Helsinki University of Technology Institute of Mathematics Research Reports
Espoo 2008 A545

THE INTEGRATED VOLATILITY IMPLIED BY OPTION PRICES A BAYESIAN APPROACH

Ruth Kaila

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology Institute of Mathematics Research Reports
Espoo 2008 A545

THE INTEGRATED VOLATILITY IMPLIED BY OPTION PRICES A BAYESIAN APPROACH

Ruth Kaila

Dissertation for the degree of Doctor of Science in Technology to be presented, with due permission of the Faculty of Information and Natural Sciences, for public examination and debate in auditorium E at Helsinki University of Technology, Espoo, Finland, on the 6th of May 2008, at 12 o'clock noon.

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Mathematics and Systems Analysis

Ruth Kaila
Department of Mathematics and Systems Analysis
Helsinki University of Technology
P.O. Box 1100, FI-02015 TKK, Finland
E-mail: ruth.kaila@tkk.fi

ISBN 978-951-22-9315-5 (print)
ISBN 978-951-22-9316-2 (electronic)
ISSN 0784-3143
TKK Mathematics, 2008

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Mathematics and Systems Analysis
P.O. Box 1100, FI-02015 TKK, Finland
email: math@tkk.fi http://math.tkk.fi/

Ruth Kaila: *The integrated volatility implied by option prices, a Bayesian approach*; Helsinki University of Technology Institute of Mathematics Research Reports A545 (2008).

Abstract: In this thesis, we present the new concept of implied integrated volatility. When the stock price volatility is stochastic, the integrated volatility is the time-average of the stock price variance. This volatility is a fundamental quantity in option theory, as the stock price returns depend on the stock price volatility only via the integrated volatility.

The implied integrated volatility is the integrated volatility implied by option Hull-White prices. It is a stochastic extension of the Black-Scholes implied volatility. Unlike the latter, however, it is independent of the strike price of options. We suggest that this volatility can be used in volatility estimation, in pricing illiquid options consistently with corresponding liquid ones, and in hedging options.

Estimating the implied integrated volatility is an ill-posed inverse problem. We present methods to estimate it within a Bayesian framework. This approach provides us with not only a point estimate, but also the possibility to gauge the reliability of this estimate.

AMS subject classifications: 91B

Keywords: implied integrated volatility, stochastic volatility, quadratic variation, Bayesian inference, MCMC sampling, hypermodel

Ruth Kaila: *Option hinnoista määräytyä integroitu volatiliteetti, Bayesilainen lähestymistapa*

Tiivistelmä: Väitöskirjassa esitetään uusi käsite implisiittinen integroitu volatiliteetti. Kun osakkeen hinnan volatiliteetti on stokastinen, integroitu volatiliteetti on osakkeen hinnan varianssin aikakeskiarvo. Osakkeen tuotto riippuu osakkeen hinnan volatiliteetista vain integroidun volatiliteetin välityksellä, minkä johdosta tämä volatiliteetti on merkittävä suure optioteoriassa.

Implisiittinen integroitu volatiliteetti on option Hull-White hinnoista määräytyvä integroitu volatiliteetti. Se on stokastinen laajennus Black-Scholesin implisiittiselle volatiliteetille. Toisin kuin viimeksi mainittu, implisiittinen integroitu volatiliteetti ei kuitenkaan riipu option lunastushinnasta. Sitä voidaan käyttää volatiliteetin estimoimiseen, epälikvidien optioiden hinnoitteluun yhtenevästi likvidien kanssa sekä suojaukseen.

Implisiittisen integroidun volatiliteetin estimointi on huonosti asetettu käänteisongelma. Väitöskirjassa esitetään menetelmiä sen estimoimiseksi bayesiläisessä paradigmassa. Yksittäisten piste-estimaattien lisäksi kyseinen lähestymistapa tarjoaa mahdollisuuden arvioida kyseisten estimattien luotettavuutta.

Avainsanat: implisiittinen integroitu volatiliteetti, stokastinen volatiliteetti, kvadraattinen variaatio, bayesiläinen päättely, MCMC otanta, hypermalli

You must rise above
The gloomy clouds
Covering the mountaintop
Otherwise, how will you
Ever see the brightness?

–Ryokan to Teishin (1758–1831)

Contents

Preface

It is not known exactly when the first derivatives were traded, but contracts similar to options were used by Phoenicians and Romans for trading. In the early 17th century, trading in tulip options blossomed in Holland. However, for several decades, growth in option trading remained modest.

The year 1973 was an important one for the option markets in both practical and theoretical ways: The first options exchange was opened in Chicago, and Black and Scholes introduced their famous formula for option pricing. Since then, a multi-trillion-dollar derivative industry has sprung up, and products based on financial derivatives have become an essential tool for risk managers and investors. The range of security instruments continues to increase, allowing diverse risks to be hedged in ways closely tailored to the specific needs of investors and companies.

One of the central problems in financial mathematics is the pricing and hedging of derivatives, such as options. The accuracy of correct pricing and hedging depends, among other things, on how well the volatility of the underlying asset has been estimated. The quest for quantitative tools for reliably dealing with the financial instruments has given birth to mathematical finance. Mathematical finance is, rather than a discipline in its own right, an area of application where results from probability theory, statistics, optimal control, functional analysis, and partial differential equations can be implemented in real life problems. In this thesis, we consider one particular instance in this vast field: the problem of volatility estimation at the interface of option theory and statistical inverse problems.

Market data differ from traditional data in natural and engineering sciences in one fundamental way. While the data in the latter is always inaccurate, approximate, and noisy, the observed market data of asset prices such as stock and options contain no error. In other words, a price is a price. When comparing observed values to theoretical predictions, the possible discrepancy between the two must therefore be accounted for in the insufficiency of the model. We can assume that the observed market data on asset prices such as stocks and options contains no error. Therefore, we must focus on how we model the relevant quantity. At least two prevalent approaches in modeling volatility based on this data can be distinguished, the idealistic approach and the instrumentalist approach.

According to the idealistic point of view, there exists a volatility model that represents a mapping of reality. This model can be inferred from market prices that fully reflect information on this process. This approach has resulted in an increasing range of parametric and semi-parametric volatility models. Even if the models are based on many unverifiable assumptions, huge calibrations are often done without taking into account the reliability of these basic assumptions. The possibility of easily accessing and processing extensive quantities of market data has strengthened this paradigm. The classical piece of advice presented in Nicomachean Ethics is often forgotten - Aristotle recommends that precision used should fit reasonably with the subject under consideration.

Alternatively, volatility models can be seen in an instrumental way: as tools to ensure that all derivative instruments are consistently priced with respect to each other in order to avoid any possibility of arbitrage. A commonly known tool for this purpose is the Black-Scholes implied volatility. However, the Black-Scholes formula is deterministic, and as such it is in conflict with the stochastic nature of markets. In this thesis, we introduce a stochastic extension of this tool, the integrated volatility implied by option prices. We call this quantity the implied integrated volatility.

L'homme moyen, the average man, is a fictional character of the social sciences introduced by Quetelet in 1835. Quetelet constructed this average man from a sample of about 100,000 French compatriots by measuring attributes like height and weight, and even computing, using arrest records, the average man's propensity to commit a crime. This fictitious average man has played a central role in financial mathematics until recent years.

The probability space for an economy can be interpreted as a natural measure expressing individual or collective subjective beliefs concerning the market events. The market prices of assets are then a reflection of this collective subjectivity. However, even if it is understood that investors have heterogeneous levels of wealth, risk aversion, patience, and belief, many theories and models on assets and options are set as if there were one average man or investor whose wealth, risk aversion, patience, and belief are aggregates of those of all investors. Latent parameters such as volatility would then have a fixed value, which can be estimated from market data fairly accurately.

In the Bayesian paradigm, randomness means a lack of information, and our subjective belief in a random event is expressed in terms of probability. Instead of an absolute truth, this approach searches for evidence meant to be consistent or inconsistent with a given hypothesis. As stock and option market prices are reflections of collective subjectivity, it is quite natural to propose a Bayesian approach to estimating volatilities from these prices. In this thesis, we present methods to estimate the implied integrated volatility within a Bayesian framework.

This work is intended to be self-contained. It can be divided conceptually into two parts. The first theoretical part, including Chapters 1-7, provides a framework for the second part, Chapters 8-12, which is devoted to the implied integrated volatility.

A brief introduction on the probability theory needed later is provided in the first chapter. It should be viewed as a glossary for notations and concepts used later on in this thesis, not as a comprehensive treatment of the topic. In Chapters 2-4, we present fundamental concepts related to stocks, options, and option pricing. We also introduce the concept of integrated volatility and discuss stock price processes with different volatilities.

The implied integrated volatility is based on the Hull-White paradigm. This paradigm, as well as various ways to derive the Hull-White pricing formula, is considered in Chapter 5. One explanation for the volatility smile, based on the Hull-White formula, is presented at the end of that chapter. Chapters 6 on hedging and 7 on inverse problems close the first part of the thesis.

Chapter 8 discusses the concept of implied integrated volatility on a general level. In Chapters 9 and 10, we present two methods to estimate this volatility from option prices, using a Bayesian approach. Three computed examples on volatility estimation, pricing options, and hedging options are presented in Chapter 11. Chapter 12 contains concluding remarks and suggestions for further research. Each chapter ends with a number of references for suggested further reading, although by no means are these lists exhaustive or all-inclusive. They are provided to the reader for both more comprehensive background information, and as a springboard into more detailed and specialized works.

Acknowledgements

First and foremost, I would like to thank my advisor and supervisor, Professor Erkki Somersalo, for his guidance and continuous support for the duration of this thesis. I am also most grateful to Professor Olavi Nevanlinna, Head of the Institute of Mathematics at Helsinki University of Technology, for providing and fostering a truly enriching academic environment. Marita Katavisto, department secretary, deserves special recognition for her unfailing logistical help, kind words, and encouragement. I am particularly indebted to the inverse problem research group for an enjoyable and encouraging research atmosphere. Professor Somersalo and this group's camaraderie and open-minded approach to research and scholarship benefited the development of this thesis in ways too numerous to list. Also, I would like to thank Professor Daniela Calvetti from Case Western Reserve University for her invaluable comments and advice. Finally, heartfelt thanks go to the rest of my colleagues at the Institute of Mathematics at Helsinki University of Technology for our inspiring and stimulating discussions.

1 Preliminaries

The aim of this section is to provide a short introduction to the probability theory needed in this thesis, in particular to stochastic processes and to Monte Carlo sampling. In addition to the theory directly used in this work, we willl present concepts that are useful for understanding the interface of this thesis and quantitative finance. In addition, for fixing the notations, we will review some basic facts concerning these fields. As the main focus of this work is not in probability theory and stochastic processes, we will present the formulas and theorems without proofs in an accessible way, and will give references to more detailed and rigorous presentations in the end of the chapter.

1.1 General probability theory

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a triple, where Ω is a probability space, $\mathcal{F}$ a σ-algebra of the subsets of Ω and $\mathrm{P} : \mathcal{F} \to [0,1]$ a σ-additive probability measure. With σ-algebra we mean that i) $\Omega \in \mathcal{F}$, ii) if $F \in \mathcal{F}$ then $\Omega \backslash F \in \mathcal{F}$, and iii) if $F_i \in \mathcal{F}, i \in \mathbb{N}$, then $\cup_{i=1}^{\infty} F_i \in \mathcal{F}$.

A measurable mapping $X : \Omega \to \mathbb{R}$ is a *random variable*. The probability distribution of X is $\mu_X(B) = \mathrm{P}\big\{X^{-1}(B)\big\}$, where $B \subset \mathbb{R}$ is a Borel set. A *multivariate random variable* is a mapping $X = [X_1, X_2, \ldots, X_n] : \Omega \to \mathbb{R}^n$, where each component is a real-valued random variable. If X is an n-variate, and Y is an m-variate random variable, then $Z = [X, Y]$ is an $(n+m)$-variate random variable and the joint probability distribution of X and Y is defined as the probability distribution of Z.

Let P and P^* be two probability measures defined on $(\Omega, \mathcal{F})$. We say that the measure P is *absolutely continuous* with respect to P^* if and only if $P^*(A) = 0 \Rightarrow P(A) = 0$ for all $A \in \mathcal{F}$. If P is absolutely continuous with respect to P^*, and if P^* is absolutely continuous with respect to P, we say that P and P^* are *equivalent measures*.

According to the *Radon-Nikodym Theorem*, if P absolutely continuous with respect to P^*, then there exists a non-negative $\mathcal{F}$-measurable random variable $\xi = \xi(\omega)$ such that for any $A \in \mathcal{F}$,

$$P(A) = \int_A \xi(\omega) dP^*(\omega). \tag{1}$$

In addition, ξ is unique as a measurable function in the sense that if there were another random variable ξ' satisfying (1) for all $A \in \mathcal{F}$, then $P(\xi = \xi') = 1$.

The random variable ξ is called the *Radon-Nikodym derivative*. It is commonly denoted by

$$\xi(\omega) = \frac{dP}{dP^*}(\omega).$$

The Radon-Nikodym theorem says that if P is absolutely continuous with respect to P^*, an integral with respect to P is a weighted integral with respect to P^*, the weight being the Radon-Nikodym derivative.

If the probability distribution μ_X of a random variable $X \in \mathbb{R}^n$ is absolutely continuous with respect to the Lebesgue measure, there is a measurable density $\pi_X : \mathbb{R}^n \to \mathbb{R}_+$ so that

$$\mu_X(B) = \int_B \pi_X(x)dx,$$

$B \subset \mathbb{R}^n$ measurable. We adopt the convention that uppercase letters denote random variables and lowercase letters their realizations. Furthermore, when there is no danger of confusion, we will omit the subindex from the densities and write $\pi(x) = \pi_X(x)$, i.e., the variable x indicates that $\pi(x)$ is the probability density of X.

The joint probability density for two real valued random variables X and Y, $\pi_{XY}(x,y) = \pi(x,y)$, is defined via their probability distribution, i.e.,

$$P\{X \in A, Y \in B\} = P(X^{-1}(A) \cap Y^{-1}(B)) = \int\int_{A\times B} \pi(x,y)dxdy,$$

and, assuming that such a density exists, the *marginal density* of X is recovered as

$$\pi(x) = \int_{\mathbb{R}} \pi(x,y)dy. \tag{2}$$

We denote the expectation of X by $\mathrm{E}\big\{X\big\}$, the variance by $\mathrm{Var}\{X\}$ the covariance of X and Y by $\mathrm{Cov}\{X,Y\}$, and define the *correlation coefficient* of X and Y by

$$\rho\{X,Y\} = \frac{\mathrm{Cov}\{X,Y\}}{\sqrt{\mathrm{Var}\{X\}\mathrm{Var}\{Y\}}}. \tag{3}$$

The *conditional expectation* is defined in the following way. Assume that $\|X\|^2 : \Omega \to \mathbb{R}_+$ is integrable. Denote $\|X\|_2^2 = \mathrm{E}\big\{\|X\|^2\big\} < \infty$, and let $L^2(\Omega, \mathcal{F}, \mathrm{P})$ denote the space of square integrable random variables.

Assume that $\mathcal{S} \subset \mathcal{F}$ is a sub-σ-algebra and denote by $L^2(\Omega, \mathcal{S}, \mathrm{P})$ the subspace of $\mathcal{S}$-measurable functions, $L^2(\Omega, \mathcal{S}, \mathrm{P}) \subset L^2(\Omega, \mathcal{F}, \mathrm{P})$. This subspace is closed. If $\mathcal{P}$ is the orthogonal projection $L^2(\Omega, \mathcal{S}, \mathrm{P}) \to L^2(\Omega, \mathcal{F}, \mathrm{P})$, we denote $\mathrm{E}\big\{X \mid \mathcal{S}\big\} = \mathcal{P}X$.

Let $X \in \mathbb{R}^n$ and $Y \in \mathbb{R}^m$ are two random variables, and consider the mapping $\mathcal{B}(\mathbb{R}^n) \to \mathbb{R}_+$, $B \mapsto \mathrm{E}\{\chi_B(X) \mid \sigma(Y)\}$, where $\mathcal{B}(\mathbb{R}^n)$ denotes the Borel sets of $\mathbb{R}^n$ and χ_B is the characteristic function of $B \subset \mathbb{R}^n$. It can be shown that there exists a *regular version of the conditional measure*, i.e., a mapping $\mu : \mathcal{B}(\mathbb{R}^n) \times \mathbb{R}^m \to \mathbb{R}_+$, so that $B \mapsto \mu(B \mid y)$ is a measure and $y \mapsto \mu(B \mid y)$ is measurable, and $\mathrm{E}\{\chi_B(X) \mid \sigma(Y)\} = \mu(B \mid Y(\omega))$, i.e., the conditional expectation depends only on the realization of Y. If the joint distribution of X and Y is absolutely continuous with respect to the Lebesgue measure of $\mathbb{R}^n \times \mathbb{R}^m$, with the density $\pi(x, y)$, then the *conditional measure* $B \mapsto \mu(B \mid y)$ is absolutely continuous with respect to the Lebesgue measure of $\mathbb{R}^n$ and the density $\pi(x \mid y)$ is obtained by

$$\pi(x \mid y) = \frac{\pi(x, y)}{\pi(y)}, \quad \text{if } \mu(y) \neq 0. \tag{4}$$

By the symmetry of the roles of X and Y, we have

$$\pi(x, y) = \pi(x \mid y)\pi(y) = \pi(y \mid x)\pi(x). \tag{5}$$

The following types of convergence are important in a stochastic environment, an will be used in the sequel.

1. The sequence $X_1, X_2, \ldots$ *converges in probability* to X if for every $\epsilon > 0$, we have
$$\lim_{n\to\infty} P\{|X_n - X| > \epsilon\} = 0, \tag{6}$$
and we denote $X \overset{P}{=} \lim_{n\to\infty} X_n$.

2. The sequence $X_1, X_2, \ldots$ *converges with probability one* to X if
$$P\{\lim_{n\to\infty} X_n = X\} = 1. \tag{7}$$
In this case, we write $X \overset{\text{a.s.}}{=} \lim_{n\to\infty} X_n$ or say that the sequence *converges almost surely*, that is *a.s.* to X.

3. The sequence $X_1, X_2, \ldots$ *converges in mean square sense* to X if
$$\lim_{n\to\infty} \mathbb{E}\{|X_n - X|^2\} = 0, \tag{8}$$
which is denoted with $X \overset{\text{m.s.}}{=} \lim_{n\to\infty} X_n$.

4. The sequence $X_1, X_2, \ldots$ *converges in distribution* to X if
$$\lim_{n\to\infty} \mathbb{E}\{f(X_n)\} = \mathbb{E}\{f(X)\} \tag{9}$$
for every bounded continuous function $f : \mathbb{R}^{\mathrm{n}} \to \mathbb{R}$. We then write $X \overset{d}{=} \lim_{n\to\infty} X_n$.

The *Law of Large Numbers* and the *Central Limit Theorem* will be used when analyzing the convergence of averages of random variables. Suppose that $X_1, X_2, \ldots$ are independent and identically distributed real valued random variables with mean μ and variance σ^2 which we assume to be finite. A version of the Law of Large Numbers says that

$$\lim_{n\to\infty} \frac{1}{n}(X_1 + X_2 + \cdots + X_n) \overset{\text{a.s.}}{=} \mu. \tag{10}$$

Let $\hat{\mu}_n$ denote the sample mean

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Then a version of the Central Limit Theorem states that the random variables $\hat{\mu}_n$ are asymptotically *normally distributed* in the following sense:

$$\lim_{n\to\infty} P\left\{\frac{\hat{\mu}_n - n\mu}{\sigma\sqrt{n}} < x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt = \mathcal{N}(x). \tag{11}$$

1.2 Stochastic processes

A stochastic process is a collection of real valued random variables $(X_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, P)$; the time parameter t can be either discrete or continuous. The mapping $t \mapsto X_t(\omega)$, $\omega \in \Omega$ fixed, is a sample path of X_t.

A *filtration* on $(\Omega, \mathcal{F})$ is a nested family $\mathcal{M} = \{\mathcal{M}_t\}_{t\geq 0}$ of σ-algebras $\mathcal{M}_t \subset \mathcal{F}$,

$$0 \leq s < t \Rightarrow \mathcal{M}_s \subset \mathcal{M}_t.$$

The filtration refers to the information available at time t. A natural filtration associated with a stochastic process $(X_t)_{t\geq 0}$ is $\mathcal{F}_t^X = \sigma(X_s \mid s \leq t)$, i.e., $\mathcal{F}_t^X$ is the σ-algebra generated by the stochastic process $(X_s)_{s\leq t}$.

A stochastic process $(X_t)_{t\geq 0}$ is *adapted* to the filtration $(\mathcal{F}_t)_{t\geq 0}$, if the stochastic process X_t is $\mathcal{F}_t$-measurable for every t. Observe that $(X_t)_{t\geq 0}$ is always $\mathcal{F}_t^X$-adapted. Two important classes of adapted processes are *Markov processes* and *martingales.*

A Markov process is a stochastic process for which only the present value of the random variable is relevant for forecasting the future, that is, the past history of the random variable and the way in which the present has emerged from the past are irrelevant in forecasting the future behavior of the variable. In other words, an $(\mathcal{F}_t)_{t\geq 0}$-adapted process $(X_t)_{t\geq 0}$ is a Markov process if

$$\mathbb{E}\{X_t \mid \mathcal{F}_s\} = \mathbb{E}\{X_t \mid X_s\} \quad \text{for all } 0 \leq s \leq t. \tag{12}$$

We call (12) the *Markov property.*

A real-valued $(\mathcal{F}_t)_{t\geq 0}$-adapted stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is a $(\mathcal{F}_t)_{t\geq 0}$-martingale if

1. $\mathbb{E}\{|X_t|\} < \infty$ for every $0 \leq t < \infty$
2. $\mathbb{E}\{X_t|\mathcal{F}_s\} = X_s$ for every $0 \leq s < t < \infty$.

If a stochastic process is a martingale, its expected value at any future time t conditioned on all information of the past equals the current value of the process. Therefore, the best guess for the future value of the process is its current value. This property is often expressed by saying that for a game that is a martingale there is no winning strategy.

Given a not-necessarily equispaced partition Π of an interval $[t_0, t]$,

$$0 = t_0 < t_1 < t_2 < \cdots < t_n = t, \tag{13}$$

denote by $\|\Pi\|$ the maximum step size of the partition

$$\|\Pi\| = \max_{i=0,\dots,n-1}(t_{i+1} - t_i).$$

By the *total variation* of a stochastic process $(X_t)_{t\geq 0}$ we mean

$$\mathrm{TV}(X)_t \stackrel{\mathrm{P}}{=} \lim_{\|\Pi\|\to 0} \{\sum_{i=0}^{n-1} |X_{t_{i+1}}(\omega) - X_{t_i}(\omega)|\},$$

and the *quadratic variation* of $(X_t)_{t\geq 0}$ is defined as

$$\langle X\rangle_t \stackrel{\mathrm{P}}{=} \lim_{\|\Pi\|\to 0} \{\sum_{i=0}^{n-1} (X_{t_{i+1}}(\omega) - X_{t_i}(\omega))^2\}. \tag{14}$$

The *quadratic covariation* of two stochastic processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ is defined as

$$\langle X, Y\rangle_t \stackrel{\mathrm{P}}{=} \lim_{\|\Pi\|\to 0} \{\sum_{i=0}^{n-1} (X_{t_{i+1}}(\omega) - X_{t_i}(\omega))(Y_{t_{i+1}}(\omega) - Y_{t_i}(\omega))\}. \tag{15}$$

It can be shown that if a continuous stochastic process has finite total variation, it has zero quadratic variation. Conversely, this means that the sample paths of a process with non-vanishing quadratic variation are non-rectifiable with probability one.

A real-valued stochastic process $(W_t)_{t\geq 0}$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ is called a *Brownian motion* if it satisfies the following properties:

1. $W_0 = 0$, almost surely;
2. for all $0 \leq s < t$, the increment $W_t - W_s$ is independent of $\mathcal{F}_s$ and is normally distributed with

$$\mathbb{E}\{W_t - W_s\} = 0 \quad \text{and} \quad \text{Var}\{W_t - W_s\} = t - s;$$

3. $(W_t)_{t\geq 0}$ has continuous sample paths almost surely.

A Brownian motion is a Markov process and a martingale. It will be used later to model the random fluctuations of stock prices.

Almost all sample paths of the Brownian motion have infinite total variation, but the quadratic variation of the Brownian motion over a time interval $[s,t]$ is finite, and is given by

$$\langle W\rangle_{s,t} \stackrel{\text{P}}{=} \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t - s, \tag{16}$$

where we have used the time partition (13) of the interval $[s,t]$. We denote $\langle W\rangle_t = \langle W\rangle_{0,t}$. In the subsequent sections, we will denote the differential of the quadratic variation $\langle X\rangle_t$ of a stochastic process $(X_t)_{t\geq 0}$ by $d\langle X\rangle_t$. It is defined as

$$d\langle X\rangle_t = \lim_{h\to 0} \frac{\langle X\rangle_{t,t+h}}{h} dt, \tag{17}$$

provided that the limit exists. The differential of the quadratic variation of the Brownian motion (16) is

$$d\langle W\rangle_t = dt.$$

Since the total variation is infinite for almost all paths of the Brownian motion, the integral of stochastic processes with respect to the Brownian motion cannot be defined as a Lebesgue-Stjeltjes integral. However, integrals for these processes can be defined as *Itô integrals*. Let (13) be a partition of $[0,t]$ and let $(X_t)_{0\leq t\leq T}$ be a stochastic process adapted to $(\mathcal{F}_t)_{0\leq t\leq T}$ for a fixed T, so that

$$\mathbb{E}\{\int_0^T X_t^2 dt\} < \infty,$$

i.e., the paths are square integrable almost surely. If $(W_t)_{0\leq t\leq T}$ is a $(\mathcal{F}_t)_{0\leq t\leq T}$-adapted Brownian motion, it can be shown that the sequence of random variables Z_n, where $Z_n = \sum_{i=0}^{n-1} X_{t_i}(W_{t_{i+1}} - W_{t_i})$ converges in the mean square sense (8). We omit the proof, which is presented in Oksendal [92]. The limit of this sequence, denoted by

$$I_t = \int_0^t X_s dW_s \stackrel{\text{m.s.}}{=} \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} X_{t_i}(W_{t_{i+1}} - W_{t_i}), \tag{18}$$

is called the *Itô integral* of X_t over $[0,t]$, $t \leq T$.

The Itô integral has several important features. As a function of time, it defines a continuous square integrable process such that

$$\mathbb{E}\{\big(\int_0^t X_s dW_s\big)^2\} = \mathbb{E}\{\int_0^t X_s^2 ds\} < \infty. \tag{19}$$

The Itô integral is an $\mathcal{F}_t$-martingale, i.e.,

$$\mathbb{E}\{\int_0^t X_u dW_u \mid \mathcal{F}_s\} = \int_0^s X_u dW_u, \quad \text{almost surely}, \quad s \leq t.$$

Since $W_{t_{i+1}} - W_{t_i}$ is independent of $\mathcal{F}_{t_i}$ and therefore of X_{t_i}, we have

$$\mathbb{E}\{X_{t_i}(W_{t_{i+1}} - W_{t_i})\} = \mathbb{E}\{X_{t_i}\}\underbrace{\mathbb{E}\{W_{t_{i+1}} - W_{t_i}\}}_{=0} = 0,$$

and thus the expectation $\mathbb{E}\{I_t\} = 0$ for all $t \geq 0$ and the variance is $\mathrm{Var}\{I_t\} = \mathbb{E}\{I_t^2\}$. Using (19), the variance of the Itô integral is therefore

$$\mathrm{Var}\{I_t\} = \mathbb{E}\{\int_0^t X_s^2 ds\}.$$

The quadratic variation of the Itô integral I_t over the interval $[0,t]$ is

$$\langle I \rangle_t = \int_0^t X_s^2 ds,$$

and its differential is

$$d\langle I \rangle_t = X_t^2 dt.$$

Observe that the quadratic variation of the Itô integral is a random variable, its expectation is the variance $\mathrm{Var}\{I_t\}$ of the Itô integral.

Let W_t be a $(\mathcal{F}_t)$-adapted Brownian motion on $(\Omega, \mathcal{F}, P)$. The stochastic process $(X_t)_{t\geq 0}$ is said to be an *Itô process* if it can be represented as

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \tag{20}$$

where X_0 is non-random, and b_s and σ_s are real-valued stochastic processes adapted to the filtration $(\mathcal{F}_s)_{s\geq 0}$ that satisfy the following conditions:

$$\int_0^t |b_s| ds < \infty \quad \text{a.s. and} \quad \mathbb{E}\left\{\int_0^t \sigma_s^2 ds\right\} < \infty, \quad \text{for all } t > 0.$$

The Itô process can be written in the differential form as

$$dX_t = b_t dt + \sigma_t dW_t. \tag{21}$$

It can be shown (see Shreve 2004 [105] that the quadratic variation of the Itô process (20) on the interval $[0,t]$ is

$$\langle X\rangle_t = \int_0^t \sigma_s^2 ds, \quad 0 \le t \le T. \tag{22}$$

Consider another Itô process $(Y_t)_{0\le t\le T}$

$$Y_t = Y_0 + \int_0^t c_s ds + \int_0^t \xi_s dZ_s,$$

such that the Brownian motions W_t and Z_t are correlated and the correlation coefficient $\rho = \rho\{W_t, Z_t\}$ is independent of t.

The quadratic covariation of the Itô processes X_t and Y_t is

$$\langle X, Y\rangle_t = \int_0^t \sigma_s \xi_s \rho ds. \tag{23}$$

An *Itô diffusion* is a stochastic process $(X_t)_{0\le t\le T}$ that satisfies a stochastic differential equation of the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t \ge s; \quad X_s = x, \tag{24}$$

where W_t is a Brownian motion and $b, \sigma : [0,T]\times\mathbb{R} \to \mathbb{R}$ satisfy the Lipshitz condition

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le D|x-y|,$$

for a positive constant D independent of t.

A typical Itô diffusion used in financial mathematics is

$$dX_t = b_t X_t dt + \sigma_t X_t dW_t, \tag{25}$$

which can be expressed in integral form as

$$X_t = X_0 + \int b_s X_s ds + \int_0^t \sigma_s X_s dW_s,$$

and whose quadratic variation over the interval $[0,t]$ is

$$\langle X\rangle_t = \int_0^t \sigma_s^2 X_s^2 ds. \tag{26}$$

In stochastic calculus, the counterpart of the chain rule in ordinary calculus is called the *Itô formula*. This formula provides a rule for computing differentials of stochastic processes of the form $f(W_t)$, where $f(x)$ is a differentiable function and W_t is a Brownian motion. We present here the one-dimensional and the two-dimensional versions of this formula.

Let X_t be an Itô process and $g(t,x)$ is twice continuously differentiable on $[0,\infty)\times\mathbb{R}$, i.e., $g\in C^2([0,\infty)\times\mathbb{R})$. Then it can be shown (see Oksendal 2003 [92]) that

$$Y_t = g(t,X_t)$$

is also an Itô process satisfying

$$dY_t = dg(t,X_t) = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)d\langle X_t\rangle, \quad (27)$$

where $x = X_t$ and $d\langle X\rangle_t$ is the derivative of the quadratic variation of X_t. We call (27) the *one-dimensional Itô formula.*

Given the two Itô processes X_t and Y_t and a function $g\in C^2([0,\infty)\times\mathbb{R}^2)$, again the process $g(t,X_t,Y_t)$ is an Itô process, and the *two-dimensional Itô formula* is

$$\begin{aligned} dg(t,X_t,Y_t) &= \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dX_t + \frac{\partial g}{\partial y}dY_t \\ &\quad + \frac{1}{2}\Big(\frac{\partial^2 g}{\partial x^2}d\langle X\rangle_t + 2\frac{\partial^2}{\partial x\partial y}d\langle X,Y\rangle_t + \frac{\partial^2}{\partial y^2}d\langle Y\rangle_t\Big), \quad (28)\end{aligned}$$

where y denotes Y_t and $d\langle X,Y\rangle_t$ is the derivative of the joint variation $\langle X,Y\rangle_t$, given in (23).

The *Feynman-Kac formula* relates stochastic differential equations and partial differential equations. It provides a solution to a parabolic partial differential equation as the expectation of a certain functional of a Brownian motion.

Consider the Itô diffusion

$$dX_t = b_t(X_t)dt + \sigma_t(X_t)dW_t.$$

Let f be a twice continuously differentiable function with a compact support, $f\in C_0^2(\mathbb{R})$ and q a continuous function, $q\in C(\mathbb{R})$. Assume that q is bounded from below. Put

$$v(t,x) = \mathbb{E}\Big\{\exp\Big(-\int_0^t q(X_s)ds\Big)f(X_t)|X_0 = x\Big\}. \quad (29)$$

The Feynman-Kac formula states that v thus defined satisfies the initial value problem

$$\begin{aligned} \frac{\partial v}{\partial t} &= Av - qv, \quad t>0, x\in\mathbb{R}, \\ v(0,x) &= f(x), \quad x\in\mathbb{R}, \end{aligned}$$

where A is the generator of this diffusion, mapping $g\in C_0^2(\mathbb{R})$ as

$$A : g(x)\mapsto b_t(x)\frac{\partial g(x)}{\partial x} + \frac{1}{2}\sigma_t(x)\frac{\partial^2 g(x)}{\partial x^2}.$$

Financial instruments are modeled using different probability measures, such as the *natural probability measure* and the *risk-neutral probability measure*, to be defined in the next chapter. The Radon-Nikodym Theorem and the *Girsanov Theorem* are important tools to cope with different probability measures.

Let $Y_t \in \mathbb{R}$ be an Itô process of the form

$$dY_t = \theta_t(\omega)dt + dW_t; \quad t \leq T, Y_0 = 0,$$

where W_t is a Brownian motion. Define a stochastic process $(M_t)_{t\geq 0}$ by

$$M_t = \exp\big\{ -\int_0^t \theta_s(\omega)dW_s - \frac{1}{2}\int_0^t \theta_s(\omega)^2 ds\big\}, \quad 0 \leq t \leq T,$$

and assume that M_t is a martingale with respect to $\mathcal{F}_t$ and P. Define the measure P^* on $\mathcal{F}_T$ by

$$dP^*(\omega) = M_T(\omega)dP(\omega).$$

Then a version of the Girsanov Theorem states that P^* is a probability measure on $\mathcal{F}_T$ and Y_t is a Brownian motion with respect to P^*, for $0 \leq t \leq T$.

The following *Novikov condition* is sufficient to guarantee that $(M_t)_{0\leq t\leq T}$ is a martingale:

$$\mathbb{E}\big\{ \exp\big(\frac{1}{2}\int_0^T \theta_s^2 ds\big)\big\} < \infty. \tag{30}$$

Proofs for the theorem and for the Novikov condition are presented in Karatzas and Shreve 1999 [80].

The Girsanov Theorem describes how the stochastic processes change with respect to the changes in the underlying measure. It is important in many applications in financial mathematics and econometrics and is used, for example, to remove trends from certain stochastic processes.

1.3 Monte Carlo sampling

A commonly encountered computational problem is the numerical approximation of the integral

$$\int_D f(x)\pi(x)dx, \tag{31}$$

where $D \subset \mathbb{R}^n$ and $\mathbb{R}^n$ is a high-dimensional space, $f(x)$ is the function of interest, and π is a density. One possibility is to use numerical quadrature methods, i.e., to define a set of support points in D, give weights to these points, and to approximate the integral as a sum of the weighted point evaluations of f. However, if the dimensionality of the space is high or if the support is poorly known or unknown, this method is not feasible. An alternative is integration based on random sampling, called *Monte Carlo integration*.

In Monte Carlo integration, instead of evaluating the integral at given points of the density, we draw a sample from the density $\pi(x)$ in such a way that the sample optimally represents the distribution. The sample point evaluations of f, possibly weighted, are then used to approximate the integral.

Suppose that we have a random variable X whose probability density is $\pi(x)$ and we want to calculate the integral

$$\mathbb{E}\{f(X)\} = \int_{\mathbb{R}^n} f(x)\pi(x)dx, \tag{32}$$

where the function f is measurable and integrable over $\mathbb{R}^n$ with respect to $\pi(x)dx$.

A natural approach is to approximate (32) by empirical average using a sample $\{x_1, x_2, \ldots, x_m\}$ which is independently sampled from the distribution of X, that is,

$$\hat{f}_m = \frac{1}{m}\sum_{j=1}^{m} f(x_j), \tag{33}$$

where $\hat{f}_m$ converges almost surely to $\mathbb{E}\{f(X)\}$ according to the Law of Large Numbers (10). In addition, when $\mathbb{E}\{f(X)\}^2) < \infty$, also the variance

$$\mathrm{Var}\{f(X)\} = \int_{\mathbb{R}^n} (f(x) - \mathbb{E}\{f(X)\})^2\pi(x)dx = v$$

can be estimated from the sample $\{x_1, \ldots, x_m\}$ by

$$v_m = \frac{1}{m}\sum_{j=1}^{m}(f(x_j) - \hat{f}_m)^2.$$

Now, according to the Central Limit Theorem (11),

$$d_m = \frac{\hat{f}_m - \mathbb{E}\{f(X)\}}{\sqrt{v}} \tag{34}$$

is approximately normally distributed for large m, with a mean of zero and a variance of order $1/m$. This provides the speed of convergence of $\hat{f}_m$, as well as confidence bounds for the approximation of $\mathbb{E}\{f(X)\}$ by $\hat{f}_m$.

A fundamental question is how to pick realizations of a random variable X in such a way that (33) converges to (32). It is possible that the probability density $\pi(x)$ of X is presented only abstractly, lacking analytical formulation, or that evaluating $\pi(x)$ numerically is possible, albeit time consuming. The sample points can then be generated using *Markov Chain Monte Carlo (MCMC) methods*.

Consider a sequence $\{X_n\}_{n\geq 0}$ of random variables. This sequence is a *Markov process*, if for all integers $0 \leq k \leq n$, we have

$$\pi(x_{k+1}|x_0, x_1, \ldots, x_k) = \pi(x_{k+1}|x_k). \tag{35}$$

If there exists a fixed function

$$q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

such that we can write

$$\pi(x_{k+1}|x_k) = q(x_k, x_{k+1}) \quad \forall k,$$

we say that the chain is *time-homogeneous* and call q the *transition kernel* and (35) the *Markov chain* generated by q.

Consider an arbitrary transition kernel q. Let X be a random variable with probability density $p(x) = \pi(x)$. Here, we emphasize that p is a known function of x, while $\pi(x)$ denotes the generic density of X. We can generate a new random variable Y by applying the transition kernel $q(x, y)$, i.e.,

$$\pi(y \mid x) = q(x, y).$$

The probability density $\pi(y)$ of the new variable is found by marginalization:

$$\begin{aligned}
\pi(y) &= \int \pi(y \mid x)\pi(x)dx \\
&= \int q(x, y)p(x)dx.
\end{aligned}$$

If $\pi(y) = p(y)$, then p is an invariant distribution of the transition kernel q. The basic idea of the MCMC methods is to define a transition kernel, or transition rule, such that the density p of interest is its invariant density. The sample distributed according to this density is then generated by this transition rule.

Two commonly used algorithms to generate a Markov chain are the *Metropolis-Hastings algorithm* and the *Gibbs sampler*. We now briefly present the latter, which will be used later in this work. References for more detailed presentations are given in Notes on references.

The Gibbs sampler, introduced by Geman and Geman in 1984 [54], is a Markov Chain Monte Carlo scheme that updates the sample points component-wise. Suppose that a random variable X with a probability density π can be presented component-wise as $X = (X^{(1)}, X^{(2)}, \ldots, X^{(n)})$, where $n > 1$. In addition, suppose that we can simulate X from the corresponding conditional univariate densities π_i,

$$\begin{aligned}
&(X^i \mid x^{(1)}, x^{(2)}, \ldots, x^{(i-1)}, x^{(i+1)}, \ldots x^{(n)}) \\
&\qquad\qquad \sim \pi_i(x^{(i)} \mid x^{(1)}, x^{(2)}, \ldots, x^{(i-1)}, x^{(i+1)}, \ldots x^{(n)})
\end{aligned}$$

for $i = 1, 2, \ldots, n$. The Gibbs sampling algorithm generates a Markov chain $\{X_n\}_{n \geq 0}$ using the following steps.

Given $x_j = [x_j^{(1)}, x_j^{(2)}, x_j^{(3)}, \ldots, x_j^{(n)}]$, generate

1. $X_{j+1}^{(1)}$ from $t \mapsto \pi_1(t \mid x_j^{(2)}, x_j^{(3)}, \ldots, x_j^{(n)})$
2. $X_{j+1}^{(2)}$ from $t \mapsto \pi_2(t \mid x_{j+1}^{(1)}, x_j^{(3)}, \ldots, x_j^{(n)})$
3. ...
4. $X_{j+1}^{(n)}$ from $t \mapsto \pi_j(t \mid x_{j+1}^{(1)}, x_{j+1}^{(2)}, \ldots, x_{j+1}^{(n-1)})$.

The Gibbs sampler accepts every proposal, and every single conditional update leaves the probability density π invariant. It can be shown that, under certain appropriate regularity conditions, the Gibbs sampler chain converges geometrically, and its convergence rate is related to the mutual correlation of the variables (see for example Schervish and Carlin 1992 [102]). All of the simulations may be univariate even in high-dimensional problems, which is usually an advantage.

1.4 Notes on references

Classical references on probability theory include Gihman and Skorohod 1974 [56], Billingsley 1979 [16], and Breiman 1992 [23], and on stochastic calculus include Oksendal 2003 [92], Karatzas and Shreve 1999 [80], and Protter 2004 [96].

Stochastic calculus for finance is considered in Shreve 2003 [105]. Many books on finance, financial modeling, or quantitative finance provide an introduction to the topic, including Cont and Tankov 2004 [39] and Platen and Heath 2006 [93].

Comprehensive books on Monte Carlo Methods include Robert and Casella 2004 [99], Liu 2001 [87], and Bremaud 1999 [24]. Applications to financial engineering are considered in Glasserman 2004 [57], a shorter presentation is provided in Boyle, Broadie and Glasserman 1997 [21], and applications to financial econometrics are treated by Johannes and Polson 2004 [74].

2 Options

In recent years, risk management has become a central question in finance. Options are tools to cope with various risks, as they can be used to some extent to protect current or anticipated positions in the asset. At the price of an option, the asset holder can obtain downside protection and still preserve upside potential reduced by the option price. Options can also be used to take advantage of the anticipated price movements on some financial asset.

Even if options have been around for centuries, it was only in 1973 that organized option markets were created by the Chicago Board Options Exchange. Nowadays, there are over 50 exchanges worldwide on which options are traded, and option markets are among the fastest growing financial markets.

In this chapter, we will present fundamental concepts related to options and option pricing.

2.1 Bonds and stocks

We start by defining a probability space for an economy, and by reviewing the basic concepts such as *bonds* and *stocks*, and the corresponding mathematical models. We model the stock price process based on Samuelson 1965 [101], who suggested using exponential Brownian motion to describe stock price movements. Even earlier, motivated by an attempt to model the fluctuations of asset prices and to price derivatives, Bachelier 1900 [8] introduced Brownian motion.

Consider a continuous-time economy with a finite trading interval $[0, T]$. We assume that trading can take place continuously and that there are no taxes. Let $(\Omega, \mathcal{F}, P)$ be a probability space, equipped with a filtration $(\mathcal{F}_t)_{0 \le t \le T}$. This filtration refers to the information available in the economy at time t, increasing as a function of time. We assume that there is a natural underlying probability measure which expresses the individual or collective subjective belief concerning the market events. We denote this measure by P and call it the *natural probability measure*. The market risk and profit expectations are expressed with respect to this measure.

A bond is a riskless asset. The price B_t of a bond at time t is governed by the ordinary differential equation

$$\frac{dB_t}{B_t} = r dt,$$

where r is a non-negative number denoting the riskless rate of return, the interest rate. If the initial investment $B_0 = 1$, we have $B_t = e^{rt}$ for $t \ge 0$.

A stock is an asset representing ownership in a company. It is common to present the stock price process as an Itô process, given by the stochastic differential equation

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t, \tag{36}$$

where X_t denotes the price of the stock at time t, μ is the *expected rate of return* of the asset, σ is the *stock price volatility*, assumed here to be constant, and W_t is a Brownian motion. The first term on the right-hand side of (36) represents the riskless part of the relative rate of change of the asset price. The second term represents the uncertainty of the asset, the volatility being the instantaneous variance of the stock price return. It is called the *martingale term* of the process.

In the thesis, we assume that the interest rate r and the expected rate of return μ are constants. Also, we assume that the asset does not pay dividends during the time scale we are considering.

When the volatility σ is constant, the stock price process (36) can be written as a *geometric Brownian motion*:

$$\frac{X_t}{X_0} = \exp\big((\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\big). \tag{37}$$

This equation is obtained from (36) applying the Itô formula (27) with $g(t,x) = \log x$, $x > 0$, yielding

$$d(\log X_t) = \frac{dX_t}{X_t} - \frac{1}{2} \cdot \frac{1}{X_t^2} d\langle X_t \rangle. \tag{38}$$

From (26), we have

$$d\langle X_t \rangle = \sigma^2 X_t^2 dt, \tag{39}$$

and by substituting (36) and (39) into (38) we get

$$d(\log X_t) = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t. \tag{40}$$

Integrating all terms from zero until time t gives

$$\log X_t = \log X_0 + (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t,$$

which leads to (37).

We can see that if the volatility is constant, the logarithm of the *stock price return* X_t/X_0 is normally distributed with mean $(\mu - \frac{1}{2}\sigma^2)t$ and variance $\sigma^2 t$. In fact, as the stock price process is Markovian, the logarithm of the stock price return for any X_t/X_s, $0 \leq s \leq t \leq T$, is normally distributed with mean $(\mu - \frac{1}{2}\sigma^2)(t - s)$ and variance $\sigma^2(t - s)$.

2.2 Stock price volatility and integrated volatility

Stock price volatility is modeled in various ways. In the simplest models, it is treated as a constant. There is empirical evidence that this is not a sufficient model in the real world. The volatility can be modeled as a function of time, $\sigma = \sigma(t)$, or as a function of time and the instantaneous stock price, $\sigma = \sigma(t, X_t)$. The latter is called *deterministic volatility*. In more complex models, the volatility is modeled, for example, as a a stochastic process or a diffusion with jumps. Various stochastic volatility models will be considered in Chapter 4.

Consider the stock price process $(X_t)_{0 \leq t \leq T}$ with a stochastic volatility σ_t, both adapted to the filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$. Let the stock price process be given by

$$\frac{dX_t}{X_t} = \mu dt + \sigma_t dW_t, \tag{41}$$

where μ denotes the expected rate of return and σ_t satisfies the condition

$$\mathbb{E}\left\{ \int_0^T \sigma_s^2 ds \right\} < \infty.$$

In the same way as in (37), we apply the Itô formula (28) to $g(t, x, y) = \log x$, $x > 0$, which leads to

$$d(\log X_t) = (\mu - \frac{1}{2}\sigma_t^2)dt + \sigma_t dW_t, \tag{42}$$

and further, by integrating all terms from zero until time t, to the logarithm of the stock price return, given by

$$\log \frac{X_t}{X_0} = \mu t - \frac{1}{2} \int_0^t \sigma_s^2 ds + \int_0^t \sigma_s dW_s. \tag{43}$$

Now, according to (22), the quadratic variation of (43) is finite and given by

$$\langle \log \frac{X_t}{X_0} \rangle_t = \int_0^t \sigma_s^2 ds, \quad 0 \leq t \leq T. \tag{44}$$

In financial mathematics and econometrics, the average of the quadratic variation of the stock price process over the time interval $[t, T]$, $0 \leq t < T$ is called the *integrated volatility*. This integral, denoted by $\bar{\sigma}_t^2$, is given by

$$\bar{\sigma}_t^2 = \frac{1}{T - t} \int_t^T \sigma_s^2 ds. \tag{45}$$

It should be noticed that even if we call (45) the integrated volatility, the integral is taken of the square of the instantaneous volatility, i.e, of the instantaneous variance, and not of the instantaneous volatility itself.

Consider a stock price process on the time interval $[t, T]$. Applying (45) to (43) leads to a formula for the stock price return, given by

$$\frac{X_T}{X_t} = \exp\left(\int_t^T (\mu - \frac{1}{2}\sigma_s^2)ds + \int_t^T \sigma_s dW_s \right),$$

We can see from this formula that the stock price return is related to the stock price volatility only through the integrated volatility $\bar{\sigma}_t^2$. This is one leading idea in our work.

2.3 Options and option markets

Derivatives are financial instruments whose value is derived from the value of some underlying asset, e.g. stock index, stock, interest rate, or foreign currency. *Options* are derivatives that give their owner the right but not the obligation to a transaction at a specific price, *strike price* K, and at or before a specific time, *maturity* T. A *call option* $U(t, X_t)$ gives the owner the right to buy, and a *put option* $P(t, X_t)$ gives the right to sell.

A *European option* is determined by the terminal price of the underlying asset X_T. Its value is independent of the path of the underlying before the maturity. In this paper, when talking about options we refer to European options. At maturity, the value of a European call is given by

$$U(T, X_T) = (X_T - K)^+ = \begin{cases} X_T - K, & \text{if } X_T > K, \\ 0, & \text{if } X_T \leq K. \end{cases}$$

We denote $h(X_T) = U(T, X_T)$ and call $h(X_T)$ the *payoff function*. The payoff of a European put is given by

$$P(T, X_T) = (K - X_T)^+ = \begin{cases} K - X_T, & \text{if } X_T < K, \\ 0, & \text{if } X_T \geq K. \end{cases}$$

The call option price $U_t = U(t, X_t)$ and the put option price $P_t = P(t, X_t)$ are related to each other with the *put-call -parity*

$$U(t, X_t) - P(t, X_t) = X_t - Ke^{-r(T-t)}, \quad 0 \leq t \leq T,$$

where r is the riskless rate of return. Because the price of one can be deduced from that of the other one, we will concentrate on call options in this paper. In general, when explaining general features of options, we call the underlying asset a stock, even though it could as well be, for example, a stock index or foreign currency.

The relation between the strike price of an option and the value of the underlying is often described in terms of *moneyness*, which we define in this paper as the ratio $K/(X_t e^{r(T-t)})$. Further, a call will be referred to as in-the-money, at-the-money, or out-of-the-money, if the strike price is less than, approximately equal to, or greater than $X_t e^{r(T-t)}$.

Simple options with normal maturities and strike prices, without special features, are called *plain vanilla options*. Other options are called *exotic options*. Plain vanilla options are traded in *liquid markets* where the options are priced by the markets. More complex options are traded in *over-the-counter (OTC) markets* where the price has been fixed by a market-maker. These options can be tailor made and have special features. They are often priced based on the market prices of plain vanilla options. Nowadays, one can also buy derivatives that are written on the volatility itself. The definition and the measurement of the volatility are then specified in the derivative contract.

In liquid option markets, there are two different prices for options, one for bids and another for offers. We call this phenomenon the *bid-offer spread*. The size of the spread depends on the liquidity of the markets, among other things, and is usually of the order of some percents of the option prices.

2.4 Replicating the value of an option

Replicating portfolios are a basic tool for pricing and hedging options. In its simplest form, the concept of *replicating an option* refers to holding a portfolio of stocks and bonds in such a way that, at the maturity of the option, the value of this replicating portfolio equals the payoff of the option $h(X_T) = (X_T - K)^+$, regardless of the path the stock price has taken.

If, at maturity, the value of the replicating portfolio equals the payoff of the option, and if no funds are added after the initial investment, then, at any time $t, 0 \leq t \leq T$, the value of the option must equal the value of the replicating portfolio. To understand this claim, consider a case in which, at some moment, the value of either the option or the replicating portfolio is less than the value of the other one. We can then sell the more expensive one and buy the cheaper one, and still get the same amount of money at maturity. In this situation, there is an opportunity for arbitrage, i.e., a way to systematically outperform the market by trading in the marketed assets. The right price of the option refers to one that leads to no arbitrage.

Holding an option does not require any extra funds after the initial investment. The replicating portfolio should not, either; it should, therefore, be *self-financing*. The price of the stock varies continuously, and, according to it, the value of the option does as well. If, for example, the stock price rises, then the value of the option usually rises also. Due to these variations, the amount of stocks and bonds in the replicating portfolio must be continuously adjusted.

If, in an economy, the payoff of any derivative can be completely replicated, we say that the economy is *complete*; otherwise, it is *incomplete*. Incompleteness can arise from many sources such as transaction costs, constraints in the trading strategies, or stochastic volatility. In a complete market, all risks can be perfectly hedged. Theoretically, markets on options with constant volatility are complete, and stochastic volatility renders these markets incomplete. In Chapter 5.3, we will show how the value of an option can be replicated when the volatility is stochastic.

2.5 Risk-neutral pricing of options with constant volatility

Instead of using replicating portfolios, the price of European options can be derived by evaluating the expectation of the discounted payoff of the underlying at maturity. We call this approach the *risk-neutral pricing* of options.

Suppose that the stock price process follows the stochastic dynamics described by equation (36). The value of a European option on the stock does not depend on the path of the underlying. We can then write its value as the expectation of the discounted payoff of the underlying at maturity, that is

$$U(t, X_t; K, T; \sigma^2) = \mathbb{E}\{e^{-r(T-t)}h(X_T)\} = \mathbb{E}\{e^{r(T-t)}(X_t - K)^+\}. \qquad (46)$$

However, unless $\mu = r$, this is not a fair price for the option. The discounted price $\tilde{X}_t = e^{-rt}X_t$, $0 \leq t \leq T$ is not a martingale, as it contains a non-zero drift term $\mu - r$, that is

$$\begin{aligned} d\tilde{X}_t = d(e^{-rt}X_t) &= (\mu - r)(e^{-rt}X_t)dt + \sigma(e^{-rt}X_t)dW_t \qquad (47) \\ &= (\mu - r)\tilde{X}_t dt + \sigma\tilde{X}_t dW_t. \end{aligned}$$

Holding an option valued with (46) and replicating it with stocks and bonds leads to an opportunity of arbitrage.

Harrison and Kreps 1979 [60] and Harrison and Pliska 1981 [61] have suggested that option prices are given as an expectation, but not as the one with respect to the natural probability measure implied by stock markets. Instead, the expectation should be taken with respect to a measure under which the discounted price of the underlying asset is a martingale. We call such a measure an *equivalent martingale measure*. Under this measure, the option price does not depend on the expected rate of return μ, which may be different for each investor. For this reason, we say that the measure is *risk-neutral*.

Let us construct an equivalent martingale measure for (46). We first absorb the drift term from (47) into the martingale term, getting

$$d\tilde{X}_t = \sigma \tilde{X}_t \Big[dW_t + \Big(\frac{\mu - r}{\sigma} \Big) dt \Big],$$

and define the *market price of asset risk* as

$$\theta = \frac{\mu - r}{\sigma}.$$

We write then

$$W_t^* = W_t + \int_0^t \theta ds = W_t + \theta t,$$

so that

$$d\tilde{X}_t = \sigma \tilde{X}_t dW_t^*.$$

We define the random variable ξ_t^θ by

$$\xi_t^\theta = \exp(-\frac{1}{2}\theta^2 t - \theta W_t), \quad 0 \le t \le T.$$

Clearly, the Novikov condition (30) is fulfilled, i.e.,

$$\mathbb{E}\big\{ \exp\big(\frac{1}{2}\int_0^t \theta^2 ds\big)\big\} < \infty$$

and ξ_t^θ is a martingale.

Then, according to the Girsanov Theorem, the probability measure P^* given by

$$P^*(\omega) = \xi_T^\theta(\omega) dP(\omega)$$

is an equivalent martingale measure to P and the process $W^* \in \mathbb{R}$ is a Brownian motion. The stock price $\tilde{X}_t$ a then a martingale under the measure P^*.

In conclusion, if we assume that μ, r, and σ are constants, without the loss of generality, we may assume that $\mu = r$, by possibly passing from the natural probability measure P to the risk neutral measure P^* of the option markets.

2.6 Notes on references

The main references on option theory used in this thesis are Fouque, Papanicolaou, and Sircar 2000 [51], Cont ja Tankov 2004 [39], and Rebonato 2004 [97]. Common references on option theory include Hull 2006 [65], Duffie 2001 [45], and Wilmott 1998 [114]. More mathematical approaches are given in Steele 2001 [106], Lamberton and Lapeyre 1997 [83] and Musiela and Rutkowski 2005 [90].

3 Black-Scholes paradigm

In 1973, Fisher Black and Myron Scholes [18] introduced the *Black-Scholes equation* for pricing European options. Assuming that the stock price process follows (36), they showed that the price $U(t, X_t)$ of an option with maturity T, $0 \leq t \leq T$, must satisfy the following partial differential equation:

$$\frac{\partial U}{\partial t}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 U}{\partial x^2}(t,x) + rx\frac{\partial U}{\partial x}(t,x) - rU(t,x) = 0, \tag{48}$$

where $x = X_t$ denotes the stock price, r the riskless rate of return, and σ the stock price volatility.

The Black-Scholes equation (48) is a general tool for pricing options. When solved using the final condition $h(X_T) = (X_T - K)^+$ of a European call option, this equation leads to the *Black-Scholes formula* providing the price $U^{\text{BS}}(t,x) = U^{\text{BS}}(t,x;K,T;\sigma^2)$ of a European call:

$$U^{\text{BS}}(t,x) = x\mathcal{N}(d_1) - Ke^{-r(T-t)}\mathcal{N}(d_2), \tag{49}$$

where

$$d_1 = \frac{\log(x/K) + (r+\sigma^2/2)(T-t)}{\sqrt{\sigma^2(T-t)}}, \qquad d_2 = d_1 - \sqrt{\sigma^2(T-t)}), \tag{50}$$

and $\mathcal{N}(z)$ is

$$\mathcal{N}(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-y^2/2}dy, \tag{51}$$

the cumulative probability distribution function for a standard normal distribution. The corresponding price for a European put option $P^{\text{BS}}(t,x)$ is given by

$$P^{\text{BS}}(t,x) = Ke^{-r(T-t)}\mathcal{N}(-d_2) - x\mathcal{N}(-d_1),$$

where d_1 and d_2 are as in (50).

One should notice that the Black-Scholes equation is independent of the risk preferences of the investor. It is related by the Feynman-Kac formula (29) with $q(X_s) = r$ to the expectation of the discounted payoff of the underlying at maturity, which is given by

$$U(t, X_t) = \mathbb{E}\{e^{-r(T-t)}h(X_T)\}. \tag{52}$$

The Black-Scholes formula can be derived from (52), as presented in Shreve 2004 [105].

Summarizing, the Black-Scholes equation is based on the following assumptions:

1. The markets are complete;
2. The price of the asset follows an Itô process, given by (36);
3. The stock price volatility is constant;
4. The security trading is continuous;
5. There are no transaction costs or taxes.

Even though, in practice, these assumptions are too restrictive and are violated by the markets, the Black-Scholes model is still by far the most used option pricing model. The model is simple and practitioners often see more complex models as complicated, costly, and risky.

3.1 Black-Scholes implied volatility, volatility smile, and term structure

The Black-Scholes formula is often represented as a tool to price prejudiciously the plain vanilla options. However, these options are priced by markets according to offer and demand; hence, there is no need to calculate their prices. In fact, this formula is merely used in an inverse way to estimate the volatility implied by option market prices, called the *implied volatility.* This volatility is used in risk management and in pricing options other than liquid vanilla options.

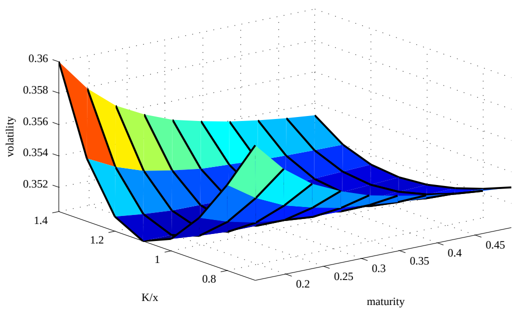
The Black-Scholes option price $U_t^{\mathrm{BS}} = U^{\mathrm{BS}}(t, X_t; K, T; \sigma^2)$ depends on various parameters. The implied volatility I is defined to be the value of the volatility parameter with which the Black-Scholes option price U_t^{BS} equals the market price of the option u_t^{obs} with the given strike price K and maturity T,

$$U_t^{\mathrm{BS}}(t, x; K, T; I^2) = u_t^{\mathrm{obs}}(K, T). \tag{53}$$

The solution of the implied volatility is unique because the Black-Scholes pricing formula (49) is strictly increasing in σ.

The implied volatility can be seen as a kind of markets forecast for the average development of the underlying during the remaining lifetime of the option. This forecast should not depend on the strike price of the option. There is, however, empirical evidence that the implied volatility is a function of both strike price and maturity; i.e., for t and X_t fixed, $I = I(K, T)$. The mapping

$$K \mapsto I(K, T)$$



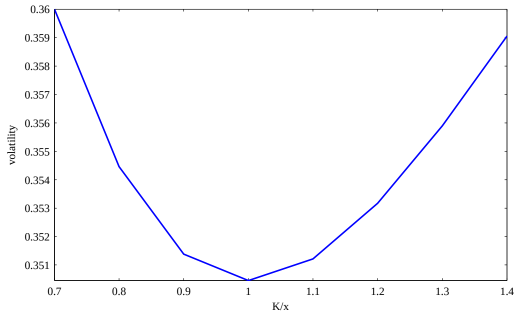


Figure 1: Simulated implied volatility surface and the corresponding volatility smile when the maturity is $T = .15$ years.

gives a *volatility smile*, and the mapping

$$T \mapsto I(K, T)$$

a *term structure*. Finally, the mapping

$$(K, T) \mapsto I(K, T)$$

gives a *volatility smile surface*. In other words, the volatility smile refers to the curve one can notice when plotting implied volatilities as a function of strike prices, and the term structure to the curve one can notice when plotting them against different maturities. One gets a smile surface when considering a collection of smiles with different maturities. In general, the smile tends to attenuate as the time to expiration increases. The smile surface and volatility smiles are illustrated in Figure 1.

The existence of volatility smiles and term structures suggests that the Black-Scholes prices are not consistent with option market prices. In particular, the assumption of constant volatility is unrealistic. Then, unless the volatility is a function of time only, the implied volatility is not a market forecast for the average development of the underlying during the remaining lifetime of the option.

Prior to the October 1987 market crash, Black-Scholes implied volatility smiles typically were more or less symmetric with respect to the implied volatility of an at-the-money option, and in-the-money and out-of-the-money options had higher implied volatilities than at-the-money options. Since the crash, the implied volatilities of call options typically increase monotonically as the call goes deeper in-the-money, and the volatility smiles are rather *skews* with respect to the strike price. In the same way, put options increase monotonically as the put goes deeper out-of-the-money. The change of the shape of the smile into a skew suggests that the implied volatility contains more information than just the expected average of the stock price volatility such as a risk premium due to incomplete markets. Despite problems with the volatility smile, volatilities implied by the market prices of liquidly traded vanilla options are used when forecasting future stock price volatility, when pricing exotic options, and in risk management.

The risk of arbitrage makes pricing exotic options a difficult task. When sharing the same underlying, options with different strike prices are not mutually independent. The problem is then what volatility to use when pricing exotic options. The implied volatility varies as a function of the strike price and maturity. When using it, one question is whether one should use a single volatility to price all options with different strikes, a weighted average of different volatilities, or separate volatilities for options with different strike prices. The situation is further complicated by the fact that the shape of the implied volatility surface changes as a function of time. Then, even if certain exotic options are priced today so that there is no possibility of arbitrage between them, the situation may be different in the future.

According to empirical studies, at-the-money implied volatilities are the less biased ones with respect to the volatilities of the underlying. At-the-money options are typically the most traded ones and probably they reflect best the traders' opinions on the future stock price volatility. For these reasons, they are often used for volatility forecasts. A commonly used alternative to at-the-money implied volatilities is a weighted average of implied volatilities with different strikes.

Three different lines have dominated the research on implied volatilities: searching alternative option pricing models to explain the existence of the volatility smile and the term structure, exploiting the implied volatility as a predictor for stock price movements and exploiting the information content of the option prices and implied volatilities to price more complicated options. We will discuss various stochastic volatility models in the next chapter, and explain in Chapter 5.4 how the stochastic volatility induces a volatility smile. In Chapter 8 we will explain how the information content of the option prices can be, to some extent, encoded in the distribution of the integrated volatility inferred from these prices.

3.2 Notes on references

Empirical evidence on the smile curve of implied volatilities is presented in many papers, for example Rubinstein 1994 [100] and Jackwerth and Rubinstein 1996 [67]. The term structure has been examined, among others, in Poterba and Summers 1986 [95], Stein 1989 [108], and Das and Sundaram 1998 [42].

The implied volatility as a predictor for stock price movements is considered in several papers, such as Christensen and Prabhala 1998 [31]. A review of explanations of the volatility smile is given in Poon and Granger 2003 [94] and Hentschel 2003 [63].

Classical references on the information content of the implied volatility and the volatility smile include Latane and Rendleman 1976 [85], Chiras and Manaster 1978 [30], Lamoureux and Lastrapes 1993 [84] and Canina and Figlewski 1993 [29], and Amin and Ng 1997 [2]. Hentschel 2003 [63] contains a lot of references on the topic.

In order to incorporate information from the whole volatility smile, many early works on implied volatility focused on finding optimal schemes to aggregate implied volatilities across strike. These weighting schemes are surveyed in Bates 1995 [11]. A newer class of models try to specify directly the dynamics of implied volatilities and implied volatility surfaces. Papers on this topic include Schönbucher 1999 [103], Cont and da Fonseca 2002 [36], and Cont, da Fonseca, and Durrleman 2002 [37].

4 Stochastic volatility

One main focus in financial mathematics during the last years has been finding optimal ways to model and estimate the stock price volatility. The time-varying character of volatility has been known for a long time. Early comments include Mandelbrot 1963 [88], Officer 1973 [91], and Black and Scholes 1972 [17], and early models have been presented by Clark 1973 [32] and Taylor 1982 [111]. Nowadays, there is a wide range of research papers on the topic.

Methods to model and estimate the volatility from stock prices can be divided in three groups, which are

1. parametric models such as stochastic volatility models and *autoregressive conditional heteroskedasticity ARCH/GARCH*-models. These models explicitly parametrize the volatility process. Stochastic volatility models are the basis of this paper and will be considered in the next section. The discrete ARCH-models by Engle 1982 [49] and the GARCH-models by Bollerslev 1986 [19] estimate the volatility σ_t as a weighted average of a constant long-run average rate V, the previous discretely sampled variance rate σ_{t-1}^2 and stock price x_{t-1}. These widely used models are out of our scope, and we refer the reader to related literature listed in the end of this chapter.

2. direct market-based *realized volatilities* which are constructed by summing historical intra-period high-frequency stock return data, i.e., data whose frequency is typically less than five minutes. These volatility estimates, suggested in the unrelated papers by Andersen and Bollerslev 1998 [3], Barndorff-Nielsen and Shephard 2001 [10], and Comte and Renault 1998 [34], are sometimes called integrated volatilities.

3. forward-looking market-based volatilities inferred from option prices. These models include implied and *local volatilities*. The local volatility, introduced by Dupire 1994 [48] and Derman and Kani 1994 [43], is a unique state-dependent diffusion coefficient $\sigma = \sigma(t, X_t)$, which is consistent with the market prices of European options. This volatility is out of our scope, and we refer to the end of this chapter for related literature.

The integrated volatility as a realized volatility is based on historical information. Later in this paper, we will introduce the concept of *integrated volatility implied by option market prices*. One should notice the different structure of these two integrated volatilities.

4.1 Modeling the continuous time stochastic volatility process

If the stock price volatility is constant, then, according to (37), the distribution of the stock price returns is log-normal. However, empirical studies on stock price data show that this is not the case in reality. The returns distributions tend to be fat-tailed and left-skewed compared to the log-normal distribution. This has been interpreted as a sign that the volatility is a stochastic process or a diffusion process with jumps.

Stochastic volatility has dominated the option pricing literature in mathematical finance and financial economics since the 1980s. From the late 1990s, stochastic volatility models have taken center stage in the econometric analysis of volatility forecasting. Nowadays, there is a wide range of stock price models with stochastic volatility. An important principle is that the model should be able to fit to historical price data and have the ability to explain option price smiles both over strike and over maturity. It should also give superior performance in risk management over the one given by the Black-Scholes model. Finally, it should be such that the estimation of model parameters is straightforward.

Important early papers with stochastic volatility are Cox and Ross 1976 [40], Geske 1979 [55], Hull and White 1987 [66], Wiggins 1987 [112], Scott 1987 [104], Johnson and Shanno 1987 [76], Melino and Turnbull 1990 [89], Stein and Stein 1991 [107] and Heston 1993 [64]. In these models, as well as in several others, the stock price process is modeled as

$$dX_t = \mu X_t dt + \sigma_t X_t dW_t, \tag{54}$$

$$\sigma_t = f(Y_t), \tag{55}$$

where μ is the expected rate of return, σ_t is the volatility, W_t is a Brownian motion, and Y_t is a stochastic process. The function f as well as the process Y_t vary in different models. It could be, for example, $f(y) = \sqrt{y}$ (used, for example, in Hull and White 1987 [66]) or $f(y) = e^y$ (used, for example, in Scott 1987 [104]), but there are many other possibilities. The weakness in using a volatility function f is that inference on the process Y_t becomes very difficult. In many models, there is correlation between the Brownian motions of the stock price process and the volatility process. A table on the characteristics of the models mentioned above is presented in Cont and Tankov 2004 [39] and in Fouque, Papanicolaou, and Sircar 2000 [51].

In general, with the exceptions of Heston 1993 [64] and Stein and Stein 1991 [107], the models mentioned above require either the use of Monte-Carlo simulation or the numerical solution of a two-dimensional parabolic differential equation. Estimating the development of stock prices or option prices with these models is computationally demanding.

From our point of view, the work of Hull and White 1987 [66] is of particular interest. Hull and White showed how stochastic stock price volatilities induced smiles and skews in the implied volatilities. They also derived a pricing function for options with stochastic volatility. These results are a bedrock of our work.

According to historical stock price data, stock price volatility seems to be *mean-reverting*, that is, oscillating around a long-term average. A commonly used model for the volatility process is the mean-reverting Ornstein-Uhlenbeck process introduced by Scott 1987 [104] and Stein and Stein 1991 [107]. The stock price process with this volatility is given by

$$\begin{aligned} dX_t &= \mu X_t dt + \sigma_t X_t dW_t, && (56)\\ \sigma_t &= f(Y_t), && (57)\\ dY_t &= \alpha(m - Y_t)dt + \beta d\hat{Z}_t, && (58)\\ \hat{Z}_t &= \rho W_t + \sqrt{1-\rho^2} Z_t, && (59) \end{aligned}$$

where $\alpha > 0$ is the rate of mean reversion, m is the long-run mean level of Y_t, β is the volatility of Y_t, Z_t is the Brownian motion of Y_t, and $\rho \in [-1, 1]$ is the correlation coefficient between $\hat{Z}_t$ and W_t, assumed to be constant, and given by (3). In the subsequent chapters, we will model the stock price process as (56)-(59). Figure 2 represents paths of two mean reverting volatilities and the respective stock price returns. The paths are generated with a discretized version of (56)-(59), where $f(y) = \sqrt{y}$, the discrete time step $\Delta t = .005$, $\mu = 0$, $m = .12$, $\beta = .3$, and $\rho = 0$. A similar simulation is presented in Fouque, Papanicolaou, and Sircar 2000 [51].

According to empirical studies the stock price model with stochastic volatility driven by one Brownian motion is unable to fit the volatility smiles implied by options with short maturity. Bates 2000 [12], Duffie, Pan, and Singleton 1999 [46] and Eraker, Johannes and Polson 2000 [75], among others, suggest to include a jump component in the stochastic volatility processes Y_t to enable rapid moves for the volatility, for example,

$$dY_t = \mu dt + \sigma_t dW_t + \kappa_t dq_t,$$

where κ_t denotes the jump size and dq_t is a counting process with $dq_t = 1$ corresponding to a jump at t and $dq_t = 0$ to the case without a jump.

Several papers, such as Fouque, Papanicolaou, and Sircar 2000 [51] and Jones 2003 [77], suggest that the volatility σ_t in (54) should be driven by two stochastic processes, $\sigma_t = f(Y_t, V_t)$. The first process Y_t could, for example, model the instantaneous moves of the volatility, and the second process V_t longer-term trends.

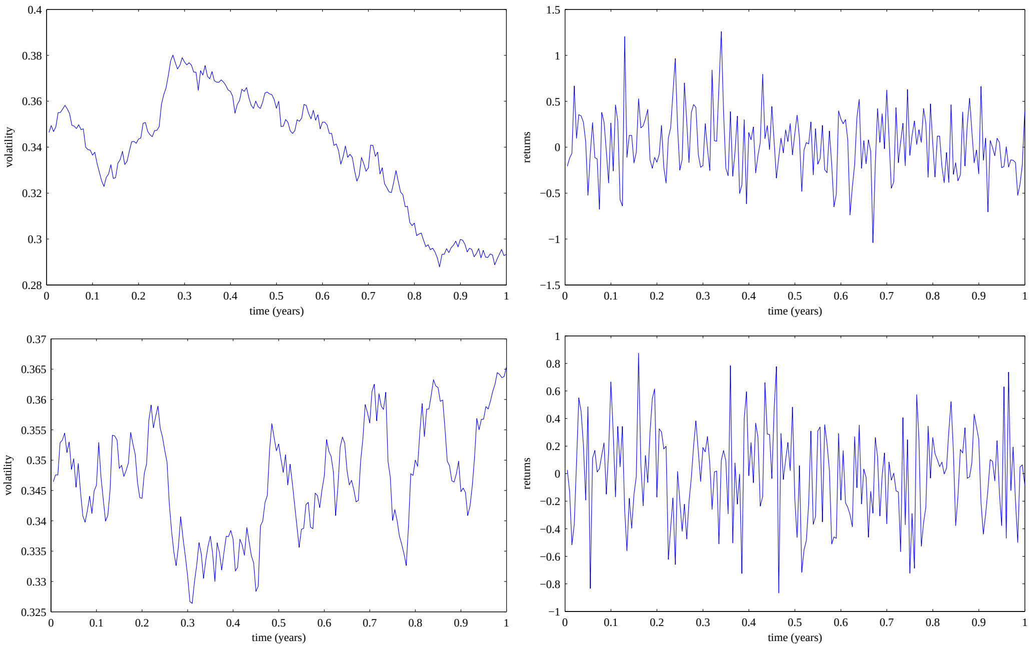


Figure 2: Simulated mean-reverting volatilities when the rate of mean reversion of the volatility is either small, $\alpha = .5$, (top left) or large, $\alpha = 200$, (bottom left), and corresponding returns paths.

4.2 Estimating continuous-time stochastic volatility models

Except for a few special cases, the estimation of continuous time stochastic volatility models is complicated and computationally intensive. As the volatility is a latent variable, one must first decide how to model the volatility process and possibly the stock price process, and then choose a method to estimate the parameters of the models. Jacquier and Jarrow 2000 [69] point out that especially when models are complex, the estimation method becomes crucial.

One possibility is to estimate the parameters of the stochastic volatility model from historical time series data of stock price returns. Common methods are, among others, maximum likelihood methods or various methods of moments. Using these methods, one tries to optimize a chosen criterion computed from the observed data. In maximum likelihood methods, one tries to optimize a likelihood function. In the methods of moments, this criterion is generally based on some difference of moments of the distribution and the corresponding empirical moments. These methods are beyond the scope of this work, and we refer the reader to the end of the chapter for related literature.

An alternative to using historical stock price data is to use option market data and calculate an output least squares (OLS) estimate for the volatility parameters from option prices. For example, when estimating the parameters

$\theta \in \mathbb{R}^n$, one tries to minimize

$$\theta = \arg\min_{\theta \in \mathbb{R}^n} \sum (U(K,T;\theta) - u^{\text{obs}}(K,T))^2,$$

where $U(K,T;\theta)$ is the option price predicted by a pricing model.

A third possibility is to use data from both stock price returns and option prices. One must then cope with the problem of two different probability measures, the natural measure implied by stock prices and the risk-neutral measure implied by option market prices. We have derived in Chapter 2.5 this risk-neutral measure in the case of constant volatility. We will next consider this measure in the case of stochastic volatility.

4.3 Market price of volatility risk

If the stock price volatility is stochastic or has jumps, the markets are incomplete and it is not possible to completely replicate options on the stock. This introduces a *volatility risk* affecting the prices of options. In this situation, there is no universal scheme for pricing options.

Two approaches for modeling the volatility risk have been suggested in the literature. In certain cases, e.g. Hull and White 1987 [66], Johnson and Shanno 1987 [76] and Scott 1987 [104], the risk is assumed to be non-systematic and, therefore, it has zero price. In other cases, e.g. Melino and Turnbull 1990 [89], the risk premium is modeled in a tractable functional form, with extra parameters to be estimated from observed option prices. However, even if the risk premium is non-zero, often there is no need to know explicitly it as long as all assets are defined in the same probability space.

The price of European options can be given as the expected discounted payoff with respect to an equivalent martingale measure which incorporates a compensation for systematic asset, volatility, interest rate, and jump risk. We now show how such a measure is constructed.

Consider a stock price process $(X_t)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, P)$, equipped with the filtration $(\mathcal{F}_t)_{t\geq 0}$ and let $(X_t)_{t\geq 0}$ satisfy the stochastic differential equation (56)-(59). As explained in Chapter 2.5, unless $\mu = r$, the discounted price $\tilde{X}_t = e^{-rt}X_t$ is not a martingale under the probability measure P as it contains a drift term $\mu - r$.

Let us absorb the drift term in (56) in the martingale term, so that

$$\begin{aligned}
d\tilde{X}_t &= (\mu - r)\tilde{X}_t dt + f(Y_t)\tilde{X}_t dW_t \\
&= f(Y_t)\tilde{X}_t \left[\frac{(\mu - r)}{f(Y_t)} dt + dW_t \right] \\
&= f(Y_t)\tilde{X}_t dW_t^*,
\end{aligned}$$

where

$$dW_t^* = \frac{(\mu - r)}{f(Y_s)} + dW_t.$$

Then, according to the Girsanov Theorem, dW_t^* is a Brownian motion under a probability measure P^*, which is an equivalent martingale measure to P. The discounted stock price $\tilde{X}_t = e^{-rt}X_t$ is then a martingale under P^* and, for any European option with payoff $h(X_T)$ and maturity T, an arbitrage-free price option V_t, $0 \leq t \leq T$, is given by the formula

$$V_t = \mathbb{E}^*\{e^{-r(T-t)}h(X_T) \mid \mathcal{F}_t\} \tag{60}$$

where $\mathbb{E}^*$ denotes the expectation with respect to the measure P^*.

Option market prices are possibly affected by a premium for the volatility risk, reflected as a trend in the volatility process. To cope with this phenomenon we make a shift to the Brownian motion of the volatility , i.e., we define

$$Z_t^* = Z_t + \int_0^t \gamma_s ds,$$

where γ_s is any adapted and suitably regular process. We call γ_t the *market price of volatility risk.*

The measure $P^{*(\gamma)}$ is defined by its Radon-Nikodym derivative as

$$\frac{dP^{*(\gamma)}}{dP} = \exp\left(-\frac{1}{2}\int_0^T \left(\frac{(\mu - r)}{f(Y_s)}^2 + \gamma_s^2\right)ds - \int_0^T \frac{(\mu - r)}{f(Y_s)}dW_s - \int_0^T \gamma_s dZ_s\right),$$

where we assume the pair $(\frac{(\mu-r)}{f(Y_s)}, \gamma_s)$ to be such that $P^{*(\gamma)}$ is a well-defined equivalent martingale measure. This will be the case if f is bounded away from zero and γ_t is bounded.

According to Girsanov's theorem, W^* and Z^* are independent standard Brownian motions under the measure $P^{*(\gamma)}$. Then, under this measure, the stochastic differential equation (56)-(59) becomes

$$\begin{aligned} dX_t &= rX_t dt + f(Y_t)X_t dW_t^*, && (61)\\ dY_t &= \left[\alpha(m - Y_t) - \beta\left(\rho\frac{(\mu - r)}{f(Y_t)} + \sqrt{1-\rho^2}\gamma_t\right)\right]dt + \beta d\hat{Z}_t^*. && (62)\\ \hat{Z}_t^* &= \rho W_t^* + \sqrt{1-\rho^2}Z_t^*. && (63) \end{aligned}$$

The corresponding option price without possibility to arbitrage is given by

$$V_t = \mathbb{E}^{*(\gamma)}\{e^{-r(T-t)}h(X_T) \mid \mathcal{F}_t\}. \tag{64}$$

We will use this equation in the next chapter, when deriving prices to options with stochastic volatility. If we write $\gamma_t = \gamma(t, X_t, Y_t)$, the process (61)-(62)

remains a Markovian process. In the sequel, we will assume that this is the case.

Because the volatility is stochastic, these is no one unique martingale measure under which (64) is a martingale. According to the Girsanov theorem any adapted and suitably regular choice of γ_t provides a possible martingale measure, so that γ_t parametrizes the space of equivalent martingale measures $\{P^{*(\gamma)}\}$. In practice, γ_t is calibrated from market data to fit a chosen stock price model with the corresponding market data.

4.4 Notes on references

The litterature on various volatility models grows at an increasing rate. We recommend the recent paper with a large amount of references by Andersen, Bollerslev, Christoffensen, and Diebold 2005 [4] as an introduction to the topic.

Later in this thesis, we will use a Bayesian approach to estimate volatilities. Bayesian methods have been used to estimate the stock price volatility by Jacquier, Polson and Rossi 1994 [70], Jacquier and Jarrow 1996 [68], and Jacquier, Polson and Rossi 2004 [71]. Jones 2003 [77] infer a stochastic volatility model from implied volatilities using Bayesian MCMC methods while Forbes, Martin and Wright 2002 [50] suggest a Bayesian approach to estimate a stochastic model using both the prices of the underlying and the option prices.

5 Hull-White paradigm

In 1987, Hull and White [66] showed how the price of a European option with stochastic volatility can be given as the expectation of Black-Scholes prices integrated over the distribution of the integrated volatility, provided that the correlation between the Brownian motions of the stock price and the volatility vanishes. In other words, they suggested that the price of an option with stochastic volatility would be given by

$$U^{\mathrm{HW}}(t, X_t; K, T; \sigma_t^2) = \int U^{\mathrm{BS}}(t, X_t; K, T; \bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2 \mid \sigma_t^2) d\bar{\sigma}_t^2, \qquad (65)$$

where U^{BS} is the Black-Scholes option price (49), X_t is the stock price, K is the strike price, T is the maturity, σ_t^2 is the instantaneous variance, and $\bar{\sigma}_t^2$ is the integrated volatility, given by

$$\bar{\sigma}_t^2 = \frac{1}{T-t}\int_t^T \sigma_s^2 ds, \qquad 0 \leq t < T. \qquad (66)$$

The equation (65) is the bedrock of our paper. It gives the option price as a function of the *distribution of the integrated volatility*. In Chapter 8, we will formulate a related inverse problem and present a method to estimate the distribution of the *implied integrated volatility*, that is the distribution of the integrated volatility implied by option market prices. In this very chapter, we consider the Hull-White paradigm more generally. We first give alternative derivations for the Hull-White formula and then give one explanation for the volatility smile, based on this formula. We also show how different volatility smile patterns can be explained by different distributions of the implied integrated volatility.

5.1 Hull-White formula

In this chapter, we derive the Hull-White formula based on the derivation in the seminal paper of Hull and White 1987 [66]. Assume that the stock price process is given by (61)-(63) under a martingale measure $P^{*(\gamma)}$ equivalent with the natural measure P, and that the correlation between the Brownian motions of the stock price process and the volatility process vanishes. As presented in (40)-(43), the logarithm of the stock price return on a time interval $[t, T]$ is obtained by applying the Itô formula (28) to the logarithm of the stock price (61) and by integrating all terms from t to T, which leads to

$$\log\frac{X_T}{X_t} = r(T-t) - \frac{1}{2}\int_t^T \sigma_s^2 ds + \int_t^T \sigma_s dW_s^*. \qquad (67)$$

In particular, the logarithm of the stock price return is normally distributed conditional on the integrated volatility $\bar{\sigma}_t^2$, so that

$$\begin{aligned}
\mathbb{E}^{*(\gamma)}\big\{\log\frac{X_T}{X_t}\mid\mathcal{F}_t,\bar{\sigma}_t^2\big\} &= \big(r-\frac{1}{2}\bar{\sigma}_t^2\big)(T-t),\\
\mathrm{Var}^{*(\gamma)}\big\{\log\frac{X_T}{X_t}\mid\mathcal{F}_t,\bar{\sigma}_t^2\big\} &= \bar{\sigma}_t^2(T-t),
\end{aligned}$$

where $\mathrm{Var}^{*(\gamma)}$ denotes the variance with respect to the probability measure $P^{*(\gamma)}$.

As explained in Chapter 4.3, a possible no-arbitrage price for an option on X_t is given as the discounted expectation of the payoff $h(X_T)$ under the measure $P^{*(\gamma)}$, that is

$$U(t,X_t;K,T;\sigma_t^2) = \mathbb{E}^{*(\gamma)}\big\{e^{-r(T-t)}h(X_T)\mid\mathcal{F}_t,X_t,\sigma_t^2\big\}, \tag{68}$$

$$= e^{-r(T-t)}\int h(X_T)\pi(X_T\mid X_t,\sigma_t^2)dX_T, \tag{69}$$

where $\pi(X_T\mid X_t,\sigma_t^2)$ is the density of the risk neutral measure.

To simplify this equation, we use the fact that if three random variables a, b and c are related such that a depends on c only through b, the conditional density functions are related by

$$\pi(a\mid b)=\int\pi(a\mid c)\pi(c\mid b)dc.$$

As X_T depends on σ_t^2 only through $\bar{\sigma}_t^2$, we can write

$$\pi(X_T\mid\sigma_t^2)=\int\pi(X_T\mid\bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2\mid\sigma_t^2)d\bar{\sigma}_t^2. \tag{70}$$

We substitute (70) into (69), which gives us

$$\begin{aligned}
U_t(X_t,\sigma_t^2) &= e^{-r(T-t)}\int\int h(X_T)\pi(X_T\mid\bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2\mid\sigma_t^2)dX_T d\bar{\sigma}_t^2, &(71)\\
&= \int\Big[e^{-r(T-t)}\int h(X_T)\pi(X_T\mid\bar{\sigma}_t^2)dX_T\Big]\pi(\bar{\sigma}_t^2\mid\sigma_t^2)d\bar{\sigma}^2. &(72)
\end{aligned}$$

Now, the inner term in (72) is the Black-Scholes price of an option on a stock with a constant volatility equal to $\sqrt{\bar{\sigma}_t^2}$, that is

$$U_t^{\mathrm{BS}}(X_t,\bar{\sigma}_t^2)=\Big[e^{-r(T-t)}\int h(X_T)\pi(X_T\mid\bar{\sigma}_t^2)dX_T\Big]. \tag{73}$$

We substitute (73) into (72), getting the Hull-White formula (65):

$$U_t(X_t,\sigma_t^2)=\int U_t^{\mathrm{BS}}(X_t,\bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2\mid\sigma_t^2)d\bar{\sigma}_t^2. \tag{74}$$

The use of this formula to option pricing is limited by the fact that the distribution of the integrated volatility is unknown and has to be simulated using some chosen volatility model. The accuracy of the option prices depends on how well the volatility process has been modeled. On the other hand, as explained in Chapter 3.1, the prices of liquid vanilla options, including European options, are defined by markets, according to offer and demand; therefore, there is no need to calculate their prices using, for example, the Hull-White formula.

We are interested in the Hull-White formula for another reason. If option market prices coincide with the corresponding Hull-White prices, the Hull-White formula can be used to extract information on the integrated volatility implied by option market prices. We will discuss this idea more deeply in Chapter 8.

5.2 Hull-White formula with correlated volatility

Renault and Touzi 1996 [98] have shown that the Hull-White formula produces a symmetric volatility smile when the implied volatilities are plotted as a function of *log-moneyness* given by $\log(K/(xe^{r(T-t)}))$. Asymmetric smiles could then be explained by a correlation between the Brownian motions of the stock price and the volatility. In 1996, Willard [113] derived a variation of the Hull-White equation, with a non-zero correlation between the Brownian motions of the stock price process and of the volatility process. For completeness of the Hull-White paradigm, we derive the Hull-White formula for options with correlated volatility, even if it will not be used in this paper.

We begin by writing the stock price process given by (61)-(63) as

$$
\begin{aligned}
dX_t &= rX_t dt + f(Y_t)X_t(\sqrt{1-\rho^2}d\hat{W}_t^* + \rho d\hat{Z}_t^*),\\
dY_t &= \left[\alpha(m-Y_t) - \beta\left(\rho\frac{(\mu-r)}{f(Y_t)} + \gamma_t\sqrt{1-\rho^2}\right)\right]dt + \beta d\hat{Z}_t^*\\
\hat{W}_t^* &= \sqrt{1-\rho^2}W_t^* - \rho Z_t^*,
\end{aligned}
$$

where $\hat{W}_t^*$ and $\hat{Z}_t^*$ are two independent Brownian motions.

We apply the Itô formula (28) to the logarithm of X_t, getting

$$
d\log X_t = (r - \frac{1}{2}\sigma_t^2)dt + \sigma_t(\sqrt{1-\rho^2}dW_t^* + \rho dZ_t^*), \tag{75}
$$

and integrate both sides of (75) from t to T, which gives us

$$
\begin{aligned}
\log X_T &= \log X_t + \left(\int_t^T (r - \frac{1}{2}\sigma_s^2)ds + \rho\int_t^T \sigma_s dZ_s^* + \sqrt{1-\rho^2}\int_t^T \sigma_s dW_s^*\right)\\
&= \log(\xi_{t,T}X_t) + \left(\int_t^T (r - \frac{1-\rho^2}{2}\sigma_s^2)ds + \sqrt{1-\rho^2}\int_t^T \sigma_s dW_s^*\right).
\end{aligned}
$$

where

$$\xi_{t,T} = \exp\left(- \frac{1}{2}\rho^2 \int_t^T \sigma_s^2 ds + \rho \int_t^T \sigma_s dZ_s^* \right). \tag{76}$$

The logarithm of the stock price return is normally distributed conditional on $\bar{\sigma}_t^2$ and $\xi_{t,T}$, with

$$\begin{aligned} \mathbb{E}^{*(\gamma)}\{ \log \frac{X_T}{X_t} \mid \mathcal{F}_t, \bar{\sigma}_t^2, \xi_{t,T}\} &= \log(\xi_{t,T}) + (r - \frac{1}{2}(1-\rho^2)\bar{\sigma}_t^2)(T-t), \\ \mathrm{Var}^{*(\gamma)}\{ \log \frac{X_T}{X_t} \mid \mathcal{F}_t, \bar{\sigma}_t^2, \xi_{t,T}\} &= (1-\rho^2)\bar{\sigma}_t^2(T-t). \end{aligned}$$

It is important to notice that here, due to the stochastic integral in (76), the logarithm of the stock price return is not normally distributed conditional on the integrated volatility. Instead, it is normally distributed conditional on the entire path of $\hat{Z}_t^*$.

Now, similarly as in Chapter 5.1, an arbitrage free option price is given by (72) where the inner term equals the Black-Scholes option price on a stock with value $x\xi_{t,T}$ and a constant volatility equal to $\sqrt{\bar{\sigma}_t^2}$. We write

$$U_t^{BS}(x\xi_{t,T}; K, T; \bar{\sigma}_t^2) = \left[e^{-r(T-t)} \int h(X_T)\pi(X_T \mid \bar{\sigma}_t^2)dX_T \mid X_t = x\xi_{t,T} \right], \tag{77}$$

substitute (77)into (72), and get

$$U_t(x; K, T; \sigma_t^2) = \int U_t^{\mathrm{BS}}(x\xi_{t,T}; K, T; \bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2 \mid \sigma_t^2)d\bar{\sigma}_t^2. \tag{78}$$

It is easy to check from the Black-Scholes formula (49) that the identity

$$U^{\mathrm{BS}}(t, x\xi_{t,T}; K, T; \sigma^2) = \xi_{t,T} U^{\mathrm{BS}}(t, x; K\xi_{t,T}^{-1}, T; \sigma^2)$$

holds. Therefore, we can write the Hull-White equation for correlated volatility as

$$U_t^{\mathrm{HW},\rho}(x; K, T; \sigma_t^2) = \xi_{t,T} \int U_t^{\mathrm{BS}}(x; K\xi_{t,T}^{-1}, T; \bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2 \mid \sigma_t^2)d\bar{\sigma}_t^2, \tag{79}$$

where we have denoted

$$U_t^{\mathrm{HW},\rho}(x; K, T; \sigma_t^2) = U_t(x; K, T; \sigma_t^2).$$

According to empirical studies, there is usually a negative correlation between the Brownian motions of the stock price process and the volatility process. Precise estimates vary widely according to the method used. It seems that the negative correlation is more pronounced for indices than for stocks. In some markets, like foreign currency option markets, the assumption of zero correlation might be tenable. In general, the practical importance of correlation is unclear for the prices of at-the-money options. According to Jarrow and Rudd 1982 [72], the additional error caused by non-zero correlation is likely to be small.

5.3 Hull-White formula by replicating portfolios

In this section, we present an alternative way to derive Hull-White formula. We start from the stock price process under the natural measure P and, using replicating portfolios, derive an arbitrage-free price for options with stochastic volatility. This price is further developed in the form of the Hull-White formula. The reason to present this approach is that the replicating portfolios are a central tool in risk management and financial *hedging* and will be needed later, when discussing hedging in Chapters 6 and 11.

Suppose that the stock whose price process is given by (56)-(59), that is

$$\begin{aligned} dX_t &= \mu X_t dt + \sigma_t X_t dW_t \\ \sigma_t &= f(Y_t), \\ dY_t &= \alpha(m - Y_t)dt + \beta dZ_t, \\ \hat{Z}_t &= \rho W_t + \sqrt{1-\rho^2} Z_t, \end{aligned}$$

where $\alpha > 0$ is the rate of mean reversion, m is the long-run mean level of Y_t, β is the volatility of Y_t, and ρ is the correlation coefficient between $\hat{Z}_t$ and W_t. Suppose that we want to define the value of a European call option on this stock. We denote this option by $U_t^{(1)} = U^{(1)}(t, X_t, Y_t)$.

In order to price the option correctly, we construct a portfolio replicating its value until maturity T. As both the stock price and the volatility are stochastic, there will be two sources of randomness. To replicate $U^{(1)}(t, X_t, Y_t)$, we need, in addition to stocks and bonds, a certain amount of another option on the same underlying but with longer maturity $T_2, T_2 > T$. We denote this *replicating option* by $U_t^{(2)} = U^{(2)}(t, X_t, Y_t)$ and the instantaneous amounts of stocks, bonds, and replicating options to hold by a_t, b_t, and c_t, respectively. To secure well-defined stochastic integrals, we assume that $\mathbb{E}\{\int_0^T a_t^2 dt\} < \infty$, and to secure well-defined original integrals, we assume that $\int_0^T |b_t| dt < \infty$. The value of the bond is $B_t = e^{rt}$, where we have assumed that $B_0 = 1$.

The value $V(t, X_t, Y_t)$ of the replicating portfolio is now

$$V(t, X_t, Y_t) = a_t X_t + b_t e^{rt} + c_t U^{(2)}(t, X_t, Y_t),$$

and the option is replicated at any time $0 \le t \le T$, if, almost surely,

$$\begin{aligned} U^{(1)}(t, X_t, Y_t) &= V(t, X_t, Y_t) \\ &= a_T X_T + b_T e^{rt} + c_T U^{(2)}. \end{aligned} \tag{80}$$

The portfolio should be self-financing, so that

$$dV_t = a_t dX_t + b_t r e^{rt} dt + c_t dU_t^{(2)}. \tag{81}$$

A change in the value of the portfolio is due to changes in the prices of the instruments, not to changes in the amounts. If such a portfolio exists, to avoid the opportunity of arbitrage, we must have

$$U_t^{(1)} = a_t X_t + b_t e^{rt} + c_t U_t^{(2)} \text{ for every } 0 \leq t \leq T. \tag{82}$$

To calculate the instantaneous amounts a_t, b_t, and c_t of stocks, bonds, and replicating options in the replicating portfolio, we apply the two-dimensional version of the Itô formula (28) to $U_t^{(1)} = U^{(1)}(t, X_t, Y_t)$, getting

$$\begin{aligned} dU^{(1)}(t, X_t, Y_t) = \bigg(& \frac{\partial U^{(1)}}{\partial t} + \frac{1}{2} f(y)^2 x^2 \frac{\partial^2 U^{(1)}}{\partial x^2} + \rho \beta x f(y) \frac{\partial^2 U^{(1)}}{\partial x \partial y} \\ & + \frac{1}{2}\beta^2 \frac{\partial^2 U^{(1)}}{\partial y^2} \bigg) dt + \frac{\partial U^{(1)}}{\partial x} dX_t + \frac{\partial U^{(1)}}{\partial y} dY_t, \end{aligned} \tag{83}$$

where we have denoted $x = X_t$ and $y = Y_t$. We then equate $U(t, X_t, Y_t) = V(t, X_t, Y_t)$ and apply the Itô formula (28) to the right-hand side of (81), getting

$$\begin{aligned} dU^{(1)}(t, X_t, Y_t) &= \left(a_t + c_t \frac{\partial U^{(2)}}{\partial x} \right) dX_t + c_t \frac{\partial U^{(2)}}{\partial y} dY_t \\ + \bigg[c_t \bigg(& \frac{\partial U^{(2)}}{\partial t} + \frac{1}{2} f(y)^2 x^2 \frac{\partial^2 U^{(2)}}{\partial x^2} + \rho \beta x f(y) \frac{\partial^2 U^{(2)}}{\partial x \partial y} + \frac{1}{2}\beta^2 \frac{\partial^2 U^{(2)}}{\partial y^2} \bigg) + b_t r e^{rt} \bigg] dt, \end{aligned} \tag{84}$$

where the derivatives of $U^{(1)}$ and $U^{(2)}$ are evaluated at (t, X_t, Y_t).

We solve c_t by equating the dY_t terms from (83) and (84), which gives us

$$c_t = \frac{\partial U^{(1)}/\partial y}{\partial U^{(2)}/\partial y}, \tag{85}$$

the amount of the replicating option $U_t^{(2)}$ we should hold in our portfolio.

Equating the dX_t terms in (83) and (84) leads to

$$a_t = \frac{\partial U^{(1)}}{\partial x} - c_t \frac{\partial U^{(2)}}{\partial x}, \tag{86}$$

the amount of the underlying stock we should hold in our portfolio, and solving b_t from (82) gives the amount of the bond we should hold in our portfolio, that is

$$b_t = e^{-rt}(U_t^{(1)} - a_t X_t - c_t U_t^{(2)}).$$

We equate the dt terms in (83) and (84) where we have substituted a_t, b_t, and c_t. This gives us

$$\left(\frac{\partial U^{(1)}}{\partial t}+\frac{1}{2}f(y)^2x^2\frac{\partial^2 U^{(1)}}{\partial x^2}+\rho\beta x f(y)\frac{\partial^2 U^{(1)}}{\partial x\partial y}+\frac{1}{2}\beta^2\frac{\partial^2 U^{(1)}}{\partial y^2}\right)dt$$
$$=\frac{\partial U^{(1)}/\partial y}{\partial U^{(2)}/\partial y}\left(\frac{\partial U^{(2)}}{\partial t}+\frac{1}{2}f(y)^2x^2\frac{\partial^2 U^{(2)}}{\partial x^2}+\rho\beta x f(y)\frac{\partial^2 U^{(2)}}{\partial x\partial y}+\frac{1}{2}\beta^2\frac{\partial^2 U^{(2)}}{\partial y^2}\right)$$
$$+r\left(U_t^{(1)}-\left(\frac{\partial U^{(1)}}{\partial x}-\frac{\partial U^{(1)}/\partial y}{\partial U^{(2)}/\partial y}\frac{\partial U^{(2)}}{\partial x}\right)x-\frac{\partial U^{(1)}/\partial y}{\partial U^{(2)}/\partial y}\frac{\partial U^{(2)}}{\partial x}\right) \quad (87)$$

We denote

$$\mathcal{A}: u\rightarrow \frac{\partial u}{\partial t}+\frac{1}{2}f(y)^2x^2\frac{\partial u^2}{\partial x^2}+\rho\beta x f(y)\frac{\partial u^2}{\partial x\partial y}$$
$$+\frac{1}{2}\beta^2\frac{\partial u^2}{\partial y^2}+r\left(x\frac{\partial u}{\partial x}-u\right),\quad u=u(t,x,y),$$

and reorganize the terms in (87) so that it can be written as

$$\left(\frac{\partial U^{(1)}}{\partial y}\right)^{-1}\mathcal{A}U^{(1)}(t,X_t,Y_t)=\left(\frac{\partial U^{(2)}}{\partial y}\right)^{-1}\mathcal{A}U^{(2)}(t,X_t,Y_t). \quad (88)$$

Now, the left-hand side of equation (88) contains terms depending on the maturity T but no terms depending on the maturity T_2. The reverse is true for the right-hand side. Accordingly, both sides must be independent of the maturity. In addition, they must be independent of the strike price of the options, as we have not defined it for $U_t^{(1)}$ and $U_t^{(2)}$.

We write both sides of (88) equal to a function γ, depending on the state variables t, X_t, and Y_t, and on the parameters μ and r, but not depending on K and T. In our case, the exact form of this function is not important.

The pricing function $U(t,X_t,Y_t)$ for $U=U^{(1)}$ and $U=U^{(2)}$ must satisfy

$$\frac{\partial U}{\partial t}+\frac{1}{2}f(y)^2x^2\frac{\partial^2 U}{\partial x^2}+\rho\beta x f(y)\frac{\partial^2 U}{\partial x\partial y}+\frac{1}{2}\beta^2\frac{\partial^2 U}{\partial y^2}$$
$$+rx\frac{\partial U}{\partial x}-rU=-\gamma(t,x,f(y);r)\frac{\partial U}{\partial y}. \quad (89)$$

Neither (89) nor the final conditions $h(X_T)=(X_T-K)^+$ depend on the risk preferences of the investor. Furthermore, we have assumed that there is no correlation between the Brownian motions of the stock price process and of the volatility process. Then, at time t, the option price $U(t,X_t,\sigma_t^2)$ must equal to the present value of the discounted expectation of the payoff $h(X_T)$ at maturity, conditional on the stock price X_t and on the variance σ_t^2, given by

$$U(t,X_t;K,T;\sigma_t^2)=\mathbb{E}^{*(\gamma)}\{e^{-r(T-t)}h(X_T)\mid \mathcal{F}_t,X_t,\sigma_t^2\}, \quad (90)$$

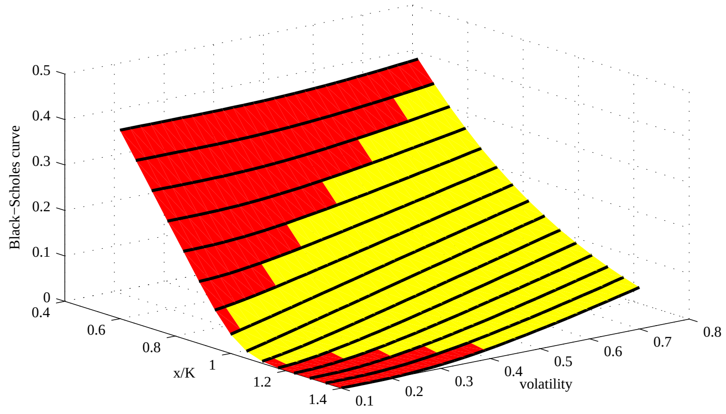


Figure 3: Black-Scholes surface where yellow indicates convexity and red concavity.

where $\mathbb{E}^{*(\gamma)}$ denotes the expectation with respect to the measure $P^{*(\gamma)}$.

This is Equation (69). In fact, Equation (89) with its final condition is related to (69) by the Feynman-Kac formula (29). To proceed and obtain the Hull-White formula, we continue from (90) as in Chapter 5.1.

5.4 Hull-White formula and volatility smile

Various explanations for the volatility smile have been proposed in the literature. We present one explanation for the smile, based in the distribution of the integrated volatility, in the convexity of the Black-Scholes formula, and in the Hull-White formula (65). To begin, we define the *expectation of the integrated volatility* $\hat{\sigma}_t^2$ as the expectation of the conditional distribution $\pi(\bar{\sigma}_t^2 \mid \sigma_t^2)$, given by

$$\hat{\sigma}_t^2 = \mathbb{E}\{\bar{\sigma}_t^2 \mid \sigma_t^2\} = \int \bar{\sigma}_t^2 \pi(\bar{\sigma}_t^2 \mid \sigma_t^2) d\bar{\sigma}_t^2.$$

The convexity of the Black-Scholes formula depends on the value of the volatility. When the Black-Scholes formula U^{BS} is a convex function of $\bar{\sigma}_t^2$, then, by Jensen's inequality, $\mathbb{E}\{U^{BS}(\bar{\sigma}_t^2)\} > U^{BS}(\mathbb{E}\{\bar{\sigma}_t^2\})$. For concave U^{BS} the reverse is true, $\mathbb{E}\{U^{BS}(\bar{\sigma}_t^2)\} < U^{BS}(\mathbb{E}\{\bar{\sigma}_t^2\})$. The Hull-White price is the expectation of the Black-Scholes prices with respect to the integrated volatility, $U^{\mathrm{HW}} = \mathbb{E}\{U^{BS}(\bar{\sigma}_t^2)\}$. This means that when U^{BS} is convex, $U^{BS}(\hat{\sigma}_t^2) = U^{BS}(\mathbb{E}\{\bar{\sigma}_t^2\})$ underprices options when compared to the Hull-White option prices, and when it is concave, $U^{BS}(\hat{\sigma}_t^2)$ overprices them. In the first case, the implied volatility is higher than the square root of the expectation of the integrated volatility, $I_t > \sqrt{\hat{\sigma}_t^2}$, the reverse being true in

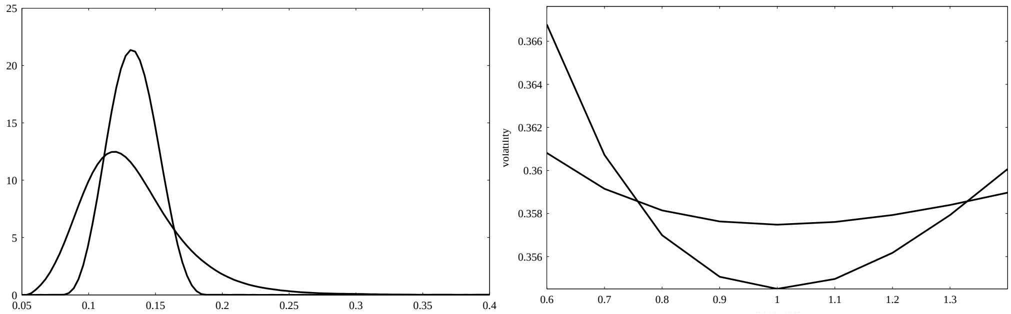


Figure 4: The left panel represents two simulated distributions of the integrated volatility with the same expectation. The right panel represents the volatility smiles that have been calculated from the Hull-White prices of these distributions.

the second case. When we say that U^{BS} is convex or concave with respect to $\bar{\sigma}_t^2$, the whole distribution of this random variable should be in the range of the convexity or concavity of U^{BS}, i.e, $\sigma^2 \mapsto U^{\mathrm{BS}}(\sigma^2)$ should be concave or convex on the whole support of $\pi(\bar{\sigma}_t^2)$.

The inflection points of the Black-Scholes formula can be calculated from the second derivative of U^{BS} with respect to the integrated volatility $\bar{\sigma}_t^2$, given by

$$(U^{\mathrm{BS}})''(\bar{\sigma}_t^2) = \frac{x\sqrt{T-t}}{4(\bar{\sigma}_t^2)^{3/2}}\mathcal{N}'(d_1)(d_1 d_2 - 1),$$

where d_1 and d_2 are as in (50).

The convexity of U^{BS} depends on the sign of $(U^{\mathrm{BS}})''(\bar{\sigma}_t^2)$, which equals the sign of $d_1 d_2 - 1$. The inflection points, denoted by $\mathrm{IP_t}$, is given when $d_1 d_2 = 1$, so that

$$\mathrm{IP_t} = \frac{2}{T-t}\left[\sqrt{1 + (\ln(x/K) + r(T-t))^2} - 1\right].$$

When $\bar{\sigma}_t^2 < \mathrm{IP_t}$, $(U^{\mathrm{BS}})''(\bar{\sigma}_t^2) > 0$, and U^{BS} is a convex function of $\bar{\sigma}_t^2$, while when $\bar{\sigma}_t^2 > \mathrm{IP_t}$, $(U^{\mathrm{BS}})''(\bar{\sigma}_t^2) < 0$, and U^{BS} is a concave function of $\bar{\sigma}_t^2$. If $x = Ke^{-r(T-t)}$, then $\mathrm{IP_t} = 0$ and U^{BS} is concave everywhere, for all values and distributions of $\bar{\sigma}_t^2$. When $\log(x/K) \to \pm\infty$, $\mathrm{IP_t}$ becomes arbitrarily large and U^{BS} is convex everywhere.

The convexity of the Black-Scholes formula is illustrated in terms of the Black-Scholes surface in Figure 3. In practice, this formula overprices options that are at-the-money or near-at-the money. Most option trading is done in this range.

Consider then the existence of volatility smiles from the point of view of integrated volatility. Because the Hull-White price U^{HW} is a function of the distribution of the integrated volatility $\bar{\sigma}_t^2$, not a function of its expectation $\hat{\sigma}_t^2$, different distributions with the same expectation $\hat{\sigma}_t^2$ may give different

Hull-White prices. Assuming that the Hull-White prices coincide with observed option prices, these differences are reflected in implied volatilities and in the shapes of the corresponding volatility smiles.

The phenomenon explained above is illustrated in Figure 4, where we have plotted two distributions of the integrated volatility and the corresponding implied volatilities calculated from Hull-White prices. The distributions have the same expectation but different shapes which results to two different volatility smiles. It is possible that the shapes of volatility smiles calculated from option market prices contain information on the support or skewness of the corresponding integrated volatilities.

We have here restricted the explanations to the case where the correlation between the Brownian motions of the stock price process and the volatility process vanishes. The relation between the volatility smile and the distribution of the integrated volatility is much more complicated in the case that the correlation is non-zero, and is out of our scope.

5.5 Notes on references

The derivation of the Hull-White formula in Chapter 5.1 is based on the one presented in the seminal paper of Hull and White 1987 [66], while the derivation with replicating portfolios in Chapter 5.3 is close to the one presented in Wilmott 1998 [114]. The Hull-White formula with a correlated volatility is derived in Willard 1996 [113] and considered by Roger W Lee 2001 [86]. The comprehensive book of Fouque, Papanicolaou, and Sircar 2000 [51] discusses all of these topics. Section 5.4 is based on Hull and White 1987 [66].

6 Dynamic hedging

The term *hedging* refers to the reduction of randomness using correlations between various risky investments. One can hedge investment portfolios as well as outcomes of horse race bets.

We can divide the hedgers related to options roughly in two categories. The first ones are those having an investment portfolio of, for example, stocks, and wanting to hedge this portfolio with options. The second ones are those selling options to the first ones, and to markets more generally. They hedge changes in the value of the *short option*, i.e., the option that has been sold, with a portfolio consisting of some other traded assets. These hedgers make profit if they can hedge the options better than other hedgers in the markets. In the sequel, when talking about hedgers and hedging, we refer to this latter type.

In static hedging, the goal is to construct a portfolio of liquidly traded standard assets that perfectly replicates the payoff of a given over-the-counter option. This hedge needs no readjustment. The problem with static hedges is that the requirements for the liquidly traded assets are seldom reached.

In dynamic hedging, the attempt is to hold and readjust a portfolio consisting of a short option, of an instantaneous amount of the underlying stock, and possibly of an instantaneous amount of replicating options so that the portfolio is immune to small changes in the price of the underlying asset in the next small interval of time. We call the amounts of stocks and replicating options *hedging ratios*. Re-hedging can not be done too frequently due to transaction costs and the fact that one always pays the higher bid-prices when buying and the corresponding lower offer-prices when selling. It would be valuable to find hedge ratios that do not need too frequent adjusting.

Hedgers are interested in the *greeks*, that is, the sensitivities of the Black-Scholes option prices U_t^{BS} describing how sensitive the value of an option is to changes in the value of the underlying, in the volatility, in time, or in the interest rates. Here, we will use the sensitivity of U_t^{BS} with respect to the stock price X_t, called the *delta*, and the sensitivity with respect to the stock price volatility σ_t, called the *vega*.

The delta, denoted by Δ_t, is given by

$$\Delta_t = \frac{\partial U_t^{\mathrm{BS}}}{\partial x} = \mathcal{N}(d_1), \tag{91}$$

and the vega, denoted by κ_t, is given by

$$\kappa_t = \frac{\partial U_t^{\mathrm{BS}}}{\partial \sigma_t} = \frac{x e^{-d_1^2/2}\sqrt{T-t}}{\sqrt{2\pi}} = x\sqrt{(T-t)}\mathcal{N}'(d_1), \tag{92}$$

where $\mathcal{N}'(d_1) = e^{-d_1^2/2}/\sqrt{2\pi}$, and d_1 is as in (50).

6.1 Hedging performance

The hedging performance depends, among other things, on the *hedging strategy*, on the type of volatility used with the hedging strategy, and on the hedging frequency. Hedging strategies in incomplete markets depend on some dynamic risk-measure that has to be minimized. Common strategies include *delta hedging* and *delta-vega hedging* as well as *minimum-variance hedging*, i.e., hedging so that the variance of the portfolio is minimized.

In delta hedging, changes in the value of the short option U_t^{H} are hedged by trading on the underlying. When the stock price volatility is constant and trading is continuous, it is theoretically possible to replicate the option by holding an instantaneous amount $\Delta_t = \partial U_t^{\mathrm{BS}}/\partial x$ of the underlying and to eliminate in this way all risk of loss if the option is executed. In practice, complete elimination is not possible as trading is done discretely.

When the stock price volatility is stochastic, perfect hedging is not possible even theoretically. One can then attempt to hedge the short option following, instead of the delta strategy, the delta-vega strategy. According to this strategy, the option is replicated by holding an instantaneous amount α_t of the underlying stock and an instantaneous amount γ_t of the replicating option U_t^{Rep} which has the same underlying and strike price than the option to be hedged U_t^{H}, but a longer maturity.

According to equations (86) and (85), the amount α_t of stocks to hold is

$$\alpha_t(\tilde{\sigma}_t) = \frac{\partial U_t^{\mathrm{H}}}{\partial x} - \frac{\partial U_t^{\mathrm{H}}/\partial \tilde{\sigma}_t}{\partial U_t^{\mathrm{Rep}}/\partial \tilde{\sigma}_t} \frac{\partial U_t^{\mathrm{Rep}}}{\partial x} \tag{93}$$

and the amount γ_t of the replicating options to hold is

$$\gamma_t(\tilde{\sigma}_t) = \frac{\partial U_t^{\mathrm{H}}/\partial \tilde{\sigma}_t}{\partial U_t^{\mathrm{Rep}}/\partial \tilde{\sigma}_t}, \tag{94}$$

where $\tilde{\sigma}_t$ denotes some kind of volatility, for example the implied volatility or the local volatility.

The hedging ratios α_t and γ_t can be written in terms of the greeks delta $\Delta_t(\tilde{\sigma}_t)$ and vega $\kappa_t(\tilde{\sigma}_t)$ as

$$\alpha_t(\tilde{\sigma}_t) = -\Delta_t^{\mathrm{H}}(\tilde{\sigma}_t) + \frac{\kappa_t^{\mathrm{H}}(\tilde{\sigma}_t)}{\kappa_t^{\mathrm{Rep}}(\tilde{\sigma}_t)} \Delta_t^{\mathrm{Rep}}(\tilde{\sigma}_t) \tag{95}$$

and

$$\gamma_t(\tilde{\sigma}_t) = -\frac{\kappa_t^{\mathrm{H}}(\tilde{\sigma}_t)}{\kappa_t^{\mathrm{Rep}}(\tilde{\sigma}_t)}, \tag{96}$$

where Δ_t^{H} and κ_t^{H} refer to the delta and vega of the hedged option U_t^{H}, and Δ_t^{Rep} and κ_t^{Rep} refer to those of the replicating option U_t^{Rep}.

The hedging ratios presented above were derived in Chapter 5.3 using the instantaneous stock price volatility. This volatility, unfortunately, is not known. Bender, Sottinen, and Valkeila 2007 [15] have shown that the hedging prices depend essentially only on the quadratic variation of the underlying. This quadratic variation, however, is not directly related to the implied integrated volatility, which is defined under a risk neutral probability measure.

Common choices for the volatility are the implied volatility and the local volatility. We suggest that, in the same way that the Hull-White option price is given as the expectation of Black-Scholes prices integrated over the distribution of the integrated volatility, the hedging ratios could be given as the distributions. The expectations of the hedging ratios, $\hat{\Delta}_t$, $\hat{\alpha}_t$, and $\hat{\gamma}_t$, would then be given by

$$\hat{\Delta}_t(\bar{\sigma}_t^2) = \int \Delta(\bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2, \tag{97}$$

$$\hat{\alpha}_t(\bar{\sigma}_t^2) = \int \alpha(\bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2, \tag{98}$$

$$\hat{\gamma}_t(\bar{\sigma}_t^2) = \int \gamma(\bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2. \tag{99}$$

The delta and vega strategy being robust hedging tools, it possible that the hedging performance of the implied integrated volatility does not outperform the one of other volatilities. However, as distributions, the hedging ratios are more informative than simple point estimates. We present a computed example on the topic in Chapter 11.4.

6.2 Notes on references

Comprehensive presentations on dynamic hedging are provided in the books by Taleb 1997 [109] and by Rebonato 2004 [97]. Early papers include Breeden and Litzenberger 1978 [22] and Green and Jarrow 1987 [58], other important papers are El Karoui, Jeanblanc-Piqué, and Shreve 1998 [81], Avellaneda, Levy, and Paràs 1995 [7], Bates 2003 [13] and 2005 [14]. A lot of references can be found in the recent paper of Alexander and Nogueira 2007 [1] on model-free hedging ratios.

The hedging performance of different volatilities is compared in Bakshi, Cao, and Chen 1997 [9], Bates 1995 [11], and Dumas, Fleming, and Whaley 1998 [47], Coleman, Kim, Li, and Verma 2001 [33]. The paper of Dudenhausen 2002 [44] contains a lot of references on hedging effectiveness under model misspecification, on hedging in incomplete markets, and on hedging with transaction costs.

7 Inverse problems

A commonly encountered problem is that we want to get information about some quantity that is not directly observable. However, this quantity may depend on some other, directly observable quantity. We have an *inverse problem* when, starting from data on this observable quantity, we try to get information about the quantity of interest. This is the situation when we try to get information on the stock price volatility from option market prices.

Consider the inverse problem of solving or estimating f from

$$\mathcal{A}(f) = d$$

when d is observed and the mapping $\mathcal{A}$ is given. We say that the inverse problem is *well-posed* if

1. There exists a solution for any data d in the data space;
2. The solution is unique;
3. The inverse mapping $d \mapsto f$ is continuous.

If small errors in the data d propagate to large errors in f, we say that the problem is *ill-conditioned*. If an inverse problem is not well-posed or if it is numerically ill-conditioned, we say that it is *ill-posed*. Traditionally, the ill-posedness of inverse problems is dealt with by using regularization: The ill-posed problem is replaced with a nearby problem that is well-posed. Another possibility is to state the problem in a new way, using a statistical, *Bayesian approach*: instead of a deterministic problem, the inverse problem is recast in the form of statistical inference on the distribution of the unknown. This approach has several advantages. First, it allows us to integrate additional prior knowledge into our estimation process. Secondly, even if the deterministic problem is ill-posed and lacking a unique solution, there always exists a probability density of the unknown, the variance of which may be large or small.

We now proceed by presenting direct and inverse problems in option pricing with both constant and stochastic volatility on a very general level. Then, before going deeper to a particular inverse problem, we review basic concepts on statistical inverse problems.

7.1 Direct and inverse problems in option pricing

Suppose that we want to price a European call option. We have at least three possibilities to do it, depending on our prior knowledge on the price process and on the volatility of the underlying stock.

1. Assume first that the stock price volatility σ is constant. The option can then be priced with the Black-Scholes formula (49).

2. Alternatively, assume that the stock price process and the volatility process are known. In this case, we can estimate the option prices by simulation, using the fact that, under a risk neutral probability measure P^*, the price of a European option is the expectation of the discounted payoff at maturity. Having chosen a proper model for the stock price process, we calculate the discounted payoffs $h_K(x_T) = (x_T - K)^+$ from a simulated sample $\{x_T^1, x_T^2, \ldots, x_T^N\}$ and approximate the option price with

$$U_t(x; K, T; \sigma_t^2) \approx \frac{1}{N} \sum_{j=1}^{N} e^{-r(T-t)} h_K(x_T^j). \tag{100}$$

3. As a third case, suppose that, instead of the volatility process, we know the distribution of the integrated volatility $\bar{\sigma}_t^2$. Then, according to the Hull-White paradigm, the option price can be calculated as the expectation of Black-Scholes prices conditional on the distribution of the integrated volatility, with respect to a risk-neutral probability measure P^*, that is

$$U_t(x; K, T; \sigma_t^2) = \int U_t^{\mathrm{BS}}(x; K, T; \bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2,$$

where U_t^{BS} denotes the Black-Scholes price given by (49).

In all these cases, pricing the option is a direct problem: We know the option pricing function as well as all input variables and parameters; what we compute is the option price.

Let us then consider the corresponding inverse problem. Suppose that we know the market price of a stock as well as the prices of options with different strike prices. Using this data, we want to estimate the stock price volatility implied by these option prices, or a quantity related to it: for example, the average squared volatility during the remaining lifetime of an option.

When the stock price volatility is constant, the Black-Scholes implied volatility (53) equals the stock price volatility and it corresponds to the average squared volatility during the remaining lifetime of the option. The inverse problem reduces then to estimating the implied volatility from option prices.

A common inverse problem considered is to estimate from option market prices the local volatility $\sigma = \sigma(t, X_t)$. The computation of this volatility is an ill-posed inverse problem and various regularization methods have been proposed. As this volatility is out of our scope, we refer to the end of the chapter for related literature.

If the stock price volatility is stochastic, the option price is a function of the distribution of the integrated volatility under a risk-neutral measure P^*. If the price process of the underlying stock is an Itô process, the integrated volatility equals the quadratic variation of the logarithm of the stock price return during the remaining lifetime of the option, given by (44). The inverse problem would then be to estimate the distribution or the expectation of this volatility. We will consider this problem in a Bayesian framework, which allows incorporating information from both historical and actual stock and option market data. As a prelude, we will briefly present some core ideas and tools of *statistical Bayesian inverse problems* to be used later in this thesis.

7.2 Statistical inverse problems

Consider a basic inverse problem with an unknown hidden variable $x \in \mathbb{R}^n$ and an observation $y \in \mathbb{R}^m$ related to the hidden variable by some known model. The inverse problem is to get information of x by measuring y when these variables are related by the function f,

$$y = f(x, e), \tag{101}$$

where $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$ is the model function and $e \in \mathbb{R}^k$ is referred to as *noise*, which accounts for the measurement noise and other poorly known parameters.

A stochastic extension of the model (101) is obtained by replacing x, y, and e with the random variables X, Y, and E. The relation between these variables is then

$$Y = f(X, E). \tag{102}$$

We call the hidden random variable X the *unknown*, the observable random variable Y a *measurement*, and its realization $Y = y$ the *measurement data* of the process. The random variable E is referred to as noise.

In the Bayesian paradigm, prior information of the random variable X, presented in a *prior density* $\pi_{\text{pr}}(x)$, is combined with the information provided by an observed realization of Y via a *likelihood function* $\pi(y \mid x)$, resulting to the joint probability density $\pi(x, y) = \pi_{\text{pr}}(x)\pi(y \mid x)$, according to formula

The conditional density $\pi(x \mid y)$, obtained by formula (4), is called the *posterior density*. Combining (4) and (5) gives the Bayes formula

$$\pi(x \mid y) = \frac{\pi_{\text{pr}}(x)\pi(y \mid x)}{\pi(y)}, \quad y = y_{\text{observed}} \tag{103}$$

for the posterior density. The posterior distribution refers to the information that we have about x if $Y = y$. Oftentimes, it is not available in a closed form.

There are three main steps in solving an inverse problem in the Bayesian framework:

1. Encoding the possible prior knowledge about the hidden variable x in the prior density $\pi_{\text{pr}}(x)$;
2. Forming the likelihood function $\pi(y \mid x)$, which describes how probable each observation y is conditional on a realization x of X;
3. Developing methods to interpret the posterior distribution $\pi(x \mid y)$.

These parts form the framework of our inverse problem.

Usually, the marginal density $\pi(y)$ has the role of a normalizing constant, with minor importance.

As an example, consider a stochastic model $Y = f(X) + E$, where $X \in \mathbb{R}^n, Y \in \mathbb{R}^m$, and X and E are mutually independent. Assume that the probability distribution of E is $\pi_{\text{noise}}(e)$. As fixing $X = x$ does not alter the probability density of E, we can deduce that Y conditioned on $X = x$ is distributed like E and the likelihood function is

$$\pi(y \mid x) = \pi_{noise}(y - f(x)). \tag{104}$$

If the noise is normally distributed with zero mean and a variance of $\text{Var}^{\text{noise}}$, this likelihood function is given by

$$\pi(y \mid x) = \left(\frac{1}{\sqrt{2\pi \text{Var}^{\text{noise}}}} \right)^m \exp\big(- \frac{1}{2\text{Var}^{\text{noise}}} \|y - f(x)\|^2 \big). \tag{105}$$

In the Bayesian paradigm, the posterior density is the solution of the inverse problem. However, it is often difficult to visualize as the dimensionality of problems tends to be very high. What can be done is to calculate from the posterior distribution point estimates of x.

Possibly the most popular statistical point estimate in the frequentist statistics is the *maximum likelihood (ML)* estimate. The corresponding estimator x_{ML} is defined as

$$x_{\text{ML}} = \arg\max_{x \in \mathbb{R}^n} \pi(y \mid x), \quad y = y^{\text{obs}},$$

provided that such a maximizer exists. This estimate tells which value of x maximizes the probability of the outcome y. It is based only on observed data with no prior information included. This estimator is often very sensitive to noise and other inaccuracies and quite useless for ill-posed inverse problems. In fact, solving the maximum likelihood estimate is tantamount to solving an inverse problem without regularization.

A common Bayesian estimate is the *Maximum-A-Posteriori (MAP) estimate*, telling which value of x maximizes the posterior distribution of this unknown. The corresponding estimator is defined as

$$x_{\mathrm{MAP}} = \arg\max_{x \in \mathbb{R}^n} \pi(x \mid y),$$

provided that such a maximizer exists.

Another Bayesian estimate is the *conditional mean (CM)* or posterior mean, defined as

$$x_{\mathrm{CM}} = E\{x \mid y\} = \int_{\mathbb{R}^n} x\pi(x \mid y)dx,$$

provided that the integral converges. This estimate provides information of the point of mass of the posterior distribution.

The two major classes of numerical problems arising in inverse estimation are optimization problems and integration problems. The Maximum-A-Posteriori estimate leads to an optimization problem and the conditional mean estimate to an integration problem. When the posterior density is Gaussian, as it will be in our case, these estimates coincide.

The conditional mean estimation is an integration problem, requiring often numerical evaluation of a high-dimensional integral. The integration usually needs to be done by Monte Carlo sampling as explained in Chapter 1.3.

7.3 Notes on references

The section on statistical inverse problems is based on Kaipio and Somersalo 2005 [78] and on Calvetti and Somersalo 2007 [27]. Other references related to the field include Jeffrey 2004 [73] and Tarantola 2005 [110].

Papers related to the inverse problem of implied volatility are given in the end of Chapter 3. The inverse problem of local volatility is considered in various papers as well as in the book by Gatheral 2006 [53]. Early papers on the topic include Avellaneda, Friedman, Holmes, and Samperi 1997 [6], Bouchouev and Isakov 1997 [20], and Lagnado and Osher 1997 [82], more recent papers with a lot of references are Coleman, Kim, Li, and Verma 2001 [33], Crépey 2003 [41] and Hein and Hofman 2003 [62]. The topic is further developed in Friz and Gatheral 2005 [52] and Cont and Hamida 2005 [38]. The latter propose a probabilistic approach for the problem.

8 The implied integrated volatility

The remaining four chapters in this thesis center on the integrated volatility implied by option market prices; they are the core of this thesis. This chapter introduces the topic. In the following two chapters, we will present two methods to estimate this volatility through a Bayesian approach. Finally, in the last chapter, we will present three simplified examples of the use of the integrated volatility.

The Hull-White option prices are functionals of the distribution of the integrated volatility. If these prices coincide with the corresponding option market prices, it should be possible to estimate, via the Hull-White formula, the distribution of the integrated volatility implied by option market prices. This distribution, which reflects to some extent the market's assumptions about the average volatility during the remaining lifetime of the option, could be used, for example, to hedge and price illiquid options.

Before delving more deeply into these ideas, we will discuss advantages related to the use of the implied integrated volatility. We then will construct a relationship between option market prices and the corresponding Hull-White prices, which will be used when estimating this volatility.

8.1 Model risk and the implied integrated volatility

In recent years, many empirical studies have pointed out the importance of the model risk in derivative industry. For instance, Figlewski and Green 1998 [59] showed empirical evidence that the pricing and hedging errors due to imperfect models and inaccurate volatility forecasts create important risk exposure for option deliverers. The model risk is considered in-depth in Cont 2006, [35].

As it is a latent variable, stock price volatility cannot be directly observed. In Chapter 4, the methods to estimate stochastic volatility were arranged in three groups, which are parametric volatility models, volatility estimates based on direct market-based integrated volatilities, and forward-looking market-based volatility estimates inferred from option prices. The choice between these estimates depends on the purpose of use.

As is common for all models, parametric volatility models are only more or less successful attempts to describe the volatility process. Besides including model errors, they might include errors due to calibration and to discretization. Estimating the parameters of a volatility model can be computationally expensive. This is the case, for example, when one tries to estimate a risk premium for the stochasticity of the volatility. This risk premium can be estimated only from option prices, and the procedure tends to be slow and computationally intensive.

Being non-parametric, the volatility estimation methods from the realized high-frequency data avoid, to some extent, the risk of model error. They are, however, severely affected by the noise of the data, as shown in Andersen, Bollerslev, and Meddahi 2006 [5]. Using realized volatilities in option pricing is not straightforward because, in addition to the volatility of the underlying, option prices can be affected by a premium for the volatility risk or a premium for a negative jump, to mention a few factors.

The implied volatility is a straightforward estimate of the stock price volatility implied by option prices, based on the assumption that the volatility is constant. This assumption is not solid in reality, and, as a result, the implied volatilities vary according to strike prices. To cope with this problem, the volatility information provided by option markets is transferred in pricing and hedging problems using different implied volatilities for different strike prices. An opportunity of arbitrage might occur if the mutual levels of the implied volatilities vary as a function of time. In addition, pricing and hedging options with strike prices different from those provided by markets can result in inaccuracies.

We suggest that, instead of using the Black-Scholes formula and the implied volatility, the information of the stock price volatility implied by option markets should be coded, transferred, and decoded using the distribution of the integrated volatility and the Hull-White formula. This volatility is independent of strike prices, which renders the pricing of options with strike prices other than those observed in the markets straightforward.

We need not always know the exact volatility process or even the instantaneous volatilities during the remaining lifetime of an option. Oftentimes, an estimate of the average integrated volatility during this period is enough. Instead of the instantaneous stock price volatility calculated with a parametric model, using a *meta-level quantity*, namely the integrated volatility, reduces the modeling and calibrating errors to a less significant error due to the assumption that the Hull-White formula adequately describes the option prices.

The implied integrated volatility is robust in that it is not based on a specific stochastic volatility model. Instead of a single valued estimate for the volatility, it provides information on the *distribution* of the volatility and thus information about the *sensitivity* of the estimate. Options with existing and non-existing strike prices are priced the same way, decoding with the Hull-White formula information from the implied distribution of the integrated volatility.

In addition to the stochasticity of the volatility, the volatility smile has been explained by many other factors, such as the risk premium for negative jumps.

When transferring information about the volatility using the distribution of the implied integrated volatility, we do not necessarily need to know the different factors affecting the volatility implied by option market prices. The latent information is embedded in the distribution, and transferred and decoded as such.

8.2 Basic setting to estimate the integrated volatility

This section sets the basis for the inverse problem of implied integrated volatility $\bar{\sigma}_t^2$. We will first model a relationship between the observed option market prices and theoretical Hull-White prices, to be be used later in the likelihood function. We then collect the prior knowledge on the integrated volatility.

As there are two different option prices in the markets, one for bids, u_t^{bid}, and another for offers, u_t^{offer}, we must choose which one to use. Suppose that we want to use the bid price as the observed price, so that $u_t^{\text{obs}} = u_t^{\text{bid}}$. We model the option prices through the Hull-White formula and expect that there is little error e_t^{obs} between this model-based value and the observed price. We assume that this error is independent of the integrated volatility and that it is a normally distributed random variable, with mean zero and a variance $\text{Var}_t^{\text{obs}}$. We denote $e_t^{\text{obs}} \sim \mathcal{N}(0, \text{Var}_t^{\text{obs}})$. The variance reflects our confidence on the Hull-White model, not the assumption of the market price distribution.

For a model to tie together the observed option market price and the integrated volatility implied by option prices, we substitute the Hull-White formula (65) into (120) , that is,

$$u_t^{\text{obs}} = \int U_t^{\text{BS}}(X_t, K, T; \bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2 \mid \sigma_t^2) d\bar{\sigma}_t^2 + e_t^{\text{obs}}. \tag{106}$$

We will present two methods to estimate the integrated volatility from equation (106). The first method is based on very general prior assumptions about the integrated volatility, such as positivity of distribution. We calculate a MAP estimate for the distribution of the integrated volatility, and study the reliability of this estimate using MCMC techniques. The second method is a Bayesian extension of the more traditional estimation of implied volatilities. In addition to an estimate for the expectation of the integrated volatility, this method provides a bridge between this quantity and the commonly used implied volatility. Before going to these methods, we collect our prior knowledge on the distribution of the integrated volatility.

We have two kinds of prior knowledge on the distribution of the integrated volatility: one based on general properties of the integrated volatility and of probability distributions, and another based on a specific volatility process. In general, we know that

i) the integrated volatility is non-negative,

$$\pi(\bar{\sigma}^2) \geq 0, \quad \text{and}$$

ii) the cumulative distribution of $\bar{\sigma}^2$ equals one,

$$\int \pi(\bar{\sigma}^2) d\bar{\sigma}^2 = 1. \tag{107}$$

Depending on a specific case, we can make assumptions about the distribution of the integrated volatility, for example,

a) on the support $[a, b]$ of the distribution,

b) on the skewness of the distribution, and

c) on the smoothness and oscillation of the distribution.

A natural source of information is historical market data on the implied integrated volatility. We explained in Chapter 5.4 how different distributions of the integrated volatility with the same expectation are reflected via the Hull-White formula to different volatility smiles. If this is the case, it should be possible to extract some information on the implied integrated volatilities from the corresponding volatility smiles.

9 Estimating the distribution of the implied integrated volatility

In this chapter, we discuss Bayesian statistical methods to estimate the distribution of the integrated volatility $\bar{\sigma}_t^2$ using indirect information provided by option market prices and very general prior information. We calculate a MAP estimate as well as an estimate based on MCMC sampling, a single-component full scan Gibbs sampler. The advantage of the first method is that it is quick and can be used on-line, while the second method provides not only an estimate of the distribution of interest, but also information on the reliability of this estimate. Information on the uncertainty of the estimate can then be used in prediction reliability analysis.

We assume that the stock price volatility is stochastic and model the option prices via the Hull-White formula (65). Furthermore, we assume that the observed option prices u_t^{obs} correspond to these prices up to a small normally distributed error, so that

$$u_t^{\mathrm{obs}} = \int U_t^{\mathrm{BS}}(x; K, T; \bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2 + e_t^{\mathrm{obs}}, \qquad (108)$$

where $e_t^{\mathrm{obs}} \sim \mathcal{N}(0, \mathrm{Var}_t^{\mathrm{obs}})$. As pointed out earlier, the role of the error is to indicate our belief on the model, since the market data obviously contain no observation error. We have market data on a stock and on several options on it, with different strike prices and the same maturity.

To begin, we fix the time t and discretize the distribution of the integrated volatility. We assume that $\pi(\bar{\sigma}^2)$ takes on positive values in the interval $[a, a+M]$, divide this interval into n volatility points,

$$\bar{\sigma}_j^2 = a + \frac{j-1}{n-1}M, \quad 1 \le j \le n,$$

and denote the values of the corresponding discretized distribution by $z \in \mathbb{R}^n$, where

$$z_j = \pi(\bar{\sigma}_j^2), \quad 1 \le j \le n.$$

We then discretize the Hull-White equation (108), thus obtaining a forward model given by

$$\begin{aligned} u_i^{\mathrm{obs}} &= \int_a^{a+M} U^{\mathrm{BS}}(x; K_i, T; \bar{\sigma}^2)\pi(\bar{\sigma}^2)d\bar{\sigma}^2 + e_i \\ &\approx \frac{M}{n}\sum_{j=1}^{n} U^{\mathrm{BS}}(x; K_i, T; \bar{\sigma}_j^2)\pi(\bar{\sigma}_j^2) + e_i \\ &= \sum_{j=1}^{n} a_{ij}z_j + e_i, \quad a_{ij} = \frac{M}{n}U^{\mathrm{BS}}(x; K_i, T; \bar{\sigma}_j^2), \end{aligned}$$

where e_i denotes the uncertainty corresponding the strike price K_i, $1 \leq i \leq L$ and the maturity T is fixed. In matrix form, we have

$$u^{\text{obs}} = Az + e^{\text{obs}}. \tag{109}$$

Since $\pi(\bar{\sigma}^2)$ is a probability density, its integral over the support equals one,

$$1 = \int_a^{a+M} \pi(\bar{\sigma}^2) d\bar{\sigma}^2 \approx \frac{M}{n} \sum_{j=1}^{n} z_j = q_n^T z,$$

where $q_n = (M/n)[1\ 1\ \dots\ 1]^T$. Using this condition, we can fix one coordinate by performing a linear transformation.

We define a basis of $\mathbb{R}^n$, $\{q_1, q_2, \dots, q_n\}$, where

$$q_1 = \frac{M}{n}\begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix}, \quad q_2 = \frac{M}{n}\begin{bmatrix}1\\1\\0\\\vdots\\0\end{bmatrix}, \quad \dots, \quad q_n = \frac{M}{n}\begin{bmatrix}1\\1\\1\\\vdots\\1\end{bmatrix},$$

and denote by Q the matrix with columns q_j, $1 \leq j \leq n$. We then define the vector $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$ as

$$\alpha = Q^T z.$$

This transformation maps the discrete probability distribution z onto the discrete cumulative distribution α where $\alpha_n = 1$ is fixed. We denote $V = Q^{-T}$, where

$$V = \frac{n}{M}\begin{bmatrix}1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1\end{bmatrix} \in \mathbb{R}^{n\times n}, \tag{110}$$

so that the discretized probability distribution z is recovered from α by

$$z = V\alpha. \tag{111}$$

Substituting (111) into (109) leads to a likelihood function of u^{obs} conditioned on α, given by

$$P_{\text{li}}(u^{\text{obs}} \mid \alpha, \alpha_n = 1) \propto \exp\big(-\frac{1}{2\text{Var}^{\text{obs}}}\|u^{\text{obs}} - AV\alpha\|^2\big), \tag{112}$$

where the notation $\propto$ means "up to a proportionality constant." To avoid confusion with the distribution $\pi(\bar{\sigma}^2)$, which we are estimating, we denote the likelihood function, the prior density, and the posterior density by P_{li}, P_{prior}, and P_{post}, respectively.

We model the prior density so that α_j is a realization of the random variable A_j, $1 \leq j \leq n-1$, with

$$A_j = \frac{1}{2}A_{j-1} + \frac{1}{2}A_{j+1} + \sqrt{\theta}_j W_j, \quad A_0 = 0, \quad A_n = 1,$$

where W_j is a time invariant Gaussian innovation process, $W_j \sim \mathcal{N}(0,1)$, and θ_j is the time varying unknown variance of this process. This model corresponds to the qualitative information that the discretized probability distribution is smooth. The corresponding matrix form is

$$LA = D^{1/2}W, \quad W \sim \mathcal{N}(0, I),$$

where $D^{1/2} = D_\theta^{1/2} = \mathrm{diag}(\theta_1^{1/2}, \theta_2^{1/2} \ldots, \theta_n^{1/2}) \in \mathbb{R}^{n\times n}$ and L is a second order smoothness matrix, i.e.,

$$L = \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{n\times n}. \tag{113}$$

The prior density of α conditional on θ is now given by

$$P_{\mathrm{prior}}(\alpha \mid \theta) \propto P_<(\alpha) \exp\big(-\frac{1}{2}\|D^{-1/2}L\alpha\|^2 - \frac{1}{2}\sum_{j=1}^{n} \log\theta_j\big), \tag{114}$$

where

$$P_<(\alpha) = \begin{cases} 1, & \text{if } \alpha_{j+1} \geq \alpha_j, \quad 1 \leq j \leq n-1 \\ 0 & \text{elsewhere.} \end{cases}$$

The second term in the exponent of (114) comes from the normalizing constant, which cannot be neglected when the variance vector θ is modeled as an unknown.

We model the variance vector $\theta = [\theta_1, \theta_2, \ldots, \theta_n]$ as a random vector and estimate it together with α. Assuming that the random variables θ_j are mutually independent and independent of A_j, the prior density of the pair (α, θ) is given by the *Bayesian hypermodel*

$$P_{\mathrm{prior}}(\alpha, \theta) = P_{\mathrm{prior}}(\alpha \mid \theta) P_{\mathrm{hyper}}(\theta), \tag{115}$$

where the *hyperprior* $P_{\mathrm{hyper}}(\theta)$ reflects our belief on θ. We model the hyperprior with a truncated gamma distribution, given by

$$P_{\mathrm{hyper}}(\theta) \propto P_+(\theta) \prod_{j=1}^{n} \exp\bigg(-\frac{\theta_j}{\theta_0} + (\beta-1)\log\theta_j \bigg), \tag{116}$$

where $P_+(\theta)$ is a characteristic function of the hypercube,

$$P_+(\theta) = P_+(\theta, \theta^{\max}) = \begin{cases} 1, & \text{if } 0 \le \theta_j \le \theta^{\max} \quad \text{for every } j \\ 0 & \text{otherwise.} \end{cases}$$

The choices of the hypermodel parameters θ_0, β, and $\theta^{\max}$ are discussed in connection with the computed examples.

We have chosen as hyperprior the gamma distribution because it favors values of θ_j that are close to the mean $\theta_0/(\beta - 1)$, but being a heavy-tailed distribution, allows large outliers. Hence, the hyperprior corresponds to the qualitative information that the discretized cumulative distribution is second order smooth with possibly few points of first order non-smoothness.

The posterior density $P_{\text{post}}(\alpha, \theta \mid u^{\text{obs}}, \alpha_n = 1)$, obtained by substituting (112) and (115) in the Bayes formula,

$$P_{\text{post}}(\alpha, \theta \mid u^{\text{obs}}, \alpha_n = 1) \propto P_<(\alpha) P_+(\theta) \exp\Bigg(- \frac{1}{2\text{Var}^{\text{obs}}} \|u^{\text{obs}} - AV\alpha\|^2 - \frac{1}{2}\|D^{-1/2}L\alpha\|^2 - \sum_{j=1}^{n} \frac{\theta_j}{\theta_0} + (\beta - \frac{3}{2}) \log \theta \Bigg). \quad (117)$$

We propose here an alternating iteration algorithm for finding an approximation for the MAP estimate of the pair (α, θ). Similar algorithms have previously been used in other contexts, such as signal and image processing applications; see Calvetti and Somersalo 2007 [27], [25], [28], and [26]. We initialize $\alpha = \alpha_0$ and $\theta = \theta_0$, set a limit for the prior variance by defining an upper bound $\theta^{\max}$ for θ, and alternate the following two steps until convergence:

1. update α by maximizing the posterior density, conditional on the current value of θ,

$$\alpha = \arg \min_{\alpha \in \mathbb{R}^n} \left\{ \frac{1}{2\text{Var}} \|u^{\text{obs}} - AV\alpha\|^2 + \frac{1}{2}\|D^{-1/2}L\alpha\|^2 \right\}, \quad \alpha_{j+1} \ge \alpha_j,$$

   where $\text{Var} = \text{Var}^{\text{obs}}$. To respect the ordering $\alpha_{j+1} \ge \alpha_j$, the minimization is done component-wise.

2. update θ component-wise by maximizing the posterior density conditional on the current value of α,

$$\theta_j^M = \arg \max_{\theta \in [0, \theta^{\max}]} \left\{ \exp(-\frac{1}{2}\|D^{-1/2}(L\alpha)_j\|^2 - \frac{\theta_j}{\theta_0} + (\beta - \frac{3}{2}) \log \theta_j \right\},$$

   and define $\theta_j = \min(\theta_j^M, \theta^{\max})$.

The discrete probability distribution z is recovered from the MAP estimate of α by (111).

The reliability of the MAP estimate for α can be assessed with MCMC sampling. We set the MAP estimates of α and θ as the initial points, $\alpha^1 = \alpha_{\mathrm{MAP}}$ and $\theta^1 = \theta_{\mathrm{MAP}}$, and set $k = 1$. We then apply the following block form Gibbs sampler N times, where N is the sample size.

1. sample α^{k+1} from the distribution $\alpha \mapsto P(\alpha \mid \theta^k) \propto P_{\mathrm{post}}(\alpha, \theta^k)$, where

$$P(\alpha \mid \theta^k) \propto P_<(\alpha) \exp\left(- \frac{1}{2\mathrm{Var}} \|u^{\mathrm{obs}} - AV\alpha\|^2 - \frac{1}{2}\|D^{1/2}L\alpha\|\right),$$

   where $D^{1/2} = \mathrm{diag}\sqrt{\theta^k}$.

2. sample θ^{k+1} from the distribution $\theta \mapsto P(\theta \mid \alpha^{k+1}) \propto P_{\mathrm{post}}(\alpha^{k+1}, \theta)$ component-wise, so that the component θ_j^{k+1} is drawn from the density

$$\theta_j \mapsto P_+(\theta) \exp\left(- \frac{(V\alpha^{k+1})_j^2}{2\theta_j} - \frac{\theta_j}{\theta_0} + (\beta - \frac{3}{2}) \log \theta_j \right).$$

3. If $k < N$, increase by k by one and repeat from 1.

The inverse cumulative distribution method is used for sampling both α and θ. We can calculate from the corresponding samples the conditional means α_{CM}, θ_{CM}, and $z_{\mathrm{CM}} = V\alpha_{\mathrm{CM}}$.

We now demonstrate the algorithms through a computed example.

Let z_{true} denote a true discrete probability density. We generate the test data with

$$u^{\mathrm{obs}} = Az_{\mathrm{true}}.$$

To avoid the obvious inverse crime, we generate the data using a denser discretization than the one, $n = 50$, used in the model for solving the inverse problem. The true distribution z_{true} and the corresponding cumulative distribution $\alpha_{\mathrm{true}} = Q^{\mathrm{T}} z_{\mathrm{true}}$ are shown in Figure 5.

Although the likelihood model assumes a white noise modeling error, we do not add any noise to our test data. Instead, we treat the variance $\mathrm{Var}^{\mathrm{obs}}$ as a quantity indicating our confidence in the Hull-White formula with respect to the true option prices.

To apply the MAP estimation algorithms and MCMC sampling, we fix the prior parameters as $\beta = 3$ and $\theta_0 = 10^{-11}$ in the gamma distribution, $\theta^{\max} = 10^{-9}$ and $\mathrm{Var}^{\mathrm{obs}} = 10^{-9}$. We run the MAP estimation algorithm, which converges after $N = 600$ iteration rounds. We have plotted the results α_{MAP} and $z_{\mathrm{MAP}} = V\alpha_{\mathrm{MAP}}$ in Figure 5.

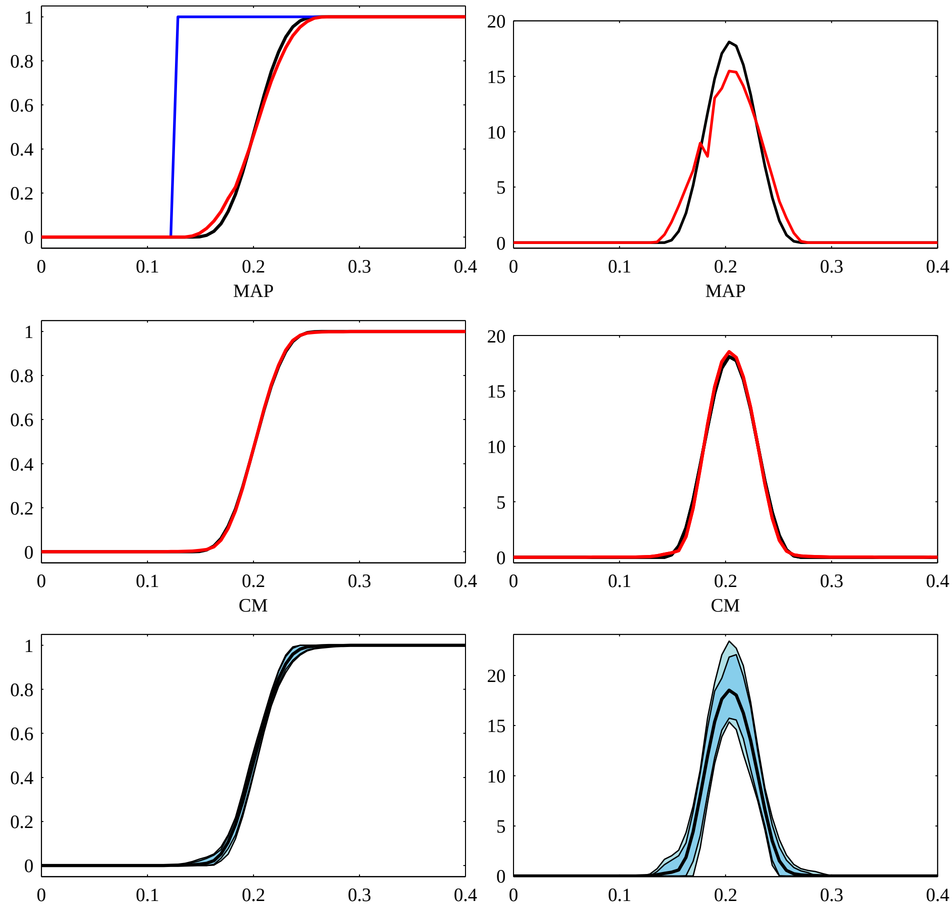


Figure 5: Estimates for α and z. The true distributions are plotted in black and the estimates are plotted in red. The MAP estimate α_{MAP} is presented in the left panel of the top row and the corresponding estimate z_{MAP} in the right panel. The initiating distribution α_0 is plotted in blue. The second row represents the conditional means, α_{CM} being presented in the left panel and z_{CM} in the right panel. On the bottom row, we have plotted the pointwise predictive output envelopes of 75% (lighter) and 90% (darker) of the estimated α and z.

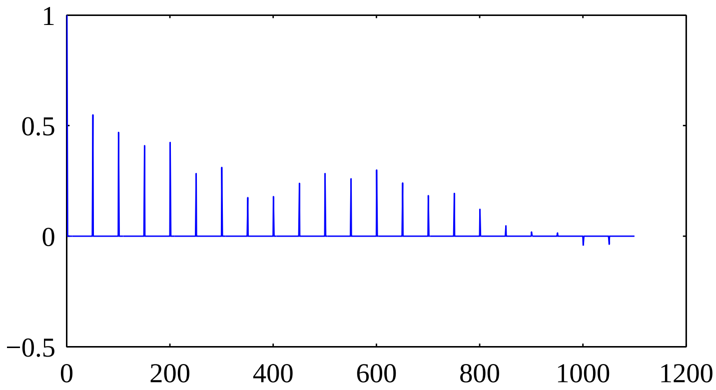


Figure 6: The autocorrelations from the sample history of the $k = 20$th component of α.

We now generate a sample $\{\alpha^1, \alpha^2, \ldots, \alpha^N\}$ by the Gibbs sampler with $N = 200000$. The posterior means $\alpha_{\rm CM}$ and $z_{\rm CM}$ are shown in Figure 5. We also present the point-wise predictive output envelopes for both α and z. The point-wise envelope of α indicates at each support point $\bar{\sigma}_j^2$ of α_j the interval containing a given percentage of the values $\alpha_j^1, \ldots, \alpha_j^N$.

To assess the convergence of the MCMC sampling, we compute the autocorrelation function (ACF) of the sample history of a given component of α. By fixing k, $1 \leq k \leq n$, define

$$x^j = \alpha_k^j - \frac{1}{N}\sum_{i=1}^{N} \alpha_k^i, \quad 1 \leq j \leq N$$

and further, the l'th lagged ACF of the k'th component as

$$h_l = \frac{1}{\|x\|^2}\sum_{j=1}^{N-l} x_j x_{j+l}, \quad \|x\|^2 = \sum_{j=1}^{N}(x^j)^2$$

where $l = 0, 1, 2, \ldots$. We have plotted the ACF of the component α_k, $k = 20$ in Figure 6.

The performance of the MAP estimation algorithm in this example gives quite good results, although numerical experiments indicate that a grossly off-target starting point for the iteration may converge to a solution that markedly differs from the true one. This indicates that the posterior density may have local minima, or that the posterior density is flat. The slow convergence of the MCMC chain, as indicated by the autocorrelation functions of various components of the vector α, enhance the impression that the posterior density has a long and narrow valley that allows the chain to move. To remove the non-uniqueness, further information, based either on prior belief or on complementary data, should be used.

As mentioned earlier, we are interested in the implied integrated volatility as a tool for pricing and hedging options, rather than as a quantity in and of itself. For this reason, we test the performance of our MAP and CM estimates by calculating option Hull-White prices $U^{\mathrm{MAP}} = Az_{\mathrm{MAP}}$ and $U^{\mathrm{CM}} = Az_{\mathrm{CM}}$ with the estimates z_{MAP} and z_{CM}, and comparing these prices with the original option prices u^{obs}. The relative pricing errors e_{MAP} and e_{CM} are computed with

$$\begin{aligned} e_{\mathrm{MAP}} &= (u^{\mathrm{obs}} - U^{\mathrm{MAP}})/u^{\mathrm{obs}}, &(118)\\ e_{\mathrm{CM}} &= (u^{\mathrm{obs}} - U^{\mathrm{CM}})/u^{\mathrm{obs}}. &(119)\end{aligned}$$

The result is that this test is very good: the relative pricing errors e_{MAP} being of size 10^{-5} for in-the-money options and out-of-the-money options, and of size 10^{-4} for at-the-money options. The relative errors e_{CM} are even smaller, being of size 10^{-6} for in-the-money options and at-the-money options, and of size 10^{-5} for out-of-the-money options.

10 Black-Scholes formula and the systematic model error

Theoretically, when the stock price volatility is constant, the average volatility during the remaining lifetime of an option is given by the implied volatility. When the stock price volatility is stochastic, according to the Hull-White paradigm, this average volatility is the *square root of the expectation of the implied integrated volatility.* However, as it is robust and easy to implement, the implied volatility is also a commonly used estimator in this case.

In this chapter, we show how the expectation of the implied integrated volatility can be estimated similarly to the implied volatility: with the Black-Scholes formula, but here adding a stochastic correction based on the Hull-White formula and its Bayesian inference. As earlier, we assume that the stock price volatility is stochastic, that the Hull-White formula gives the correct option price, and that the observed option price u_t^{obs} equals the Hull-White price U_t^{HW} with possibly a small normally distributed error, i.e.,

$$u_t^{\text{obs}} = U_t^{\text{HW}} + e_t^{\text{obs}}, \tag{120}$$

where $e_t^{\text{obs}} \sim \mathcal{N}(0, \text{Var}_t^{\text{obs}})$.

Suppose that we want to estimate from option market prices the average variance during the remaining lifetime of an option. A simple, classical solution is to approximate the observed price with the Black-Scholes price,

$$u_t^{\text{obs}} \approx U_t^{\text{BS}}(I_t^2), \tag{121}$$

and to calculate the corresponding implied volatility I_t. The expectation of the integrated volatility can then be approximated by $I_t^2 \approx \hat{\sigma}_t^2$. There is, however, a systematic error between these quantities, as explained in Chapter 5.4. To compensate for this bias, we add to (121) an error term $\hat{e}_t^{\text{BS}}$ and write

$$\begin{aligned} U_t^{\text{HW}}(\sigma_t^2) &= U_t^{\text{BS}}(\hat{\sigma}_t^2) + [U_t^{\text{HW}}(\sigma_t^2) - U_t^{\text{BS}}(\hat{\sigma}_t^2)] \\ &= U_t^{\text{BS}}(\hat{\sigma}_t^2) + \hat{e}_t^{\text{BS}}, \end{aligned} \tag{122}$$

where the error term is given by

$$\begin{aligned} \hat{e}_t^{\text{BS}} &= \int U_t^{\text{BS}}(\bar{\sigma}_t^2)\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2 - U_t^{\text{BS}}(\hat{\sigma}_t^2)d\bar{\sigma}_t^2 \\ &= \int (U_t^{\text{BS}}(\bar{\sigma}_t^2) - U_t^{\text{BS}}(\hat{\sigma}_t^2))\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2. \end{aligned} \tag{123}$$

Stated in another way, $\hat{e}_t^{\text{BS}}$ is the expectation of the random model error e_t^{BS}, that is

$$\begin{aligned} e_t^{\text{BS}} &= U_t^{\text{BS}}(\bar{\sigma}_t^2) - U_t^{\text{BS}}(\hat{\sigma}_t^2) \quad \text{and} \\ \hat{e}_t^{\text{BS}} &= \int e_t^{\text{BS}}\pi(\bar{\sigma}_t^2)d\bar{\sigma}_t^2. \end{aligned}$$

The distribution of the model error e_t^{BS} depends on the variables affecting the implied volatility smile, i.e., the distribution of the integrated volatility, the ratio X_t/K and the remaining time to maturity $T-t$. Clearly, the expectation of this error also depends on these variables. In the sequel, when talking about model error, we mean $\hat{e}_t^{\mathrm{BS}}$, the expectation of the model error, not the random model error e_t^{BS}.

The available possibilities for estimating the model error depend on our prior knowledge of the distribution of the integrated volatility. If we have an estimate for the distribution of the integrated volatility, we can calculate the model error directly from (123), using

$$\hat{\sigma}_t^2 = \int \bar{\sigma}_t^2 \pi(\bar{\sigma}_t^2) d\bar{\sigma}_t^2.$$

Suppose, then, that we do not know the distribution of the integrated volatility but do have a model for the volatility process. However, some of the model parameters are poorly known, so we have only qualitative information on them, and thus model them as random variables. In this case, we can pick M realizations of the volatility process, propagate each of them N times in the time interval $[t, T]$ and calculate the corresponding distributions of the integrated volatility. Having performed $[N \times M]$ simulations, we obtain a sample of M distributions, $\{\pi^1(\bar{\sigma}_t^2), \pi^2(\bar{\sigma}_t^2), \ldots, \pi^M(\bar{\sigma}_t^2)\}$. Using (123), we calculate a model error $\hat{e}_{mt}^{\mathrm{BS}}$ for each distribution, that is,

$$\hat{e}_{mt}^{\mathrm{BS}} = \int (U_t^{\mathrm{BS}}(\sqrt{\bar{\sigma}_t^2}) - U_t^{\mathrm{BS}}(\hat{\sigma}_t^2))\pi^m(\bar{\sigma}_t^2) d\bar{\sigma}_t^2, \quad 1 \le m \le M.$$

The average error $\hat{e}_t^{\mathrm{BS}}$ and the variance $\mathrm{Var}_t^{\mathrm{BS}}$ of the error are then given by

$$\hat{e}_t^{\mathrm{BS}} = \frac{1}{M}\sum_{m=1}^{M} \hat{e}_{mt}^{\mathrm{BS}}, \quad \mathrm{Var}_t^{\mathrm{BS}} = \frac{1}{M-1}\sum_{m=1}^{M}(\hat{e}_t^{\mathrm{BS}} - \hat{e}_{mt}^{\mathrm{BS}})^2. \tag{124}$$

For simplicity, we assume that the distribution of the model errors can be approximated by a Gaussian distribution. The applicability of these estimates depends on how much and how quickly the integrated volatility implied by option markets changes. When this volatility is estimated as an implied volatility with a systematic correction $\hat{e}_t^{\mathrm{BS}}$, the volatility estimates for each strike price should coincide. If this is not roughly the case, the estimate for the model error $\hat{e}_t^{\mathrm{BS}}$ is not accurate.

Let us then return to our pricing model. We substitute (122) in (120), which gives

$$u_t^{\mathrm{obs}} = U_t^{\mathrm{BS}}(\hat{\sigma}_t^2) + \hat{e}_t^{\mathrm{BS}} + e_t^{\mathrm{obs}}. \tag{125}$$

As we assumed that the error e_t^{obs} is normally distributed and approximated the distribution of $\hat{e}_t^{\text{BS}}$ via a normal distribution, then, by (105), the likelihood function $P_{\text{li}}(u_t^{\text{obs}} \mid \hat{\sigma}_t^2)$ is given by

$$P_{\text{li}}(u_t^{\text{obs}} \mid \hat{\sigma}_t^2) \propto \exp\big(- \frac{1}{2(\text{Var}_t^{\text{obs}} + \text{Var}_t^{\text{BS}})} (u_t^{\text{obs}} - U_t^{\text{BS}}(\hat{\sigma}_t^2) - \hat{e}_t^{\text{BS}})^2 \big). \quad (126)$$

We encode the prior information about the expectation of the integrated volatility in the prior density $P_{\text{prior}}(\hat{\sigma}_t^2)$, given by

$$P_{\text{prior}}(\hat{\sigma}_t^2) \propto \exp\left(- \frac{1}{2\text{Var}_t^{\text{pr}}} \big(\hat{\sigma}_t^2 - \tilde{\sigma}_t^2\big)^2 \right), \quad (127)$$

where $\tilde{\sigma}_t^2$ denotes the expected value of $\hat{\sigma}_t^2$ and Var_t^{pr} reflects our confidence on the prior information.

Combining information provided by the prior density (127) and the likelihood function (126) provides us, via the Bayes formula, the posterior distribution $P_{\text{post}}(\hat{\sigma}_t^2 \mid u_t^{\text{obs}})$, given by

$$\begin{aligned} P_{\text{post}}(\hat{\sigma}_t^2 \mid u_t^{\text{obs}}) &= P_{\text{li}}(u_t^{\text{obs}} \mid \hat{\sigma}_t^2) P_{\text{prior}}(\hat{\sigma}_t^2) \\ &\propto \exp\big(- \frac{1}{2(\text{Var}_t^{\text{obs}} + \text{Var}_t^{\text{BS}})} (u_t^{\text{obs}} - U_t^{\text{BS}}(\hat{\sigma}_t^2) - \hat{e}_t^{\text{BS}})^2 - \frac{1}{2\text{Var}_t^{\text{pr}}} \big(\hat{\sigma}_t^2 - \tilde{\sigma}_t^2)^2 \big) \\ &\propto \exp\big(- \delta_t^2 \big(\hat{\sigma}_t^2 - \tilde{\sigma}_t^2)^2 - (u_t^{\text{obs}} - U_t^{\text{BS}}(\hat{\sigma}^2) - \hat{e}_t^{\text{BS}})^2 \big), \end{aligned} \quad (128)$$

where

$$\delta_t = \sqrt{\frac{\text{Var}_t^{\text{BS}} + \text{Var}_t^{\text{obs}}}{\text{Var}_t^{\text{pr}}}}.$$

The MAP estimate for $\hat{\sigma}_t^2$ is the minimizer of the log-posterior:

$$\begin{aligned} \hat{\sigma}_{\text{MAP},t}^2 &= \arg \min_{\hat{\sigma}_t^2 \in \mathbb{R}^n} \Bigg\| \underbrace{\begin{bmatrix} U_t^{\text{BS}}(\hat{\sigma}_t^2) + \hat{e}_t^{\text{BS}} \\ \delta \hat{\sigma}_t^2 \end{bmatrix}}_{F(\hat{\sigma}_t^2)} \underbrace{\begin{bmatrix} u_t^{\text{obs}} \\ \delta \tilde{\sigma}_t^2 \end{bmatrix}}_{z} \Bigg\|^2 \\ &= \arg \min_{\hat{\sigma}_t^2 \in \mathbb{R}^n} \| F(\hat{\sigma}_t^2) - z \|. \end{aligned} \quad (129)$$

If $\text{Var}_{\text{pr}} \to \infty$, i.e., no prior information is available, the formula reduces to the Black-Scholes implied volatility formula, with the *significant difference* that the systematic bias $\hat{e}_t^{\text{BS}}$ due to the smile has been removed. The problem reduces to the non-linear least squares problem of minimizing the objective function $\|F(\hat{\sigma}_t^2) - z\|^2$, which can be done with standard minimization techniques.

We illustrate the idea of the model error $\hat{e}$ in Figure 7, and a computed example is presented in Chapter 11.2. We can see from Figure 7 that, as

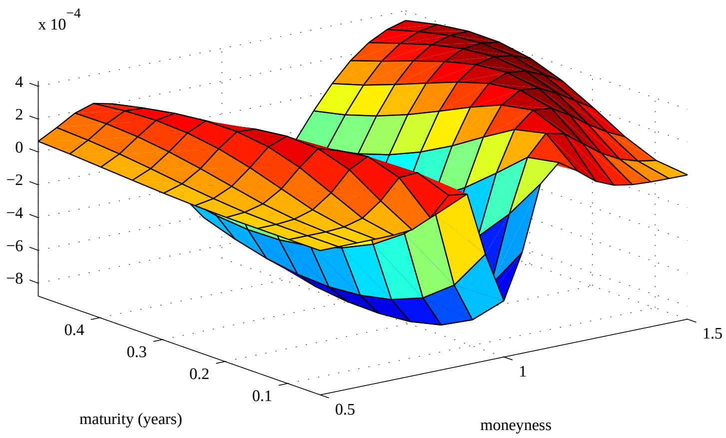


Figure 7: Simulated model error surface

expected from Section 5.4, the model error is very little for at-the-money options, positive for in-the-money and out-of-the-money options, and negative for deep-in-the-money and deep-out-of-the-money options. The option market prices are calculated with the Hull-White formula using a distribution of the integrated volatility generated by a discretized version of (58)-(59), with $f(y) = \sqrt{y}$, $\alpha = 1$, $m = .12$, $\beta = .6$, and $\rho = 0$. The model errors are calculated with (123).

In this chapter, we have used ideas presented in the paper by Kaipio and Somersalo 2007 [79].

11 Computed examples with the implied integrated volatility

This chapter presents three examples of the use of the implied integrated volatility. First, we test how well the volatility smile can be removed when, instead of the implied volatility, we estimate the expectation of the integrated volatility. We then consider a simplified example of pricing illiquid European options using the integrated volatility. Finally, we consider a hedging problem where a European option is hedged using the inference on the volatility. In the pricing and hedging problems, the use of the implied integrated volatility provides, in addition to an estimate, information on the reliability of this estimate.

We use simulated market data in all the three examples. We assume that all prices are given in a risk-neutral martingale measure $P^{*(\gamma)}$, which is equivalent to the natural measure of the markets P. Since we are working with volatilities implied by option prices, not with real stock price volatilities, this assumption would be suitable even if we were using real market data. When the stock price volatility is stochastic, the option prices include a premium for the volatility risk. This premium is reflected in different volatilities implied by option prices, and via these volatilities, it is transferred to new prices or to hedging ratios. We can estimate the implied volatility, the distribution of the integrated volatility, and the expectation of the integrated volatility without knowing the exact value of this risk premium merged in our estimates.

11.1 Generating the market data

The simulated market data consist of the prices of one stock and of options on this stock with several strike prices and two maturities on a time interval $[0, T]$. The discrete market data on stock prices and stock price volatilities are generated using the time grid

$$v_j = \frac{j-1}{V-1}T, \quad 1 \leq j \leq V, \tag{130}$$

with $\Delta v = T/V$ as the time step. We choose from this grid W evenly spaced points as observation points from which we calculate the integrated volatilities and option prices. We denote the observation times by t_j, $1 \leq j \leq W$.

We assume that the stock price process is given by (61)-(62), and generate stock price data using a discretized version of this process, given by

$$X_{v+1} = X_v + rX_v\Delta v + \sigma_v X_v \nu_v \sqrt{\Delta v}, \quad (131)$$

$$\sigma_v = f(Y_v) \quad (132)$$

$$Y_{v+1} = Y_v + \alpha(m - Y_v)\Delta v + \beta\epsilon_v\sqrt{\Delta v}, \quad (133)$$

where ν_v and ϵ_v are mutually uncorrelated, $(\nu_v, \epsilon_v) \propto \mathcal{N}(0,1)$, and $1 \le v \le V$. We have here discretized (131) and (133) using the Euler method, considered, for example, in Glasserman 2004 [57].

We assume that the observed option prices u_t^{obs} equal the Hull-White prices up to a small, normally distributed error. This error accounts for the uncertainty of the Hull-White formula as a model for the market price. Although there is no guarantee that the market price coincides with the theoretical one, we use this value as a model for the market value. We use the discretized version of the Hull-White formula to calculate the observed option prices, and approximate the option market prices by

$$u_j^{\text{obs}} = \frac{M}{n} \sum_{k=1}^{n} U_k^{\text{BS}}(t_j, x_j; K, T; \bar{\sigma}_{jk}^2)\pi(\bar{\sigma}_{jk}^2) + e_j, \quad (134)$$

where M is the length of the support interval of the volatility distribution, n is the number of discrete volatility points on this support, and the error $e_j \sim \mathcal{N}(0, \text{Var}_j^{\text{obs}})$. The integrated volatilities $\bar{\sigma}_{jk}^2$ are generated using the volatility model (132)-(133) and the discretized version of the integrated volatility (66).

When estimating the hedging performance of different volatilities, we need a replicating option with a longer maturity T_2, $T < T_2$ than the one of the option to be hedged. We denote this option with $u_j^{\text{Rep}} = u^{\text{Rep}}(t_j, x_j; K, T_2; \sigma_j^2)$.

The parameters used when simulating the stock price process are $r = 0$ $\alpha = 1$, $m = .14$, $\beta = .4$, $\Delta t = 2$ days, $x_0 = 1$, $\sigma_0^2 = .12$, and $f(x) = \sqrt{x}$. The parameters related to option prices will be presented separately with each example.

11.2 Removing the volatility smile

In this first example, we present how the expectation of the integrated volatility can be estimated from option market prices. The average volatility during the remaining lifetime of an option is sometimes approximated with the implied volatility. When the volatility is constant, this is the proper estimate. However, if the stock price volatility is stochastic, the implied volatility depends on the strike price, a phenomenon reflected as a volatility smile. In this case, theoretically, the average volatility is given by the expectation of the implied integrated volatility. This expectation is not usually a straightforward forecast for the future stock price volatility, but it probably reflects market expectations on stock price fluctuations.

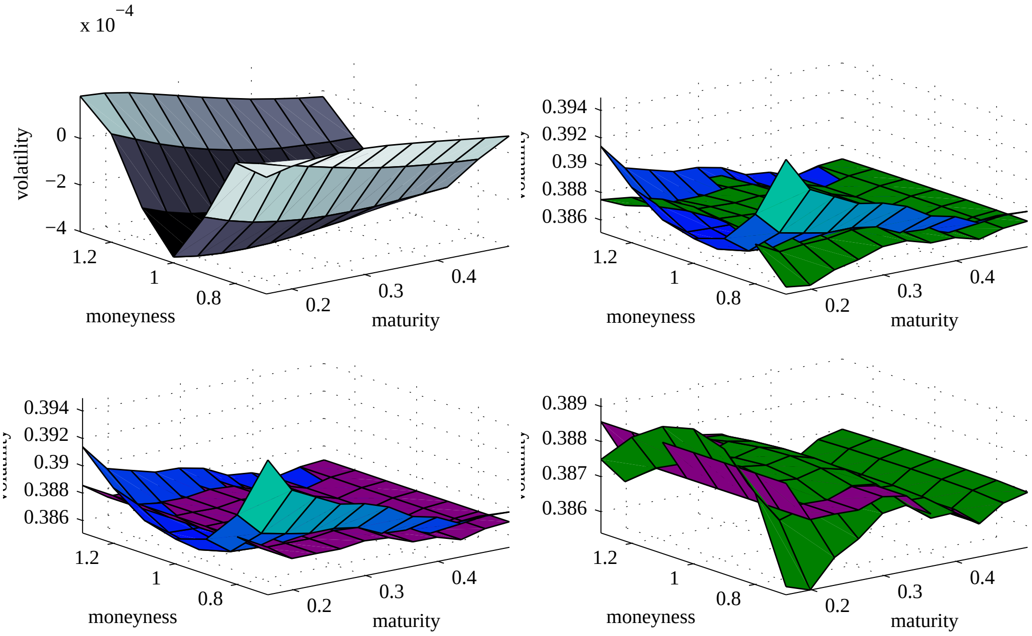


Figure 8: The model error surface and volatility surfaces implied by option market prices. The top left panel represents the surface of the systematic model error $\hat{e}$. In the top right panel, we have plotted the surfaces of the implied volatility I (blue) and of the square root of the expectation of the implied integrated volatility (green). The left panel of the bottom row represents the real volatility surface (violet) and the implied volatility surface (blue); the right panel represents the real volatility surface (violet) and the surface of the square root of the expectation of the implied integrated volatility (green).

Even though the implied volatility can be a biased estimator, it has its strengths - it is a commonly known and robust estimator that is easy to calculate. We now show how, using this estimator, one can calculate an estimate for the expectation of the integrated volatility, an estimate independent of the strike price. The correction is done by simply adding a systematic model error to the observed option prices and then calculating the corresponding implied volatilities, as discussed in Chapter 10.

Suppose that we want to estimate the expectation of the integrated volatility implied by the market prices of certain options with the same underlying and maturity but different strike prices. In this example, we assume that we have an accurate model for the volatility process but have only qualitative information on some of its parameters.

As a preliminary task, we estimate the systematic model error $\hat{e}_t^{\mathrm{BS}}$ by simulation. We model the volatility process with

$$\sigma_j = f(Y_j) \tag{135}$$

$$Y_{j+1} = \alpha(M - Y_j)\Delta t + B\epsilon_j\sqrt{\Delta t}, \tag{136}$$

where $\Delta t = T/W$, $\epsilon_j \propto \mathcal{N}(0,1)$ and the long run level of mean reversion M and the volatility of the volatility B are modeled as random variables. In the forward simulations, we draw from uniform distributions $M \sim \text{Uniform}_{[m_1,m_2]}$ and $B \sim \text{Uniform}_{[\beta_1,\beta_2]}$.

We pick L realizations of the volatility process (135)-(136), propagate each of them N times in the time interval $[t,T]$ and calculate the corresponding distributions of the integrated volatility. Using (123), we calculate a model error $\hat{e}^{\text{BS}}_{mt}$ for each distribution, as well as the average error $\hat{e}^{\text{BS}}_t$ and the variance Var^{BS}_t, given by (124).

An estimate for $\hat{\sigma}^2_j$ is now obtained from the option market prices u^{obs}_j by solving this volatility from

$$u^{\text{obs}}_j = U^{\text{BS}}_j(x;K,T;\hat{\sigma}^2_j) + \hat{e}^{\text{BS}}_j$$

with the MAP algorithm presented in Chapter 10 with prior variance $\text{Var}^{\text{pr}} \to \infty$. To compare the performance of this estimate, we also calculate the corresponding squared implied volatilities I^2_j from

$$u^{\text{obs}}_j = U^{\text{BS}}_j(x;K,T;I^2_j).$$

We estimate the model error from a sample of $L = 100$ realizations of the volatility process (135)-(136), where M and B are drawn from the uniform distribution $M \sim \text{Uniform}_{[0.13,0.15]}$ and $B \sim \text{Uniform}_{[0.3,0.45]}$. To avoid an inverse crime, we have used a different time discretization than the one used when simulating the market data.

The results of the simulation are presented in Figures 8 and 9. Figure 8 represents the model error surface, the estimated surfaces of implied volatility and of the square root of the implied integrated volatility, and finally the true volatility surface. Clearly adding a systematic model error to observed option prices and calculating the corresponding implied volatilities removes some of the volatility smile. The original volatility surface and the surface of the square root of the implied integrated volatility are compared in the left panel of the bottom row. The effect of adding a systematic model error is particularly drastic due to the fact that we knew the volatility process up to the two unknown parameters modeled as random variables.

Figure 9 represents the implied volatilities, square roots of the implied integrated volatilities, and the true volatility at maturities $T = 0.5$ and $T = 0.3$. The square root of the expectation of the implied integrated volatility is fairly flat for both maturities.

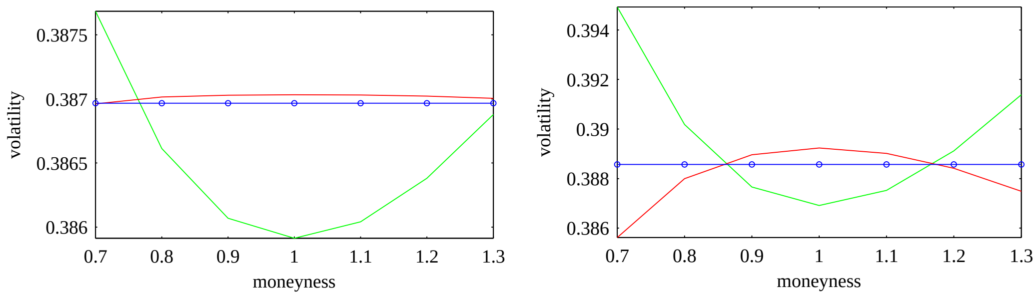


Figure 9: Implied volatility (green), square root of the expectation of the implied integrated volatility (red), and the true volatility (blue). estimated from simulated option market data. In the left panel, the maturity is $T = 0.5$ years and in the right panel, it is $T = 0.3$ years.

11.3 Pricing illiquid European options

Illiquid options on a certain underlying can be priced using a volatility implied by some liquidly traded option on the same underlying. As implied volatilities depend on the option strike prices, illiquid options with different strike prices are often priced using different implied volatilities. We suggest the use of the distribution of the integrated volatility implied by option prices as an alternative for the use of the implied volatility. Unlike the Black-Scholes-based implied volatility, the implied integrated volatility is independent of the strike prices, which reduces the risk of arbitrage between illiquid options with different maturities.

In this simulation, we compare the pricing performances of the distribution of the implied integrated volatility and the implied volatility in a very simplified example of pricing illiquid options. We assume that we know the prices for certain liquid European options with the same underlying and maturity, but different strike prices. Based on volatilities implied by the prices of these options, we price two similar illiquid options, one having a lower and the other a higher strike price than those of the liquid options. The illiquid options with the lowest and highest strike price are denoted by $U_{t,-}(\tilde{\sigma}_t^2)$ and $U_{t,+}(\tilde{\sigma}_t^2)$, respectively, where $\tilde{\sigma}_t^2$ refers to the square of the volatility. To be able to compare the pricing performance of the two different volatilities, we must also know the real prices $u_{t,-}^{\text{real}}$ and $u_{t,+}^{\text{real}}$, i.e., the simulated prices of these options. Of course, the estimated option prices are calculated without this information.

To price the illiquid options, we proceed as follows. We estimate from the market data the distribution of the integrated volatility $\bar{\sigma}_j^2$ as a posterior mean, using the Gibbs sampler presented in Chapter 9, and price the illiquid options $U_{j,-}^{\text{HW}}(x, \sigma_j^2)$ and $U_{j,+}^{\text{HW}}(x, \sigma_j^2)$ using this estimate and the Hull-White formula. We then calculate the implied volatilities and price $U_{j,-}^{\text{BS}}(x, I_{j,-}^2)$ and $U_{j,+}^{\text{BS}}(x, I_{j,+}^2)$ using the Black-Scholes formula and a volatility which has been

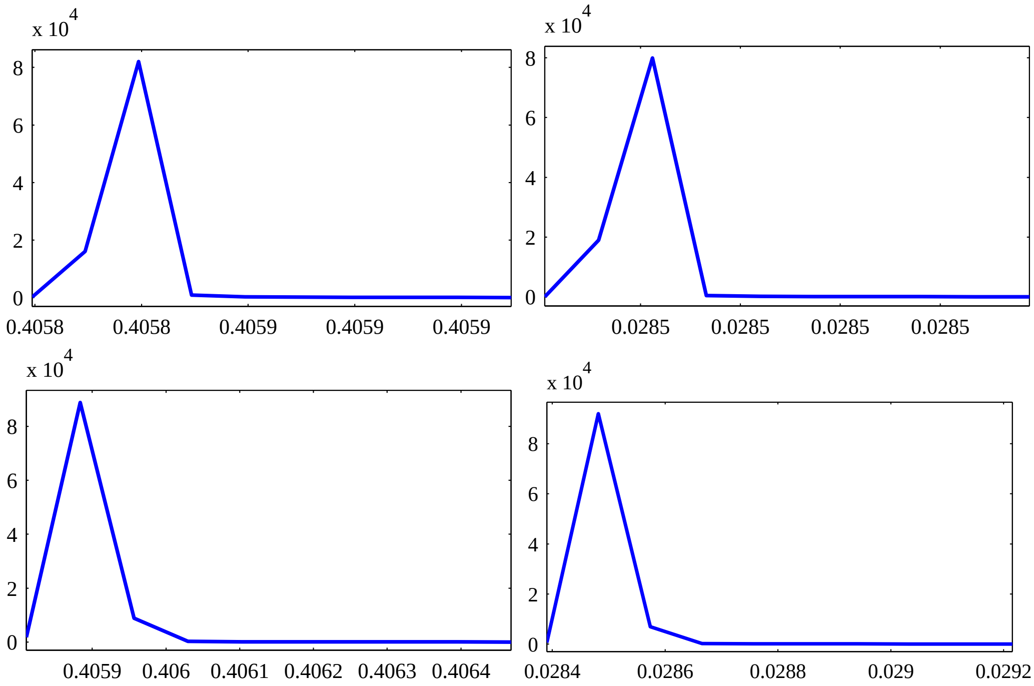


Figure 10: Histograms on the prices price estimates $U^{\mathrm{HW}}_{j,-}(x,\sigma_j^2)$ and $U^{\mathrm{HW}}_{j,+}(x,\sigma_j^2)$ provided by the Gibbs sampler. The left panel represents the histograms of prices for $U_{t,-}(\tilde{\sigma}_t^2)$ and the right panels those for $U_{t,+}(\tilde{\sigma}_t^2)$. The variance $\mathrm{Var}^{\mathrm{obs}}$ on the upper row is $\mathrm{Var}^{\mathrm{obs}}_1 = 10^{-9}$, on the bottom row it is $\mathrm{Var}^{\mathrm{obs}}_2 = 10^{-8}$.

linearly interpolated from the implied volatilities corresponding to the two nearest strike prices.

We finally compute the relative errors between the estimated and real prices of the illiquid options, given by

$$e_{K\pm}(\tilde{\sigma}_j^2) = (u^{\mathrm{real}}_{K\pm} - U_{K\pm}(\tilde{\sigma}^2))/u^{\mathrm{real}}_{K\pm},$$

where $\tilde{\sigma}^2$ denotes the square of the volatility that has been used. The Gibbs sampler provides a sample of prices for the illiquid options. We present these samples as histograms in Figure 10.

The strike prices of the illiquid options used in this example are $K = [.6, 1.4]$, and the maturity is half a year. To apply the the Gibbs sampler, we fix the prior parameters as $\beta = 3$ and $\theta_0 = 10^{-9}$ in the gamma distribution and $\theta^{\max} = 10^{-9}$. We estimate the prices of the illiquid options using two different variances, $\mathrm{Var}^{\mathrm{obs}}_1 = 10^{-9}$ and $\mathrm{Var}^{\mathrm{obs}}_2 = 10^{-8}$.

Figure 10 represents as histograms the price estimates of the illiquid options provided by MCMC sampling. It is interesting to see that the variance $\mathrm{Var}^{\mathrm{obs}}$ has a remarkable effect on the accuracy of the price estimates. The prices with nearly no variation on the upper row are calculated with $\mathrm{Var}^{\mathrm{obs}}_1 = 10^{-9}$, and prices with a larger variation on the bottom row using $\text{Var}_2^{\text{obs}} = 10^{-8}$. In our case, a natural explanation is that the market data was generated with the Hull-White formula, and the variance reflects confidence on this formula.

The effect of Var^{obs} is remarkable also when comparing the pricing accuracy of the pricing estimates. In general, the relative errors between the correct prices of the illiquid options and the prices computed with interpolated implied volatilities are of the size $e_{K_-}(I^2) \approx 10^{-4}$ for the lower illiquid option and of the size $e_{K_+}(I^2) \approx 10^{-3}$ for the upper one. The corresponding posterior mean estimates calculated with the Gibbs sampler are of the same size for both options when the variance used is $\text{Var}_2^{\text{obs}} = 10^{-8}$. The performance is clearly better using $\text{Var}_1^{\text{obs}} = 10^{-9}$, the relative errors being of the size $e_{K_-}(I^2) \approx 10^{-6}$ for the lower illiquid option and of the size $e_{K_+}(I^2) \approx 10^{-5}$ for the upper one.

11.4 Hedging a European option

The value of an option is sensitive to changes in the value of the underlying. One can try to hedge changes in the value of a short option, i.e., an option that has been sold, by holding a replicating portfolio consisting of a certain amount of the underlying and possibly of a certain amount of a replicating option, i.e., of an option similar to the short option, but with longer maturity. The hedging performance depends on how well changes in the value of the replicating portfolio replicate changes in the value of the short option. In practice, due to the fact that hedging is done discretely, to transaction costs, market restrictions etc., perfect hedging is impossible even if the stock price volatility is constant and known. Some degree of hedging is, however, possible.

Rebonato 2004 [97] has discussed practical hedging of a call option when the volatility of the underlying is stochastic, but the hedger lives in a Black-Scholes world and hedges a short option either by following the delta strategy or the delta-vega strategy described in Chapter 6. Rebonato has shown that if the trader knows exactly the correct average quadratic variation of the underlying, both hedging strategies reduce significantly the standard deviation of the terminal payoff of the portfolio consisting of the short option and the replicating portfolio compared to that one of the naked option, at least if the re-hedging frequency is greater than once a week. However, if he guesses wrong the average quadratic variation, the delta-vega strategy performs better than the delta strategy.

In this example, we illustrate how using the implied integrated volatility in hedging provides, in addition to a point estimate for a hedging ratio, information on the reliability of this estimate. We assume that option prices are given by the Hull-White formula and try to replicate changes in option prices with either the delta strategy or the delta-vega strategy, modelling the instantaneous hedging ratios delta (91), alpha (93), and gamma (94) as distributions. These ratios are calculated with the integrated volatility implied by the Hull-White option prices. To have a benchmark, we hedge the short options also using the implied volatility. We estimate the implied integrated volatility as a posterior mean, using MCMC techniques presented in Chapter 9, and the implied volatility with optimization.

Suppose that at the initial moment $t = t_1$, we sell an option with price U_1^{H}. To hedge this option, we first estimate from the market data the discrete distribution of the integrated volatility $\bar{\sigma}_1^2$ and calculate the corresponding discrete distributions for delta, $\Delta_1(\bar{\sigma}_1^2)$, alpha, $\alpha_1(\bar{\sigma}_1^2)$, and gamma, $\gamma_1(\bar{\sigma}_1^2)$. The expectations of the initial values of the delta portfolio $V^{\mathrm{D}}(\bar{\sigma}^2)$ and of the delta-vega portfolio $V^{\mathrm{DV}}(\bar{\sigma}^2)$ are then given by

$$V_1^{\mathrm{D}}(\bar{\sigma}_1^2) = \int \Delta_1(\bar{\sigma}_1^2)\pi(\bar{\sigma}_1^2)d\bar{\sigma}_1^2 X_1$$

and

$$V_1^{\mathrm{DV}}(\bar{\sigma}_1^2) = \int \alpha_1(\bar{\sigma}_1^2)\pi(\bar{\sigma}_1^2)d\bar{\sigma}_1^2 X_1 + \int \gamma_1(\bar{\sigma}_1^2)\pi(\bar{\sigma}_1^2)d\bar{\sigma}_1^2 U_1^{\mathrm{Rep}},$$

where X denotes the value of the underlying and U^{Rep} the value of the replicating option.

We then estimate the instantaneous implied volatility I_1 and calculate the corresponding hedging ratios. The initial values of the delta portfolio $V^{\mathrm{D}}(I^2)$ and the delta-vega portfolio $V^{\mathrm{DV}}(I^2)$ are then

$$V_1^{\mathrm{D}}(I_1^2) = \Delta_1(I_1^2)X_1$$

and

$$V_1^{\mathrm{DV}}(I_1^2) = \alpha_1(I_1^2)X_1 + \gamma_1(I_1^2)U_1^{\mathrm{Rep}}.$$

At $n-1$ evenly spaced moments of time, we calculate new estimates for the discrete distribution of the implied integrated volatility and for the implied volatility, and calculate the corresponding hedging ratios. We then re-hedge the portfolios V_j^{D} and V_j^{DV} using either the distribution of the ratios $\Delta_j(\bar{\sigma}_j^2)$, $\alpha_j(\bar{\sigma}_j^2)$, and $\gamma_j(\bar{\sigma}_j^2)$, or the ratios $\Delta_j(I_j^2)$, $\alpha_j(I_j^2)$, and $\gamma_j(I_j^2)$, $2 \leq j \leq n$.

To know the cost of re-hedging, we also calculate the changes ΔV_j^{D} and ΔV_j^{DV} in the value of the replicating portfolio, i.e,

$$\begin{aligned}\Delta V_j^{\mathrm{D}} &= (\Delta_j - \Delta_{j-1})X_j \\ \Delta V_j^{\mathrm{DV}} &= (\alpha_j - \alpha_{j-1})X_j + (\gamma_j - \gamma_{j-1})U_j^{\mathrm{Rep}}.\end{aligned}$$

At the last moment of time t_n, we calculate how much selling the option and hedging it has cost to us, following each of the two hedging strategies and

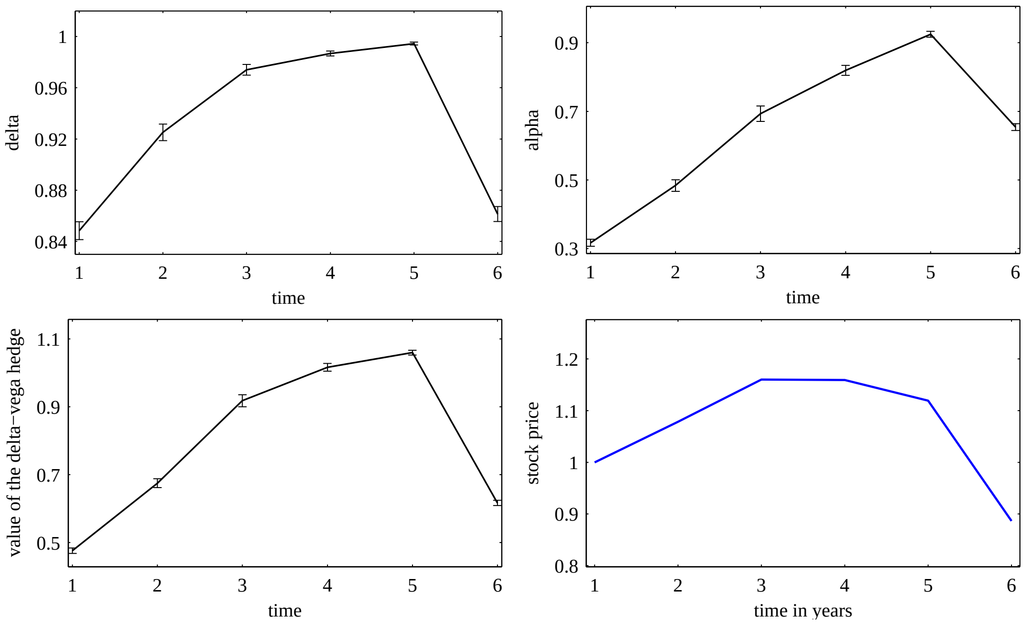


Figure 11: Hedging ratios for the option with strike price $K = 0.8$. The top left panel represents the expectation $\hat{\Delta}_j(\bar{\sigma}_j^2)$, as well as error bars calculated using integrated volatilities within one standard deviation from the expectation of this volatility. The top right panel represents the corresponding values of alpha, and the bottom left panel the corresponding values of the whole delta-vega portfolio V_j^{DV}. Finally, the bottom right panel represents the value of the underlying stock.

using each of the two volatilities. The price $C(\tilde{\sigma}^2)$ of the whole procedure is given by

$$C(\tilde{\sigma}^2) = U_1^{\mathrm{H}} - V_1(\tilde{\sigma}_1^2) - \sum_{j=2}^{n} \Delta V_j(\tilde{\sigma}_j^2) - U_n^{\mathrm{H}} + V_n,$$

where V refers to either V^D or V^{DV} and $\tilde{\sigma}_j^2$ to either $\bar{\sigma}_j^2$ of I_j^2.

In this example, we have options with strike prices $K = [.6, .8, 1, 1.2, 1.4]$ and a maturity of half a year, $T = .5$, the maturity of the corresponding replicating options being $T = .9$. Hedging and re-hedging is done $n = 6$ times, once a month. We estimate the implied integrated volatility from the same option prices using the Gibbs sampler, and fix the variance to $\mathrm{Var}_1^{\mathrm{obs}} = 10^{-9}$ as well as the prior parameters to $\beta = 3$ and $\theta_0 = 10^{-9}$ in the gamma distribution, and define $\theta^{\max} = 10^{-9}$.

In Figure 11, we present the expectations of the hedging ratios delta and alpha, given by

$$\begin{aligned}\hat{\Delta}_j(\bar{\sigma}_j^2) &= \int \Delta_j(\bar{\sigma}_j^2)\pi(\bar{\sigma}_j^2)d\bar{\sigma}_j^2 \quad \text{and} \\ \hat{\alpha}_j(\bar{\sigma}_j^2) &= \int \Delta_j(\bar{\sigma}_j^2)\pi(\bar{\sigma}_j^2)d\bar{\sigma}_j^2,\end{aligned}$$

for an option with strike price $K = 0.8$. The error bars indicate the ratios calculated with implied integrated volatilities within one standard deviation from the expectation of this volatility. We denote by $\Delta_j^{\mathrm{STD}}(\bar{\sigma}_j^2)$ and $\alpha_j^{\mathrm{STD}}(\bar{\sigma}_j^2)$ the ratios calculated with the volatilities exactly one standard deviation from the expectation.

In Figure 11, we also present the corresponding values of the replicating portfolio V^{DV} of the delta-vega strategy and the prices of the underlying. The hedging ratios calculated with the implied volatility are approximately the same as the expectations $\hat{\Delta}_j(\bar{\sigma}_j^2)$ and $\hat{\alpha}_j(\bar{\sigma}_j^2)$.

Obviously, when the option is deep in-the-money and the time to maturity decreases, the error bars decrease as well. However, at the last moment of time t_6, the value of the stock has fallen, the option is not so deep in-the-money, and the delta ratio is again more sensitive to the volatility.

Figure 12 represents the ratios $\Delta_i(\bar{\sigma}_i^2)$ and $\alpha_i(\bar{\sigma}_i^2)$, $i = [1, 5]$ of different strike prices. The ratios calculated with integrated volatilities within 1 and 1.5 standard deviation from the expectations of this volatility are presented in darker and lighter green, respectively, an the corresponding expectations are presented in red. In this example, the delta and alpha ratios calculated with the implied volatility coincide with the latter ones.

In the same way than the Black-Scholes formula, also the sensitivity delta (91) is linear with respect to the volatility for an at-the-money options and non-linear for other options. This phenomenon that can be observed in the left panel of the first and second rows by comparing the ratios $\Delta_j(\hat{\sigma}_j^2)$, $\Delta_j^{\mathrm{STD}}(\bar{\sigma}_j^2)$, and $\Delta_j^{1.5\mathrm{STD}}(\bar{\sigma}_j^2)$ of the at-the-money options. At the first moment of time t_1 the option with strike price $K = 1$ is at the money, and at t_5 the corresponding strike price is approximately $K = 1.1$.

At t_5, the option is close to maturity, and the changes in the value of the underlying are reflected directly in the value of the options that are deep-in-the-money, with strike prices $K = 0.6$ and $K = 0.8$. This is then reflected in the delta ratios, which, regardless the volatility used, are approximately one.

The alpha ratios of the delta-vega strategy are presented in the right panels of the first two rows in Figure 12. Now, as the short option is hedged with both an instantaneous amount of the underlying and of a replicating options, the relation between alpha and the value of the underlying is not as straightforward as earlier. However, these ratios clearly have similarities with the delta ratios.

The bottom row represents the price $C(\bar{\sigma}^2)$ of the whole hedging procedure of both delta and delta-vega strategies. The expectation of the cost is presented in red while the green refers to the price of hedging when the ratios are calculated with integrated volatilities within one standard deviation from

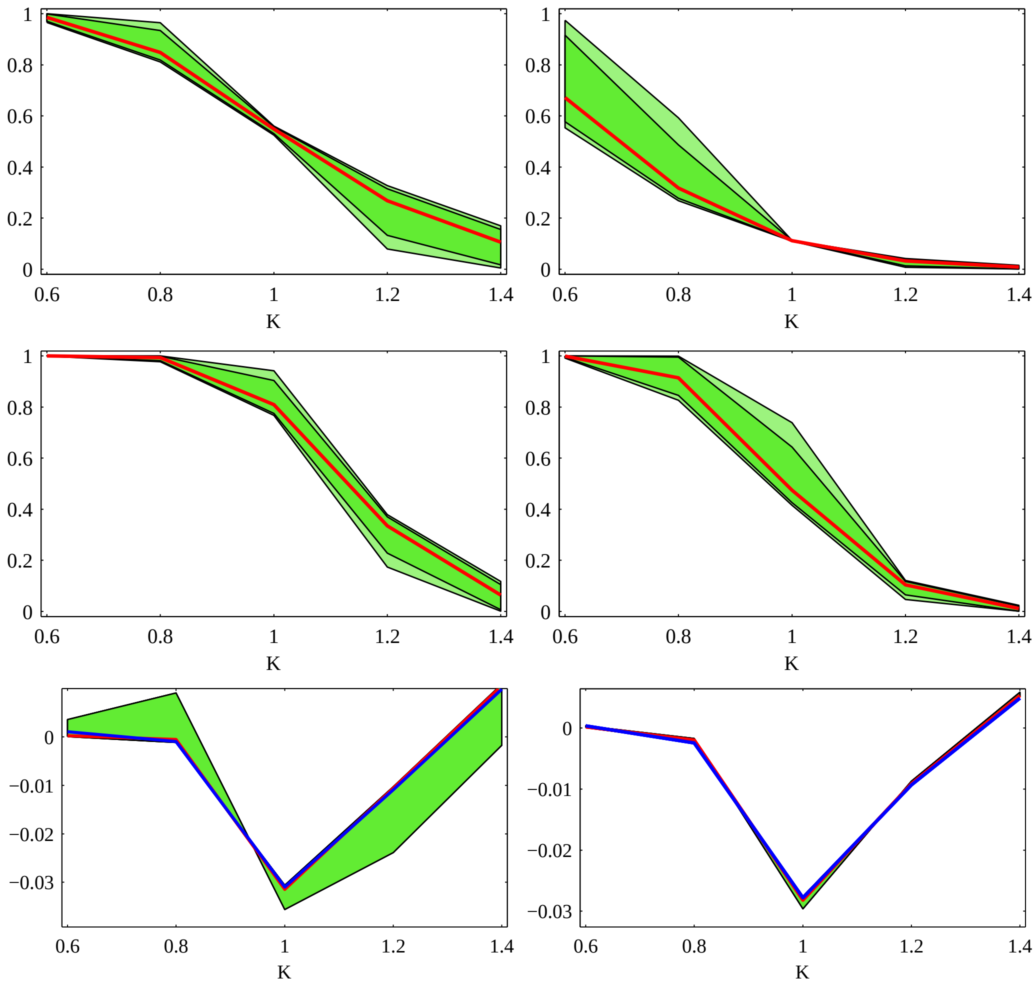


Figure 12: Expectations and distributions of the delta and alpha ratios. On the top row, the left panel represents delta ratios and the right panel alpha ratios at the initial moment t_1. The delta ratios calculated with integrated volatilities within 1 and 1.5 standard deviation from the expectation of these volatilities are presented in darker and lighter green, respectively, and the expectation $\hat{\alpha}_j(\bar{\sigma}_j^2)$ is presented in red. The left panel of the second row represents the corresponding delta ratios and the right panel the alpha ratios at t_5. The bottom row represents the expectation and distribution of the total cost of the delta strategy (left) and of the delta-vega strategy (right), as well as corresponding performances of the implied volatility.

the expectation of these volatilities. We can see that the delta-vega strategy is less sensitive to wrong volatilities, as the error from a wrong volatility is partly cancelled by the replicating option. In this particular example, the average cost of both strategies is approximately the same. Repeating the simulation several times reveals that the cost of the delta-vega strategy is in the average less than the cost of the delta strategy. In this example, possible differences between the performance of the implied integrated volatility and the implied volatility are dominated with errors due to the low re-hedging frequency, which is once a month.

12 Concluding remarks

The integrated volatility is the time-average of the stock price variance. It is a fundamental quantity related to the stock price returns, as these returns depend on the stock price volatility only via the integrated volatility. Hull-White option prices are functionals of the distribution of the integrated volatility.

In this thesis, we presented the new concept of implied integrated volatility, i.e., the integrated volatility implied by option Hull-White prices. The commonly known Black-Scholes implied volatility is based on the assumption of constant stock price volatility. As it is well known, this assumption is non-tenable, which is reflected as a volatility smile, an inconsistency of volatilities implied by option prices with different strikes. In contrast, the integrated volatility is based on the assumption that stock price volatility is a stochastic process.

The Hull-White formula (65) is a bedrock of our work. We have derived this formula in various ways, and, based on it, explained how the different shapes of volatility smiles are related to different distributions of the implied integrated volatility.

The Bayesian approach was used for the ill-posed problem of estimating the integrated volatility implied by option prices. This approach allowed us to integrate additional prior information into the estimation process. We presented two methods to estimate the implied integrated volatility, calculated MAP estimates for this volatility and studied their reliability with MCMC methods.

We suggested that the implied integrated volatility can be used in volatility estimation, in pricing illiquid options consistently with corresponding liquid ones, and in hedging options. Three computed examples on these topics using the integrated volatility were presented. We showed how, in addition to a point estimate, the Bayesian approach provides information on the reliability of these estimates.

Numerous possibilities for future research predicated on the topics considered in this thesis exist. For option prices, we used simulated data based on the Hull-White pricing formula. We did not add noise to the data, as in our paradigm, observed market data contains no noise. The real test of the theory remains to be done: to test the theory and algorithms with real market data.

We have presented only a very simplified example of pricing illiquid options with the integrated volatility. Further research should be done on the possibilities of present and future pricing of different exotic options.

Hedging with integrated volatility was discussed and demonstrated only on a general level. In a computed example, we demonstrated what kind of information on the hedging ratios and hedging performance is provided by our volatility. Comparing the hedging performance of this volatility to the performance of other volatilities remains to be done.

In addition to the topics presented in this thesis, other possible fields to apply the implied integrated volatility should be surveyed, such as, for example, the fast growing fields of volatility options and the Value at Risk.

References

[1] C Alexander and L M Nogueira. Model-free hedge ratios and scale-invariant models. *Journal of Banking & Finance*, 31(6):1839–1861, June 2007.

[2] K I Amin and V K Ng. Inferring future volatility from the information in implied volatility in eurodollar options: A new approach. *The Review of Financial Studies*, 10(2):333–367, 1997.

[3] T G Andersen and T Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905, November 1998.

[4] T G Andersen, T Bollerslev, P Christoffersen, and F X Diebold. Volatility forecasting. *SSRN eLibrary*, 2005.

[5] T G Andersen, T Bollerslev, and N Meddahi. Realized volatility forecasting and market microstructure noise. *working paper*, 2006.

[6] M Avellaneda, C Friedman, R Holmes, and D Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Applied Mathematical Finance*, 4:37–64, March 1997.

[7] M Avellaneda, A Levy, and A Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2(2):73–88, 1995.

[8] L Bachelier. Théorie de la spéculation. *Annales de l'École Normale Supérieure (3)*, 17:21–86, 1900.

[9] G S Bakshi, C Q Cao, and Z Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52(5):2003–2049, December 1997.

[10] O E Barndorff-Nielsen and N Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B*, 63(2):167–241, 2001.

[11] D S Bates. Testing option pricing models. NBER Working Papers 5129, National Bureau of Economic Research, Inc, May 1995.

[12] D S Bates. Post-'87 crash fears in the S&P 500 futures option market. *J. Econometrics*, 94(1–2):181–238, 2000.

[13] D S Bates. Empirical option pricing: A retrospection. *Journal of Econometrics*, 116(1–2):387–404, 2003.

[14] D S Bates. Hedging the smirk. *Financial Research Letters 2*, 4:195–200, 2005.

[15] C Bender, T Sottinen, and E Valkeila. No-arbitrage pricing beyond semimartingales. *WIAS Preprint No. 1110*, 2006.

[16] P Billingsley. *Probability and measure.* John Wiley and Sons, New York-Chichester-Brisbane, 1979.

[17] F Black and M Scholes. The valuation of option contracts and a test of market efficiency. *The Journal of Finance*, 27(2):399–417, May 1972.

[18] F Black and M Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54, May-June 1973.

[19] T Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, April 1986.

[20] I Bouchouev and V Isakov. The inverse problem of option pricing. *Inverse Problems*, 13:L11–L17, 1997.

[21] P Boyle, M Broadie, and P Glasserman. Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21(8–9):1267–1321, June 1997.

[22] D T Breeden and R H Litzenberger. Prices of state-contingent claims implicit in option prices. *The Journal of Business*, 51(4):621–651, October 1978.

[23] L Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[24] P Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics.* Springer-Verlag, New York, 1999.

[25] D Calvetti and E Somersalo. A gaussian hypermodel to recover blocky objects. *Inverse Problems*, 23:733–754, 2007.

[26] D Calvetti and E Somersalo. Hypermodels in the bayesian imaging framework. *Submitted*, 2007.

[27] D Calvetti and E Somersalo. *Introduction to Bayesian Scientific Computing – Ten Lectures on Subjective Computing.* Springer Verlag, 2007.

[28] D Calvetti and E Somersalo. A unified bayesian frameworkfor algorithms to recover blocky signals. *Advanced Signal Processing Algorithms, Architectures and Implementations XVII, (Ed) Franklin T. Luk, Proc. of SPIE Vol. 6697 (SPIE 2007), paper no. 6697–4 (10 pages)*, 2007.

[29] L Canina and S Figlewski. The informational content of implied volatility. *Review of Financial Studies*, 6(3):659–681, 1993.

[30] D P Chiras and S Manaster. The information content of option prices and a test of market efficiency. *Journal of Financial Economics*, 6(2–3):213–234, 1978.

[31] B J Christensen and N R Prabhala. The relation between implied and realized volatility. *Journal of Financial Economics*, 50:125–150, 1998.

[32] P K Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, January 1973.

[33] T F Coleman, Y Li, Y Kim, and A Verma. Dynamic hedging with a deterministic volatility function model. *Journal of Risk*, 4:64–90, 2001.

[34] F Comte and E Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4):291–323, 1998.

[35] R Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3):519–547, 2006.

[36] R Cont and J da Fonseca. Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60, 2002.

[37] R Cont, V Durrleman, and J da Fonseca. Stochastic models of implied volatility surfaces. *Economic Notes*, 31(2):361–377, July 2002.

[38] R Cont and S B Hamida. Recovering volatility from option prices by evolutionary optimization. *Journal of Computational Finance*, 8(4), Summer 2005.

[39] R Cont and P Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[40] J C Cox and S A Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3(1-2):145–166, 1976.

[41] S Crépey. Calibration of the local volatility in a generalized Black-Scholes model using tikhonov regularization. *SIAM Journal on Mathematical Analysis*, 34(5):1183–1206 (electronic), 2003.

[42] S R Das and R K Sundaram. Of smiles and smirks: A term-structure perspective. *The Journal of Financial and Quantitative Analysis*, 34(2):211–239, June 1999.

[43] E Derman and I Kani. Riding on the smile. *Risk*, (7):32–39, 1994.

[44] A Dudenhausen. How to avoid a hedging bias. Bonn Econ Discussion Papers bgse34/2002, University of Bonn, Germany, November 2002.

[45] D Duffie. *Dynamic asset pricing theory, 3rd ed.* Princeton (NJ): Princeton University Press, 2001.

[46] D Duffie, J Pan, and K Singleton. Transform analysis and asset pricing for affine jump-diffusions. NBER Working Papers 7105, National Bureau of Economic Research, Inc, April 1999.

[47] B Dumas, J Fleming, and R E Whaley. Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6):2059–2106, 1998.

[48] B Dupire. Pricing with the smile. *Risk*, (7):8–12, 1994.

[49] R F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, July 1982.

[50] C S Forbes, G M Martin, and J Wright. Bayesian estimation of a stochastic volatility model using option and spot prices. Monash Econometrics and Business Statistics Working Papers 2/02, Monash University, Department of Econometrics and Business Statistics, February 2002.

[51] J-P Fouque, G Papanicolaou, and K R Sircar. *Derivatives in financial markets with stochastic volatility.* Cambridge University Press, Cambridge, 2000.

[52] P Friz and J Gatheral. Valuation of volatility derivatives as an inverse problem. *Quantitative Finance*, 5(6):531–542, December 2005.

[53] J Gatheral. *The volatility surface.* John Wiley and Sons, New Yersey, 2006.

[54] S Geman and D Geman. Stochastic relaxation, gibbs distribution, and the bayesian restoration of images. *IEEE Trans. on Pattern Anal. Machine Intell.*, (6):721–741, 1984.

[55] R Geske. The valuation of compound options. *Journal of Financial Economics*, (7):63–81, 1979.

[56] I I Gihman and A V Skorohod. *The theory of stochastic processes.* Springer-Verlag, Berlin, 1974.

[57] P Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics, New York*. Springer-Verlag, New York, 2004.

[58] R C Green and R A Jarrow. Spanning and completeness in markets with contingent claims. *Journal of Economic Theory*, 41(1):202–210, February 1987.

[59] T C Green and S Figlewski. Market risk and model risk for a financial institution writing options. *Journal of Finance*, 54(4):1465–1499, 08 1999.

[60] J M Harrison and D M Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, 1979.

[61] J M Harrison and S R Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3):215–260, 1981.

[62] T Hein and B Hofmann. On the nature of ill-posedness of an inverse problem arising in option pricing. *Inverse Problems*, 19(6):1319–1338, 2003.

[63] L Hentschel. Errors in implied volatility estimation. *Journal of Financial and Quantitative Analysis*, (38):779–810, 2003.

[64] S L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[65] J C Hull. *Options, futures, and other derivatives, 6 ed.* Upper Saddle River, NJ, Pearson, Prentice Hall, 2006.

[66] J C Hull and A D White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, June 1987.

[67] J C Jackwerth and M Rubinstein. Recovering probability distributions from option prices. *The Journal of Finance*, 51(5):1611–1631, December 1996.

[68] E Jacquier and R Jarrow. Model error in contingent claim models dynamic evaluation. Cirano working papers, CIRANO, March 1996.

[69] E Jacquier and R Jarrow. Bayesian analysis of contingent claim model error. *Journal of Econometrics*, 94(1–2):145–180, 2000.

[70] E Jacquier, N G Polson, and P E Rossi. Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, 12:371–389, 1994.

[71] E Jacquier, N G Polson, and P E Rossi. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212, September 2004.

[72] R Jarrow and A Rudd. Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10(3):347–369, November 1982.

[73] R Jeffrey. *Subjective probability : The real thing*. Cambridge University Press, Cambridge, 2004.

[74] M Johannes and N G Polson. MCMC methods for financial econometrics. *Handbook of Financial Econometrics, Ait-Sahalia Y and Hansen L eds.*, 2004.

[75] M S Johannes, B Eraker, and N Polson. The impact of jumps in volatility and returns. *SSRN eLibrary*, 2000.

[76] H Johnson and D Shanno. Option pricing when the variance is changing. *The Journal of Financial and Quantitative Analysis*, 22(2):143–151, June 1987.

[77] C S Jones. The dynamics of stochastic volatility: Evidence from underlying and options markets. *Journal of Econometrics*, 116(1–2):181–224, 2003.

[78] J Kaipio and E Somersalo. *Statistical and computational inverse problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2005.

[79] J Kaipio and E Somersalo. Statistical inverse problems: discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, 198:493–504, 2007.

[80] I Karatzas and S E Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[81] N El Karoui, M Jeanblanc-Picqué, and S Shreve. Robustness of the Black and Scholes formula. *Mathematical Finance*, 8:93–126, 1998.

[82] Ronald Lagnado and Stanley Osher. A technique for calibrating derivative security pricing models: Numerical solution of an inverse problem. Computing in Economics and Finance 1997 101, Society for Computational Economics, 1997.

[83] D Lamberton and B Lapeyre. *Introduction au calcul stochastique appliqué à la finance*. Ellipses Édition Marketing, Paris, second edition, 1997.

[84] C G Lamoureux and W D Lastrapes. Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *Review of Financial Studies*, 6(2):293–326, 1993.

[85] H A Latane and R J Jr Rendleman. Standard deviations of stock price ratios implied in option prices. *The Journal of Finance*, 31(2):369–381, May 1976.

[86] R Lee. Implied and local volatilities under stochastic volatility. *International Journal on Theoretical and Applied Finance*, 4(1):45–89, 2001.

[87] J S Liu. *Monte Carlo strategies in scientific computing*. Springer Series in Statistics. Springer-Verlag, New York, 2001.

[88] B Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419, October 1963.

[89] A Melino and S M Turnbull. Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45(1–2):239–265, 1990.

[90] M Musiela and M Rutkowski. *Martingale methods in financial modelling*, volume 36 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, second edition, 2005.

[91] R R Officer. The variability of the market factor of the New York Stock Exchange. *Journal of Business*, 46(3):434–53, July 1973.

[92] B Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003.

[93] E Platen and D Heath. *A benchmark approach to quantitative finance*. Springer Finance. Springer-Verlag, Berlin, 2006.

[94] S-H Poon and C W J Granger. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539, June 2003.

[95] J M Poterba and L H Summers. The persistence of volatility and stock market fluctuations. *American Economic Review*, 76(5):1142–1151, December 1986.

[96] P E Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004.

[97] R Rebonato. *Volatility and correlation*. John Wiley and Sons, Chichester, UK, 2004.

[98] E Renault and N Touzi. Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance*, 6(3):279–302, December 1996.

[99] C P Robert and G Casella. *Monte Carlo statistical methods*. Springer Texts in Statistics. Springer-Verlag, New York, second edition, 2004.

[100] M Rubinstein. Implied binomial trees. *The Journal of Finance*, 49(3):771–818, July 1994.

[101] P A Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, (6):41–49, Spring 1965.

[102] M J Schervish and B P Carlin. On the convergence of successive substitution sampling. *Journal of Computational and Graphical Statistics*, 1(2):111–127, 1992.

[103] P J Schönbucher. A market model for stochastic implied volatility. Series A 357, University of Bonn, Germany, 1999.

[104] L O Scott. Option pricing when the variance changes randomly: Theory, estimation, and an application. *The Journal of Financial and Quantitative Analysis*, 22(4):419–438, December 1987.

[105] S E Shreve. *Stochastic calculus for finance. II.* Springer Finance. Springer-Verlag, New York, 2004.

[106] J M Steele. *Stochastic calculus and financial applications*, volume 45 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 2001.

[107] E M Stein and J C Stein. Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies*, 4(4):727–52, 1991.

[108] J Stein. Overreactions in the options market. *Journal of Finance*, 44(4):1011–1023, September 1989.

[109] N Taleb. *Dynamic hedging : Managing vanilla and exotic options.* Wiley, New York, 1997.

[110] A Tarantola. *Inverse problem theory.* SIAM, Philadelphia, 2005.

[111] S J Taylor. Financial returns modelled by the product of two stochastic processes – a study of daily sugar prices 1961–79. *In O. D. Anderson (Ed), Time Series Analysis: Theory and Practice, 1*, pages 203–226, 1982.

[112] J B Wiggins. Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics*, 19(2):351–372, December 1987.

[113] G Willard. Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *The Journal of Derivatives*, (5):45–61, 1997.

[114] P Wilmott. *Derivatives.* John Wiley & Son Ltd, New York, 1998.

(continued from the back cover)

A540 Riikka Korte
Geometric properties of metric measure spaces and Sobolev-type inequalities
January 2008

A539 Aly A. El-Sabbagh, F.A. Abd El Salam, K. El Nagaar
On the Spectrum of the Symmetric Relations for The Canonical Systems of Differential Equations in Hilbert Space
December 2007

A538 Aly A. El-Sabbagh, F.A. Abd El Salam, K. El Nagaar
On the Existence of the selfadjoint Extension of the Symmetric Relation in Hilbert Space
December 2007

A537 Teijo Arponen, Samuli Piipponen, Jukka Tuomela
Kinematic analysis of Bricard's mechanism
November 2007

A536 Toni Lassila
Optimal damping set of a membrane and topology discovering shape optimization
November 2007

A535 Esko Valkeila
On the approximation of geometric fractional Brownian motion
October 2007

A534 Jarkko Niiranen
A priori and a posteriori error analysis of finite element methods for plate models
October 2007

A533 Heikki J. Tikanmäki
Edgeworth expansion for the one dimensional distribution of a Lévy process
September 2007

A532 Tuomo T. Kuusi
Harnack estimates for supersolutions to a nonlinear degenerate equation
September 2007

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The reports are available at *http://math.tkk.fi/reports/* .
The list of reports is continued inside the back cover.

A545 Ruth Kaila
The integrated volatility implied by option prices, a Bayesian approach
April 2008

A544 Stig-Olof Londen, Hana Petzeltová
Convergence of solutions of a non-local phase-field system
March 2008

A543 Outi Elina Maasalo
Self-improving phenomena in the calculus of variations on metric spaces
February 2008

A542 Vladimir M. Miklyukov, Antti Rasila, Matti Vuorinen
Stagnation zones for A-harmonic functions on canonical domains
February 2008

A541 Teemu Lukkari
Nonlinear potential theory of elliptic equations with nonstandard growth
February 2008

ISBN 978-951-22-9315-5 (print)
ISBN 978-951-22-9316-2 (electronic)
ISSN 0784-3143
TKK Mathematics, 2008