ = K\left(\frac{S_T}{K}-1\right)^+ = Kf\left(\frac{F_T e^z}{K}\right)$$

where $f(x) = (x-1)^+$.

Starting from the expression of the price $P_0$ in the Black-Scholes model:

$$P_0 = e^{-rT}\int_{-\infty}^{\infty}\rho_0(z)\,Kf\left(\frac{F_T e^z}{K}\right)dz = e^{-rT}\int_{-\infty}^{\infty}\rho_0(u-\ln F_T)\,Kf\left(\frac{e^u}{K}\right)du$$

and taking the derivative with respect to $\ln S$ – remembering that $F_T = Se^{(r-q)T}$ – yields:

$$\frac{d^nP_0}{d\ln S^n} = \frac{d^nP_0}{d\ln F_T^n} = e^{-rT}\int_{-\infty}^{\infty}(-1)^n\frac{d^n\rho_0}{dz^n}(z)\,Kf\left(\frac{F_T e^z}{K}\right)dz \tag{5.87}$$

The perturbation $\delta\rho$ in (5.85) then results in a price variation $\delta P$ given by:

$$\delta P \;=\; \sum_{n=3}^{\infty} \frac{\delta\kappa_n}{n!}\left(\frac{d^n}{d\ln S^n} - \frac{d^2}{d\ln S^2}\right) P_0 \tag{5.88}$$

If we do not require that the implied volatility of the log contract be unchanged and use formula (5.81) for $\delta\rho$ we get:

$$\delta P \;=\; \sum_{n=3}^{\infty} \frac{\delta\kappa_n}{n!}\left[\left(\frac{d^n}{d\ln S^n} - \frac{d^2}{d\ln S^2}\right) + \theta_n\left(\frac{d^2}{d\ln S^2} - \frac{d}{d\ln S}\right)\right] P_0 \tag{5.89}$$

To obtain the perturbation of implied volatilities $\delta\widehat{\sigma}$ at order one in the $\delta\kappa_n$ simply divide $\delta P$ by the option's vega $\frac{dP_0}{d\widehat{\sigma}}$, which is related to its dollar gamma through equation (5.66):

$$\frac{dP_0}{d\widehat{\sigma}} \;=\; S^2\frac{d^2P_0}{dS^2}\widehat{\sigma}T \;=\; \left(\frac{d^2P_0}{d\ln S^2} - \frac{dP_0}{d\ln S}\right)\widehat{\sigma}T$$

Equation (5.89) translates into:

$$\delta\widehat{\sigma} \;=\; \frac{1}{\widehat{\sigma}_0 T}\sum_{n=3}^{\infty}\frac{\frac{\delta\kappa_n}{n!}\left(\frac{d^nP_0}{d\ln S^n} - \frac{d^2P_0}{d\ln S^2}\right)}{\frac{d^2P_0}{d\ln S^2} - \frac{dP_0}{d\ln S}} \;+\; \frac{1}{\widehat{\sigma}_0 T}\sum_{n=3}^{\infty}\frac{\delta\kappa_n}{n!}\theta_n \tag{5.90}$$

As (5.90) makes it plain, choosing $\theta_n \neq 0$ has the effect of simply adding a constant shift to $\delta\widehat{\sigma}$, independent of the option's strike.

## B.4 The ATMF skew

We now derive the expression of the ATMF skew $\mathcal{S}_T = \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{F_T}$ at order one in the $\delta\kappa_n$. Because the second piece in the right-hand side of (5.90) generates a uniform shift of implied volatilities, it does not contribute to $\mathcal{S}_T$, hence $\mathcal{S}_T$ does not depend on the $\theta_n$.

Starting from (5.90), using (5.87) to express derivatives of $P_0$ in terms of derivatives of $\rho_0$, and using the Black-Scholes expression of the gamma yields, after some tedious algebra, the following result:

$$\mathcal{S}_T \;=\; \frac{1}{\sqrt{T}}\sum_{n=3}^{\infty}\frac{\delta\kappa_n}{n!}\,\frac{\int_0^{\infty}\Big((-1)^n\Big(\frac{1}{2}\frac{d^n\rho_0}{dz^n} + \frac{d^{n+1}\rho_0}{dz^{n+1}}\Big) - \Big(\frac{1}{2}\frac{d^2\rho_0}{dz^2} + \frac{d^3\rho_0}{dz^3}\Big)\Big)(e^z-1)dz}{\Sigma\rho_0\left(0\right)} \tag{5.91}$$

Remember that $\Sigma$ is not the volatility, but the unperturbed standard deviation of $\ln S_T$: $\Sigma = \widehat{\sigma}_0\sqrt{T}$.

Using formula (5.91) for $\mathcal{S}_T$ at order one in the $\delta\kappa_n$ is advantageous in situations when it is easier to approximately compute cumulants than carry out an expansion in powers of a given parameter. Let us concentrate on the contribution of $\delta\kappa_3$ and $\delta\kappa_4$. They are usually expressed in terms of the skewness $s$ and kurtosis $\kappa$ of $\delta\rho$:

$$\delta\kappa_3 = s\Sigma^3, \, \delta\kappa_4 = \kappa\Sigma^4$$

Straightforward, though tedious, calculation of the numerator in (5.91) yields:

$$\mathcal{S}_T \;=\; \frac{1}{\sqrt{T}}\left(\frac{s}{6} \;+\; \frac{\kappa}{12}\widehat{\sigma}_0\sqrt{T} \;+\; \cdots\right) \tag{5.92}$$

This recovers the result in [5].

If we only consider the contribution of the third-order cumulant, $\delta P$ is given by:

$$\delta P \;=\; \frac{\delta\kappa_3}{6}\left[\left(\frac{d^3}{d\ln S^3} - \frac{d^2}{d\ln S^2}\right) + \theta_3\left(\frac{d^2}{d\ln S^2} - \frac{d}{d\ln S}\right)\right] P_0$$

with $\theta_3 = 0$ in case we keep the log contract implied volatility unchanged. (5.92) supplies the following simple relationship relating the ATMF *skew* to the *skewness* of $\ln S_T$:

$$\mathcal{S}_T \;=\; \frac{s}{6\sqrt{T}} \tag{5.93}$$

We have focused on the ATMF skew, but could have derived from (5.90) an expression for the ATMF implied volatility as well.

While approximations of the ATMF volatility at order one in $\delta\kappa_3$ and $\delta\kappa_4$ are usually insufficiently accurate for practical use, relationship (5.93) is remarkably robust – presumably because it does not involve any volatility reference level, and only depends on skewness – a dimensionless number.

## Chapter's digest

### 5.1 Variance swap forward variances

▶ Variance swap volatilities are strikes of variance swap contracts. $\widehat{\sigma}_{\mathrm{VS},T}(t)$ is such that, at inception, a contract delivering at $T$ the following payoff is worth zero.

$$\frac{1}{T-t}\sum_{i=0}^{N-1}\ln^2\left(\frac{S_{i+1}}{S_i}\right) \;-\; \widehat{\sigma}^2_{\mathrm{VS},T}(t)$$

Forward VS variances $\xi_t^T$ are defined as:

$$\xi_t^T \;=\; \frac{d}{dT}\left((T-t)\,\widehat{\sigma}^2_{\mathrm{VS},T}(t)\right)$$

They are positive and driftless.

❧ ❧ ❧ ❧ ❧

### 5.2 Relationship of variance swaps to log contracts

▶ Up to second order in $S_{i+1}-S_i$, VSs can be replicated by delta-hedging a static position in European payoff $-2\ln S_T$, called the log contract. In the absence of cash-amount dividends, log-contract, hence VS, implied volatilities are given by formula (5.17):

$$\begin{aligned}\widehat{\sigma}^2_{\mathrm{VS},T} \;&=\; \int_{-\infty}^{+\infty} dy\, \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\widehat{\sigma}^2_{K(y)T}\\ y(K) \;&=\; \frac{\ln\left(\frac{F_T}{K}\right)}{\widehat{\sigma}_{KT}\sqrt{T}} - \frac{\widehat{\sigma}_{KT}\sqrt{T}}{2}\end{aligned}$$

❧ ❧ ❧ ❧ ❧

### 5.3 Impact of large returns

▶ The property that a delta-hedged log contract replicates the payoff of a VS holds at second order in $S_{i+1}-S_i$.

▶ In a diffusive model, the VS volatility $\widehat{\sigma}_{\mathrm{VS},T}$ and the log-contract implied volatility $\widehat{\sigma}_T$ match: any diffusive model calibrated to a given market smile yields the same value for $\widehat{\sigma}_{\mathrm{VS},T}$.

▶ In a jump-diffusion model, the difference between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ is given by: (5.28):

$$\widehat{\sigma}^2_{\mathrm{VS},T} - \widehat{\sigma}^2_T \;=\; \lambda\,\overline{\ln^2(1+J) + 2\ln(1+J) - 2J}$$

This is expressed, as a function of the ATMF skew of the smile for maturity $T$, $\mathcal{S}_T$, at order one in $\mathcal{S}_T$, through (5.32):

$$\widehat{\sigma}_{\mathrm{VS},T} \simeq \widehat{\sigma}_T \left(1 - \widehat{\sigma}_T \mathcal{S}_T T\right)$$

The difference between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ is independent on $T$ only if $\mathcal{S}_T$ decays as $\frac{1}{T}$.

▶ In model-free fashion the difference between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ is given by (5.38), which involves the skewness $s_{\Delta t}$ of daily returns.

$$\frac{\widehat{\sigma}_{\mathrm{VS},T}}{\widehat{\sigma}_T} - 1 \simeq -\frac{s_{\Delta t}}{6}\widehat{\sigma}_T\sqrt{\Delta t}$$

▶ Inferring the skewness of returns at short time scales from market smiles is very model-dependent and unreasonable.

▶ The realized skewness of daily returns is such that the adjustment it warrants for $\widehat{\sigma}_{\mathrm{VS},T}$ is minute. The implied value of this adjustment, however, could be arbitrarily large.

▶ Are there variance payoffs that can be exactly replicated, even for large returns? There is only one, whose payoff is:

$$\sum_i (S_{i+1} - S_i)^2$$

It is replicated by delta-hedging a parabolic profile.

## 5.4 Impact of strike discreteness

▶ The fact that, in practice, only discrete strikes – rather than continuous ones – can be traded, further adds to the imperfection of the replication of the VS.

## 5.5 Conclusion

▶ In practice one can take $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ to be equal, in effect setting $\widehat{\sigma}_T = \widehat{\sigma}_{\mathrm{VS},T}$. VS and log-contract forward variances are identical objects. The model is simulated according to SDE (5.43):

$$\begin{cases} dS_t = \sqrt{\xi_t^t} S_t dW_t^S \\ d\xi_t^T = \lambda_t^T dW_t^T \end{cases}$$

▶ For very liquid indexes, for which $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ are both observable, a practical solution is to stay within a diffusive model, driven by SDE (5.43). The spread between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$ is taken into account by adjusting the realized variance according to:

$$\ln^2\left(\frac{S_{i+1}}{S_i}\right) \rightarrow \ln^2\left(\frac{S_{i+1}}{S_i}\right) + (\lambda\Delta)\,\overline{\ln^2(1+J)+2\ln(1+J)-2J}$$

and adjusting the implied realized variance – or the forward VS variance – according to:

$$\zeta_t^T = \xi_t^T + \lambda\overline{\ln^2(1+J)+2\ln(1+J)-2J}$$

when the payoff calls for observation of implied VS volatilities. $\lambda$, $J$ are chosen to match market values for the spread between $\widehat{\sigma}_{\mathrm{VS},T}$ and $\widehat{\sigma}_T$. The term structure of this spread is captured by making $\lambda$ time-dependent.

❧❧❧❧❧

### 5.6 Dividends

▶ The impact of dividends on the VS payoff itself is zero for stocks, whose returns are corrected for the dividend impact, and minute for indexes.

▶ Fixed cash-amount dividends impact the replication of VSs. The log contract is supplemented with additional European payoffs with maturities matching the dividend schedule.

❧❧❧❧❧

### 5.7 Pricing variance swaps with a PDE

▶ VS volatilities for indexes are most easily calculated by solving PDE (5.49):

$$\frac{dU}{dt} + (r-q)\,S\frac{dU}{dS} + \frac{\sigma^2(t,S)}{2}S^2\frac{d^2U}{dS^2} = -\sigma^2(t,S)$$

▶ The adjustment for large returns is performed by adding an extra term – one solves PDE (5.53):

$$\frac{dU}{dt} + (r-q)\,S\frac{dU}{dS} + \frac{\sigma^2(t,S)}{2}S^2\frac{d^2U}{dS^2} = -\left(\sigma^2(t,S) - \frac{1}{3}\varepsilon J^3\right)$$

❧❧❧❧❧

### 5.8 Interest-rate volatility

▶ $\widehat{\sigma}_T$ is the implied volatility of the log-contract, a European payoff. Thus it really is the implied volatility of the forward for maturity $T$, $F_t^T$. VSs, on the other hand, pay the realized variance of $S_t$. Interest-rate volatility creates a difference between realized volatilities of $S_t$ and $F_t^T$ – at order one in interest-rate volatility, the resulting adjustment for $\widehat{\sigma}_{\mathrm{VS},T}$ is given in (5.55):

$$\widehat{\sigma}_{\mathrm{VS},T} = \widehat{\sigma}_T - \frac{\rho}{2}\sigma_r T$$

❧❧❧❧❧

## 5.9 Weighted variance swaps

▶ In weighted VSs, realized variance is weighted with a function of the spot $w(S)$. At order two in $\delta S$, weighted VSs can be replicated exactly. Standard examples of weighted VSs include the Gamma swap, the arithmetic swap, for which replication is exact, and corridor variance swaps. The latter require an additional adjustment to their strike to take into account barrier crossings.

❧❧❧❧❧

## Appendix A – timer options

▶ Timer options expire when a given quadratic-variation budget $\mathcal{Q}$ is exhausted. The price of a timer option is a function of $t,\ S$ and the current quadratic variation $Q$, measured using daily log-returns of the underlying stock or index:

$$Q_{i+1} - Q_i \ = \ \ln^2\left(\frac{S_{i+1}}{S_i}\right)$$

▶ For vanishing interest rate and repo, no dividends, and if the process for $S_t$ is a diffusion, timer options are model-independent. They are replicated by a plain delta strategy and their prices are given by a Black-Scholes formula with an effective volatility equal to 1 and an effective maturity equal to $\mathcal{Q} - Q$. Physical time disappears.

▶ In practice, timer options are not exactly model-independent. While order-two contributions in $\delta S$ vanish, higher-order terms contribute to the carry P&L. Moreover, the presence of non-vanishing interest rate and repo as well as dividends reintroduces the dependence on physical time. Two additional effects need to be priced-in: the overshoot in realized quadratic variation with respect to the budget, and the provision of a maximum maturity in the term sheet.

▶ Leveraged ETFs are another breed of quadratic-variation-based payoffs that are model-independent, in a diffusive setting, as they are replicated by a plain delta strategy. Starting from a value $I_0$, the NAV $I_t$ at time $t$ of the ETF is $I(t, S_t, Q_t)$ where $I(t, S, Q)$ is given by (5.75):

$$I(t, S, Q) \ = \ I_0 e^{rt}\left(\frac{S}{S_0 e^{(r-q)t}}\right)^{\beta} e^{-\frac{\beta(\beta-1)}{2}Q}$$

where $\beta$ is the ETF's leverage.

❧❧❧❧❧

**Appendix B – perturbation of the lognormal distribution**

▶ The smile of vanilla options for maturity $T$ is generated by the non-lognormality of the distribution of $S_T$. Many types of models can be collapsed onto the Black-Scholes by setting a parameter to zero – stochastic volatility and jump-diffusion models are two examples. It is useful to have an expansion of implied volatilities at order one in such parameter.

▶ The non-lognormality of the distributon of $S_T$ is quantified by the cumulants $\kappa_n$ of the distribution of $\ln S_T$. For a lognormal distribution of $S_T$, $\kappa_n = 0, \ \forall n \geq 3$. We carry out an expansion of implied volatilities in the $\kappa_n$, $n \geq 3$, at order one.

▶ As we perturb the cumulant-generating function, we require that prices of log contracts be unaffected, so that VS volatilities, in diffusive models, be unchanged.

▶ The perturbation of implied volatilities around a lognormal density of implied volatility $\widehat{\sigma}_0$ is given, at order one in the cumulants, by: (5.90):

$$\delta\widehat{\sigma} \; = \; \frac{1}{\widehat{\sigma}_0 T} \sum_{n=3}^{\infty} \frac{\frac{\delta\kappa_n}{n!}\left(\frac{d^n P_0}{d\ln S^n} - \frac{d^2 P_0}{d\ln S^2}\right)}{\frac{d^2 P_0}{d\ln S^2} - \frac{dP_0}{d\ln S}} \; + \; \frac{1}{\widehat{\sigma}_0 T} \sum_{n=3}^{\infty} \frac{\delta\kappa_n}{n!}\theta_n$$

where $\theta_n = 0$ if we demand that log-contract implied volatilities stay unchanged.

▶ At order one in the third cumulant – expressed in terms of the skewness $s$ of $\ln S_T$ through $\delta\kappa_3 = s^3(\widehat{\sigma}_0\sqrt{T})^3$ – the ATMF skew is given by (5.93):

$$\mathcal{S}_T \; = \; \frac{s}{6\sqrt{T}}$$

This page intentionally left blank