Università della Svizzera Italiana
Facoltà di Scienze Economiche

HEDGING AND RISK MEASUREMENT FOR OPTION PORTFOLIOS

A thesis submitted by
Francesco Fierli

for the degree of

Dottore in Scienze Economiche

February 2004

Members of the jury:

Prof. F. Trojani, president of the jury

Prof. G. Barone-Adesi, thesis director

Prof. H. Geman, external member

Acknowledgments

The writing up of a PhD thesis is the result of a personal effort of will. However, the final goal could never have been achieved without the support of a number of people who contributed in one way or another. To mention all these people in these acknowledgments is the minimum that the author can do.

First of all, I would like to thank Professor Giovanni Barone-Adesi (the director of my thesis) and Professor Fabio Trojani for their unfailing assistance and their valuable advice. This thesis is the most evident result of their support.

Second, I would like to thank Professor Hélyette Geman for the trust and interest that she showed in my work. I was very glad when she accepted to be a member of the jury of my thesis.

Third, I would like to thank BSI-Gamma Foundation and Banca del Gottardo for their financial support during the past four years. In particular I would like to thank everybody in the Risk-management Group of Banca del Gottardo who motivated me greatly during the last two years.

I would also like to mention the priceless contribution of all those who somehow helped me to write this work. In an alphabetic order they are: Simona Cain, who proofread most of the final draft; Professor Jérôme Detemple, for his useful comments during the Swiss Doctoral Workshop in Ascona; Patrik Gagliardini, for his ready and sharp answers to my questions; Gianluca Marcoli, for his crucial moral support; Gisèle Nodiroli, for organizing the PhD viva and for handling the related bureaucratic aspects; Claudio Ortelli, for his significant help in the estimation part; Paolo Porchia, for our challenging discussions.

Finally, my gratitude goes to my father, my mother, my sister, and, most of all, to my wife Inna for her boundless understanding. This work would not have been completed without her precious support.

I dedicate this work to my grandfather, who is no longer, and to my first child, who is soon to come into the world.

Contents

Preface

The unceasing increase of option transactions makes this kind of contract very popular among practitioners. Nowadays options constitute the most important component of the trading book for many financial institutions. A further incentive was given by the introduction of the electronic trading platforms to access to derivative markets. For instance, in 2002 more than 800m option contracts were done just on EUREX market.

The large familiarity that traders generally have with options does not prevent the problems induced by large option positions to cover. First of all, a pricing problem has to be solved. The value of an option can be computed as the expectation under some risk neutral measure of the discounted future cash flows. Hence, the central issue becomes the definition of the risk neutral measure to get the expected value.

Closely related to the pricing problem, the hedging problem appears each time an investor has an open option position. This could be the case of an institution selling an option to gain on the edge paid by the customer. In order to avoid unexpected losses, the institution has to build a portfolio that replicates the option value in order to cover the risk when the position is closed.

The third problem is inherent to the nature of option contracts. The non linearity of the relation that links option prices to underlying asset prices generates some problems for risk measurement, too. In particular, all the distribution based risk measures, such as the value at risk, become very difficult to estimate also under the standard assumption of the underlying asset return normality.

In spite of the introduction of a consistent number of realistic pricing models, the most used pricing model is still the one introduced by Black-Scholes [9] and Merton [51] about thirty years ago. The popularity of this model is primarily due to ease of implement. Indeed, only one unobservable parameter has to be estimated and all the other variables are observable in the market.

However, the Black-Scholes model cannot be considered the optimal choice for pricing and hedging options. The most debatable issue is the assumption that the underlying process has a constant volatility parameter. The high kurtosis on stock return distribution, originally pointed out by

Fama [23] and Mandelbrot [49], and the volatility smile effect clash with the assumption of constant volatility.

Moreover, even if one accepts the constant volatility assumption, an estimation problem still remains. Indeed, the underlying volatility is not observable and has to be estimated. Since option prices are generally very sensitive with respect to volatility, the error made by substituting the true parameter with the estimated one could not be considered negligible either for pricing or for hedging purposes.

In order to relax the constant volatility assumption, in the last two decades a large number of stochastic volatility models have been proposed. Some seminal papers on this subject are those by Hull and White [41], Wiggins [62], Stein and Stein [60] and Heston [40][1]. The main feature of these models is that the volatility is no longer constant, but its dynamics is described by a stochastic differential equation. Some of these models provide a closed-form solution so that their implementation difficulty is similar to the Black-Scholes model.

However, stochastic volatility models are not lacking drawbacks. The introduction of a second source of uncertainty produces market incompleteness. This problem is generally solved by assuming a functional form for the volatility risk premium and implicitly by making some hypothesis on investor preferences.

A more practical problem concerns the estimation of model parameters. Indeed, stochastic volatility models are generally based on one (or more) latent variables and a set of unobservable parameters. Empirical investigations[2] show that stochastic volatility models are very difficult to estimate. Hence, the parameter misspecification problem due to estimation error cannot be neglected either for the option pricing or for the hedging.

To reduce the impact of the parameter misspecification, the super-hedging approach proposed by Avellaneda Levy and Parás [5] (hereafter ALP) can be generalized to a stochastic volatility framework. The super-hedging under ALP approach provides two bounds for the price of an option portfolio by assuming that the volatility is unknown but bounded between two values. Defining a set of possible volatility values is equivalent to defining a set of equivalent martingale measures for the no-arbitrage option price. Under this approach the two bounds for the option price are generally very distant from each other and cannot be conveniently used. By assuming a model for the volatility dynamic and by imposing some bounds on the parameter set, it is possible to reduce the set of equivalent martingale measures and therefore the distance between the two price bounds.

In Chapter 2 the ALP approach is extended to a stochastic volatility

[1] See also the detailed surveys by Ball and Roma [7] and Frey [29].

[2] See Bakshi et al. [6], Chernov and Ghysels [12], Andersen at al. [3] and Fiorentini et al. [26].

framework. The two bounds for the price of an option portfolio are obtained by numerically solving a non-linear PDE. The application to Heston's model shows that the distance between the two price bounds is lower than the distance between the price bounds obtained in the ALP approach.

Within the ALP approach, the super-hedging price of an option portfolio with a delta function monotonic with respect to the underlying price is given by the Black-Scholes formula with volatility equal to one of the two bounds. Under the stochastic volatility approach, a closed-form solution cannot be obtained even in the case of monotonic delta function. However, a closed form approximation of the two super-hedging bounds can be used whenever a closed form solution can be obtained for the stochastic volatility model considered.

In order to obtain a useful tool for pricing and hedging, a way to determine the parameter bounds has to be defined. In Chapter 3 some bounds for the volatility process parameters are obtained by performing an interval estimation of a class of stochastic volatility models. To have a consistent estimator, the indirect inference introduced by Gourieroux, Monfort and Renault [34] and the EMM introduced by Gallant and Tauchen [31] are used. Both the estimation methods provide an asymptotic distribution for the estimator that can be used to assess a confidence interval for model parameters. In the same chapter, a Monte Carlo study is performed to verify whether the asymptotic distribution could be considered a good approximation also in the finite sample.

The estimation framework is divided into two parts. The first one is devoted to estimating the model under the real-world probability measure. For this purpose, the underlying asset returns are used. The parameters estimated in this way cannot be used for pricing and hedging purposes, where a risk adjusted probability measure has to be considered. In order to estimate the model parameters under the risk adjusted probability measure, a time series of option implied volatility is used. Moreover, in the risk adjusted framework, the distribution of option implied volatility can be estimated.

Chapter 4 is devoted to verifying pricing and hedging reliability of the super-hedging method proposed. The pricing properties are examined in a real data experiment. The option market price is almost always between the two super-hedging bounds. Compared to the super-hedging bounds of the ALP approach, the super-hedging bounds under stochastic volatility get closer to each other and could be conveniently used to define a bid-ask spread.

In a standard Black-Scholes framework, a self-financing replicating portfolio can be found by using the underlying asset and a riskless bond. Since in a stochastic volatility model there is a new source of randomness, the two assets are no longer sufficient. In order to ensure both the replicating and the self-financing properties, it is necessary to introduce a second option, i.e. a new asset whose value depends on the volatility.

A Monte Carlo experiment shows that the super-hedging under stochastic volatility solves the parameter misspecification problem. Moreover, the proposed super-hedging approach performs better than the ALP approach. The same results are obtained under model misspecification.

In the second part of this work, we aim to show the impact of large option positions on risk measurement methods. Chapter 5 provides an overview of some value at risk (hereafter VaR) estimation methods. Among all VaR estimation methods, only those able to manage non-linear positions are considered. The first method considered is completely parametric and it is based on a quadratic approximation. The non-parametric method based on historical simulations is also described. The option portfolio revaluation made for each past scenario makes this method suitable to manage large option positions.

Two generalizations of the standard historical simulation method are presented. The first one is that introduced by Boudoukh, Richardson and Whitelaw [10] and it improves the sensitivity of VaR measure to sudden changes in market risk. A more complex generalization is that introduced by Barone-Adesi, Bourgoin and Giannopoulos [8], which is known as filtered historical simulation method. This approach is primarily devoted to solving the inconsistency problem of the standard historical simulation VaR estimator. A generalization of the filtered historical simulation method is also proposed.

The different VaR estimation methods are tested in Chapter 6 by using an unconditional and a conditional test. The test refers to a time horizon of one day and to accepted loss probabilities of 1% and 5%. Four portfolios are considered: two with linear positions only and two with relevant option positions.

The test results show that the filtered historical simulation approach performs better than the other VaR estimation methods. The generalization proposed performs better than the standard filtered historical simulation method for the accepted loss probability of 5%. The worst performances are obtained with the parametric method. It seems completely inadequate to manage non-linear positions.

Part I

Hedging problems with uncertain volatility

Chapter 1

Complete and incomplete markets

In a complete market all the contingent claims can be replicated by a portfolio composed of some basic risky assets (stocks or currencies) and a riskless asset. In this market, the no-arbitrage argument was used to price contingent claims firstly by Black and Scholes [9] and Merton [51]. A more mathematically precise framework for the application of the no-arbitrage argument to the option pricing was given later by Harrison and Kreps [36] and Harrison and Pliska [37], [38].

Although market completeness is a desirable property to price a claim, it is not even a realistic assumption. However, also in an incomplete market, the investor can be safe if he accepts to pay a higher hedging cost. The subject was developed by Delbaen [17], El Karoui and Quenez [21] and Kramkov [45] who introduced the idea of super-hedging strategy.

In this chapter, we stress the differences between complete and incomplete markets. The basic setting throughout will be a continuous time economy. However, to give a clearer interpretation to the results of the continuous case, the discrete time economy will be described too.

In Section 1.1 the most important results in a complete market are described. The section is a survey of the work by Harrison and Kreps [36] for the discrete time economy and of the works by Harrison and Pliska [37], [38] for the continuous time economy. In Section 1.2 we try to link the issue of the uncertain volatility structure with the market incompleteness while in Section 1.3 we describe the super-hedging strategy idea.

1.1 The complete market

In this section we adopt the distinction, originally made by Harrison and Pliska [37], between the discrete-time and the continuous-time economies. Although we are interested in the continuous time setting, the discrete time economy with finite sample space enables us to give a more comprehensible interpretation of the results. The continuous time results will be compared to the discrete time ones whenever possible[1].

Later on, the market is assumed to be frictionless, i.e. trading can be continuous, with neither transaction costs nor short selling restrictions. Moreover, stocks pay no dividends.

1.1.1 The discrete time economy

Given a time horizon $T > 0$, where $T \in \mathbb{N}$, let $(\Omega, \mathbb{F}, \mathbb{P})$ be a filtered probability space where the filtration $\mathbb{F} = \{\mathcal{F}_t : t = 1, 2, \ldots, T\}$ is right-continuous, $\mathcal{F}_0 := \{\emptyset, \Omega\}$, $\mathcal{F}_T := \mathcal{F}$ and Ω has a finite number of elements. Moreover, assume $\mathbb{P}(\omega) > 0$ for all $\omega \in \Omega$.

Remark 1.1.1 The last statement implies that the investors agree only on which state of the world are possible and not also on what the probability of the states is. Thereafter the following definitions and results are true also with a substitution of $\mathbb{P}$ by an equivalent probability measure. $\square$

The economy is composed by $K + 1$ securities whose price processes $S = \{S_t : t = 1, 2, \ldots, T\}$ has strictly positive and $\mathbb{F}$-adapted components. The first security is riskless with $S_0^0 = 1$. To be coherent with the notation of the following sections we denote $S_t^0 = B_t$. Moreover we choose the riskless security as numéraire asset.

In the above market, a *trading strategy* is a predictable $(K+1)$-dimensional process $\phi = \{\phi_t : t = 1, 2, \ldots, T\}$, where predictable means that $\phi_t \in \mathcal{F}_{t-1}$ for $t = 1, 2, \ldots, T$. The interpretation of each component is the number of security units detained by the investor at time t.

Let us define the wealth process $V_t(\phi)$ at time t as

$$V_t(\phi) := \begin{cases} \phi_t S_t = \phi_t^0 B_t + \sum_{k=1}^K \phi_t^k S_t^k & \text{if} \quad t = 1, 2, \ldots, T \\ \phi_1 S_0 = \phi_1^0 B_0 + \sum_{k=1}^K \phi_1^k S_0^k & \text{if} \quad t = 0 \end{cases}$$

and the gain process $G_t(\phi)$ as

$$G_t(\phi) := \sum_{i=1}^t \phi_i \Delta S_i = \sum_{i=1}^t \phi_i^0 \Delta B_i + \sum_{k=1}^K \sum_{i=1}^t \phi_i^k \Delta S_i^k \ ,$$

where $\Delta S_i = S_i - S_{i-1}$ and $\Delta B_i = B_i - B_{i-1}$. The gain process can be interpreted as the sum of the capital gains realized by the investor up through time t. It seems natural to set $G_0(\phi) = 0$.

[1]Note that a similar reason motivates the description of the discrete time framework in Harrison and Pliska [37].

Definition 1.1.1 *A trading strategy ϕ is said to be* ***self-financing*** *if and only if no funds are added or withdrawn at any period $t = 1, 2, \ldots, T$, that is to say if $\phi_t S_t = \phi_{t+1} S_t$.* □

Remark 1.1.2 Recalling the definition of gain process, one can state that

$$\begin{aligned} G_t(\phi) &= \sum_{i=1}^{t} \phi_i (S_i - S_{i-1}) \\ &= \phi_1 S_1 - \phi_1 S_0 + \phi_2 S_2 - \phi_2 S_1 + \ldots + \phi_T S_T - \phi_T S_{T-1} \\ &= \phi_T S_T - \phi_1 S_0 \end{aligned}$$

where the last line is obtained by using the definition of self-financing trading strategy. It follows that a trading strategy is self financing if and only if $V_t(\phi) = V_0(\phi_t) + G_t(\phi)$, i.e. all changes in the portfolio value are due only to the realized capital gains. □

Definition 1.1.2 *A trading strategy ϕ is said to be* ***admissible*** *if it is self-financing and $V_t(\phi) \geq 0$, where the last statement means that $V_t(\phi)$ is a positive process.* □

The set of all admissible trading strategies is denoted by Φ.

A T-maturity contingent claim φ is a $\mathcal{F}$-measurable nonnegative random variable. The value of φ can be interpreted as a payoff at time T of a derivative asset such that it is determined by the realization of the S price path. Without any other statement, only European style contingent claims are considered. An European contingent claim is a derivative contract which enables the investor to exercise it only at the expiration date T. It differs from the American style contingent claims where the exercise can take place during all the life of the claim.

Definition 1.1.3 *A contingent claim with time T payoff equal to φ is said to be* ***attainable*** *if there exist an admissible trading strategy ϕ such that $V_T(\phi) = \varphi$.* □

The main question is to determine the "rational" price of such a claim. The most common approach is to use the so called arbitrage argument. An arbitrage opportunity is defined to be an admissible strategy such that $V_0(\phi) = 0$ and $E[V_T(\phi)] > 0$, i.e. a strategy which can make profit without any investment and any risk. The rational claim price is obtained by assuming the absence of arbitrage opportunity in the considered market.

First of all we have to define a price system for the contingent claims as a function π which maps from the set of all integrable contingent claims to $\mathbb{R}^+ \cup \{0\}$ and satisfies the following conditions:

$$\begin{aligned} &\pi(\varphi) = 0 \quad \text{if and only if} \quad \varphi = 0 \\ &\pi(a\varphi + b\varphi') = a\pi(\varphi) + b\pi(\varphi') \quad \forall\, a, b \geq 0 \ , \end{aligned}$$

where φ and φ' are two different arbitrary contingent claims. Note that the last condition must be true for all the contingent claims. The above price system is said to be *consistent* with the market model if $\pi[V_T(\phi)] = V_0(\phi)$ for all $\phi \in \Phi$.

Definition 1.1.4 *An equivalent martingale measure $\mathbb{Q}$ is a probability measure defined on the space $(\Omega, \mathcal{F})$ such that*

- *it is equivalent to $\mathbb{P}$, i.e. $\mathbb{P}(A) = 0$ if and only if $\mathbb{Q}(A) = 0$ for all $A \in \mathcal{F}$;*
- *the Radon-Nicodym derivative $d\mathbb{Q}/d\mathbb{P} \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, that is*

$$\int (d\mathbb{Q}/d\mathbb{P})^2 \, d\mathbb{P} < \infty \ ;$$

- *the discounted process $B_t^{-1} S_t$ is a $\mathbb{Q}$-martingale.*

□

The set of all the equivalent martingale measures is denoted by $\mathcal{Q}$. Moreover, one can say that if $\phi \in \Phi$, then each $\mathbb{Q} \in \mathcal{Q}$ is a martingale measure also for the discounted value process $B_t^{-1} V_t(\phi)$. Indeed, by using again the definition of self-financing trading strategy, it is possible to verify that $\Delta[B_t^{-1} V_t(\phi)] = \Delta(\phi_t B_t^{-1} S_t)$. The process $\{\phi_t\}$ is predictable such that $E^{\mathbb{Q}}[\Delta(\phi_t B_t^{-1} S_t)] = 0$ and $B_t^{-1} V_t(\phi)$ is a $\mathbb{Q}$-martingale.

Harrison and Pliska [37] show the relationship between $\mathcal{Q}$ and the set of price systems by proving the following proposition.

Proposition 1.1.1 *There is a one-to-one correspondence between price systems π consistent with the market model and the probability measure $\mathbb{Q} \in \mathcal{Q}$ via*

$$\begin{aligned} \pi(\varphi) &= E^{\mathbb{Q}}\left[B_T^{-1} \varphi\right] \\ \mathbb{Q}(A) &= \pi(B_T \mathbf{1}_A) \qquad\qquad A \in \mathcal{F} \end{aligned}$$

where $\mathbf{1}_A$ is the indicator function of the set A and $E_{\mathbb{Q}}[.]$ is the expectation operator under the probability $\mathbb{Q}$. □

See Harrison and Pliska [37] pp. 227-228 for a proof.

Note that the price $\pi(B_T \mathbf{1}_A)$ is the price of a digital option that gives B_T at time T if $\omega \in A$. Such a price is equal to the probability that ω belongs to A, indeed

$$\pi(B_T \mathbf{1}_A) = E^{\mathbb{Q}}\left[B_T^{-1} B_T \mathbf{1}_A\right] = \mathbb{Q}(A) \ .$$

The above proposition is used by Harrison and Pliska to prove the following theorem and the consequent corollary.

Theorem 1.1.1 *The market model contains no arbitrage opportunity if and only if $\mathcal{Q}$ is nonempty or, equivalently, if and only if there is at least one price system.* □

See Harrison and Pliska [37] pp. 228-229 for a complete proof.

Corollary 1.1.1 *If the market model contains no arbitrage opportunity (or if $\mathcal{Q}$ is nonempty or if there is al least one price system) then, for each $\mathbb{Q} \in \mathcal{Q}$, there is a single price π associated with any attainable contingent claim φ. The price satisfies $\pi = E^{\mathbb{Q}}\left[B_T^{-1}\varphi\right]$.* □

The above corollary states that in a market with no arbitrage opportunity there is a unique price for every attainable contingent claim and for every measure $\mathbb{Q} \in \mathcal{Q}$. The following proposition goes a step ahead by identifying what the price at any time t is.

Proposition 1.1.2 *If the contingent claim φ is attainable, then*

$$V_t(\phi) = B_t E^{\mathbb{Q}}[B_T^{-1}\varphi|\mathcal{F}_t] \qquad\qquad t = 0, 1, \ldots, T$$

for any $\phi \in \Phi$ which generate φ and for any $\mathbb{Q} \in \mathcal{Q}$. □

This proposition is proved by remembering that $V_T(\phi) = \varphi$ and that $B_t^{-1}V_t(\phi)$ is a martingale under each measure $\mathbb{Q} \in \mathcal{Q}$.

Above there are the conditions to have no-arbitrage opportunity and to obtain a no-arbitrage price for any attainable contingent claim, but nothing was said about market completeness.

Definition 1.1.5 *A security market model is said to be* ***complete*** *if every contingent claim is attainable.* □

Harrison and Kreps [36] characterize a complete market setting by proving the following theorem.

Theorem 1.1.2 *The security market model is free of arbitrage opportunity and complete if and only if there exist a unique equivalent martingale measure.* □

See Harrison and Kreps [36] p. 392 for a proof.

1.1.2 The continuous time economy

In the preceding section investors were allowed to trade only at discrete times. If we assume that they can trade continuously, we need a more complex setting. However, we try to facilitate the interpretation of the continuous time results by giving to the present section a structure similar to that of the previous one. Hence most of the definitions, propositions and theorems revised in the new setting keep an akin interpretation.

Given a time horizon $T > 0$ let $(\Omega, \mathbb{F}, \mathbb{P})$ be a filtered probability space where the filtration $\mathbb{F} = \{\mathcal{F}_t : t \in [0, T]\}$ is right-continuous, $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_T := \mathcal{F}$. As in the preceding section, the economy is composed by $K + 1$ securities whose price process $S = \{S_t : t \in [0, T]\}$ has strictly positive and $\mathbb{F}$-adapted components. Here the components $S_t^0, S_t^1, \dots, S_t^n$ are right continuous and left limited (hereafter *cadlag*).

In the above market, a *trading strategy* is a predictable $(K+1)$-dimensional process $\phi = \{\phi_t : t \in [0, T]\}$ whose components are locally bounded. More precisely, we assume that either $\int_0^T |\phi_t^B| dt$ or $E[\int_0^T (\phi_t^k)^2 dt]$ for $k = 1, 2, \dots, K$ are finite with probability one such that the integral involving ϕ_t^B and the stochastic integral involving ϕ_t^k are well defined[2].

Remark 1.1.3 The definition of predictability is slightly different from that in the preceding section. Indeed a process $\{S_t\}$ is said to be *predictable* if it is measurable with respect to the predictable σ-algebra. The *predictable σ-algebra* is the σ-algebra on $\Omega \times [0, T]$ generated by the simple predictable process. The process $\{S_t\}$ is a *simple predictable process* if there exist times $0 = t_0 < t_0 < \dots < t_n = T$ and bounded random variable $\xi_0, \xi_1, \dots, \xi_{n-1}$ where

$$\xi_i \in \mathcal{F}_i \qquad\qquad i = 0, 1, \dots, n-1$$

such that

$$S_i = \xi_i \quad \text{if} \quad t_i < t < t_{i+1} \qquad\qquad i = 0, 1, \dots, n-1$$

However, the meaning is similar to that in the discrete market setting. Indeed, the predictability condition means that ϕ_t^k is know immediately before the time t. □

Let us *assume* that the set $\mathcal{Q}$ contains at least one equivalent martingale measure. The above assumption enables us to define the wealth process $V_t(\phi)$ at time t as

$$V_t(\phi) = \phi_t S_t = \phi_t^0 B_t + \sum_{k=1}^n \phi_t^k S_t^k \ ,$$

and the gain process $G_t(\phi)$ as

$$G_t(\phi) = \int_0^t \phi_u dS_u = \int_0^t \phi_u^0 dB_u + \sum_{k=1}^n \int_0^t \phi_u^k dS_u^k \ .$$

[2] Note that these are sufficient conditions and that they are invariant with respect to an equivalent change of probability measure.

Definition 1.1.6 *A trading strategy ϕ is said to be* ***self-financing*** *over the interval $[0,T]$ if the wealth process $V_t(\phi)$ satisfies the following condition:*

$$dV_t(\phi) = \phi_t^0 dB_t + \sum_{k=1}^{n} \phi_t^k dS_t \ ,$$

or in the integral form

$$\begin{aligned} V_t(\phi) &= V_0(\phi) + \int_0^t \phi_u^0 dB_u + \sum_{k=1}^{n} \int_0^t \phi_u^k dS_u \\ &= V_0(\phi) + G_t(\phi) \ . \end{aligned}$$

□

This is equivalent to say that for a self-financing trading strategy the wealth process variations are not due to addition or withdrawal of money. Note that the last line of the above equation is equal to the definition of self-financing trading strategy in the discrete time economy as in remark 1.1.2.

Let us denote Φ the set of all self-financing trading strategies such that $V_t(\phi) \geq 0$. Note that this corresponds to the definition of admissible trading strategy given in the preceding section. In the case of a continuous time model one should impose a further restriction to say that a strategy is admissible.

To simplify the notation let us define the discounted price process $\{Z_t : t \in [0,T]\}$ as $Z_t := B_t^{-1} S_t^k$ for $k = 1, 2, \ldots, K$. The discounted value process is

$$V_t^*(\phi) := B_t^{-1} V_t(\phi) \qquad\qquad t \in [0,T]$$

whereas the discounted gain process is

$$G_t^*(\phi) := \int_0^t \phi_u dZ_u \qquad\qquad t \in [0,T]$$

By using the above definitions, Harrison and Pliska [37] proved the next proposition and the consequent corollary:

Proposition 1.1.3 *The trading strategy ϕ is self-financing if and only if $V_t^*(\phi) = V_0^*(\phi) + G_t^*(\phi)$. Moreover, $V_t(\phi) \geq 0$ if and only if $V_t^*(\phi) \geq 0$.* □

Corollary 1.1.2 *If $\phi \in \Phi$, then $V_t^*(\phi)$ is a positive local martingale, and also a supermartingale, under each $\mathbb{Q} \in \mathcal{Q}$.* □

By recalling the definition of arbitrage opportunity, the above corollary states that none of these are present in the market. Indeed, if $\phi \in \Phi$, then $V_t^*(\phi)$ is a positive supermartingale and by letting $V_0^*(\phi) = V_0(\phi) = 0$ also $V_T^*(\phi)$ has to be equal to zero. Moreover, being the process $\{B_t\}$ bounded, one must have $V_t(\phi) = 0$ for each $t \in [0,T]$[3].

[3]The equalities are in the $\mathbb{Q}$-a.s. sense.

Remark 1.1.4 Proposition 1.1.3 and Corollary 1.1.2 give the interpretation of the assumption made on the set $\mathcal{Q}$. Indeed, if the set $\mathcal{Q}$ is nonempty then there are no arbitrage opportunities. Hence, it is possible to see now the similarity with the discrete time setting, where no-arbitrage and nonempty $\mathcal{Q}$ are two equivalent conditions. □

In spite of the results of the preceding section, the condition that ϕ belongs to Φ does not ensure the uniqueness of the price π. A bright example made by Harrison and Pliska [37] shows how the set Φ is too large to guarantee the uniqueness of the price system. They solve the problem by introducing a reference probability measure $\mathbb{P}^* \in \mathcal{Q}$ and by defining $\mathcal{L}(Z)$ as the set of all the predictable process $H_t : t \in [0,T]$ for which the increasing process

$$\left[\int_0^t (H_u^k)^2 d\langle Z^k\rangle_u\right]^{1/2} \qquad t \in [0,T] \quad k = 1,2,\dots,K$$

is locally integrable under $\mathbb{P}^*$.

Definition 1.1.7 *A trading strategy is said to be* ***admissible*** *if $\phi \in \mathcal{L}(Z)$, $V_t^*(\phi) \geq 0$ for $t \in [0,T]$, $V_t^*(\phi) = V_0^*(\phi) + G_t^*(\phi)$ and $V^*(\phi)$ is a $\mathbb{P}^*$-martingale.* □

The set of all admissible trading strategies is denoted by Φ^*. Note that the set Φ^* now is the set of all the trading strategies for which $V^*(\phi)$ is a $\mathbb{P}^*$-martingale and not just a local martingale. Moreover, the last statement of the above definition implies that

$$E^{\mathbb{P}^*}[V_T^*(\phi)] = V_0^*(\phi) \ ,$$

that immediately proofs the following proposition.

Proposition 1.1.4 *The unique price associated with the attainable contingent claim φ is $\pi = E^{\mathbb{P}^*}[B_T^{-1}\varphi]$.* □

Harrison and Pliska [37] proved the analogous to the proposition 1.1.2.

Proposition 1.1.5 *Let φ be an integrable contingent claim and V_t^* the cadlag modification of $V_t^* = E^{\mathbb{P}^*}[B_T^{-1}\varphi|\mathcal{F}_t]$. Then φ is attainable if and only if V_t^* can be represented as $V_t^* = V_0^* + \int HdZ$ for some process H_t belonging to $\mathcal{L}(Z)$, in which case $V_t^*(\phi) = V_t^*$ for any $\phi \in \Phi^*$ which generate φ.* □

Keeping the definition of completeness of the previous section, it is now useful to introduce the following theorem.

Theorem 1.1.3 *The following statements are equivalent:*

(a) *The model is complete under $\mathbb{P}^*$.*

(b) *Every martingale* $\{M_t : t \in [0, T]\}$ *can be represented in the form*

$$M_t = M_0 + \int_0^t H_u dZ_u \ ,$$

for some H_t *belonging to* $\mathcal{L}(Z)$ *(representation property).*

(c) $\mathcal{Q}$ *has exactly one element (singleton condition).*

□

For a proof see Harrison and Pliska [38] pp. 315-316.

Note that as in the preceding section the singleton condition of the set $\mathcal{Q}$ is a necessary and sufficient condition for the completeness of the market.

1.2 Incomplete market and volatility misspecification

In spite of the convenience of the completeness condition, usually one has to work with a more realistic incomplete market. The market is incomplete when some contingent claims are not attainable, i.e. every time that it is not possible to find an admissible trading strategy to replicate the payoff of the claim at the maturity.

The sources of incompleteness are manifold. When the probability space is finite and the time is discrete, there is incompleteness if the number of the future states of the world is larger than the number of independent assets in the market. This can arise, for instance, when there is no access to certain stocks.

In the continuous time economy, Karatzas [42] shows that the market is complete if and only if the number of independent risky assets traded in the market matches the number of independent sources of uncertainty which drive the asset prices. Hence, there is incompleteness when the number of independent sources of uncertainty is larger than the number of independent risky assets. This situation arises, for instance, when the volatility moves according to a stochastic process. Indeed, the volatility is generally not traded in the market[4].

In any case, when the economy is incomplete, one has to face the problem to price the contingent claims without using the standard arbitrage argument. This is true even if the market model is arbitrage-free. As a consequence, one is no longer sure about the uniqueness condition of the set $\mathcal{Q}$.

The kind of incompleteness that we deal with is generated by the unobservability of the volatility. In the Black-Scholes model the volatility

[4]However, note that, in this case, it is sufficient to introduce a traded option to complete the market. See Section 4.3 for more details.

parameter is assumed to be known. As matter of fact, the volatility is never known and it is replaced by its estimated value. Such a substitution is the cause of the so called volatility misspecification problem that belongs to the more general field of model risk.

One of the solutions to the problem is to remove the assumption of a known volatility. This is the aim of stochastic volatility models, where volatility is assumed to follow a process whose variations are described by a stochastic differential equation. With the exception of the CEV diffusion model introduced by Cox [16], in all the stochastic volatility models there is the assumption that the sources of uncertainty driving the volatility process are not the same ones driving the price process. According to Karatzas [42], this feature generates market incompleteness.

In this case market incompleteness is generally handled by equilibrium considerations. Indeed, by exogenously assuming a risk premium structure for volatility risk, the no arbitrage argument can be used as in a complete market[5].

Even if one solves the incompleteness problem, stochastic volatility models are affected by other practical problems which reduce their attractiveness. One of the most important problem concerns the estimation of the process parameters. Indeed, while in the Black-Scholes model the only parameter to be estimated is the volatility, here the number of parameters to be estimated is generally higher than one. Hence, the misspecification problem is transferred from the volatility value to the parameters of the volatility process.

A completely different approach, introduced by Avellaneda, Lavy and Parás [5], does not make any assumption on the volatility process with the exception that the realized volatility values have to lie inside two bounds. The results are based on the idea of *super-hedging* strategy described in the next section.

1.3 The super-hedging strategy

To solve the pricing problem in an incomplete framework, Delbaen [17] and El Karoui and Quenez [21], among others, introduced the notion of super-hedging strategy in a continuous time framework. Subsequently Kramkov [45] developed the theory for a general semimartingale framework.

The super-hedging theory states that in a perfect but incomplete market, it is possible to create a trading strategy that enables the seller (buyer) of a claim to have a portfolio value higher (lower) or equal to the claim value at the expiration date.

Definition 1.3.1 *A* ***super-hedging*** *strategy for a short position in the con-*

[5] See Wiggins [62] and Lewis [48] among the others.

tingent claim φ is a self-financing trading strategy ϕ for which $V_T(\phi) \geq \varphi$. □

Definition 1.3.2 *A **super-hedging** strategy for a long position in the contingent claim φ is a self-financing trading strategy ϕ for which $V_T(\phi) \leq \varphi$.* □

A super-hedging strategy does not replicate the claim. Even so, it avoids any losses at the expiration of it. Moreover, the considered trading strategy is self-financing. This differentiates the present approach from the risk-minimization hedging method developed by Fölmer and Schweizer [27], [28], where the trading strategy allows additional transfers of funds.

According to the above definitions, the seller's and buyer's price (respectively W^+ and W^-) for the non attainable contingent claim are:

$$\begin{aligned} W_t^+ &= \inf \{V_t(\phi) | \exists \ \phi \in \Phi^* : V_T(\phi) \geq \varphi\} \ , \\ W_t^- &= \sup \{V_t(\phi) | \exists \ \phi \in \Phi^* : V_T(\phi) \leq \varphi\} \ . \end{aligned}$$

For a non negative derivative one can state also that the buyer's price is less then the seller's price. The two prices equal only when the derivative is attainable.

In an incomplete market there is no guarantee that the equivalent martingale measure is unique. In this case, the set of equivalent martingale measures defines a range of candidate contingent claim prices. The two bounds of this range are

$$B_t \sup_{\mathbb{Q} \in \mathcal{Q}} E^{\mathbb{Q}} \left[B_T^{-1} \varphi | \mathcal{F}_t \right] \ , \qquad\qquad B_t \inf_{\mathbb{Q} \in \mathcal{Q}} E^{\mathbb{Q}} \left[B_T^{-1} \varphi | \mathcal{F}_t \right] \ .$$

When the claim payoff is bounded below, El Karoui and Quenez [21] proved that the two bounds are, respectively, the seller's price and the buyer's price.

At first glance, the above theory seems a very attractive deal to price and hedge contingent claims in an incomplete market. However, when the incompleteness is generated by the uncertainty on the volatility structure the difference between seller's and buyer's price is often very high and the super-hedging is too expensive. In this situation, prices obtained by the super-hedging theory have no-practical relevance.

Chapter 2

Super-hedging strategy and stochastic volatility models

The major problem for an institution selling an option not traded in a liquid market is to determine the cost of the hedging strategy to be used in order to avoid a loss. A related problem is the computation of the price that the same institution has to propose to buy the claim back. These tasks are indeed very hard even in a standard Black-Scholes framework. Indeed, the unobservability of the volatility parameter can lead an incorrect valuation of the hedging cost and then to a misspricing of the option.

The super-hedging approach introduced in Chapter 1 addresses this problem. However, although the super-hedging setting gives a clear theoretical framework for the pricing of claims under uncertain volatility structure, it is often unable to provide prices with a practical interest for practitioners. The main problem is that the implied hedging cost is too high to be payable, so that a "pure" super-hedging strategy cannot be reasonably used. For instance, in a model where volatility follows an unbounded diffusion process, Frey and Sin [30] show that the seller's price for a European call is equal to the price of the underlying asset.

To tackle the super-hedging problem, many authors impose further restrictions on volatility behavior in order to reduce the cost of a super-hedging strategy. In this direction, Avellaneda, Levy and Parás [5] assume that volatility lies between two bounds from the date the claim is issued to maturity. Without any assumption on the parametric class of volatility process, they get a non-linear PDE whose solution is the super-hedging price of the claim under a possible volatility misspecification. However, the no-arbitrage interval is still too large.

In this paper, we apply the approach proposed by Avellaneda, Levy and Parás (hereafter ALP) to a parametric class of stochastic volatility models. The main idea is to switch from the bounds on the volatility level to some bounds on the parameters of the volatility process. The implied seller's price can be obtained by solving a PDE and, generally, it is lower than the price obtained by the ALP approach. On the other hand, the buyer's price is higher than the price obtained by the ALP approach, so that the no-arbitrage pricing interval is reduced.

In order to show some numerical results, an application to Heston [40] model is presented. However, note that the following analysis is sufficiently general and many other stochastic volatility models could be considered. In this framework we investigate the reduction effect in the no-arbitrage pricing interval. We show that the formula proposed by Heston, with a proper choice of parameters, can be considered a good approximation of the seller's and the buyer's price for plain vanilla options.

In Section 2.1 we show the main results of the work of Avellaneda, Levy and Parás [5]. Section 2.2 introduces a general parametric specification and shows how it is possible to move from a volatility misspecification problem to a misspecification problem that involves only volatility process parameters. A numerical application is done in Section 2.3 and Section 2.4 concludes.

2.1 Uncertain volatility and stochastic control

The basic idea behind the ALP approach is to consider some volatility bounds instead of the whole process. Indeed, they do not impose any volatility dynamics, but they guess that future volatility values will stay in a bounded set for all times up to the maturity of the claim.

Let us introduce a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a fixed time horizon $T \in (0, \infty)$ and a right-continuous filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ with $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_T := \mathcal{F}$. The market is assumed to be frictionless and the transactions are in continuous time. The economy is composed by two primary assets whose price processes $\{B_t\}_{t\in[0,T]}$ and $\{S_t\}_{t\in[0,T]}$ are *cadlag*, strictly positive and adapted to the filtration $\{\mathcal{F}_t\}$. The first security is the riskless asset and we choose it as numéraire asset.

In order to exclude arbitrage opportunities, we assume that the discounted price process S_t/B_t admits (at least) an equivalent local martingale measure $\mathbb{Q}$. Moreover, under $\mathbb{Q}$ the price process S_t is described by the following SDE:

$$dS_t = rS_t \; dt + \sqrt{v_t} S_t \; dw_t \; , \tag{2.1}$$

where w_t is a $\mathbb{Q}$-Brownian motion, r is the instantaneous interest rate and v_t represents the square of the volatility process at time t. The instantaneous interest rate is assumed to be constant for all relevant trading dates. Hence, $dB_t = rB_t dt$ and $B_0 = 1$.

A T-maturity European style contingent claim $\varphi(S_T)$ is an $\mathcal{F}_T$-measurable non-negative random variable. The value of $\varphi(S_T)$ can be interpreted as a payoff at time T of a derivative asset and it is assumed to be bounded below.

Since the number of independent sources of uncertainty is higher than the number of risky assets, the market is incomplete. Because of the incompleteness, the probability measure $\mathbb{Q}$ is not unique. Let us denote $\mathcal{Q}$ the set of probability measures such that (2.1) holds for some non-anticipative volatility function satisfying the following condition:

$$v_t \in [v_{min},\ v_{max}] \qquad\qquad \forall\, t \in [0,T]\ , \tag{2.2}$$

where $v_{min} \geq 0$ and $v_{max} < \infty$. In other words, the set $\mathcal{Q}$ contains all measures $\mathbb{Q}$ determined by any volatility process satisfying (2.2).

Remark 2.1.1 The assumption (2.2) implies that the Novikov condition for the process v_t is true. Hence, the set $\mathcal{Q}$ is a set of martingale measures (not just local martingales) for the discounted price process S_t/B_t[1]. □

The no-arbitrage price of the contingent claim $\varphi(S_T)$ lies between the following bounds:

$$W^+(S_t,t) \;=\; \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} E^{\mathbb{Q}}\left[e^{-r(T-t)}\varphi(S_T)|\mathcal{F}_t\right]\ , \tag{2.3}$$

$$W^-(S_t,t) \;=\; \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} E^{\mathbb{Q}}\left[e^{-r(T-t)}\varphi(S_T)|\mathcal{F}_t\right]\ . \tag{2.4}$$

As suggested by Avellaneda, Levy and Parás [5], the two bounds can be obtained by solving a dynamic programming problem and by considering (2.3) and (2.4) as two stochastic control problems with control variable v_t. The PDEs obtained for the upper and lower bounds are, respectively,

$$\begin{aligned}\frac{\partial W^+}{\partial t} + \frac{\partial W^+}{\partial S}rS + \frac{1}{2}\Gamma^+S^2v_{max} &+ \\ +\frac{1}{2}\Gamma^+S^2\mathbf{1}_{\{\Gamma^+<0\}}(v_{min}-v_{max}) &- rW^+ = 0\ ,\end{aligned} \tag{2.5}$$

$$\begin{aligned}\frac{\partial W^-}{\partial t} + \frac{\partial W^-}{\partial S}rS + \frac{1}{2}\Gamma^-S^2v_{max} &+ \\ +\frac{1}{2}\Gamma^-S^2\mathbf{1}_{\{\Gamma^->0\}}(v_{min}-v_{max}) &- rW^- = 0\ ,\end{aligned} \tag{2.6}$$

where $\Gamma^+ := \frac{\partial^2 W^+}{\partial S^2}$ and $\Gamma^- := \frac{\partial^2 W^-}{\partial S^2}$. The solutions to (2.5) and (2.6) are found by assuming the boundary condition $W^{\pm}(S_T,T) = \varphi(S_T)$.

Remark 2.1.2 When the portfolio is composed only by short or long positions on plain vanilla options, its payoff function is, respectively, always concave or always convex. In such a case the solution of the above PDE is straightforward and is equal to the Black-Scholes solution with the volatility equal to one of the two bounds in (2.2)[2]. However, in portfolios with mixed convexity the previous statement is no longer true. □

[1] See, for instance, Karatzas and Shreve [43] Corollary 5.13 on Section 3.5.
[2] See El Karoui et al. [20] for more details.

Example 2.1.1 Let us consider a bullish call spread with the two strikes equal to 80 and 100. The risk-free rate is 0.05 and the option maturities are both 6 months. Moreover, let as assume that with a certain confidence the volatility $\sqrt{v_t}$ will lie inside the interval $[0.1, 0.5]$.

By using a Crank-Nicholson scheme[3], we can solve equations (2.5) and (2.6) for different values of the underlying. The results are showed in Figure 2.1. The dotted line in the figure represents the spread price according to Black-Scholes formula.

In particular, if we assume that $S_0 = 90$, the prices according to equations (2.5) and (2.6) are 13.289 and 5.751, i.e. respectively 34.09% higher and 41.96% lower than the Black-Scholes price obtained by considering a volatility of 0.3. □

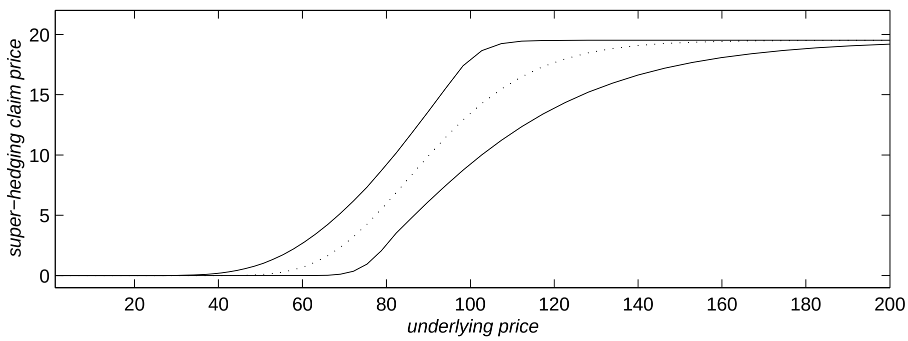


Figure 2.1: Seller's and buyer's prices according to equations (2.5) and (2.6) of a bullish call spread (solid line), compared with Black-Scholes prices for the same spread (dotted line).

2.2 Stochastic volatility with unknown process parameters

Let us assume that the stock price process is the solution of the following stochastic differential equation system:

$$\begin{aligned} dS_t &= \mu_s S_t dt + |v_t|^{1/2} S_t dw_{1,t} \ , \\ dv_t &= \mu_v(v_t;\theta)dt + \eta(v_t;\theta)[\rho dw_{1,t} + \sqrt{1-\rho^2}dw_{2,t}] \ , \end{aligned} \tag{2.7}$$

defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ where the filtration $\{\mathcal{F}\}_{t\in[0,T]}$ is generated by the two dimensional $\mathbb{P}$-Brownian motion $w_t = (w_{1,t}, w_{2,t})$. Moreover, $\theta \in \mathbb{R}^n$ is the set of process parameters whereas

[3]See Press et al. [54]. In Wilmott et al. [63] there is an application to the Black-Scholes model.

$\mu_v(v_t;\theta)$ and $\eta(v_t;\theta)$ are continuous functions of v_t. All coefficients defined are supposed to be adapted to $\{\mathcal{F}_t\}$.

Let us define two adapted and suitably regular[4] processes $\{\lambda_{1,t}\}_{t\in[0,T]}$ and $\{\lambda_{2,t}\}_{t\in[0,T]}$ such that

$$w^*_{j,t} = w_{j,t} + \int_0^t \lambda_{j,u}\, du \qquad\qquad j = 1, 2$$

and $w^*_t = (w^*_{1,t}, w^*_{2,t})$ is a $\mathbb{Q}$-Brownian motion. As usual, the new probability measure $\mathbb{Q}$ is given by the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left[-\int_0^T \lambda'_u\, dw_u - \frac{1}{2}\int_0^T \|\lambda_u\|^2\, du\right] \tag{2.8}$$

where $\lambda_t = [\lambda_{1,t}\ \lambda_{2,t}]'$. Under the new measure $\mathbb{Q}$, the system of stochastic differential equations (2.7) can be written as

$$\begin{array}{rcl} dS_t &=& rS_t dt + |v_t|^{1/2} S_t dw^*_{1,t}\ , \\ dv_t &=& \mu^{\mathbb{Q}}_v(v_t;\theta)dt + \eta(v_t;\theta)[\rho dw^*_{1,t} + \sqrt{1-\rho^2}dw^*_{2,t}]\ . \end{array} \tag{2.9}$$

where

$$\mu^{\mathbb{Q}}_v(v_t;\theta) = \mu_v(v_t;\theta) - \eta(v_t;\theta)(\rho\lambda_{1,t} + \sqrt{1-\rho^2}\lambda_{2,t})\ .$$

This implies that the discounted price process S_t/B_t is a positive local martingale under $\mathbb{Q}$. For a given value of θ, using standard arbitrage arguments we can state that in the above economy, the price of the contingent claim whose payoff in T is $\varphi(S_T)$ is the solution of the following PDE:

$$\begin{array}{l} \dfrac{\partial W}{\partial t} + \dfrac{\partial W}{\partial S}\ S_t r + \dfrac{\partial W}{\partial v}\ \mu^{\mathbb{Q}}_v(v_t;\theta) + \dfrac{1}{2}\dfrac{\partial^2 W}{\partial S^2}\ v_t S_t^2 + \\ + \dfrac{1}{2}\dfrac{\partial^2 W}{\partial v^2}\ \eta^2(v_t;\theta) + \dfrac{\partial^2 W}{\partial S\partial v}\ |v_t|^{1/2} S_t \eta(v_t;\theta)\rho - rW_t = 0\ . \end{array} \tag{2.10}$$

While the specification of $\lambda_{1,t}$ can be defined as $(\mu - r)\ v_t^{-1/2}$, to define $\lambda_{2,t}$ it is necessary to introduce some equilibrium considerations. Indeed, there is an infinite number of processes $\lambda_{2,t}$ such to produce a claim price that ensures no arbitrage opportunities.

Let us assume that the process $\lambda_{2,t}$ is a function of S_t, v_t and θ only[5]. Once defined the process $\lambda_{2,t}$, it is possible to define a pricing function for the contingent claim. However, a problem still remains. Indeed, the volatility process parameters θ are not known and they have to be estimated. Also if one is sure about the parametric specification of the volatility distribution, estimation risk remains.

[4] At least, the two processes have to satisfy Novikov condition.

[5] Note that this assumption implies that we are working in a Markovian setting.

In order to address the problem, we consider the parameter values unknown and we assume that for every parameter the following condition is true:

$$\theta_{i,min} \leq \theta_{i,t} \leq \theta_{i,max} \qquad i = 1, 2, \ldots, n \qquad \forall t \in [0, T] \tag{2.11}$$

Since the set of volatility parameters is no longer unique, equation (2.8) cannot uniquely define the probability measure $\mathbb{Q}$. Let us define $\mathcal{Q}_\theta$ the set of all probability measure consistent with condition (2.11).

Remark 2.2.1 Note that, if the process v_t satisfies condition (2.2) then $\mathcal{Q}_\theta \subseteq \mathcal{Q}$. In Reference 2.1.1 we show that under condition (2.2) the set $\mathcal{Q}_\theta$ is a set of martingale measure for the discounted price process S_t/B_t. The same result can be obtained by assuming a non-positive correlation between w_1 and w_2. To prove the last statement and for a discussion on the subject see Sin [59]. □

Assumption 1 The process v_t satisfies the Novikov condition.

Proposition 2.2.1 *If the price dynamics is described by* (2.9) *and assumption 1 is true, then the solution of the optimization problem* (2.3)*, where* $\mathbb{Q} \in \mathcal{Q}_\theta$*, is given by the following PDE*

$$\begin{aligned} &\frac{\partial W^+}{\partial t} + \frac{\partial W^+}{\partial S} r S_t + \frac{\partial W^+}{\partial v} \varphi_1 + \frac{1}{2}\frac{\partial^2 W^+}{\partial S^2} v_t S_t^2 + \\ &+ \frac{1}{2}\frac{\partial^2 W^+}{\partial v^2}\varphi_2 + \frac{\partial^2 W^+}{\partial S \partial v}|v_t|^{1/2} S_t \varphi_3 - rW^+ = 0 \end{aligned} \tag{2.12}$$

where

$$\varphi_1 := \begin{cases} \max_\theta \mu_v^{\mathbb{Q}}(v_t;\theta) & \text{if } \frac{\partial W^+}{\partial v} \geq 0 \\ \min_\theta \mu_v^{\mathbb{Q}}(v_t;\theta) & \text{if } \frac{\partial W^+}{\partial v} < 0 \end{cases} \tag{2.13}$$

$$\varphi_2 := \begin{cases} \max_\theta \eta^2(v_t;\theta) & \text{if } \frac{\partial^2 W^+}{\partial v^2} \geq 0 \\ \min_\theta \eta^2(v_t;\theta) & \text{if } \frac{\partial^2 W^+}{\partial v^2} < 0 \end{cases} \tag{2.14}$$

$$\varphi_3 := \begin{cases} \max_\theta \rho\ \eta(v_t;\theta) & \text{if } \frac{\partial^2 W^+}{\partial S \partial v} \geq 0 \\ \min_\theta \rho\ \eta(v_t;\theta) & \text{if } \frac{\partial^2 W^+}{\partial S \partial v} < 0 \end{cases} \tag{2.15}$$

The solution of the optimization problem (2.4) *is given by the same PDE with reversed inequalities* φ_1*,* φ_2 *and* φ_3*.*

Let us remark that φ_1, φ_2 and φ_3 are functions whose value at time $s \in [t, T]$ depends respectively on the sign of the first derivative with respect to instantaneous variance, the second derivative with respect to instantaneous variance and the cross derivative at time s.

2.3 An application to a specific model

The Heston model is characterized by the following definitions:

$$\mu_v(v_t;\theta) := \kappa[\bar{v} - v_t] \qquad\qquad \eta(v_t;\theta) := \delta\sqrt{v_t}\ ,$$

where $\kappa > 0$, $\bar{v} > 0$ and δ are parameters. In order to get v_t non negative a.s. for all t, one has to impose[6] $\delta^2 \le 2\kappa\bar{v}$.

By using the results of the general equilibrium model of Cox, Ingersoll and Ross [15], Heston finds a volatility risk premium proportional to v_t and such that the process $\lambda_{2,t}$ satisfies the following equation

$$\delta\left[\rho(\mu - r) + \sqrt{v_t}\sqrt{1-\rho^2}\lambda_{2,t}\right] = \lambda v_t\ , \tag{2.16}$$

where λ is a constant. The risk-adjusted process in the Heston model can then be written as

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{v_t}S_t dw^*_{1,t}\ , \\ dv_t &= [\zeta - \xi v_t]dt + \delta\sqrt{v_t}[\rho\ dw^*_{1,t} + \sqrt{1-\rho^2}\ dw^*_{2,t}]\ , \end{aligned} \tag{2.17}$$

where $\xi := \kappa + \lambda$, $\zeta := \kappa\bar{v}$ and $\theta := [\zeta,\ \xi,\ \delta,\ \rho]$ is the set of model parameters. With the risk-adjusted parameters, the restriction imposed to get v_t non negative a.s. becomes $\delta^2 \le 2\zeta$. Moreover, the functions φ_1 and φ_2 become

$$\varphi_1 := \begin{cases} \zeta_{max} - \xi_{min} v_t & \text{if } \frac{\partial W^+}{\partial v} \ge 0 \\ \zeta_{min} - \xi_{max} v_t & \text{if } \frac{\partial W^+}{\partial v} < 0 \end{cases}$$

$$\varphi_2 := \begin{cases} \delta^2_{max} v_t & \text{if } \frac{\partial^2 W^+}{\partial v^2} \ge 0 \\ \delta^2_{min} v_t & \text{if } \frac{\partial^2 W^+}{\partial v^2} < 0 \end{cases}$$

while φ_3 becomes

$$\text{for } \rho_{min} \le \rho_{max} \le 0 \qquad \varphi_3 := \begin{cases} \rho_{max}\delta_{min}\sqrt{v_t} & \text{if } \frac{\partial^2 W^+}{\partial S\partial v} \ge 0 \\ \rho_{min}\delta_{max}\sqrt{v_t} & \text{if } \frac{\partial^2 W^+}{\partial S\partial v} < 0 \end{cases}$$

$$\text{for } \rho_{min} \le 0 \le \rho_{max} \qquad \varphi_3 := \begin{cases} \rho_{max}\delta_{max}\sqrt{v_t} & \text{if } \frac{\partial^2 W^+}{\partial S\partial v} \ge 0 \\ \rho_{min}\delta_{max}\sqrt{v_t} & \text{if } \frac{\partial^2 W^+}{\partial S\partial v} < 0 \end{cases}$$

$$\text{for } 0 \le \rho_{min} \le \rho_{max} \qquad \varphi_3 := \begin{cases} \rho_{max}\delta_{max}\sqrt{v_t} & \text{if } \frac{\partial^2 W^+}{\partial S\partial v} \ge 0 \\ \rho_{min}\delta_{min}\sqrt{v_t} & \text{if } \frac{\partial^2 W^+}{\partial S\partial v} < 0 \end{cases}$$

Example 2.3.1 Let us consider again the data of the Example 2.1.1. Assume that the parameters can fluctuate inside the following intervals: $\zeta \in [0.01, 0.25]$, $\xi \in [0.5, 2]$, $\delta \in [0.01, 0.10]$ and $\rho \in [-0.5, 0]$. The instantaneous

[6] See Cox, Ingersoll and Ross [14] for more details.

variance v_t is 0.13. The above intervals are consistent with the EMM estimation of the Heston model done on the S&P 500 option prices by Chernov and Ghysels [12]. The restriction $\delta^2 \leq 2\zeta$ is respected.

The super-hedging prices according to equation (2.12) are represented in Figures 2.2 and 2.3. In plot (a) of both figures, seller's and buyer's prices are computed by letting all the parameters fluctuate within the defined bounds. In the two figures, we perform a sort of sensitivity analysis to show what parameters mostly influence the spread price. In all the cases, seller's and buyer's prices (solid line) are compared with Heston prices (dotted line) calculated by assuming that the value of the parameters is the mean of the two bounds.

Figure 2.4 shows the differences between super-hedging prices and Heston prices. In order to remove the dependence on the claim moneyness, the differences were divided by the underlying price. It is clear that the differences are higher when one of the two options of the spread is at-the-money. This shows that at-the-money options are more sensitive to parameter misspecifications.

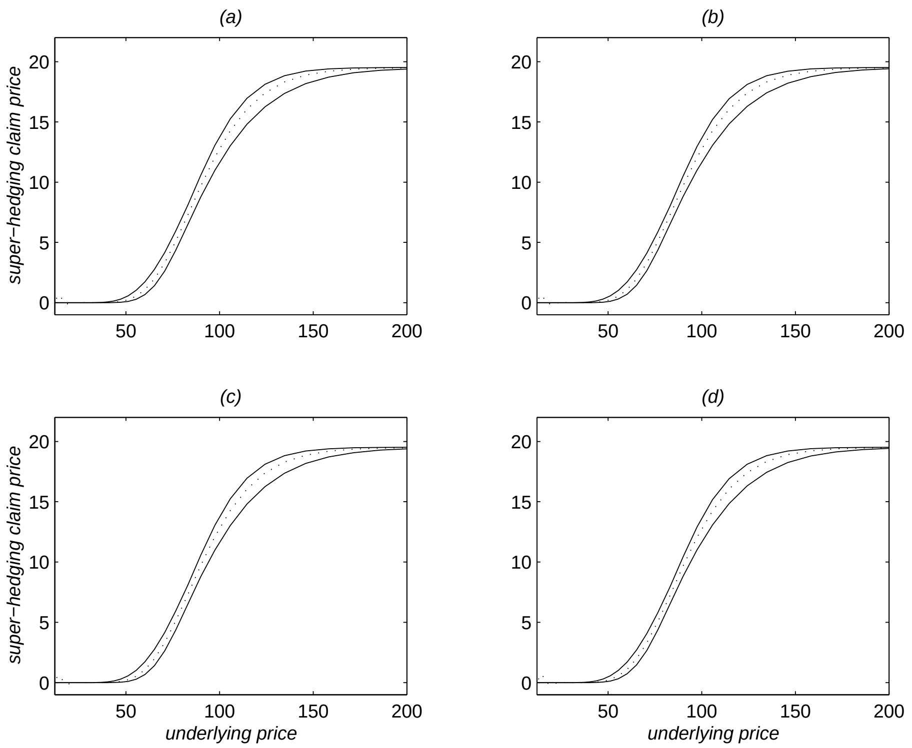
Figure 2.5 presents the differences between prices obtained by equation (2.12) and those obtained by equation (2.5), i.e. the result of the ALP approach. In order to solve equation (2.5) the two bounds v_{max} and v_{min} are fixed such that $\mathbb{Q}(v_t \geq v_{max}) = \mathbb{Q}(v_t \leq v_{min}) = 0.05\%$. The two values are $v_{max} = 0.437$ and $v_{min} = 0.259$.

The solid line represents the differences between seller's prices obtained by equation (2.12) and seller's prices according to (2.5). Since the difference is always negative, we see that using equation (2.12) seller's prices are lower than the ones obtained by the ALP approach. The opposite result is obtained for buyer's prices where the differences are always positive. The result is that the no-arbitrage pricing interval is reduces. □

In the Heston model the optimization problems (2.13), (2.14) and (2.15) can be solved in a trivial way. In spite of this, we cannot have a closed form solution for equation (2.12) even for plain vanilla options. Indeed, in this case one is sure only about the sign of $\frac{\partial W^+}{\partial v}$.

However, the function φ_1 depends on the drift parameters only, while the functions φ_2 and φ_3 depend on the diffusion parameters only. In Example 2.3.1, we showed that the misspecification of the diffusion parameters can scarcely affect the super-hedging price. By assuming no-misspecification on δ and ρ the non-linearity of the equation (2.12) depends only on φ_1.

For plain vanilla options, the first derivative with respect to v_t is not negative and the Heston formula can be used. In this case, the parameter ζ has to be set equal to ζ_{max} for a short option position and to ζ_{min} for a long option position. The parameter ξ has to be set equal to ξ_{min} for a short option position and to ξ_{max} for a long option position.

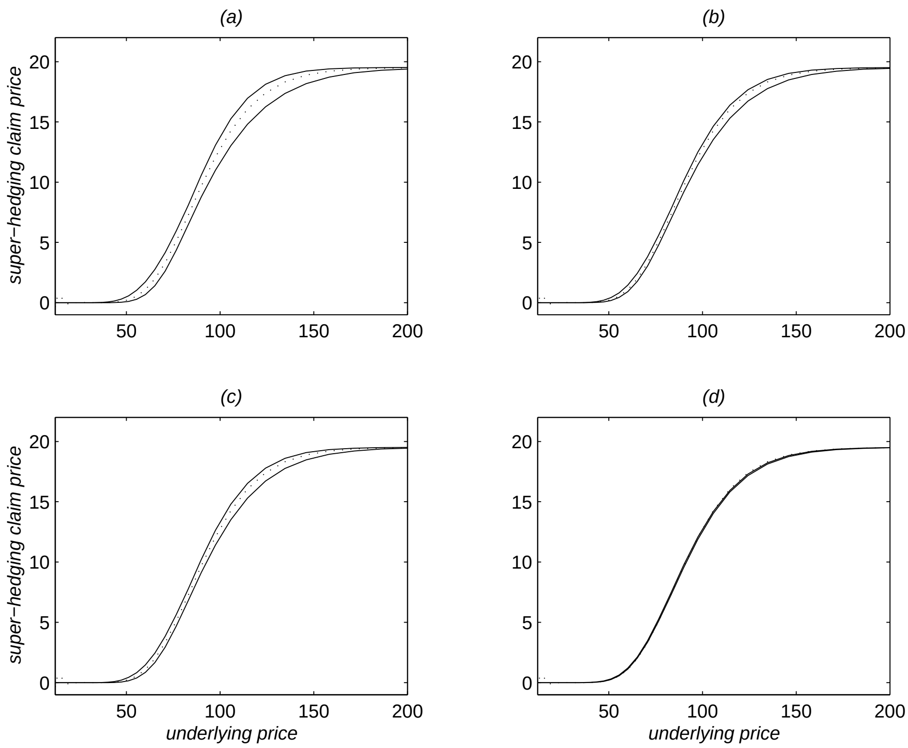


	ζ	ξ	δ	ρ
plot (a)	[0.01 0.25]	[0.5 2]	[0.01 0.1]	[−0.5 0]
plot (b)	[0.01 0.25]	[0.5 2]	0.05	[−0.5 0]
plot (c)	[0.01 0.25]	[0.5 2]	[0.01 0.1]	−0.5
plot (d)	[0.01 0.25]	[0.5 2]	[0.01 0.1]	0

Figure 2.2: Seller's and buyer's prices according to equation (2.12) of a bullish call spread (solid line) compared with Heston's prices of the same spread (dotted line).

Example 2.3.2 Let us consider a put option with a strike price 100 and maturity 6 months. Moreover, $r = 0.05$ and $v_t = 0.13$. The parameters can fluctuate inside the following intervals: $\zeta \in [0.01, 0.25]$, $\xi \in [0.5, 2]$, $\delta \in [0.01, 0.10]$ and $\rho \in [-0.5, 0]$. By assuming no-misspecification of the parameters δ and ρ, we can use the Heston formula to price the put by fixing $\zeta = 0.25$ and $\xi = 0.5$ for seller's price and $\zeta = 0.01$ and $\xi = 2$ for buyer's price.

In Figure 2.6, the prices of the put for different values of the underlying are plotted. In plot (a), seller's and buyer's prices are computed according to equation (2.12), i.e. by considering also δ and ρ misspecified. In plot (b), seller's and buyer's prices are computed according to equation (2.12) by



	ζ	ξ	δ	ρ
plot (a)	[0.01 0.25]	[0.5 2]	[0.01 0.1]	$[-0.5\ 0]$
plot (b)	0.13	[0.5 2]	[0.01 0.1]	$[-0.5\ 0]$
plot (c)	[0.01 0.25]	1.25	[0.01 0.1]	$[-0.5\ 0]$
plot (d)	0.13	1.25	[0.01 0.1]	$[-0.5\ 0]$

Figure 2.3: Seller's and buyer's prices according to equation (2.12) of a bullish call spread (solid line) compared with Heston's prices of the same spread (dotted line).

fixing $\delta = 0.05$ and $\rho = -0.25$. The same result is obtained in plot (c) by using the Heston formula where for seller's price we fix $\zeta = 0.25$, $\xi = 0.5$, $\delta = 0.05$ and $\rho = -0.25$ while for buyer's price $\zeta = 0.01$, $\xi = 2$, $\delta = 0.05$ and $\rho = -0.25$.

In plot (d) of Figure 2.6 we present the differences between the prices in plot (a) and the prices in plot (c). The solid line represents the seller's price differences while the dotted line represents the buyer's price differences. In order to remove the dependence of the differences on the option moneyness the differences are divided by the underlying price.

In Table 2.1 there is a comparison between the size of the no-arbitrage pricing interval according to equation (2.12) and the size of the no-arbitrage

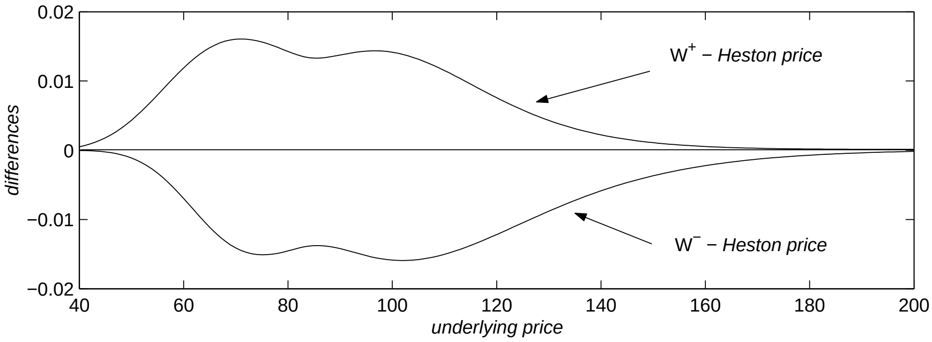


Figure 2.4: Differences between super-hedging price and Heston's price. In order to remove the dependence on the option moneyness the differences are divided by the underlying price.

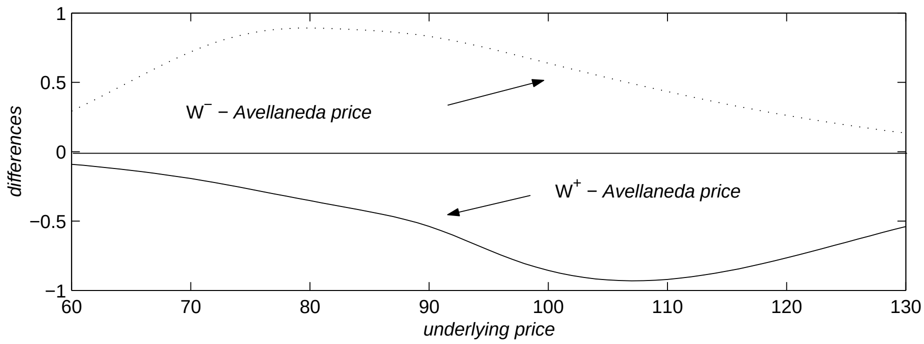


Figure 2.5: Differences between the prices obtained by equation (2.12) and prices according to Avellaneda, Levy and Paràs. Solid line represents the differences between seller's prices whereas dotted line represents the differences between buyer's prices

pricing interval according to the ALP approach. In the last two lines of the same table, we show the differences between the claim prices according to equation (2.12) and the claim prices according to the ALP approach. Note that the former was obtained by using the Black-Scholes formula (see remark 2.1.2).

The values reported refer to three levels of moneyness. In all three cases, equation (2.12) allows a remarkable reduction of the no-arbitrage pricing interval. □

Table 2.1: Comparison between prices according to equation (2.12) and prices according to Avellaneda, Levy and Paràs for three different levels of moneyness.

	moneyness					
	0.851		1		1.175	
$W^+ - W^-$	2.784		3.456		2.871	
ASP − ABP	3.857		4.894		3.942	
W^+- ASP	-0.308	(-1.69%)	-0.354	(-3.35%)	-0.277	(-5.28%)
W^-- ABP	0.765	(4.97%)	1.083	(15.21%)	0.794	(33.49%)

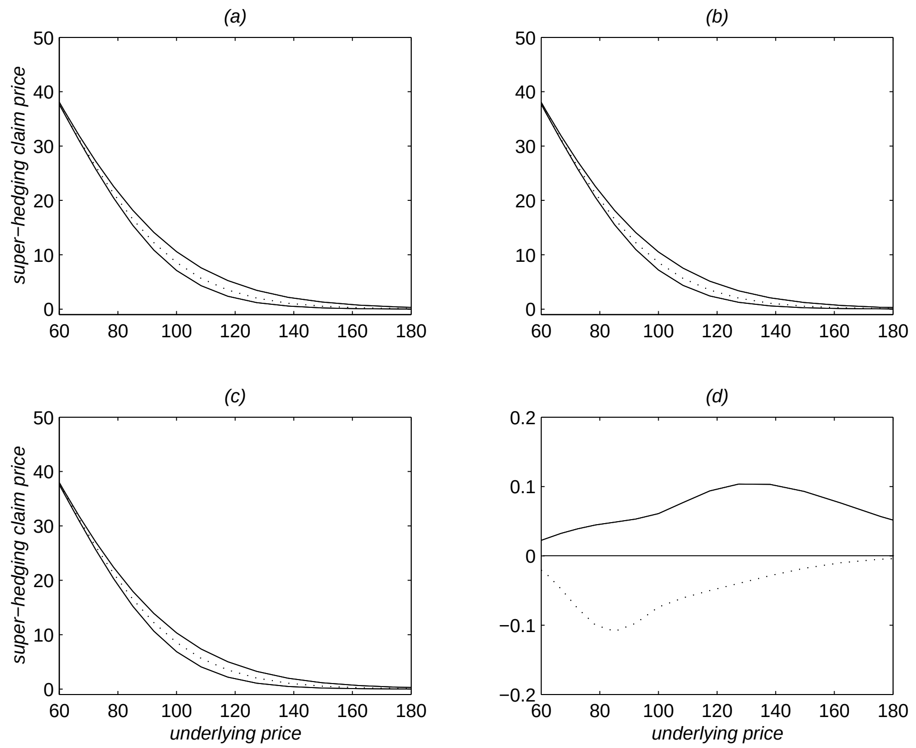
ASP is the seller's price of Avellaneda ($\sqrt{v_t} = 0.437$).
ABP is the buyer's price of Avellaneda ($\sqrt{v_t} = 0.259$).

2.4 Conclusion

In this work we try to address the problem of parameter misspecification in stochastic volatility models. To this end, we use the approach proposed by Avellaneda, Levy and Parás [5] and we move from a framework with uncertain volatility to uncertainty on the volatility process parameters.

We consider a generic class of stochastic volatility models that depends on a set of parameters. We assume that parameter values are unknown but limited between two bounds and we find a PDE whose solutions represent seller's and buyer's prices of a European contingent claim. A numerical application shows that seller's prices and buyer's prices are, respectively significantly lower and higher than those obtained by the ALP approach. Since the super-hedging bounds under stochastic volatility get closer to each other, they can be conveniently used to define a bid-ask spread.

In this framework, it is not possible to get a closed form solution either for plain vanilla options. However, we show that, at least in the case examined, seller's and buyer's prices of a plain vanilla option can be properly approximated by the Heston formula with a slight change of parameter. The same approximation can be extended to all the options whose first derivative with respect to v_t is always positive or always negative for all the possible values of S_t and v_t. This feature allows a remarkable simplification for pricing and hedging this kind of options.



	ζ	ξ	δ	ρ
plot (a)	$[0.01\ 0.25]$	$[0.5\ 2]$	$[0.01\ 0.1]$	$[-0.5\ 0]$
plot (b)	$[0.01\ 0.25]$	$[0.5\ 2]$	0.05	-0.25
plot (c)	Heston formula with $\theta = [0.25\ 0.5\ 0.13\ -0.25]$			
	Heston formula with $\theta = [0.01\ 2\ 0.13\ -0.25]$			
plot (d)	Differences between prices in (a) and prices in (c)			

Figure 2.6: Seller's and buyer's prices of a put option (solid line) compared with Heston prices of the same put (dotted line). In plot (d) there are the differences between seller's prices in plot (a) and seller's prices in plot (c) (solid line) and between buyer's prices in plot (a) and buyer's prices in plot (c) (dotted line).

Chapter 3

A consistent stochastic volatility model estimation

Option pricing has become one of the most studied features both by academics and by practitioners. During the last two decades, the Black-Scholes pricing model has become the benchmark for academics and the standard for practitioners. However, from the academic point of view, the Black-Scholes model cannot be considered consistent with observed data. Indeed, the presence of high kurtosis in log-return distributions, together with the evidence of implied volatility "smiles" and "term structures", remarkably contradicts the model hypotheses.

In spite of its limits, the Black-Scholes model is almost the only option pricing model used by practitioners. Some classes of alternative pricing models have been proposed. In the best known one, some of the Black-Scholes limits have been overcome by assuming a stochastic process for volatility.

However, by using a stochastic volatility model, pricing problems are only shifted or reduced, not removed completely. One of the most important problems, at least from the point of view of practitioners, is related to parameter estimation. Indeed, whereas the Black-Scholes model has only one unobservable parameter, stochastic volatility models are based on one (or more) latent variable and a set of unobservable parameters.

Moreover, in a stochastic volatility framework the distinction between objective probability $\mathbb{P}$ and risk adjusted probability $\mathbb{Q}$ produces a non-trivial effect on parameter estimation. Indeed, since markets are not complete, the risk premium due to stochastic volatility does not disappear in the pricing framework and it has to be estimated. This implies a distinct estimation of objective parameters (i.e. parameters estimated under $\mathbb{P}$) and risk-adjusted parameters (i.e. parameters estimated under $\mathbb{Q}$).

Under $\mathbb{Q}$, one of the most diffused methods to find some values for model parameters is to calibrate the model to the cross-section of option prices. Bakshi et al. [6] use the cross-section of options prices to calibrate the model completely disregarding the underlying return time series. Indeed, they get the parameters by minimizing the sum of the squared differences between model prices and market prices of the options.

Pure model calibration does not generally allow us to perform a confidence interval estimation for parameter values. This problem can be removed by using an estimation procedure. Indeed, the estimators distribution, at least the asymptotic one, is generally known and a confidence interval can be estimated. On the other hand, the continuous time framework of a stochastic volatility model introduces some serious problems on the estimation procedure. Standard estimation techniques, such as maximum likelihood estimation or GMM, become computationally very intensive and, in many cases, unfeasible. Moreover, there is a discretization bias that has to be taken into account.

In this framework, the indirect inference estimation method introduced by Gourieroux et al. [34] is appealing for two main reasons. First, it is able to estimate consistently a stochastic volatility model by removing the discretization bias. Second, it is always feasible or, in any case, the conditions for the implementation are typically weak.

Based on the same idea of indirect inference, the Efficient Method of Moments (hereafter EMM) is a widespread estimation method for stochastic volatility models. Introduced by Gallant and Tauchen [31], it has the same merits of indirect inference with the advantage that, in most applications, it requires less computational time.

Indirect inference and EMM estimators can be considered the two most widely used estimation methods for stochastic volatility models. They can be equally used to estimate objective parameters and risk adjusted parameters. Hence, input data can be both underlying prices and option prices. The main literature uses option data and underlying time series in different manners. For instance, Fiorentini et al. [26] use jointly the underlying return time series and the cross-section of option prices to estimate the Heston [40] model. In particular, they separately estimate the parameters under $\mathbb{P}$ (by indirect inference) and the price of volatility risk (by a calibration similar to that proposed by Bakshi et al. [6]).

With the same aim, Chernov and Ghysels [12] use underlying returns and option prices simultaneously. By using at-the-money call option prices and underlying returns, they estimate the objective and the risk adjusted densities by EMM. One of the main conclusions drawn in Chernov and Ghysels concerns hedging performances. From the hedging point of view, they show that parameters estimated by using only option prices "dominate" those estimated by using both option prices and underlying prices. Hence, the use of underlying prices does not improve hedging performances.

Especially for pricing and hedging options, an accurate parameter estimate is probably the most remarkable issue. Indeed, biases and large estimation standard errors can lead to wrong results both in pricing and hedging options. For these reasons, the main goal of this chapter is to test empirically the properties of indirect inference estimator and EMM estimator. According to this aim, the main focus will be on parameter estimation under $\mathbb{Q}$, i.e. by using option price information. Estimation according to underlying prices will be considered only as a term of comparison.

The stochastic volatility models considered are the Heston model and a model proposed by Lewis [48] (hereafter 3/2 model). The same author proposed a perturbation approach on the diffusion parameter to get an approximated option price under a class of stochastic volatility models that includes also the Heston model.

Under $\mathbb{Q}$, two estimation methods are proposed: the first one is based on an indirect inference and the second on an EMM estimator. Option prices are not used directly. For estimation purposes Black-Scholes implied volatilities (hereafter BSIV) are used. Since the process that governs the BSIV dynamics is unknown, it is convenient to approximate it by an Ornstein-Uhlenbeck process. The Ornstein-Uhlenbeck process admits an exact discretization and it can be suitably taken as auxiliary model.

Estimation of objective parameters is performed by means of a GARCH auxiliary model. This is a standard choice that seems to give good results also when sample size is not too large[1]. Since there is no closed-form expression for the auxiliary model estimator, only the EMM estimation method will be applied. Indeed, without closed form for the auxiliary parameter estimation, indirect inference is computationally very intensive. More details on the issue may be found in Section 3.1.

For both objective parameters and risk adjusted parameters, a Monte Carlo study and a real data estimation are performed. The Monte Carlo study is mainly devoted to three aims. The first one is to show whether estimation bias is completely removed for all the parameters. The second is to show what the estimation variance for each parameter is. The final goal is to verify whether the estimator's asymptotic distribution is a good approximation of the finite sample counterpart.

The real data application is based on SMI index returns and on the SMI volatility index (hereafter VSMI), respectively, for objective parameters and risk adjusted parameters. VSMI is a public volatility index obtained by at-the-money options with time to maturity of about 45 days.

Sections 3.1 and 3.2 briefly review the indirect inference and the EMM methodologies. The asymptotic properties of the two estimators are described in Section 3.3, while Section 3.4 presents an application to a class of stochastic volatility models. This application is the central topic of the work and it is divided into two parts: the objective parameter estimation (Section 3.4.1) and the risk adjusted parameter estimation (Section 3.4.2).

The results of the Monte Carlo study are presented in Section 3.5, while

[1] For more details see Andersen et al. [3].

in Section 3.6 an application to the Swiss market is described. Section 3.7 concludes.

3.1 Indirect Inference

Let us consider a process $\{S_t\}_{t\in\mathbb{N}}$, with strictly stationary increments $\{y_t\}$, generated by the dynamics:

$$\begin{aligned} S_t &= \varphi_1(S_{t-1}, \varepsilon_t, \theta) \\ u_t &= \varphi_2(u_{t-1}, \varepsilon_t, \theta) \end{aligned} \tag{3.1}$$

where $\theta \in \Theta \subset \mathbb{R}^m$ is the parameter vector, $\varphi_1(.)$ and $\varphi_2(.)$ are parametric functions, u_t is a latent variable and the p-dimension process $\{\varepsilon_t\}$ is iid distributed with mean 0_p, variance I_p and known distribution.

For a given value of the parameters θ, model (3.1) defines a probability measure[2] $\mathbb{P}_\theta$. The class of probability measures generated by every admissible $\theta \in \Theta$ is denoted by $\mathcal{P} := \{\mathbb{P}_\theta : \theta \in \Theta \subset \mathbb{R}^m\}$.

Following the main indirect inference literature, $\mathbb{P}_\theta$ will be called *structural model* and θ *structural parameter* vector. Let us assume that $\mathbb{P}_\theta$ is correctly specified, i.e. that there exists at least a value $\theta_0 \in \Theta$ such that the probability measure $\mathbb{P}_{\theta_0}$ is the true probability measure of the process $\{y_t\}$.

The structural model is possibly based on a class of density functions $f(y_t;\theta)$ which do not admit a closed form expression. In this case, maximum likelihood estimators are not feasible and an alternative estimation method has to be applied.

In order to sidestep the unfeasibility of maximum likelihood, another model is introduced. This model is called *auxiliary model* and defines a further class of density functions $h(y_t;\beta)$, where $\beta \in \mathcal{B} \subset \mathbb{R}^\ell$ is the *auxiliary parameter* vector. In order to easily obtain a consistent estimator for the auxiliary parameters, the auxiliary model should be analytically tractable. Moreover, for the structural parameter identification, the number of auxiliary parameters cannot be less than the number of the structural parameters, i.e. $\ell \geq m$.

Let us denote the observations of the process $\{y_t\}$ as $y^T := (y_1, y_2, \ldots, y_T)$. Moreover, let Q_T be a criterion function which depends on the observations y^T and on the auxiliary parameter vector β. The auxiliary parameter estimator $\hat{\beta}_T$ is defined as

$$\hat{\beta}_T = \arg\max_{\beta\in\mathcal{B}} Q_T(y^T, \beta) \ . \tag{3.2}$$

[2]The distinction between objective and risk-adjusted probability measure will be introduced in Section 3.4. Here $\mathbb{P}_\theta$ is a generic probability measure

Let us assume that Q_T converges a.s. to a non-stochastic limit function Q_∞

$$\lim_{T\to\infty} Q_T\left(y^T,\beta\right) = Q_\infty(\theta_0,\beta) \ .$$

which is continuous in β and has a unique maximum $\beta_0 = \arg\max_{\beta\in\mathcal{B}} Q_\infty(\theta_0,\beta)$. Under this assumptions, $\hat\beta_T$ is a consistent estimator of β_0.

The limit function Q_∞ can be defined for a generic value of θ. In this case the binding function $b(\theta) : \Theta \to \mathcal{B}$ is introduced:

$$b(\theta) = \arg\max_{\beta\in\mathcal{B}} Q_\infty(\theta,\beta) \ .$$

If $b(.)$ was known and one to one, the estimator $\tilde\theta_T = b^{-1}(\hat\beta_T)$ would be a consistent estimator of θ_0. The problem is that $b^{-1}(\hat\beta_T)$ has generally no-closed form. Hence, the idea in indirect inference estimation is to substitute it with an estimator $\hat b^{-1}(\hat\beta_T)$ based on the simulations of the process $\{y_t\}$.

Gourieroux, Monfort and Renault [34] propose to simulate a path of length $\tau > T$ from model (3.1) based on a given value of θ, say $\tilde\theta$. In the sequel, $y^\tau(\tilde\theta)$ denotes the simulated path $(y_1(\tilde\theta), y_2(\tilde\theta), \ldots, y_\tau(\tilde\theta))$.

For given $\tilde\theta$, the estimator of the binding function is given by:

$$\hat\beta_\tau(\tilde\theta) = \arg\max_{\beta\in\mathcal{B}} Q_\tau\left[y^\tau(\tilde\theta),\beta\right] \ . \tag{3.3}$$

To estimate the structural parameter vector, the basic idea consists in minimizing the norm of the difference $b(\tilde\theta) - \beta$. In practice, one has to find the value of $\tilde\theta$ that makes $\hat\beta_\tau(\tilde\theta)$ as close as possible to $\hat\beta_T$. Hence, the indirect inference estimator is defined as

$$\hat\theta_T^{II} = \arg\min_{\tilde\theta\in\Theta} \left[\hat\beta_\tau(\tilde\theta) - \hat\beta_T\right]' \Omega_T \left[\hat\beta_\tau(\tilde\theta) - \hat\beta_T\right] \tag{3.4}$$

where Ω_T is a positive definite matrix that converges a.s. to some positive definite matrix Ω for $T \to \infty$.

Remark 3.1.1 The path $y^\tau(\tilde\theta)$ depends both on $\tilde\theta$ and on the simulated path of the process $\{\varepsilon_t\}$, say $\tilde\varepsilon^\tau$. Note that, in order to ensure estimator consistency, the simulated path $\tilde\varepsilon^\tau$ has to be the same for all the steps of the optimization algorithm. □

Optimization problem (3.3) involves a generic criterion Q_T that has to be maximized in order to estimate the binding function. It seems reasonable for this purpose to use the quasi log-likelihood function of the auxiliary model[3]:

$$\ln h^s(y^T;\beta) := \sum_{t=1}^{s} \ln h_t(y_t;\beta) \qquad \text{for } s = T, \tau$$

[3]Sometimes, the Kullback-Leibler Information Criterion (KLIC) is used. It can be shown that minimizing the average KLIC is equivalent to maximizing the quasi log-likelihood function of the auxiliary model.

where $h_t(y_t;\beta)$ is the conditional density function of the auxiliary model at time t. Hence, equation (3.2) and (3.3) become[4]

$$\begin{aligned}\hat{\beta}_T &= \arg\max_{\beta\in\mathcal{B}} Q_T(y^T,\beta) := \frac{1}{T}\arg\max_{\beta\in\mathcal{B}} \ln h^T(y^T;\beta) \ , && (3.5)\\ \hat{\beta}_\tau(\tilde{\theta}) &= \arg\max_{\beta\in\mathcal{B}} Q_\tau(\tilde{y}^\tau,\beta) := \frac{1}{\tau}\arg\max_{\beta\in\mathcal{B}} \ln h^\tau(\tilde{y}^\tau;\beta) \ . && (3.6)\end{aligned}$$

To summarize, the indirect inference estimation method involves the following steps:

1. By taking market data, auxiliary parameters are estimated (equation (3.5));
2. At each step of the optimization algorithm, a simulation from the model (3.1) is run by using some value for θ, say $\tilde{\theta}$;
3. By taking the simulated data, auxiliary parameters are estimated (equation (3.6));
4. The distance between $\hat{\beta}_\tau(\tilde{\theta})$ and $\hat{\beta}_T$ is computed. If it is not "sufficiently" small, steps 2 and 3 are performed again.

In order to solve optimization problem (3.4), a recursive algorithm has to be applied

$$\theta_{n+1} = \theta_n + \Gamma_n\left(\hat{\beta}_T, \hat{\beta}_\tau(\theta)\right)\Delta_n\left(\hat{\beta}_T, \hat{\beta}_\tau(\theta)\right) \ ,$$

where the function $\Gamma_n(.)$ defines the step size and the function $\Delta_n(.)$ defines the direction.

Since both step size and direction function depend on $\hat{\beta}_T$, at each step of the algorithm an estimation of the binding function has to be performed. Therefore, when problem (3.6) does not have a closed form solution, the indirect inference estimation method is computationally very intensive. To mitigate this problem, Gallant and Tauchen [31] proposed a second version of indirect inference estimation, called Efficient Method of Moments.

3.2 The Efficient Method of Moments

In order to solve the optimization problem (3.6), the first order condition can be defined as

$$\nabla_\beta E^{\mathbb{P}_\theta}\left[\ln\left(h_t(\tilde{y}_t;\beta)\right)\right] = 0 \ ,$$

[4]In order to simplify the notation, hereafter we will use $\tilde{y}$ to mean $y(\tilde{\theta})$, such that $\tilde{y}^\tau := y^\tau(\tilde{\theta})$ and $\tilde{y}_t := y_t(\tilde{\theta})$.

where ∇_β is the gradient operator with respect to β. Interchanging integration and differentiation, the first order condition becomes

$$E^{\mathbb{P}_\theta}\left[\nabla_\beta \ln\left(h_t(\tilde{y}_t;\beta)\right)\right] = 0 \ .$$

The idea of Gallant and Tauchen is to select θ to fulfill, or at least to approximate, the above equality for $\beta = \beta_0$. In order to perform the estimation in practice, two issues have to be considered:

- β_0 is unknown;
- generally $\ell > m$, i.e., the number of the auxiliary parameters can be greater than the number of the structural parameters. In this case, the model is overidentified since $\dim \nabla_\beta > \dim\ \theta$.

For estimation purposes β_0 can be replaced by its quasi-maximum likelihood estimation based on real data:

$$\hat{\beta}_T = \arg\max_{\beta\in\mathcal{B}} \ln h^T(y^T;\beta) \ .$$

The EMM first order condition then becomes

$$m(\theta,\hat{\beta}_T) := E^{\mathbb{P}_\theta}\left[\nabla_\beta \ln\left(h_t(\tilde{y}_t;\hat{\beta}_T)\right)\right] = 0 \ . \tag{3.7}$$

In the overidentified case, condition (3.7) could not be satisfied. However, one can select the value of θ that minimizes the distance between $m(\theta,\hat{\beta}_T)$ and an ℓ-vector of zeros. Hence, the general form of the EMM estimator can be written as

$$\hat{\theta}_T^{EMM} = \arg\min_{\theta\in\Theta}\ m(\theta,\hat{\beta}_T)'\ \Sigma_T\ m(\theta,\hat{\beta}_T) \ , \tag{3.8}$$

where Σ_T is a positive definite matrix that converges a.s. to some positive definite matrix Σ for $T\to\infty$.

Remark 3.2.1 Gourieroux, Monfort and Renault [34] show that, for a given value of the matrix $\hat{\Sigma}_T$, the EMM estimator is asymptotically equivalent to the indirect inference estimator (3.4). Unlike indirect inference approach, EMM estimator does not require to solve a maximization problem at each step of the optimization algorithm. Indeed, it does not estimate the binding function $b(\theta)$ at each step. □

Summarizing, the EMM estimation method involves the following steps:

1. By taking market data, the auxiliary parameters are estimated (equation (3.5));
2. At each step of the optimization algorithm, a simulation from the model (3.1) is run by using some value for θ, say $\tilde{\theta}$;

3. The simulated path is used to approximate $m(\tilde{\theta}, \hat{\beta}_T)$ by its simulated sample analogous
$$m(\tilde{\theta}, \hat{\beta}_T) \approx \frac{1}{\tau} \sum_{t=1}^{\tau} \nabla_\beta \ln \left[h_t(y_t(\tilde{\theta}); \hat{\beta}_T) \right] \ ;$$
4. The value of $m(., \hat{\beta}_T)\Sigma_T m(., \hat{\beta}_T)$ is computed. If it is not "sufficiently" small, steps 2 and 3 are performed again.

3.3 Best estimator and asymptotic properties

The optimization problems (3.4) and (3.8) involve a minimization according to the metric associated with the scalar product defined by Ω_T and Σ_T, respectively. The choice of Ω_T and Σ_T has to be based on some efficiency principle.

To determine the best estimator, it is sufficient to recognize that both indirect inference and EMM estimators belong to the general class of Asymptotic Least Squares (hereafter ALS) estimators[5]. In order to review some asymptotic properties of quasi-maximum likelihood estimators, let us define the score of the quasi-likelihood function for the auxiliary model as

$$s(y^T; \beta) := T^{-1} \nabla_\beta \ln h^T(y^T; \beta) \ .$$

White [61] shows that

$$B_T^{-1/2} A_T \sqrt{T} (\hat{\beta} - \beta_0) \overset{a}{\sim} N(0, I_\ell) \ ,$$

where

$$\begin{aligned} A_T &:= E^{\mathbb{P}}[\nabla_\beta s(y^T; \beta_0)] \ , \\ B_T &:= \operatorname{var}[T^{1/2} s(y^T; \beta_0)] \ , \end{aligned}$$

and β_0 is the true value of the auxiliary parameters. Moreover, matrix A_T can be written as

$$\begin{aligned} A_T &= T^{-1} E^{\mathbb{P}} \left[\nabla_\beta^2 \ln h^L(y^L; \beta_0) \right] + T^{-1} \sum_{t=L+1}^{T} E^{\mathbb{P}} \left[\nabla_\beta^2 \ln h_t(y_t; \beta_0) \right] \\ &= T^{-1} E^{\mathbb{P}} \left[\nabla_\beta^2 \ln h^L(y^L; \beta_0) \right] + \frac{T-L}{T} E^{\mathbb{P}} \left[\nabla_\beta^2 \ln h_t(y_t; \beta_0) \right] \ , \end{aligned}$$

implying

$$\lim_{T \to \infty} A_T = E^{\mathbb{P}} \left[\nabla_\beta^2 \ln h_t(y_t; \beta_0) \right] := A \ .$$

[5] In particular, both the optimal matrixes Ω_T and Σ_T depend on the covariance matrix of the auxiliary parameter estimator.

Similarly, the limit of B_T can be written as

$$\lim_{T\to\infty} B_T = \mathrm{var}\left[\nabla_\beta \ln h_t(y_t;\beta_0)\right] := B \; .$$

Hence, the asymptotic covariance matrix of the auxiliary estimator $\hat{\beta}_T$ is $A^{-1}BA^{-1}$.

In order to get the best ALS estimator, some further notation has to be introduced. Let us denote by $M_\theta \in \mathbb{R}^{\ell\times m}$ and $M_\beta \in \mathbb{R}^{\ell\times\ell}$ the two following matrixes:

$$M_\theta = \frac{\partial}{\partial\theta'}\, g(\theta;\beta)\Big|_{\theta=\theta_0,\beta=\beta_0} \qquad M_\beta = \frac{\partial}{\partial\beta'}\, g(\theta;\beta)\Big|_{\theta=\theta_0,\beta=\beta_0} \; .$$

where $g(\theta;\beta) = b(\theta) - \beta$ for the indirect inference estimator and $g(\theta;\beta) = E^{\mathbb{P}_\theta}[\nabla_\beta \ln(h_t(y_t;\beta))]$ for EMM.

Gourieroux and Monfort [32] show that, in order to get the best ALS estimator, the sequence of matrixes $\{\Omega_T\}$ and $\{\Sigma_T\}$ has to converge to $(M_\beta A^{-1}BA^{-1}M_\beta')^{-1}$. Since for the indirect inference $M_\beta = -I$, the optimal matrixes can be then written as:

$$\begin{aligned} \Omega &:= \left(A^{-1}BA^{-1}\right)^{-1} \; , \\ \Sigma &:= \left(M_\beta A^{-1}BA^{-1}M_\beta'\right)^{-1} \; . \end{aligned}$$

Moreover, in the EMM setting, an interchange between integration and differentiation gives

$$M_\beta = E^{\mathbb{P}_\theta}\left[\nabla^2_\beta \ln h_t(y_t;\beta)\right]\Big|_{\theta=\theta_0,\beta=\beta_0} = A \; ,$$

implying $\Sigma = B^{-1}$.

Once defined the best ALS estimator, one can analyze its asymptotic properties. Under some regularity condition both indirect inference and EMM estimator are consistent. For the asymptotic distribution, Gourieroux and Monfort [33] show that both estimators are asymptotically normally distributed:

$$\sqrt{T}(\hat{\theta}_T^{II} - \theta_0) \;\overset{a}{\sim}\; N\left[0, \left(1+\frac{T}{\tau}\right)\left(M_\theta'\Omega M_\theta\right)^{-1}\right] \; , \tag{3.9}$$

$$\sqrt{T}(\hat{\theta}_T^{EMM} - \theta_0) \;\overset{a}{\sim}\; N\left[0, \left(1+\frac{T}{\tau}\right)\left(M_\theta'\Sigma M_\theta\right)^{-1}\right] \; . \tag{3.10}$$

This finding is similar to that obtained for ALS estimators. The difference is that here the asymptotic covariance matrix depends also on the ratio T/τ. The reason is that in the indirect inference and in the EMM setting, the binding function $b(\theta)$ has to be estimated by simulations. Hence, the variance of the asymptotic distribution has to depend also on the simulation error.

Note that, in practice, one has to replace β_0 by its quasi-maximum likelihood estimator. Moreover, by assuming that the quasi-score $\nabla_\beta \ln h_t(y_t; \hat{\beta})$ is a martingale difference sequence, a convenient estimator for B is obtained by the outer product of the quasi-score itself:

$$\hat{B}_T = \frac{1}{T}\sum_{t=1}^{T}\left[\nabla_\beta \ln h_t(y_t;\hat{\beta})\right]\left[\nabla_\beta \ln h_t(y_t;\hat{\beta})\right]' \ , \tag{3.11}$$

and A can be estimated as:

$$\hat{A}_T = \frac{1}{T}\sum_{t=1}^{T}\left[\nabla_\beta^2 \ln h_t(y_t;\hat{\beta})\right] \ ,$$

implying $\hat{\Omega}_T = \hat{A}_T\hat{B}_T^{-1}\hat{A}_T$ and $\hat{\Sigma}_T = \hat{B}_T^{-1}$.

From equations (3.9) and (3.10) a consistent estimator for the asymptotic covariance matrix of $\hat{\theta}_T^{II}$ and $\hat{\theta}_T^{EMM}$ is given by:

$$\mathrm{AsCov}(\hat{\theta}_T^{II}) = \left(\frac{1}{T}+\frac{1}{\tau}\right)\left[\frac{\partial\tilde{\beta}_\tau(\hat{\theta}_T^{II})}{\partial\theta'}\hat{A}_T\hat{B}_T^{-1}\hat{A}_T\frac{\partial\tilde{\beta}_\tau(\hat{\theta}_T^{II})}{\partial\theta}\right]^{-1} \ , \tag{3.12}$$

$$\mathrm{AsCov}(\hat{\theta}_T^{EMM}) = \left(\frac{1}{T}+\frac{1}{\tau}\right)\left[\frac{\partial m(\hat{\theta}_T^{EMM},\hat{\beta}_T)}{\partial\theta'}\hat{B}_T^{-1}\frac{\partial m(\hat{\theta}_T^{EMM},\hat{\beta}_T)'}{\partial\theta}\right]^{-1} \tag{3.13}$$

No explicit solution can be provided for the first derivative of $\tilde{\beta}_\tau(\hat{\theta}_T^{II})$ and $m(\hat{\theta}_T^{EMM},\hat{\beta}_T)$ with respect to θ, so that they have to be computed numerically.

3.4 The application to stochastic volatility models

In the continuous time framework of a stochastic volatility model, methods based on indirect inference can perform a consistent estimation of the model parameters. In this work two different stochastic volatility models are estimated. The first model is that proposed by Heston [40], which has an exact closed form solution for option prices. The second model was proposed by Lewis [48], who provides an approximated analytical solution based on a perturbation approach. Actually, Lewis proposes a general perturbation approach to find option price approximations for a class of stochastic volatility models that includes also the Heston model. More details on option pricing according to the perturbation approach are given in appendix B.2.

The choice of the two models is due to the closed form that they provide for option pricing. Indeed, in spite of the generality of the estimation approach proposed, in Section 3.4.2 we will see that the estimation under $\mathbb{Q}$ is unfeasible without a closed form expression for derivatives prices.

Under the real world measure $\mathbb{P}$, the dynamics for the price process $\{S_t\}_{t\in[0,T]}$ is determined by the following system of stochastic differential equations:

$$
\begin{aligned}
dS_t &= \mu S_t dt + \sqrt{v_t} S_t dw_{1,t} \ , \\
dv_t &= [\zeta - \xi v_t]dt + \delta v_t^{\gamma} dw_{2,t} \ ,
\end{aligned}
$$

where μ, ζ, ξ and δ are positive parameters and $w_{1,t}$ and $w_{2,t}$ are two $\mathbb{P}$-Brownian motions with correlation ρ. The value of γ is $1/2$ for the Heston model and $3/2$ for the model proposed by Lewis (hereafter 3/2 model).

By assuming a volatility risk premium proportional to v_t, say λv_t, one can state that the price process dynamics under $\mathbb{Q}$ is given by

$$
\begin{aligned}
dS_t &= r S_t dt + \sqrt{v_t} S_t d\tilde{w}_{1,t} \ , \\
dv_t &= [\zeta - \tilde{\xi} v_t]dt + \delta v_t^{\gamma} d\tilde{w}_{2,t} \ , \qquad (3.14)
\end{aligned}
$$

where r is the risk-free rate, $\tilde{\xi} = \xi + \lambda$ and $\tilde{w}_{1,t}$ and $\tilde{w}_{2,t}$ are two $\mathbb{Q}$-Brownian motions with correlation ρ.

While the real world process is estimated by using price return time series, for the estimation under $\mathbb{Q}$ we need something related to option prices. By considering Black-Scholes formula as a metric, we can use Black-Scholes implied volatility (hereafter BSIV) time series to estimate the risk-adjusted parameters.

In this section we will estimate both objective parameters and risk-adjusted parameters of the model. While for objective parameters we will use only EMM estimator, for risk-adjusted parameters, EMM and indirect inference estimators will be applied. Indeed, under $\mathbb{Q}$ the computational complexity of the indirect inference estimator is remarkably reduced by the characteristics of the auxiliary model.

In both the applications, we will assume a zero correlation between asset price and volatility ($\rho = 0$).

3.4.1 The auxiliary model for underlying asset returns

As previously noted, one has to be able to define a consistent estimator for the auxiliary parameters. For this reason, the auxiliary model has to be analytically tractable and it has to be defined in a discrete time framework. Two other features are requested for the auxiliary model. It cannot be too distant from the structural model (at least the number of the auxiliary parameters has to be larger or equal to the number of structural parameters) and, despite it is not the true market model, it has to be able to fit market data reasonably well.

Under $\mathbb{P}$, Andersen et al. [3] show that the choice of the auxiliary model has to be considered important for the estimation results. In particular, they show that for very large sample size, models which incorporate the semi-non parametric (SNP) Hermite polynomial expansion[6] perform very well but not remarkably better than a purely parametric GARCH(1,1) model. For smaller sample sizes there are some convergence problems because of the overparametrization of the SNP density. Indeed, for a more realistic sample size of 500 observations, standard Newton-type optimization methods are not always able to converge to a solution when the auxiliary model is based on a SNP density.

For this reason, we decided to take a GARCH(1,1) as the auxiliary model. Accordingly, we define the return dynamics as:

$$\begin{aligned} y_t &= \alpha + \sqrt{v_t}\,\varepsilon_t \\ v_t &= \beta_0 + \beta_1 v_{t-1} + \beta_2 (y_{t-1} - \alpha)^2 \end{aligned} \tag{3.15}$$

where $y_t := \ln S_t - \ln S_{t-1}$, S_t is the asset price at time t and ε_t is a standard normal disturbance. Conditions $\beta_0 > 0$, $\beta_1 \geq 0$ and $\beta_2 \geq 0$ are required to ensure a positive value of v_t. Moreover, for the stationarity of the volatility process, one has to impose the condition $\beta_1 + \beta_2 < 1$.

Let us denote the parameter vector $(\alpha\ \beta_0\ \beta_1\ \beta_2)'$ by β. To estimate β, the conditional quasi likelihood function that has to be maximized is

$$\ln h^T(y^T;\beta) = -\frac{T}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\left[\ln v_t(\beta) + \frac{e_t^2(\beta)}{v_t(\beta)}\right] ,$$

where $e_t(\beta) := y_t - \alpha$. In order to solve the expected value in equation (3.7) and to get the matrix $\hat{B}_T$, one has to compute the score of the auxiliary model. In our case it is given by:

$$\nabla_\beta \ln h_t(y_t;\beta) = \left[s_1 \frac{e_t}{v_t} + \frac{1}{2}\frac{1}{v_t}\left(\frac{e_t^2}{v_t} - 1\right)\frac{\partial v_t}{\partial \beta}\right] , \tag{3.16}$$

where s_1 is a selection vector whose first component is 1 and the others are 0. The derivative of v_t with respect to the vector parameter β can be defined in the usual recursive manner:

$$\frac{\partial v_t}{\partial \beta} = z_t + \beta_1 \frac{\partial v_{t-1}}{\partial \beta} ,$$

where

$$z_{t-1} := \begin{bmatrix} -2\beta_2 e_{t-1} & 1 & v_{t-1} & e_{t-1}^2 \end{bmatrix}'$$

Such a recursive definition needs an initial value for the conditional variance v_t. A common practice is to substitute the initial value of the conditional variance by the sample estimate of the unconditional variance. From this assumption follows that the first derivative of v_t with respect to β is zero.

[6] See Gallant and Tauchen [31].

3.4.2 The auxiliary model for BSIV

In order to estimate the risk-adjusted parameters, we need to introduce an auxiliary model for BSIV. A proper choice seems to be an Ornstein-Uhlenbeck process of the form[7]:

$$d\sigma_t = (\kappa_0 - \kappa_1 \sigma_t)dt + \kappa_2 dw_t \ , \tag{3.17}$$

where w_t is a $\mathbb{Q}$-Brownian motion. This model is consistent whit Heston's assumption on instantaneous variance dynamics. Indeed, by applying Ito's lemma to σ_t^2 we get a square root process similar to that used by Heston.

The Ornstein-Uhlenbeck processes admits as exact discretization an AR(1) process:

$$\begin{aligned} \sigma_t &= \frac{\kappa_0}{\kappa_1}\left(1 - e^{-\kappa_1 \Delta t}\right) + e^{-\kappa_1 \Delta t}\sigma_{t-\Delta t} + \kappa_2 \left(\frac{1 - e^{-2\kappa_1 \Delta t}}{2\kappa_1}\right)^{1/2} \varepsilon_t \\ &:= a_0 + a_1 \sigma_{t-\Delta t} + a_2 \varepsilon_t \ , \end{aligned}$$

where

$$\kappa_0 = -\tfrac{a_0 \ln a_1}{\Delta t(1-a_1)} \quad \kappa_1 = -\tfrac{1}{\Delta t} \ln a_1 \quad \kappa_2 = \tfrac{a_2}{\sqrt{\Delta t}} \left(-\tfrac{1-a_1^2}{2 \ln a_1}\right)^{-1/2} \ .$$

Note that parameter adjustment for the time interval Δt was done in order to take into account that the data for the estimation are available at each time interval Δt (generally one day), while the parameters to estimate are annualized.

The conditional log-likelihood of the AR(1) process and the score function can be written as

$$\ln h^T(\sigma^T; a) = c - \frac{T-1}{2}\ln(a_2^2) - \sum_{t=2}^{T}\left[\frac{(\sigma_t - a_0 - a_1\sigma_{t-\Delta t})^2}{2a_2^2}\right]$$

$$\nabla_a \ln h_t(\sigma_t; a) = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{a_2} \end{bmatrix} + \frac{1}{a_2^2}(\sigma_t - a_0 - a_1\sigma_{t-\Delta t}) \begin{bmatrix} 1 \\ \sigma_{t-\Delta t} \\ \frac{\sigma_t - a_0 - a_1 \sigma_{t-\Delta t}}{a_2} \end{bmatrix}$$

where c is a constant and $a := (a_0 \ a_1 \ a_2)'$. From the last equation, it is clear that the score is a martingale difference sequence under $\mathbb{Q}$. Hence, the outer product (see equation (3.11)) can be conveniently used to estimate the asymptotic covariance matrix of the sample quasi-score.

By maximizing the above conditional log-likelihood, we have an explicit expression for the estimators[8] of parameters a_0, a_1 and a_2. Hence, the computational difficulties inherent with the indirect inference estimator can be avoided here.

[7] A similar approach was applied in Pastorello et al. [53].

[8] See Hamilton [35] for more details.

Since the auxiliary model is based on BSIV time series, we have to simulate a path of implied volatilities (and not instantaneous variances). For this purpose, for each underlying price and each instantaneous variance, option prices are computed:

$$\tilde{W}_t := W_{SV}(S_t, K, T-t, r, v_t; \theta) \qquad \text{for } t = 1, 2, \ldots, \tau$$

and then inverted, to obtain a simulated BSIV series, say $\tilde{\sigma}_t$:

$$\tilde{\sigma}_t = W_{BS}^{-1}(S_t, K, T-t, r; \tilde{W}_t) \qquad \text{for } t = 1, 2, \ldots, \tau$$

where $W_{SV}(.)$ and $W_{BS}^{-1}(.)$ are, respectively, the stochastic volatility pricing function and the inverse with respect to volatility of the Black-Scholes pricing function.

Since the number of auxiliary parameters is equal to the number of structural parameters, for a sample size sufficiently large we have that $\hat{\theta}_T^{II}$ and $\hat{\theta}_T^{EMM}$ are equal and both are independent on the choice of the metric[9]. Hence, without loss of generality, Σ and Ω can be substituted by a three dimensional identity matrix.

However, matrixes Σ and Ω have to be computed in order to estimate the asymptotic covariance matrix of the estimator. For such a purpose, we apply equations (3.12) and (3.13).

3.5 Monte Carlo study

The aim of this Monte Carlo study is to verify the accuracy of the estimation methods proposed. For this purpose, we test three different estimation methods, one for objective parameters and two for risk-adjusted parameters:

1. EMM with GARCH(1,1) auxiliary model (objective parameters);
2. Indirect inference with Ornstein-Uhlenbeck auxiliary model (risk-adjusted parameters);
3. EMM with Ornstein-Uhlenbeck auxiliary model (risk-adjusted parameters);

All the estimation methods are applied to the Heston model and to the 3/2 model. For the Heston model estimation under $\mathbb{Q}$, both exact pricing formula and perturbation approach are considered. This allows to check whether the error due to the approximation has a remarkable impact on the estimation procedure.

Each simulation of the Monte Carlo study is based on a path of 1000 simulated observations. The used parameters are shown in Table 3.1. The parameter values are similar to those obtained in the real data estimation

Table 3.1: Parameters used to get the simulated path from the Heston model.

	μ	ζ	ξ	δ
Real world parameters (Heston model)	0.010	0.150	2.800	0.450
Real world parameters (3/2 model)	0.010	0.150	2.800	6.000
Risk adjusted parameters (Heston model)		0.150	3.000	0.450
Risk adjusted parameters (3/2 model)		0.150	3.000	6.000

(see Section 3.6). Note that for the Heston model parameters fulfill instantaneous variance positivity constraint, since $\delta^2 \leq 2\zeta$.

We perform 1000 simulations for each of the methods proposed. For the implementation of the EMM and the indirect inference algorithm, a path of 20.000 steps is simulated by using a Milstein discretization scheme[10]. Following Andersen and Lund [4], an antithetic variables technique is used to reduce the simulation error.

GARCH(1,1) maximum likelihood function is maximized by a BHHH algorithm[11]. Moreover, in order to solve the more difficult optimization problems in equations (3.4) and (3.8), we tested three of the minimization algorithms proposed by Press et al. [54]: (i) variable metric method, (ii) downhill simplex method and (iii) simulated annealing method for continuous variable. For this application downhill simplex method turns out to be the best. Indeed, despite the large number of function evaluations it requires, its results are much more stable if compared, for instance, with the variable metric method.

The simulated annealing method is useful when one deals with a local minimum problem. Here there is the advantage to know the function value at the minimum. Indeed, since the number of the auxiliary parameters is equal to the number of structural parameters, the function has to be zero at minimum. Hence, to avoid local minimum problem it is sufficient to disregard all the solutions where the function value at minimum is not close to zero.

In order to implement the estimation of methods 2 and 3, option prices are computed by assuming that options will expire in 45 days, while the risk-free rate is assumed equal to 2%. Integration in Heston's formula was done by using a Gauss-type numerical integration with a Laguerre polynomial[12] of order 12. As usual, BSIV was computed by the Newton-Raphson method.

[9] See Gourieroux and Monfort [33], Proposition 4.1.

[10] See appendix C for details.

[11] For BHHH algorithm we mean the Newton-like algorithm where Hessian is substituted by the outer product matrix.

[12] See appendix B.1 and Abramowitz and Stegun [1] for more details.

Table 3.2: Monte Carlo results. For each estimation method, in the first column there is the mean value, in the second column there is the root mean square error and in the third column there is the average of standard error computed for each estimation.

		Heston model under $\mathbb{P}$			3/2 model under $\mathbb{P}$		
		mean	rms error	std error	mean	rms error	std error
	μ	0.0086	0.0092	0.0501	0.0095	0.0141	0.0725
	ζ	0.1557	0.0139	0.0447	0.1543	0.0237	0.0471
	ξ	2.9112	0.3894	1.0518	2.7251	0.4149	1.1283
	δ	0.4364	0.0486	0.0734	5.8416	0.4183	1.6584
		Ind. inf. under $\mathbb{Q}$			EMM under $\mathbb{Q}$		
		mean	rms error	std error	mean	rms error	std error
	ζ	0.1574	0.0283	0.0353	0.1587	0.0269	0.0364
(1)	ξ	3.1476	0.6956	0.7027	3.2100	0.6461	0.7306
	δ	0.4541	0.0203	0.0209	0.4555	0.0179	0.0216
	ζ	0.1527	0.0316	0.0340	0.1581	0.0274	0.0352
(2)	ξ	3.1797	0.6609	0.6995	3.1910	0.6598	0.7224
	δ	0.4567	0.0232	0.0196	0.4550	0.0187	0.0205
	ζ	0.1434	0.0337	0.0911	0.1626	0.0331	0.0862
(3)	ξ	2.8184	0.9286	2.0749	3.2922	0.8440	2.0114
	δ	5.9749	0.2180	1.4457	6.0086	0.2227	1.2022

(1) Heston model exact formula
(2) Heston model perturbation approach
(3) 3/2 model perturbation approach

Monte Carlo results are summarized in Table 3.2. In the first column there is the mean of the estimated value, in the second the root mean square error with respect to the true parameter and in the third, the average of standard error computed for each estimation.

The first part of the table shows the estimation results under $\mathbb{P}$. In spite of the fact that the pointwise estimation seems to be slightly biased, the mode of the estimated values is equal to the values used for the simulation. The second part of the table shows the estimation results under $\mathbb{Q}$. For parameters ζ and δ, biases are small. On the contrary, ξ estimation is not very accurate. The bias value is about 5% for indirect inference estimation and 7% for EMM estimation and root mean squared error is very high. However, also in this case, the mode of the empirical distribution is equal to the true parameter value.

In Figures 3.1 and 3.2, the empirical estimator distribution is compared with a normal distribution with standard deviation equal to the root mean squared error. Figure 3.1 shows the results of the exact Heston model while

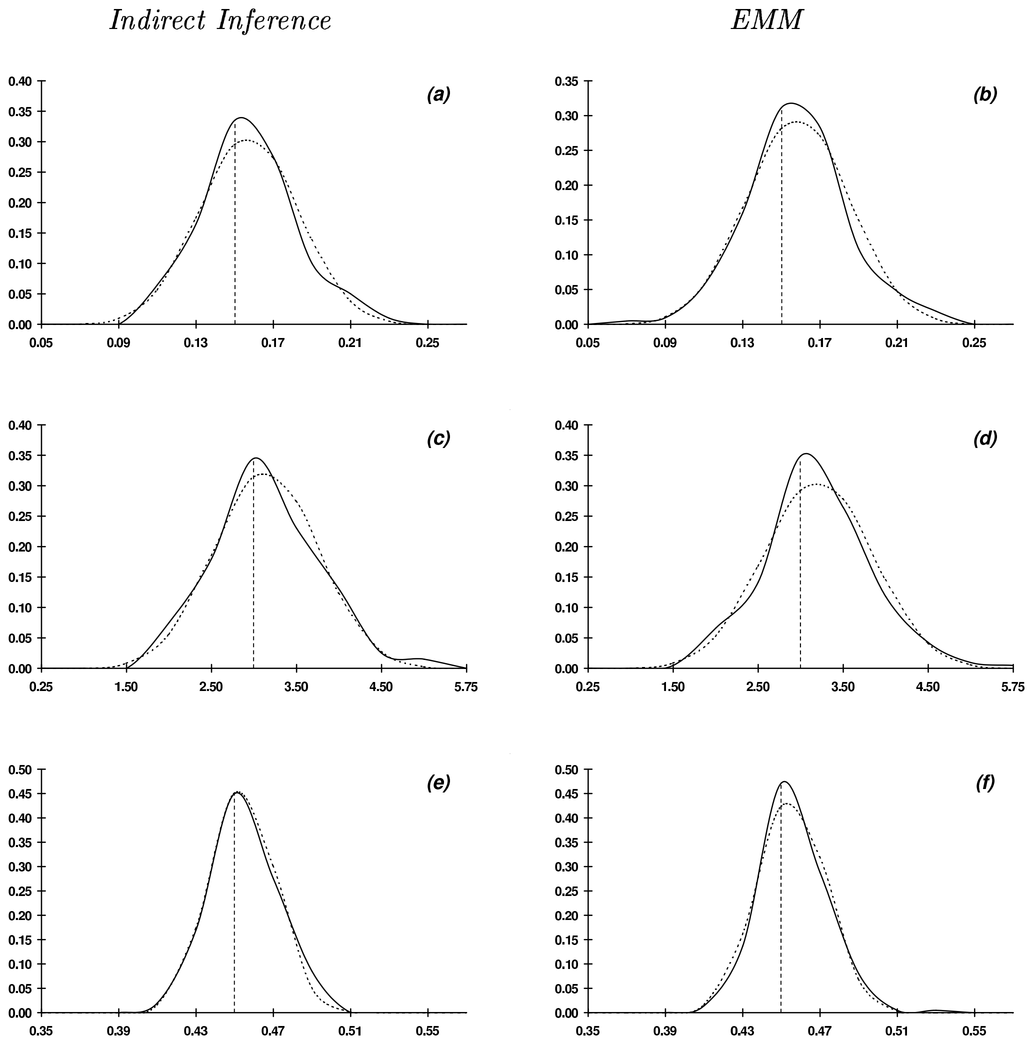


Figure 3.1: Exact Heston model. Empirical distributions (solid line) compared with normal distribution (dotted line) with mean equal to the true value of the parameter and standard deviation equal to mean standard error. ζ estimation results are in plots *(a)* and *(b)*, ξ estimation results are in plots *(c)* and *(d)* and δ estimation results are in plots *(e)* and *(f)*.

Figure 3.2 shows the results of 3/2 model. Results obtained by the Heston model with perturbation approach are not reported since they are indistinguishable from those obtained by the exact formula. In particular plots *(a)* and *(b)* refer to ζ estimation results, plots *(c)* and *(d)* refer to ξ estimation results and plots *(e)* and *(f)* refer to δ estimation results. Especially for ζ and δ, normal distribution can be considered a good approximation also in the sample.

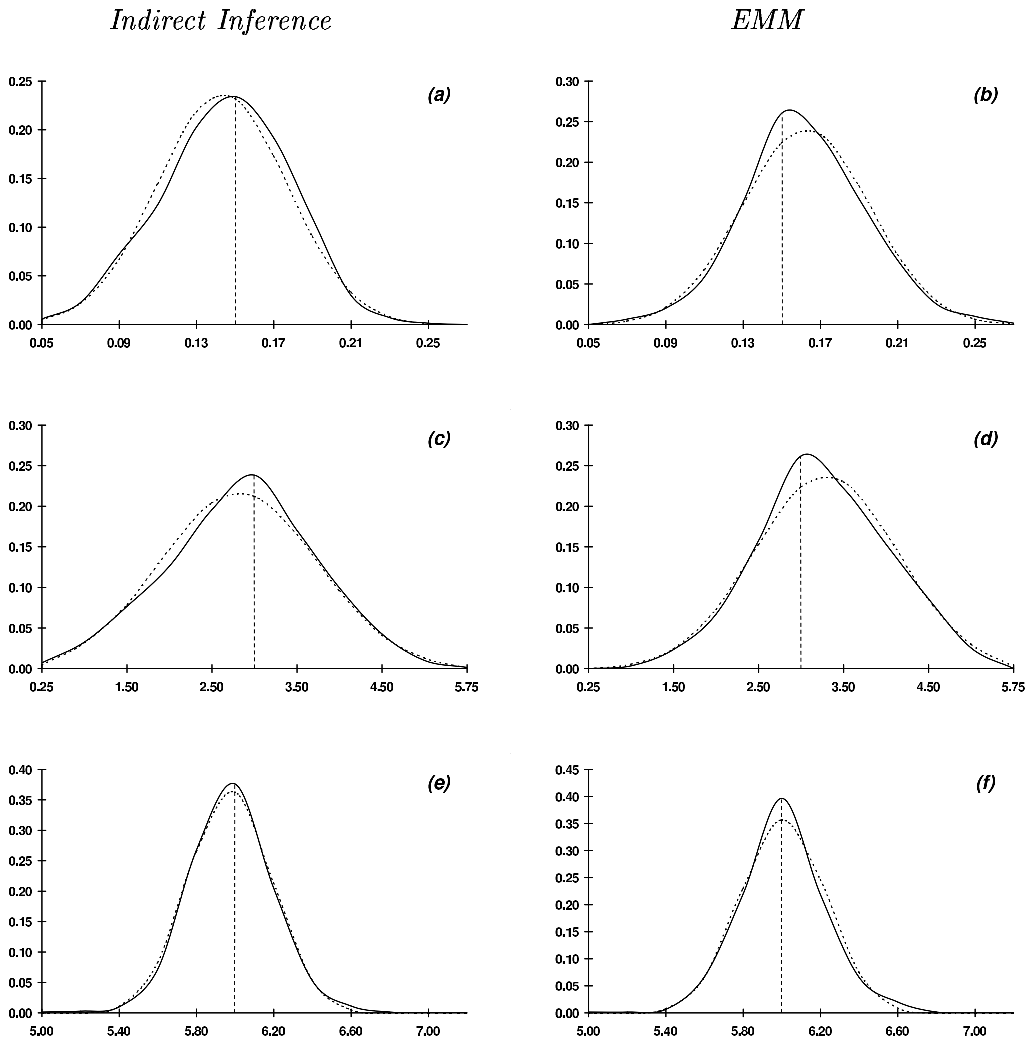


Figure 3.2: 3/2 model. Empirical distributions (solid line) compared with normal distribution (dotted line) with mean equal to the true value of the parameter and standard deviation equal to mean standard error. ζ estimation results are in plots *(a)* and *(b)*, ξ estimation results are in plots *(c)* and *(d)* and δ estimation results are in plots *(e)* and *(f)*.

3.6 Model estimation on SMI data

The real data estimation experiment involves Swiss Market Index (SMI) returns for parameter estimation under $\mathbb{P}$ and SMI options implied volatilities (hereafter VSMI) for parameter estimation under $\mathbb{Q}$. All time series consist of 1007 daily data (from 04.01.1999 to 03.01.2003). As a proxy of the risk-free rate the average of the three months LIBOR rate is considered (1.975%). BSIV is computed by using options with moneyness near to one

Table 3.3: Estimated parameters and asymptotic covariance matrix on SMI returns. The auxiliary parameter estimates (GARCH(1,1) model) are in the first panel while the structural parameter estimates are in the second panel.

Auxiliary parameters		values	std error
	α	-0.000036	(0.030954)
	β_0	0.000004	(0.013513)
	β_1	0.840867	(0.024916)
	β_2	0.138820	(0.020717)

Structural parameters				Asymptotic covariance matrix			
		values	std error	μ	ζ	ξ	δ
(1)	μ	-0.000029	(0.092693)	0.00859	–	–	–
	ζ	0.155083	(0.110104)	-0.00006	0.01212	–	–
	ξ	3.436133	(1.899103)	-0.00281	0.19086	3.60659	–
	δ	0.410676	(0.197345)	-0.00139	0.02067	0.29646	0.03894
(2)	μ	-0.000027	(0.081780)	0.00669	–	–	–
	ζ	0.059507	(0.140506)	0.00053	0.01974	–	–
	ξ	1.992169	(1.488870)	0.00457	0.64146	2.21673	–
	δ	6.262296	(2.828623)	0.13081	1.07396	3.01899	8.00111

(1) Heston model – EMM
(2) 3/2 model – EMM

and time to maturity close to 45 business days[13]. The above characteristics should guarantee the liquidity of the options taken into account.

In Table 3.3, estimation results on SMI returns are reported. The first panel shows auxiliary parameter estimates while the second panel shows the structural parameter estimates. The high standard error value has to be noted for the mean reversion parameter for both models. This is due to the low value of the Jacobian matrix M_θ for this parameter. Indeed, a low Jacobian value means that high variations of ξ have a low impact on the value of the orthogonality condition. This result is consistent with that obtained in Section 3.5 and it is a further indication that ξ estimation cannot be considered precise.

Estimation results on VSMI index are given in Table 3.4. The auxiliary parameter estimates are in the first panel, while the structural parameter estimates are in the second panel. The last two columns show the interval estimation with a 90% confidence level. As expected, indirect inference estimator and EMM estimator give quite the same results both for pointwise

[13] More details on BSIV computation can be found at the url www.gottardo-fs.com/vsmi.

Table 3.4: Estimated parameters and asymptotic covariance matrix on SMI option implied volatility. The auxiliary parameter estimates are in the first panel, while the structural parameter estimates are in the second panel.

Auxiliary parameters			Asymptotic cov. matrix		
	values	std error	a_0	a_1	a_2
a_0	0.002806	0.045596	0.002	–	–
a_1	0.986731	0.246467	-0.011	0.061	–
a_2	0.011979	0.023958	-0.000	0.001	0.001

Structural parameters				Asymptotic cov. matrix			Interval estim.	
		values	std error	ζ	ξ	δ	(conf. 90%)	
	ζ	0.166654	0.085153	0.007	–	–	0.0266	0.3067
(1)	ξ	3.199720	1.698485	0.175	2.885	–	0.4060	5.9935
	δ	0.464134	0.067830	0.005	0.130	0.005	0.3528	0.5760
	ζ	0.167045	0.084522	0.007	–	–	0.0280	0.3061
(2)	ξ	3.206658	1.839694	0.184	3.384	–	0.1806	6.2327
	δ	0.464339	0.069570	0.005	0.141	0.005	0.3499	0.5788
	ζ	0.166656	0.085018	0.007	–	–	0.0268	0.3065
(3)	ξ	3.232507	1.743532	0.177	3.040	–	0.3646	6.1004
	δ	0.486809	0.070753	0.005	0.141	0.005	0.3704	0.6032
	ζ	0.167127	0.083385	0.007	–	–	0.0300	0.3043
(4)	ξ	3.240398	1.857503	0.181	3.450	–	0.1851	6.2957
	δ	0.487002	0.073362	0.005	0.153	0.005	0.3663	0.6077
	ζ	0.039668	0.051429	0.003	–	–	-0.0449	0.1243
(5)	ξ	0.253134	1.226627	0.062	1.505	–	-1.7645	2.2707
	δ	6.056826	1.855278	0.061	1.660	3.442	3.0052	9.1085
	ζ	0.040104	0.049122	0.002	–	–	-0.0407	0.1209
(6)	ξ	0.202397	1.129433	0.055	1.276	–	-1.6553	2.0601
	δ	6.260973	2.190422	0.073	1.804	4.798	2.6580	9.8639

(1) Heston model exact formula – Indirect inference
(2) Heston model exact formula – EMM
(3) Heston model perturbation approach – Indirect inference
(4) Heston model perturbation approach – EMM
(5) 3/2 model perturbation approach – Indirect inference
(6) 3/2 model perturbation approach – EMM

estimation and for interval estimation.

Another important result is that the estimated values obtained by perturbation approach in the Heston model are very close to those obtained

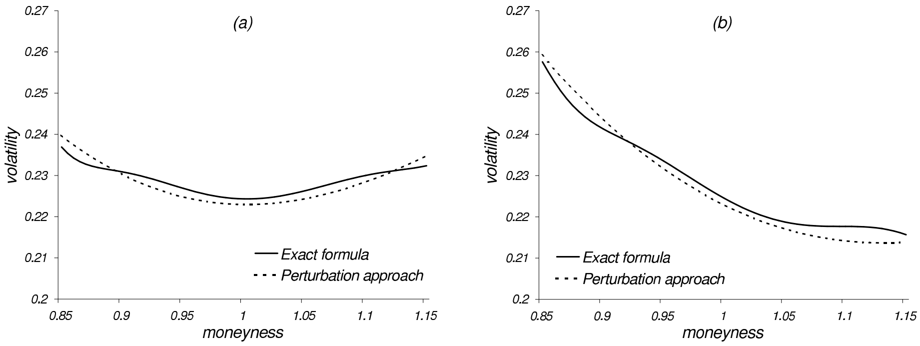


Figure 3.3: Comparison between implied volatility smile obtained by the exact Heston formula (solid line) and implied volatility smile obtained by the Heston price approximation (dotted line). In plot (a) the correlation parameter is assumed to be 0 while in plot (b) the correlation parameter is assumed to be -0.3. Parameter used are those estimated by indirect inference algorithm.

by exact formula. Figure 3.3 shows a comparison between implied volatility smile obtained by the exact Heston formula and implied volatility smile obtained by the approximated Heston price. The difference between the two volatility curves suggests that the perturbation approach approximation can be properly used only for estimation purposes. Indeed, the observed volatility differences are not negligible for pricing purposes.

Pointwise parameter estimation allows us to estimate the BSIV distribution. In Table 3.5 a comparison between the empirical distribution and the estimated distribution under $\mathbb{Q}$ is presented. The values of the first two centered moments are very close each other and all distributions seem to be positive skewed an slightly leptokurtic. The above features are more remarked for 3/2 model. Moreover, in the first column there is the long run instantaneous volatility. For the Heston model, it is close to the mean of the VSMI (0.21355).

In Figures 3.4 and 3.5, BSIV distribution is estimated according to different option times to maturity. Figure 3.4 is obtained by using the exact Heston pricing formula while Figure 3.5 is obtained by the 3/2 model with perturbation approach. The figure obtained by the Heston model with the perturbation approach is not shown since it is indistinguishable from that obtained by exact formula.

The longer the maturity the lower the distribution variance. This feature is due to the mean reversion of the variance process and it is particularly evident for the Heston model. Indeed, for $T \to \infty$, instantaneous variance converges to its unconditional mean ζ/ξ. Hence, the Black-Scholes assumption of constant volatility is not misleading for long maturity options[14].

[14]For an empirical result on the issue see Bakshi et al. [6].

Table 3.5: Centered moments of the empirical BSIV distribution compared with the estimated BSIV distribution. Estimated BSIV distribution is obtained by simulating a path of 500'000 steps.

Centered moments	$\sqrt{\zeta/\xi}$	mean	std	skew.	kurt.
Empirical distribution	–	0.21355	0.07511	1.17490	3.86224
(1)	0.22822	0.21248	0.07449	0.60876	3.15586
(2)	0.22824	0.21253	0.07441	0.60868	3.15605
(3)	0.22706	0.21088	0.07338	0.73752	3.39583
(4)	0.22710	0.21096	0.07330	0.73722	3.39573
(5)	0.39586	0.21051	0.07176	2.59113	14.00763
(6)	0.44513	0.21341	0.07504	2.39714	11.80407

(1) Heston model exact formula – Indirect inference
(2) Heston model exact formula – EMM
(3) Heston model perturbation approach – Indirect inference
(4) Heston model perturbation approach – EMM
(5) 3/2 model perturbation approach – Indirect inference
(6) 3/2 model perturbation approach – EMM

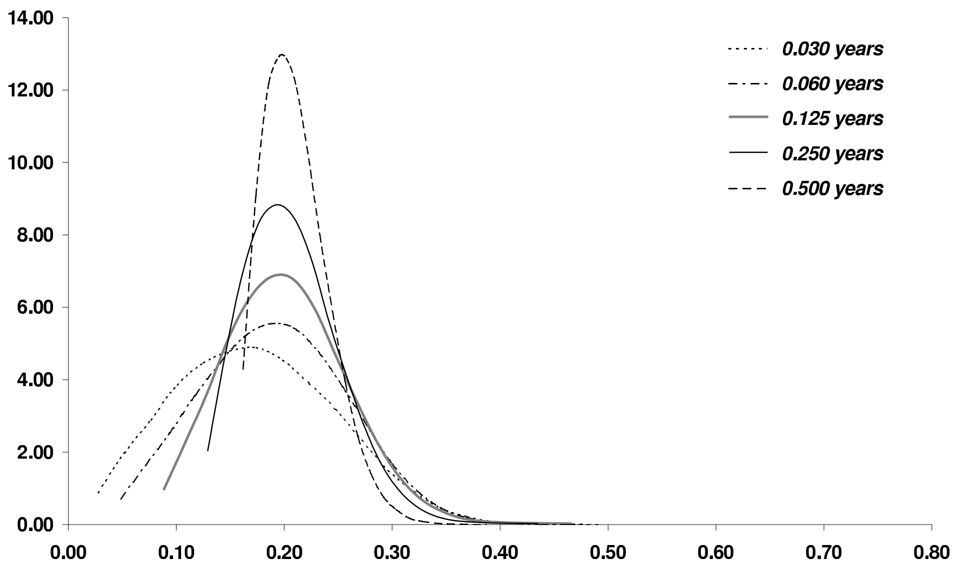


Figure 3.4: Exact Heston model. Estimated distribution of Black-Scholes implied volatility for different option time to maturity. Plot is obtained by using indirect inference parameters.

3.7 Conclusion

This chapter describes three algorithms to estimate the parameters of a class of stochastic volatility models. Two of them deal with risk adjusted param-

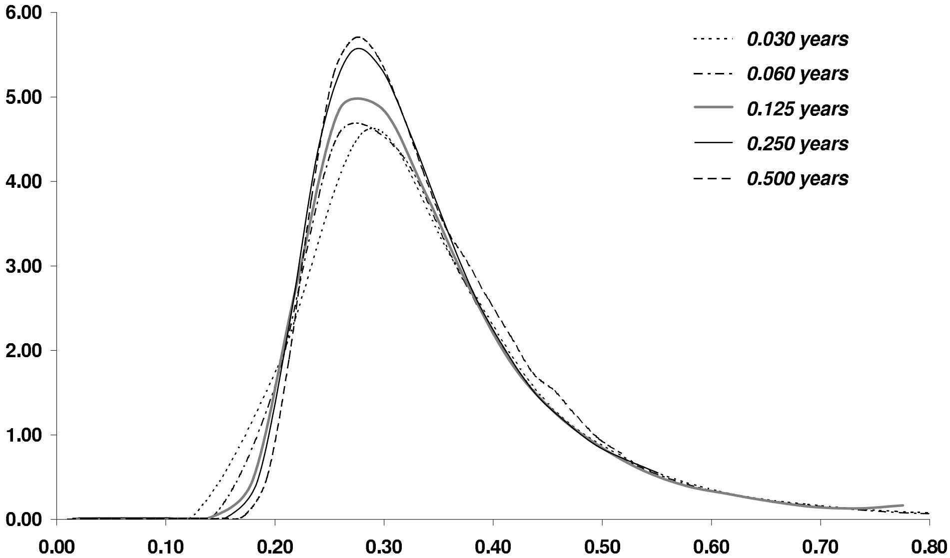


Figure 3.5: 3/2 model. Estimated distribution of Black-Scholes implied volatility for different option time to maturity. Plot is obtained by using indirect inference parameters.

eters, while the third handles objective parameter estimation. Underlying prices are used for objective parameter estimation while a BSIV time series is used for risk adjusted estimation. For objective parameter estimation the auxiliary model is a GARCH(1,1), while for risk-adjusted parameters an Ornstein-Uhlenbeck process is used.

Because of the volatility risk premium, the two sets of parameters are different from one another. In particular, for risk adjusted parameters this chapter proposes a simple but effective estimation method which is able to provide interval estimation too. The proposed algorithms are very general since they can be potentially applied to a broad class of stochastic volatility models. However, in order to reduce computational time, they are more appropriately applied to models with closed form solution for option prices. In this chapter, the Heston model and the 3/2 model are used.

A real data estimation and a Monte Carlo study are performed. Our Monte Carlo study shows that the discretization bias is small in almost all the estimates with the exception of the mean reverting parameter. Moreover, root mean square error is high for ξ and much lower for the other two parameters. The poor estimation precision of ξ is evident both for objective estimation and for risk adjusted estimation. Finally, normal distribution can be considered a good approximation especially for δ. For the other two parameters, the empirical distribution is slightly different from the normal one, though the distance is not large.

The Monte Carlo study substantially emphasizes a poor estimation precision for the mean reverting parameter. This is an interesting result since both Heston option prices and 3/2 option prices are very sensitive with respect to this parameter. Hence, simple pointwise estimation can provide considerable errors in option pricing and hedging.

For the risk adjusted estimation of the Heston model, both exact pricing formula and the perturbation approach are considered. For estimation purposes, the error due to the approximation of the perturbation approach seems to be negligible. Moreover, the approximated formula remarkably reduces the computational time during the estimation.

A confirmation of the Monte Carlo results is given by the real data application. Indeed, by applying the estimators to a BSIV time series, both a pointwise estimation and an interval estimation are performed. The confidence interval is large for ζ and ξ and much smaller for δ.

Moreover, the BSIV distribution is estimated according to different option times to maturity. Because of the mean reverting feature, distribution variance declines for long maturities.

Chapter 4

Pricing and hedging reliability

Chapter 3 points out the intricacy on stochastic volatility model estimation. The high estimation standard error implies that the parameter misspecification feature cannot be neglected. The parameter misspecification affects both pricing and hedging, producing an incorrect option pricing and some unexpected losses due to the hedging errors.

To face the problem, the super-hedging approach proposed in Chapter 2 can be applied. Indeed, the approach is primarily used to tackle the misspecification problem that can involve the parameters of the instantaneous variance process. To be applied, the approach needs the definition of a set of bounds within which each parameter has to keep. The bounds can be defined by the interval estimation described in Chapter 3.

The super-hedging approach and the estimation method compose a complete framework to price and hedge options. This chapter aims to test the approach by verifying the pricing reliability and the capacity of the hedging strategy to reduce the unexpected losses. For pricing reliability we mean the capacity of the model to fix a price consistent with the observed market price. For the model proposed in Section 2.2 the pricing reliability is obtained when the market price keeps between buyer's and seller's price.

Hedging is generally defined as the ability to compose a portfolio replicating the option value in order to reduce the possibility of unexpected losses. In a standard Black-Scholes framework, a self-financing replicating portfolio can be found by using the underlying asset and a riskless bond. Since in a stochastic volatility model there is a new source of randomness, the two assets are no longer sufficient. Indeed, the portfolio composed only of the underlying asset and the riskless bond can no longer replicate the option value or is no longer self-financing.

In order to ensure both properties, it is necessary to introduce a new asset whose value depends on volatility. For such a purpose, a candidate asset seems to be an option with the same underlying of the first one, but a longer time to maturity. Moreover, parameter misspecification has to be considered. This is done by introducing the super-hedging strategy based on the approach proposed in Chapter 2.

In order to test the super-hedging approach, some simulation experiments are performed. In particular we try to verify whether the proposed super-hedging strategy is able to properly cope with the parameter misspecification problem. Moreover, we compare the proposed super-hedging with that introduced by Avellaneda, Levy and Parás [5] (hereafter ALP).

Another set of experiments is developed to test the super-hedging strategy under model misspecification. The model misspecification is obtained by contaminating the standard error variable of the instantaneous variance process. This contamination produces some jumps in the implied volatility path and it affects the hedging error quite remarkably.

In Section 4.1 a method for estimating the unobservable instantaneous variance is described. This method is used in Section 4.2 to verify the pricing properties of the super-hedging model under stochastic volatility. Section 4.3 describes the hedging under a stochastic volatility framework and Section 4.3.1 extends the approach to a misspecified framework. The hedging reliability is verified in Section 4.4 where the different hedging approaches are compared. Section 4.5 concludes.

4.1 Unobservable instantaneous variance

In order to apply an option pricing formula under the stochastic volatility assumption, not only does one need to estimate the model parameters, but also the unobservable instantaneous variance v_t. For this purpose we propose to substitute v_t with its expected value conditional on the information in $t-1$.

In the class of stochastic volatility models considered in Chapter 3, the conditional expected value of the instantaneous variance is[1]

$$E^{\mathbb{Q}}[v_t|\mathcal{F}_s] = v_s e^{-\xi(t-s)} + \frac{\zeta}{\xi}\left(1 - e^{-\xi(t-s)}\right) \qquad s < t\ . \tag{4.1}$$

Recalling that $\frac{\zeta}{\xi}$ is the long run variance level, the above expected value can be interpreted as a weighted average between the instantaneous variance at time s and the long run variance.

According to intuition, weights depend on the distance between t and s and on mean reversion parameter ξ. In particular, the weight associated to v_s is inversely proportional to $t-s$ and to ξ. Indeed, the longer the time between t and s the less informative the value of v_s. Moreover, the higher

[1] It can be easily obtained by applying Ito's lemma to $v_t e^{\xi(t-s)}$.

ξ the faster the mean reversion process. In this case, the long run variance contains more information than the instantaneous variance at time s.

Since the conditional expected value of v_t depends on v_s, one has to compute the unobservable value of v_s. This can be done by using BSIV value in s and solving the following equation with respect to v_s:

$$W_{SV}(S_s, X, T-s, r, v_s, \theta) - W_{BS}(S_s, X, T-s, r, \mathrm{BSIV}_s) = 0 \ . \qquad (4.2)$$

In equation (4.1) and (4.2), parameter values will be replaced by their estimates. Since $\hat{\theta}^{II}$ and $\hat{\theta}^{EMM}$ are consistent estimators of the model parameters, the proposed instantaneous variance estimator will be consistent, too.

4.2 Pricing reliability

The pricing reliability feature can be tested by comparing the market price of options and the option price given by the model. Since the proposed model gives a seller's and buyer's price, the price reliability is obtained if the market price keeps between the two price bounds.

For this purpose, the parameters are estimated on the VSMI[2] time series from 04.01.1999 to 03.01.2003. In order to get an out-of-sample test, the option market prices are obtained by applying the Black-Scholes formula to VSMI data from 06.01.2003 to 12.02.2004 (275 observations). Moreover, the three months LIBOR rate is used as a proxy for risk-free rate and, consistently with VSMI computation, maturity is assumed equal to 45 working days.

The implied instantaneous variance is computed daily as described in Section 4.1. Super hedging prices are computed by assuming that the parameter values lie inside the estimated confidence interval (confidence level 90%). The value of ρ is assumed to belong to interval $[-0.8, 0.8]$.

In the most general case, the European call super-hedging price can be obtained by numerically solving a non-linear PDE. However, by assuming no misspecification for δ and ρ, the explicit Heston formula can be used.

In Figure 4.1 the super-hedging prices obtained under the assumption of stochastic volatility are compared with market prices. The same figure shows also the super-hedging prices according to the ALP approach (broken line). The uncertain volatility bounds are is obtained for each day by the 5% and the 95% percentile of the simulated BSIV distribution.

Plot (a) of Figure 4.1 shows the super-hedging prices obtained by numerically solving the PDE, while in plot (b) the super-hedging prices are obtained under the assumption that δ is constant and ρ is zero. It is clear from the figure that these assumptions have a small impact in the seller's

[2] See Section 3.6 for more details.

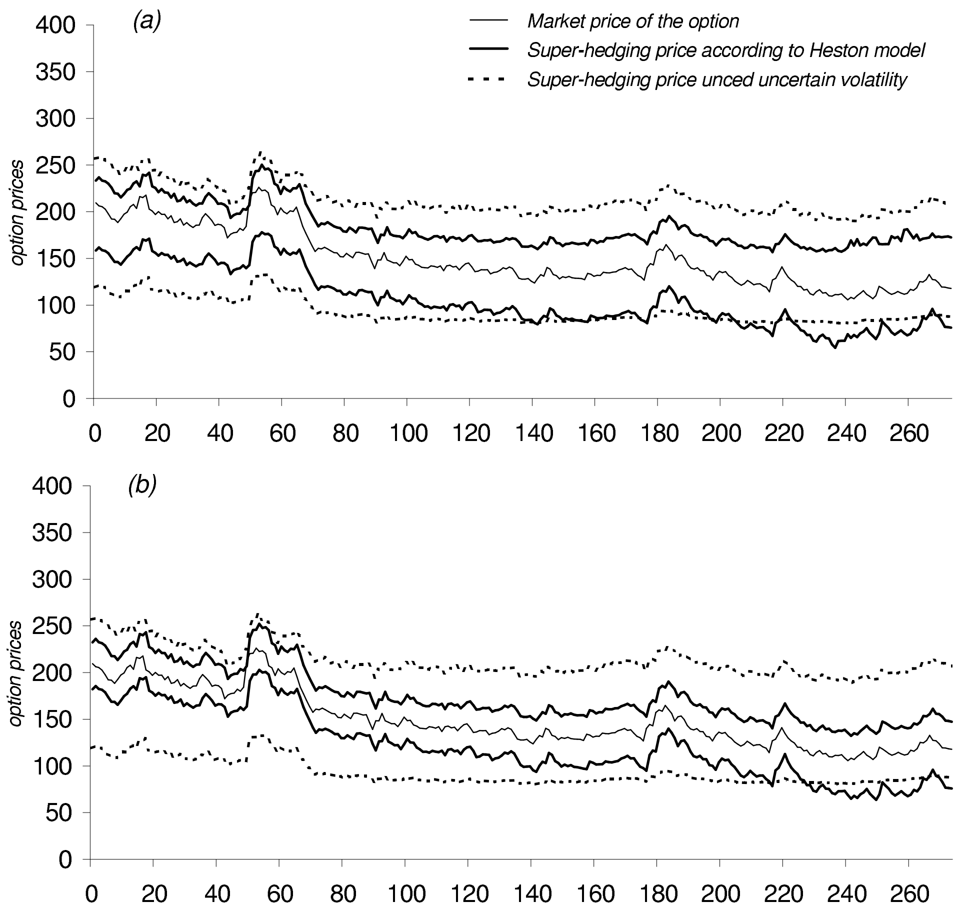


Figure 4.1: Comparison between market prices, super-hedging prices under uncertain volatility and super-hedging price according to stochastic volatility with uncertain parameters. In plot (a) the super-hedging prices according to the Heston model are obtained by numerically solving the PDE. In plot (b), by assuming δ and ρ constant, the analytic formula is used.

price and a more pronounced impact in the buyer's price. However, in both cases super-hedging prices under stochastic volatility are much closer to market price than super-hedging prices under the ALP approach.

Table 4.1 shows the main results of the comparison between prices. The second and third columns show the average percentage distances between super-hedging prices and market prices. One can notice a strong reduction in the distances due to the stochastic volatility approach with respect to the uncertain volatility approach. This suggests that under the assumption of uncertain volatility there is a strong overestimation of the seller's price and a strong underestimation of the buyer's price.

The fourth column of Table 4.1 shows the number of violations recorded. A violation is recorded each time the market price is higher than the seller's

Table 4.1: Results of the comparison between super-hedging prices and market prices.

	average distance from market prices		violations		size
	sell	buy	sell	buy	
uncertain volatility	47.43%	−36.57%	0	0	–
stochastic vol. numerical	25.39%	−30.56%	1	0	−2.67%
stochastic vol. simplified	18.02%	−20.24%	1	0	−2.71%

price or lower than the buyer's price. No violations are recorded for the uncertain volatility approach while the stochastic volatility approach shows a violation of the seller's price. However, this violation is slightly higher than 2.5% of the option market price and can be considered negligible.

4.3 Stochastic volatility and hedging

In a stochastic volatility framework, a self-financing hedging portfolio has to be composed of an asset whose value depends on the volatility as well as on the underlying and a riskless bond. Indeed, the underlying and the riskless bond only are not able to generate a portfolio that is at the same time replicating and self-financing.

One of the most common assets whose price depends on the volatility is an option on the same underlying of the first one, but with a longer time to maturity. The hedging approach that uses the underlying, the riskless bond and a second option on the same underlying is known as delta-sigma hedging[3].

Remark 4.3.1 In a correctly specified stochastic volatility framework, the introduction of a second option on the same underlying is sufficient to complete the market since the number of traded assets becomes equal to the risk factors driving the asset price[4]. □

Let us denote $W_2(t, S_t, v_t)$ the traded option written on S_t. With a slight change of notation, the option to hedge is denoted by W_1. We define $V_t(\theta)$ the hedging portfolio associated with the strategy $\phi := (a, b, c)$ such that

$$V_t(\phi) = a_t S_t + b_t B_t + c_t W_{2,t} \ ,$$

where $\{a_t\}$ and $\{c_t\}$ are two $\mathcal{F}_t$-measurable processes. In order to ensure the replicating property of the portfolio, we define the process $\{b_t\}$ as

$$b_t = B_t^{-1} \left(W_{1,t} - a_t S_t - c_t W_{2,t} \right) \ .$$

[3]See among others Scott [57].
[4]See Karatzas [42].

Moreover, the self-financing constraint imposes that

$$dV_t = a_t \, dS_t + b_t \, dB_t + c_t \, dW_{2,t} \, . \tag{4.3}$$

By applying Ito's formula to both sides of the above equation, we get

$$\begin{aligned}
&\frac{\partial W_1}{\partial t}dt + \frac{1}{2}\frac{\partial^2 W_1}{\partial S^2}d\langle S\rangle + \frac{1}{2}\frac{\partial^2 W_1}{\partial v^2}d\langle v\rangle + \frac{\partial^2 W_1}{\partial S \partial v}d\langle S, v\rangle + \\
&+ \frac{\partial W_1}{\partial S}dS + \frac{\partial W_1}{\partial v}dv = b_t dB + \left(a_t + c_t\frac{\partial W_2}{\partial S}\right) dS + c_t\frac{\partial W_2}{\partial v}dv + \\
&+ c_t\left(\frac{\partial W_2}{\partial t}dt + \frac{1}{2}\frac{\partial^2 W_2}{\partial S^2}d\langle S\rangle + \frac{1}{2}\frac{\partial^2 W_2}{\partial v^2}d\langle v\rangle + \frac{\partial^2 W_2}{\partial S \partial v}d\langle S, v\rangle\right) \ ,
\end{aligned} \tag{4.4}$$

In order to remove the two random parts in (4.4) one needs

$$c_t = \frac{\partial W_1/\partial v}{\partial W_2/\partial v} \qquad \text{and} \qquad a_t = \frac{\partial W_1}{\partial S} - c_t\frac{\partial W_2}{\partial S} \ .$$

Indeed, substituting the above values, both sides of equation (4.4) become equal to $\mu_v^{\mathbb{Q}}(v_t;\theta)$. Since $\mu_v^{\mathbb{Q}}(v_t;\theta)$ does not depend on the characteristics of the options, the self-financing portfolio constructed with the strategy ϕ perfectly replicates the claim price at each time $t \leq T$. However, a perfect replication is subject to the condition that the model parameters are correctly specified.

Remark 4.3.2 In Heston's model the first derivative with respect to the underlying price is P_1 as defined in appendix B.1. The first derivative with respect to the instantaneous variance is equal to

$$\frac{\partial W}{\partial v} = S_t\frac{\partial P_1}{\partial v} - Ke^{-r(T-t)}\frac{\partial P_2}{\partial v} \ ,$$

where

$$\frac{\partial P_j}{\partial v} = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-iu\ln K}\varphi_j(u)}{iu}D_j(u)\right] du \qquad \text{for } j = 1, 2$$

and the definition of $\varphi_j(u)$ and $D_j(u)$ are given in appendix B.1. □

4.3.1 Hedging approach in a misspecified SV framework

In order to address the risk of parameter misspecification, it is possible to use the result of Chapter 2. Hence, for instance, for an institution which sells the option, the strategy can be defined as

$$\begin{aligned}
&c_t^+ = \frac{\partial W_1^+/\partial v}{\partial W_2/\partial v} \qquad a_t^+ = \frac{\partial W_1^+}{\partial S} - c_t^+\frac{\partial W_2}{\partial S} \\
&b_t^+ = B_t^{-1}\left(W_{1,t}^+ - a_t^+ S_t - c_t^+ W_{2,t}\right) \ .
\end{aligned} \tag{4.5}$$

Hereafter, this strategy will be denoted $\phi^+ := (a^+, b^+, c^+)$.

Definition 4.3.1 *The discounted replication error[5] associated with the strategy ϕ is the discounted difference between the self-financing portfolio $V(\phi)$ and the claim price:*

$$e_t(\phi) := B_t^{-1}(V_t(\phi) - W_t) \ ,$$

where $e_0 = 0$ as long as $V_0 = W_0$. □

If the discounted replication error associated with the strategy ϕ^+ is non-negative for all $t \in [0, T]$, then the strategy ϕ^+ is a super-hedging strategy. Indeed, in such a case we will have $V_t(\phi^+) = W_t + e_t(\phi^+)B_t \geq W_t$.

Proposition 4.3.1 *If condition* (2.11) *is true, the discounted replication error associated with the strategy ϕ^+ is always non-negative inside the interval* $[0, T]$. □

4.4 Hedging reliability

The effectiveness of a hedging strategy depends on its capacity to reduce the unexpected losses of an open option position. An unexpected loss occurs when the value of the replicating portfolio is different from the option value and it appears when the investor decides to close the position. This difference is usually denoted as hedging error.

Hence, the hedging error gives an indication of the reliability of the considered hedging strategy. However, the hedging error mean cannot be the only observed variable. One has to consider also the variability of the hedging error and, in general, the characteristics of its distribution. This kind of analysis can be conducted in a simulation framework, where there is the possibility to simulate a large number of price paths in order to properly estimate the hedging error distribution.

In this framework, we consider an investor that has to hedge a short option position. Without loss of generality, we assume that the position is closed at the option maturity. At the issue date, the call option is at-the-money with a maturity of one month. In order to reduce the discretization problems, the hedging strategies are recalibrated four times a day. Moreover, since the hedging portfolio is recalibrated at discrete times, the hedging strategy is no longer self-financing. To face this problem, profits and losses obtained by the recalibration of the portfolio are compounded and added to the final hedging error.

Each experiment involves 100.000 simulation paths. All the experiments are performed by using two different sets of parameters. Table 4.3 shows the results related to the first set while Table 4.4 shows the results related to the second set. The first column shows the initial cost of the strategy

[5] Frey and Sin [30] define this error as "tracking error".

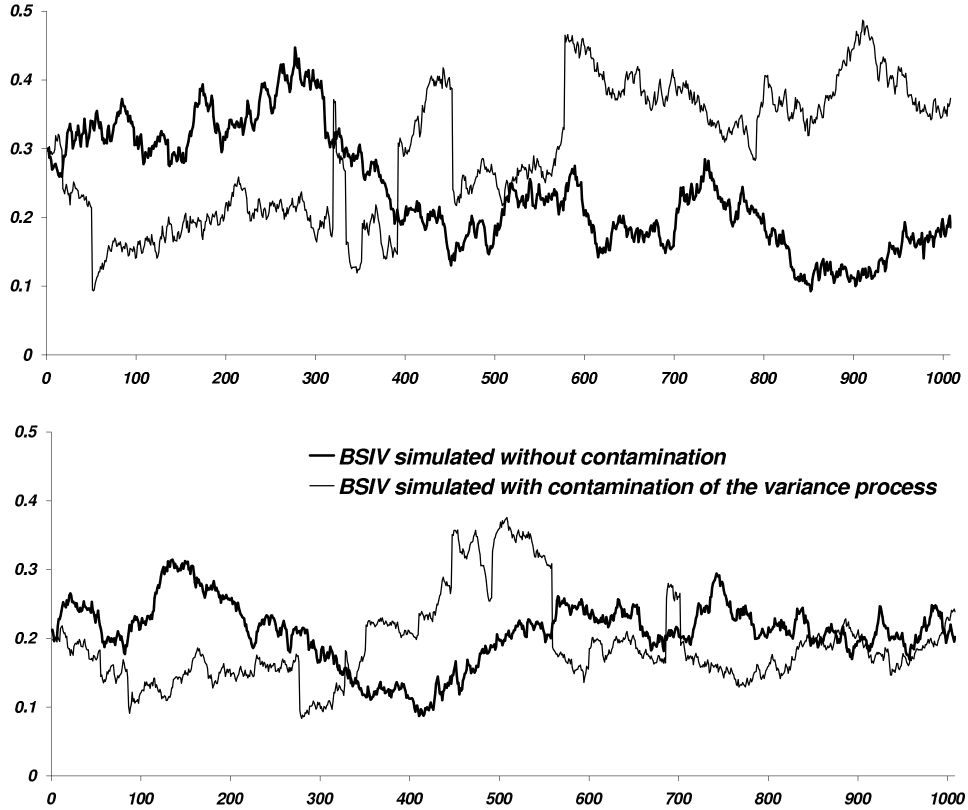


Figure 4.2: Simulated path of the Black-Scholes implied volatility. Above there are the simulations according to the first set of parameters and below those according to the second kind of parameters. The thick line shows the path simulated without contamination while the thin line shows the path simulated under contamination of the instantaneous variance process.

while the second shows the percentage difference between the initial cost and the benchmark cost, i.e. the cost obtained by Black-Scholes formula. The next columns show some information on the hedging errors distribution. In particular, hedging errors are described by their mean, the percentage mean with respect to initial cost, and their standard deviation. Moreover, to have an idea on the left tail of the hedging error distribution, two columns are devoted to the distribution percentiles. Finally, the last column shows the frequency of losses.

The first two experiments concern the Black-Scholes delta hedging. The first line of Tables 4.3 and 4.4 shows the results of the delta hedging when prices are simulated according to a geometric Brownian motion. The drift parameter is assumed equal to the risk-free rate while the volatility is assumed equal to the square root of the long-run instantaneous variance of the

Table 4.2: Parameter estimates for the hedging experiments. In the first panel there are the estimates of the first set of parameters while in the second panel there are the estimates of the second set of parameters. Upper and lower bound are computed as a confidence interval with a confidence of 80%.

	Heston model				contaminated Heston model			
	value	std error	lower bound	upper bound	value	std error	lower bound	upper bound
ζ	0.1233	0.065	0.0405	0.2061	0.2020	0.127	0.0392	0.3649
ξ	2.0418	1.260	0.4272	3.6565	2.3668	1.416	0.5520	4.1816
δ	0.3852	0.032	0.3446	0.4259	0.4316	0.083	0.3257	0.5375
ζ	0.1035	0.045	0.0456	0.1613	0.1685	0.084	0.0605	0.2765
ξ	3.4337	1.527	1.4762	5.3912	3.3104	1.435	1.4707	5.1500
δ	0.2742	0.025	0.2421	0.3063	0.4599	0.069	0.3710	0.5487

Heston model ($\sigma = \sqrt{\zeta/\xi}$). The mean and the standard deviation of the hedging error are not zero because of the discrete recalibration of the hedging portfolio. On the whole, in the Black-Scholes framework, delta hedging performs quite well and it can be taken as a benchmark for all the other experiments.

The results on the second line of Tables 4.3 and 4.4 refers to the Black-Scholes delta hedging when prices are simulated according to the Heston model. We can see a remarkable increase in the hedging error standard deviation. Moreover, also the two quantiles indicate that in this framework the losses can be more serious than in the previous experiment. This confirms the inappropriateness of the pure delta hedging in a stochastic volatility framework. Indeed, with delta hedging one completely disregards the impact of the instantaneous variance in the hedging strategy.

As described in Section 4.3, a proper hedging approach in a stochastic volatility framework can be followed by using a second option on the same underlying. According to the delta-sigma approach, the hedging portfolio has to be composed of the underlying, the money market account and the second option.

The next three experiments involve a simulation according to Heston's model and a delta-sigma hedging to hedge the short call position. In order to make the experiment easy to interpret, we assume a zero risk premium both for price uncertainty and for variance uncertainty. Hence, the drift of the price process is equal to the risk-free rate and no distinction is made between real-world and risk-adjusted parameters.

In practice, the hedging approach involves the following steps: (1) observation of the underlying price, the money market account and the option price, (2) computation of the implied instantaneous variance from the op-

Table 4.3: Results of the hedging experiments with the first set of parameters. The first column shows the initial cost of the strategy while the second column shows the variation of the initial cost with respect of the cost of the first experiment. The other columns show some results on hedging errors (mean, percentage mean with respect to initial cost and standard deviation), the frequency of losses and the percentiles of the profits and losses distribution. The following list points out the hedging strategy of the experiment, the model used to simulate the price path and the parameter used to apply the hedging strategy.

Black-Scholes delta hedging:

1. Delta hedging – geometric Brownian motion – true parameters ($\mu = r = 0.02$ and $\sigma = 0.30$).
2. Delta hedging – Heston's stochastic volatility model – true parameters ($\mu = r = 0.02$, $\zeta = 0.18$, $\xi = 2$, $\delta = 0.40$ and $\rho = 0$).

Hedging with parameter misspecification:

3. Delta-sigma hedging – Heston's stochastic volatility model – true parameters ($\mu = r = 0.02$, $\zeta = 0.18$, $\xi = 2$, $\delta = 0.40$ and $\rho = 0$).
4. Delta-sigma hedging – Heston's stochastic volatility model – estimated parameters ($\hat{\zeta} = 0.1233$, $\hat{\xi} = 2.0418$ and $\hat{\delta} = 0.3852$).
5. Delta-sigma hedging – Heston's stochastic volatility model – super-hedging parameters ($\hat{\zeta} = 0.2061$, $\hat{\xi} = 0.4272$ and $\hat{\delta} = 0.3852$).
6. Delta hedging – Heston's stochastic volatility model – super-hedging parameters according to the ALP approach ($\sigma_{\max} = 0.3097$).

Hedging with model misspecification:

7. Delta-sigma hedging – contaminated Heston's stochastic volatility model – estimated parameters ($\hat{\zeta} = 0.2134$, $\hat{\xi} = 3.1628$ and $\hat{\delta} = 0.3485$).
8. Delta-sigma hedging – contaminated Heston's stochastic volatility model – super-hedging parameters ($\hat{\zeta} = 0.3313$, $\hat{\xi} = 1.4562$ and $\hat{\delta} = 0.3485$).
9. Delta hedging – contaminated Heston's stochastic volatility model – super-hedging parameters according to the ALP approach ($\sigma_{\max} = 0.3119$).

	initial		hedging errors			percentiles		frequency of
	cost	%	mean	%	std	5%	1%	losses
1) Delta hedging	3.5349	–	0.020	0.57%	0.326	−0.511	−0.848	47.01%
2) Delta hedging	3.5349	–	0.040	1.12%	0.523	−0.829	−1.455	45.80%
3) Delta-sigma hedging	3.5160	−0.53%	0.014	0.41%	0.286	−0.404	−0.844	47.75%
4) Delta-sigma hedging	3.4710	−1.81%	−0.031	−0.88%	0.287	−0.450	−0.892	58.31%
5) Delta-sigma super-hedging	3.6468	3.16%	0.145	3.99%	0.288	−0.277	−0.722	21.90%
6) Delta super-hedging	3.6468	3.16%	0.152	4.16%	0.524	−0.690	−1.286	36.61%
7) Delta-sigma hedging	3.5050	−0.85%	−0.002	−0.07%	0.296	−0.417	−0.883	51.76%
8) Delta-sigma super-hedging	3.7533	6.18%	0.246	6.57%	0.296	−0.170	−0.640	11.78%
9) Delta super-hedging	3.7533	6.18%	0.245	6.52%	0.693	−0.862	−1.796	31.76%

Table 4.4: Results of the hedging experiments with the second set of parameters. The first column shows the initial cost of the strategy while the second column shows the variation of the initial cost with respect of the cost of the first experiment. The other columns show some results on hedging errors (mean, percentage mean with respect to initial cost and standard deviation), the frequency of losses and the percentiles of the profits and losses distribution. The following list points out the hedging strategy of the experiment, the model used to simulate the price path and the parameter used to apply the hedging strategy.

Black-Scholes delta hedging:

1. Delta hedging – geometric Brownian motion – true parameters ($\mu = r = 0.02$ and $\sigma = 0.20$).
2. Delta hedging – Heston's stochastic volatility model – true parameters ($\mu = r = 0.02$, $\zeta = 0.12$, $\xi = 3$, $\delta = 0.30$ and $\rho = 0$).

Hedging with parameter misspecification:

3. Delta-sigma hedging – Heston's stochastic volatility model – true parameters ($\mu = r = 0.02$, $\zeta = 0.12$, $\xi = 3$, $\delta = 0.30$ and $\rho = 0$).
4. Delta-sigma hedging – Heston's stochastic volatility model – estimated parameters ($\hat{\zeta} = 0.0957$, $\hat{\xi} = 2.7625$ and $\hat{\delta} = 0.2494$).
5. Delta-sigma hedging – Heston's stochastic volatility model – super-hedging parameters ($\hat{\zeta} = 0.1520$, $\hat{\xi} = 1.2833$ and $\hat{\delta} = 0.2494$).
6. Delta hedging – Heston's stochastic volatility model – super-hedging parameters according to the ALP approach ($\sigma_{\max} = 0.2090$).

Hedging with model misspecification:

7. Delta-sigma hedging – contaminated Heston's stochastic volatility model – estimated parameters ($\hat{\zeta} = 0.1783$, $\hat{\xi} = 5.2352$ and $\hat{\delta} = 0.2994$).
8. Delta-sigma hedging – contaminated Heston's stochastic volatility model – super-hedging parameters ($\hat{\zeta} = 0.2576$, $\hat{\xi} = 2.8822$ and $\hat{\delta} = 0.2994$).
9. Delta hedging – contaminated Heston's stochastic volatility model – super-hedging parameters according to the ALP approach ($\sigma_{\max} = 0.2121$).

	initial		hedging errors			percentiles		frequency of
	cost	%	mean	%	std	5%	1%	losses
1) Delta hedging	2.3853	–	0.013	0.56%	0.217	−0.340	−0.564	47.05%
2) Delta hedging	2.3853	–	0.029	1.21%	0.369	−0.585	−1.024	45.57%
3) Delta-sigma hedging	2.3703	−0.63%	0.009	0.40%	0.189	−0.266	−0.554	47.69%
4) Delta-sigma hedging	2.3355	−2.09%	−0.025	−1.09%	0.189	−0.300	−0.588	60.40%
5) Delta-sigma super-hedging	2.4875	4.29%	0.127	5.10%	0.190	−0.150	−0.440	16.17%
6) Delta super-hedging	2.4884	4.32%	0.132	5.31%	0.370	−0.458	−0.874	33.79%
7) Delta-sigma hedging	2.3908	0.23%	0.027	1.12%	0.197	−0.245	−0.562	42.07%
8) Delta-sigma super-hedging	2.5912	8.63%	0.228	8.78%	0.197	−0.044	−0.363	6.84%
9) Delta super-hedging	2.5912	8.63%	0.223	8.59%	0.494	−0.559	−1.220	27.55%

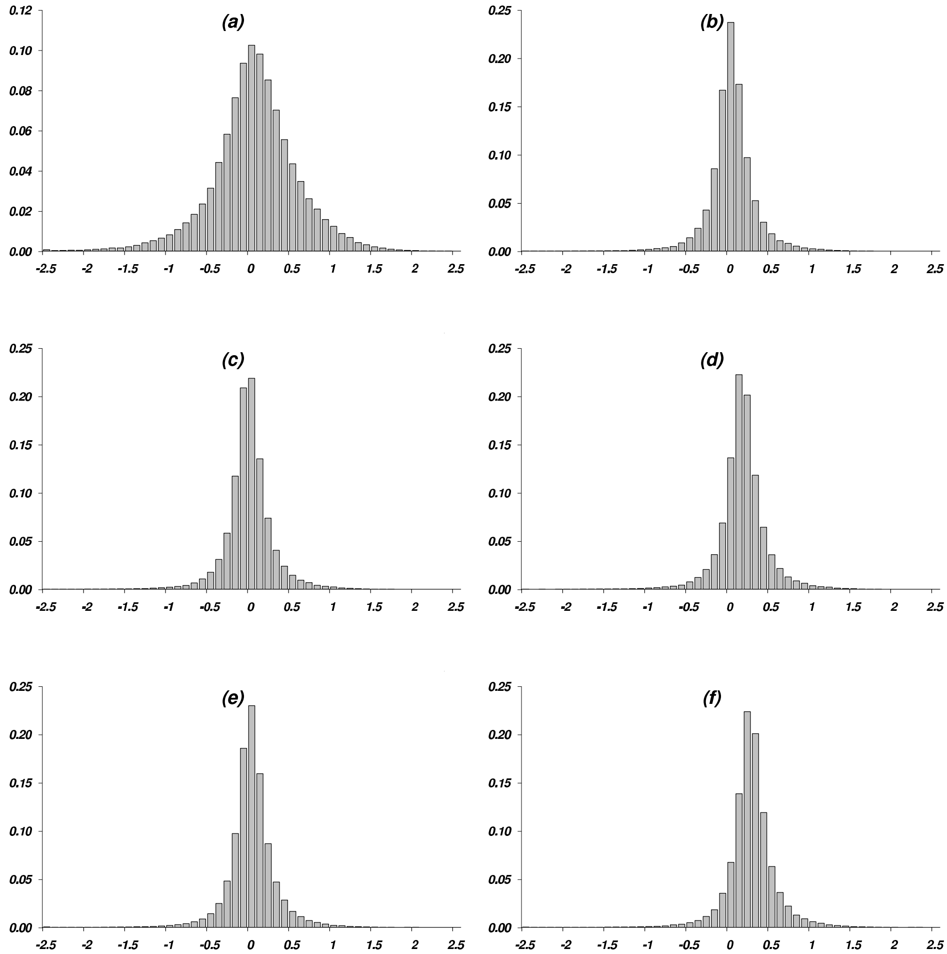


Figure 4.3: Final profits (losses) of different hedging strategies. For plots (a), (b), (c), and (d) the underlying price simulations are done by the Heston model with the first set of parameters. Plot (a) shows the results of Black-Scholes delta hedging, plot (b) shows the results of the delta-sigma strategy obtained by the true parameters. In plot (c) the delta-sigma strategy is obtained by the parameters estimated on the simulated price path. Plot (d) shows the results of the delta-sigma super-hedging. In plots (e) and (f) the underlying price simulations are done by a contaminated Heston model. Plot (e) shows the performance of the delta-sigma strategy obtained by the parameters estimated on the simulated price path while plot (d) shows the results of the delta-sigma super-hedging.

tion price and (3) computation of the hedging portfolio coefficients. At each step, the cost (profit) due to the hedging portfolio recalibration is recorded and compounded to the maturity of the option in order to get the final loss (profit) value.

In the third experiment, the hedging portfolio is built by using the true parameter of the underlying process. In this case, no parameter misspecification is considered. According to the results in Tables 4.3 and 4.4 one can say that the correctly specified delta-sigma hedging performs much better than delta hedging. Indeed, the hedging error standard deviation is substantially reduced. This result is confirmed also by the comparison between plot (a) and plot (b) of Figure 4.3.

However, the true parameters are not known by the investor that has to estimate them. This introduces a misspecification phenomenon due to a potential estimation problem. To investigate it, in the fourth experiment the parameters for the hedging portfolio are estimated in a sample path simulated by using the true parameters. The estimation results are in Table 4.2. The parameter misspecification produces a worsening especially of the frequency of losses and of the quantiles of the hedging error distribution.

The parameter misspecification impact can be slashed by the super-hedging strategy introduced in Section 4.3.1. To compute the value of a^+, b^+ and c^+ we assume that the misspecification involves only the drift parameters of the instantaneous variance process. Indeed, as shown in Chapter 3 the estimation of the diffusion parameter of the variance process is very precise. Moreover, in several early instances the influence exercised by δ and ρ on option pricing and hedging was shown to be small. This simplification allows us to get a closed-form solution for the first derivative with respect to the underlying price and to the instantaneous variance.

The fifth experiment is carried out by using the super-hedging parameters (see again Table 4.2). The super-hedging parameters are obtained by the interval estimation proposed in Chapter 3. For this purpose, an 80% confidence is used. Due to the large standard error, especially for the mean reverting parameter, the 80% confidence is a good trade-off between hedging safety and credibility of parameter's bound values. Indeed, in many cases a higher confidence could lead to a negative lower bound for the mean reverting parameter.

The delta-sigma super-hedging approach effectively reduces the frequency of losses and the percentiles of the left tail. It is clear from plot (d) of Figure 4.3 that the hedging error distribution is shifted to the right. On the whole, the delta-sigma super-hedging approach solves the parameter misspecification problem. However, to assess its real effectiveness, our approach has to be compared with alternative super-hedging strategies. The one proposed under the ALP approach is based on the solution of a non-linear PDE. However, when the option to hedge has a monotonic first derivative with respect to the underlying price (such as a call option), the PDE has the Black-Scholes solution with the volatility equal the upper bound of the possible volatility values. Let us call this value σ_{max}.

The sixth experiment involves the ALP approach. A proper comparison can be done only by using a value of σ_{max} such that the initial cost equals the initial cost of the delta-sigma super-hedging strategy. Tables 4.3 and 4.4 show that the ALP approach is clearly dominated by the delta-sigma super-hedging approach.

So far only parameter misspecification is considered. Indeed, we made the assumption that the Heston model is the true model for the underlying price dynamics. Eraker, Johannes and Polson [22] show that the Heston model is misspecified. They state that the misspecification can be remarkably reduced by introducing jumps in the instantaneous variance process.

In order to show whether the proposed hedging strategy is effective also with model misspecification, the next experiments are done by simulating the price paths by a contaminated Heston model. The contamination is caused by assuming that the error term of the instantaneous variance process is no longer distributed as a standard normal but it is distributed as a mixture of normal distributions. By assuming that $x \sim N(\mu_x, \sigma_x^2)$, $y \sim N(\mu_y, \sigma_y^2)$ and $\lambda \sim \text{Bern}(p)$, the new error term $z = \lambda y + (1 - \lambda)x$ is distributed as a mixture of normal distributions. In this application, the contamination percentage p is equal to 0.02, both y and x have zero mean and variance respectively equal to 9 and $\frac{1-9p}{1-p}$, so that the new error term z has mean 0 and variance 1.

As shown in Figure 4.2, the contamination of the instantaneous variance process produces some jumps in the implied volatility path. However, since the investor persists in considering the Heston model, both the estimation and the hedging strategy are based on this model. The estimated parameters are in Table 4.2.

The seventh experiment is conducted by using the parameter estimated on the simulated path. The hedging performances are similar to those obtained under parameter misspecification only. Also in this case the right translation produced by the delta-sigma super-hedging reduces the problem quite radically. Indeed, the frequency of losses of the eighth experiment is reduced to a level lower than that obtained without model misspecification. This is due to the rise of the estimation standard error, especially for the long run variance parameter. Hence, the value of ζ used for the super-hedging strategy with model misspecification is higher than the value of ζ considered with parameter misspecification only.

On the contrary, the super-hedging based on the ALP approach seems to give unsatisfactory results. Indeed, the value of the percentiles indicates that the left tail of the hedging error distribution is still too high.

4.5 Conclusion

This chapter pursues an analysis aimed to assess pricing and hedging reliability of the model proposed in Chapter 2 and estimated in Chapter 3. The analysis concerns real data for pricing purposes and simulations for hedging purposes.

The pricing reliability is assessed by comparing buyer's and seller's prices obtained by the model with the market price of the option. The comparison shows that the market price of the option is almost always between the two price bounds defined by the super-hedging approach under stochastic volatility assumption. Moreover, the two price bounds are very realistic since they are not too far from market price. The comparison involves also the super-hedging price obtained by Avellaneda et al. [5] under the assumption of uncertain volatility. The prices obtained in this case are very far from the market price and cannot be considered realistic.

A super-hedging approach under stochastic volatility assumption is proposed. This is based on the delta-sigma hedging introduced by Scott [57] and uses a second option written on the same underlying. The super-hedging approach is introduced primarily to solve the parameter misspecification problem due to the low precision in stochastic volatility estimation.

In a simulation framework, a hedging error distribution is obtained for different hedging approaches. Under Heston model assumptions, the delta-sigma super-hedging approach is able to solve the parameter misspecification problem. This approach dominates also the ALP approach that is unable to reduce sensibly the hedging error standard deviation.

Moreover, the different hedging approaches are compared under model misspecification. In this case, too, the delta-sigma super-hedging improves noticeable the hedging performances both with respect to the standard delta-sigma hedging and with respect to the ALP approach.

Part II

VaR computation and non-linear portfolios

Chapter 5

VaR estimation methods for non-linear portfolios

The increase in the dimension of traded portfolios and the rise in market volatility have made market risk measurement an even more significant and challenging issue for a financial institution. Value at Risk (VaR) has become one of the most used instruments to measure market risk both for regulatory purposes and for internal control motivations. Defined as the maximum portfolio loss that the institution can have with a certain probability and within a time interval, VaR is attractive because it summarizes in one single number a complex market risk exposure.

Different methods are in general used to estimate the VaR. Many of these methods are based on some assumptions on the asset returns distribution, while some others give a sort of non-parametric VaR estimate. Neither the first kind of models nor the second one is free of drawbacks.

The parametric methods strongly depend on the hypothesis made for the return distribution that is often assumed to be normal. As already noted by Mandelbrot [49] and Fama [23] the normality assumption is unrealistic since the return distribution seems to be more fat-tailed than the normal. As documented by a wide literature[1] the fat-tail problem produces an underestimation of the VaR.

Another problem emerges when the portfolio contains option positions. Indeed, in such a situation the relation between the option price and the underlying price is non-linear and it is not clear, given the distribution of the underlying price, what the distribution of the option returns is.

On the other hand, the non-parametric approaches implicitly assume that the returns are identically and independently distributed (hereafter iid). This assumption, too, is violated by the evidences that volatility changes over time. As pointed out by Hendricks [39] and by McNeil and Frey [50] the wrong assumption of iid returns leads to an inconsistent VaR estimation.

[1] See Duffie and Pan [18] for an overview on the subject.

Barone-Adesi, Bourgoin and Giannopoulos [8] propose a semi-parametric method to deal with non-iid returns. The main idea is to standardize the returns by assuming a model for their volatility, in order to obtain an iid series of standard residuals. Inside the series, the bootstrap is applied to obtain a simulated distribution for standardized returns. These returns are then scaled by a variance forecast.

Pritsker [56] shows that the standard non-parametric approach based on historical data reduces the sensitivity of the VaR measure to sudden changes of risk. To mitigate the problem Boudoukh, Richardson and Whitelaw [10] suggest to relax the standard assumption that all the past returns have the same probability to occur again. They assign a higher probability to occur again to the more recent realizations.

We intend to apply something similar to the method proposed by Barone-Adesi et al. [8]. Indeed, The bootstrap is usually applied by assuming that all the standardized residuals have the same probability to occur. This would be the best thing to do if the volatility model were correctly specified. If we take into account a possible misspecification of the volatility model, we would be considering the most recent observations as more probable.

In Section 5.1 we introduce a general definition of portfolio VaR, while Section 5.2 is devoted to describing the parametric method based on the quadratic approximation of the non-linear assets. In Section 5.3 the non-parametric method based on historical data is described. The filtered historical simulation method introduced by Barone-Adesi et al. [8] is described in Section 5.4.

5.1 A general definition of portfolio VaR

In a portfolio composed by n assets, let $X_t := [X_{1,t}, X_{2,t}, \dots, X_{n,t}]' \in \mathbb{R}^n$ be the vector of the asset prices. Moreover, let us define $q \in \mathbb{Z}^n$ the vector of the number of the assets in the portfolio between t and $t + \Delta t$ and V_t the price of the portfolio at time t such that $V_t = q'X_t$. We define the vector $a \in \mathbb{R}^n$ as the vector of the weights of each asset. For the i-th asset the weight a_i is equal to $\frac{q_i X_{it}}{V_t}$.

If α is the accepted loss probability, the VaR measure between t and $t + \Delta t$ will be defined in the following way:

$$\inf_{VaR_t} \{\mathbb{P}_t[V_{t+\Delta t} - V_t < -VaR_t] \leq \alpha\} \ ,$$

or with another notation

$$\inf_{VaR_t} \left\{\mathbb{P}_t \left[V_t a' \bar{R}_{t+\Delta t} < -VaR_t\right] \leq \alpha\right\} \ , \tag{5.1}$$

where $\bar{R}_{t+\Delta t}$ is the vector of asset returns at time $t + \Delta t$.

By the above definition, it is clear that VaR depends on the available information in t, the horizon Δt, the portfolio allocation a (that is assumed constant during the VaR horizon) and the accepted loss probability α.

One of the most diffused approaches to estimate VaR is the so called delta-normal approach, where the assets returns are considered normally distributed and all the assets are assumed to be linear with respect to their underlying.

The first problem of this approach is due to the non normality of the returns, originally noted by Mandelbrot [49] and Fama [23]. Figure 5.1 shows how the standardized returns of the Swiss Market Index and the Standard & Poor 500 cannot be considered normally distributed. The evidence accrues mostly by the QQ-plot where the plot diverges from the dashed line especially on the tails. This phenomenon is sometimes called heavy-tails or fat-tails problem and induces a VaR underestimate also in portfolios with linear positions only.

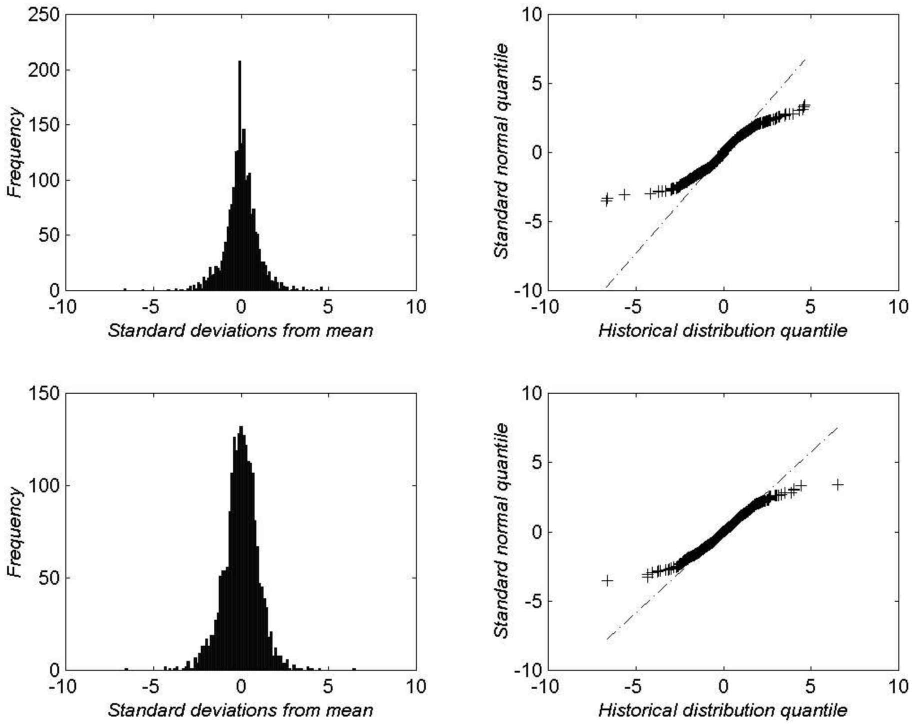


Figure 5.1: Frequency plot compared with the standard normal density and QQ-plot of the same historical distribution with respect to the standard normal distribution. The sample is composed by the standardized log-return of the Standard & Poor 500 (above) and of the Swiss Market Index (below) from May 1993 to May 2001. Source: Datastream.

The non-linearity of some positions increases the difficulty in the application of any analytical approach to estimate VaR. Indeed, even under normality of the underlying risk factors, the portfolio distribution is no longer normal.

To solve the non-linearity problem some methods can be used. Among them one can distinguish those based on approximation of the function between non-linear assets and underlying risk factors from those based on the simulations and on a full valuation of the non-linear positions.

5.2 The quadratic approximation

A frequently used hypothesis to compute VaR is to assume that the returns are jointly normally distributed. In spite of this assumption, when there are option positions in the portfolio, portfolio returns are no longer normally distributed. A way to face the problem is to substitute the non-linear relation between the option returns and the underlying returns by a quadratic approximation.

In a one-dimension Black-Scholes framework, let us call S_t the price of the underlying asset in t and X_t the price of an option at the same time. In this framework, the underlying asset is the only risk factor of the option. Hence, for a small time interval Δt, the option price at time $t + \Delta t$ can be approximated as:

$$X_{t+\Delta t} = X_t + \frac{\partial X}{\partial S}(S_{t+\Delta t} - S_t) + \frac{1}{2}\frac{\partial^2 X}{\partial S^2}(S_{t+\Delta t} - S_t)^2 + o(|S_{t+\Delta t} - S_t|^2) \ .$$

By defining $R_{t+\delta t} = \frac{X_{t+\delta t} - X_t}{X_t}$ and $\tilde{R}_{t+\delta t} = \frac{S_{t+\delta t} - S_t}{S_t}$, the above equation can be written as

$$R_{t+\Delta t} = \frac{S_t}{X_t}\frac{\partial X}{\partial S}\tilde{R}_{t+\Delta t} + \frac{1}{2}\frac{S_t^2}{X_t}\frac{\partial^2 X}{\partial S^2}\tilde{R}^2_{t+\Delta t} + o(|S_{t+\Delta t} - S_t|^2) \ . \qquad (5.2)$$

The ratios $\frac{\partial X}{\partial S}$ and $\frac{\partial^2 X}{\partial S^2}$ are called respectively *delta* and *gamma*, while the last term is the error made by the approximation.

Remark 5.2.1 In the delta-normal approach the Taylor series expansion is arrested at the first order. The advantage is that the normality assumption of the underlying returns is transferred to the option returns also. The drawback is that this assumption produces a wrong VaR value. In particular, for short option positions the VaR is strongly underestimated. Indeed, in this situation the option function is concave and the linear approximation always lies above the option function. On the contrary, for a long option position the linear approximation always lies below the option function leading to an overestimated VaR.

The delta-gamma approach replaces the option function with a quadratic function. This substitution produces an underestimated VaR for long option positions and an overestimated VaR for short option positions[2]. □

The above equations are approximately true in a neighborhood of S_t. When one faces large variations they are not in general a good approximation. Intuitively, the problem is that by estimating VaR one is interested in large variations on the risk factor while the approximation is close to be true for small variations only.

Moreover, also by assuming the conditional normality of the underlying returns, from equation (5.2) we cannot easily say anything about the distribution of the derivative returns. The most common ways to take into account the non normality of the derivatives returns are three:

- The Cornish-Fisher expansion corrects the normal critical value of the α-percentile to deal with the kurtosis and the skewness of the option return distribution. The use of this approach is explained in Zangari [64].
- The Johnson transformations approach is based on the matching of the option return first four moments with a distribution belonging to the Johnson distributions family. The use of this approach is summarized in Zangari [65].
- The Fourier transform approach is based on the inversion of the characteristic function of the approximated derivative returns. By the inversion of the Fourier transform, one can get the distribution of derivative return and obtains the required quantile.

In what follows the used approach is the last one. The advantage is that by Fourier inversion we get the exact distribution of approximated option returns. Moreover, some recent papers show the superiority of this methodology with respect to the others[3].

5.2.1 The multivariate framework

Let us define the underlying return vector as

$$\tilde{R}_{t+\Delta t} = \left[\tilde{R}_{1,t+\Delta t}, \tilde{R}_{2,t+\Delta t}, \ldots, \tilde{R}_{m,t+\Delta t}\right]' \ .$$

By equation (5.2) the portfolio return can be approximated in the following way:

$$R_{t+\Delta t} := a'\bar{R}_{t+\Delta t} \approx a'C\tilde{R}_{t+\Delta t} + \tilde{R}'_{t+\Delta t}\tilde{a}'A\tilde{R}_{t+\Delta t} \ ,$$

[2]See El-Jahel et al. [19] for more details.

[3]See among others Mina and Ulmer [52].

where $C \in \mathbb{R}^{n\times m}$ is the matrix of the first term approximation coefficients, $A \in \mathbb{R}^{n\times m}$ is the matrix of the second term approximation coefficients and $\tilde{a} \in \mathbb{R}^{n\times m}$ is a matrix composed by the portfolio weights.

Let us denote by i a generic row of a matrix and j a generic column of the same matrix. A general rule for the construction of the matrixes C, A and $\tilde{a}$ is to put the coefficients of the i-th option which depends on the j-th underlying on the element i, j of each matrix. Hence, the coefficients of the options written on the same underlying will be on the same column. Moreover, the linear part of the portfolio has to be a zero coefficient in both matrix A and $\tilde{a}$.

To be more general let us introduce a constant term in the above approximation such that[4]

$$R_{t+\Delta t} = a'K + a'C\tilde{R}_{t+\Delta t} + \tilde{R}'_{t+\Delta t}\tilde{a}'A\tilde{R}_{t+\Delta t} \ ,$$

where $K \in \mathbb{R}^n$ is a vector of constants. With a more compact notation we can write the above equation in the following way:

$$R_{t+\Delta t} = \kappa + c'\tilde{R}_t + \tilde{R}'_t B\tilde{R}_t \ , \tag{5.3}$$

where $\kappa := a'K$, $c := C'a$ and $B := \tilde{a}'A$. We assume that B is symmetric[5].

Let us assume as usual that the returns of the risk factors follow a multivariate conditional normal distribution

$$\tilde{R}_{t+\Delta t} \sim N(0, \Sigma) \ ,$$

where $\Sigma \in \mathbb{R}^{m\times m}$ is a positive definite symmetric matrix. The assumption of a zero mean is not so strong for a short horizon. Moreover, for some authors such an approximation performs better than an estimate based on historical data[6].

Definition 5.2.1 *Let Y be a random variable and u a real number, we call characteristic function of Y the expected value of e^{iuY}.* □

If we assume that $R_{t+\Delta t}$ is a continuous random variable with probability density function $f(r)$, the characteristic function can be written as

$$\varphi_R(u) = E[e^{iuR}] = \int_{-\infty}^{+\infty} e^{iur} f(r)dr \ ,$$

where the last integral is called the Fourier transform.

[4] Later on, we will use the equality by neglecting the error term.

[5] When B is not symmetric Feuerverger and Wong [24] suggest to substitute it by $\frac{1}{2}(B + B')$.

[6] See among others Figlewski [25].

Proposition 5.2.1 *If we assume that $\tilde{R}_{t+\Delta t}$ has a multivariate normal distribution with mean zero and covariance matrix Σ, then the random variable $R_{t+\Delta t}$ will have the following characteristic function*

$$\varphi_R(u) = |D|^{-1/2} \exp\left[iu\kappa - \frac{1}{2}u^2 c' D^{-1} \Sigma c\right] , \tag{5.4}$$

where $D := I - 2iu\Sigma B$. □

See appendix A.3 for a proof.

Remark 5.2.2 When $B = 0_m$ i.e. when the assets are linear with respect to the risk factors, the above characteristic function becomes

$$\varphi_R(u) = \exp\left[iu\kappa - \frac{1}{2}u^2 c' \Sigma c\right] ,$$

where $c'\Sigma c$ is the portfolio variance. As expected, in this framework we obtain the characteristic function of a multivariate normal distribution. □

The probability density function $f(r)$ of the random variable $R_{t+\Delta t}$ can be obtained by the result of the Fourier inversion theorem

$$f(r) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iur} \varphi_R(u) du .$$

Moreover, we can also obtain the distribution function $F(r)$ of the same variable by the following theorem[7].

Theorem 5.2.1 *Let $f(r)$ and $\varphi_R(u)$ be Lebesgue-integrable, if the mean and variance of the random variable $R_{t+\Delta t}$ exist, then its distribution function $F(r)$ will be*

$$F(r) = \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} Re\left[\frac{e^{-iur}\varphi_R(u)}{iu}\right] du , \tag{5.5}$$

where $Re[g(u)]$ is the real part of $g(u)$. □

Recalling that $F(r) = \mathbb{P}_t[R_{t+\Delta t} < r]$, we can guess that if $F(r) = \alpha$ then r will be equal to $-\frac{VaR}{V_t}$. Hence, one can obtain the portfolio VaR by numerically solving the following equation:

$$\frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} \mathrm{Re}\left[\frac{1}{iu} \exp\left(iu\frac{VaR}{V_t}\right) \varphi_R(u)\right] du = \alpha . \tag{5.6}$$

[7] See Shephard [58].

5.3 Non parametric method based on historical simulations

By taking a time series of data of size T as a sample of the whole population it is possible to get an empirical distribution of portfolio returns. Indeed, at every date of the sample the past risk factor returns are used to revalue the portfolio and to get the empirical distribution of its returns. VaR is the α-quantile of this distribution.

Without loss of generality, we will assume $\Delta t = 1$. The idea is to replace in equation (5.1) the theoretical probability by an empirical frequency:

$$\frac{1}{T}\sum_{k=0}^{T-1} \mathbf{1}_{\{V_t R_{t-k} < -VaR\}} = \alpha \ ,$$

such that VaR can be obtained by the following minimization problem:

$$\widehat{VaR}_t = \arg\min_{VaR} \left[\left(\frac{1}{T}\sum_{k=0}^{T-1} \mathbf{1}_{\{-(V_t R_{t-k} + VaR) > 0\}} \right) - \alpha \right]^2 \ . \tag{5.7}$$

We can see equation (5.7) as a non-linear last square regression

$$\alpha = \mathbf{1}_{\{-(V_t R_{t-k} + VaR) > 0\}} + \varepsilon_{t-k} \qquad\qquad k = 0, 1, \dots, T-1$$

where VaR is now the true VaR and $\varepsilon_{t-k} \sim \text{iid}(0, \sigma_\varepsilon)$. Summing and dividing by the sample dimension we get

$$\alpha = \frac{1}{T}\sum_{k=0}^{T-1} \mathbf{1}_{\{-(V_t R_{t-k} + VaR) > 0\}} + \frac{1}{T}\sum_{k=0}^{T-1} \varepsilon_{t-k} \ .$$

Note that the last term on the above equation goes to zero if some version of the law of large number applies. In this case, the first sum of the right hand side of the above equation defines a consistent estimator for α.

5.3.1 Generalized historical simulation method

The iid assumption allows us to say that every realized return has the same probability to occur again. Indeed, equation (5.7) does not differentiate among all the past realizations which are equally weighted independently on the time they are occurred. As noted by Pritsker [56], giving the same probability to occur for each return reduces the sensitivity of the VaR measure to the changes of risk due to market crashes.

To face the problem Boudoukh, Richardson and Whitelaw [10] suggest to give different weights to different realized returns such that equation (5.7) becomes

$$\widehat{VaR}_t = \arg\min_{VaR} \left[\left(\sum_{k=0}^{T-1} p_k \mathbf{1}_{\{-(V_t R_{t-k} + VaR) > 0\}} \right) - \alpha \right]^2 \ , \tag{5.8}$$

where p_k is the weight and $\sum_k^{T-1} p_k = 1$.

In particular they suggest to give a higher weight to the most recent realizations by using a weight that decays with time

$$p_k = \left(\sum_{i=0}^{T-1} \lambda^{i-1}\right)^{-1} \lambda^{k-1} = \frac{1-\lambda}{1-\lambda^T}\lambda^{k-1} \qquad k = 0, 1, \ldots, T-1$$

where $\lambda \in (0,1)$ is called the decay factor. Note that for λ equal to one we have equation (5.7).

Moreover, we can say that $\lambda p_k = p_{k-1}$. This shows that the lower λ the higher the decay effect on the weights associated with far returns. This should increase the VaR sensibility to market crashes or in general to risk raises.

5.3.2 Some hidden drawbacks

Because of the ease of implement and of the absence of an explicit model assumption on the risk factors returns, the historical simulation approach has become very popular. There are some drawback, though. One of the most relevant concerns the revaluation of the options. Indeed, for this purpose it is necessary to introduce some assumptions on the underlying returns distribution that partially reduce one of the two main benefits of the approach.

A second problem is that a lot of data have to be used, otherwise the empirical distribution is not properly defined on the tails. This drawback is common to all the methodology based on simulations, but it is more evident for historical simulations since the data available for some assets can be very few.

5.4 Filtered historical simulation

The assumption that sample returns are iid is violated by the evidence that the volatility changes over time. This causes an inconsistent estimation of VaR[8].

Barone-Adesi, Bourgoin and Giannopoulos [8] introduce a method to face the problem. The approach is based on historical data. Indeed, the aim is to have a sort of independence from the assumptions on the risk factors distribution. In this framework, the iid realizations are obtained by filtering the data with a predefined model for returns.

The models generally assumed belong to the ARCH class. Hereafter, we will assume that returns have some GARCH(1,1) errors. Let us describe the idea of the method in the one dimensional framework. For the j-th asset

[8]The same idea is discussed by Hendricks [39] and by McNeil and Frey [50].

the return dynamics is described by the following model:

$$\begin{aligned} R_{j,t} &= \varepsilon_{j,t} \qquad\qquad j = 1, 2, \ldots, m \ , \\ \varepsilon_{j,t} &\sim N(0, h_{j,t}) \ , \\ h_{j,t} &= \alpha_0 + \alpha_1 R^2_{j,t-1} + \alpha_2 h_{j,t-1} \ , \end{aligned} \tag{5.9}$$

To apply the method one has to consider the following steps:

1. collect a set of observed daily returns $R_{j,t}$ for $t = 1, 2, \ldots, T$;
2. estimate equation (5.9) to have the estimated variance $\hat{h}_{j,t}$ for each time t;
3. define the standardized residuals in the following way:

$$e_{j,t} := \frac{R_{j,t}}{\sqrt{\hat{h}_{j,t}}} \qquad\qquad t = 1, 2, \ldots, T$$

4. to generate the k-th simulation pick randomly (with replacement) one of the T standardized residuals (let us define it e^k);
5. forecast the variance of the period $T + 1$ by equation (5.9);
6. define the simulated innovation forecast for time $T + 1$ as

$$z^k_{j,T+1} := e^k \sqrt{\hat{h}_{j,T+1}} \ ,$$

7. define the $T + 1$ simulated risk factor price as

$$S^k_{j,T+1} := S_{j,T}(1 + z^k_{j,T+1})$$

8. by using $S^*_{j,T+1}$ calculate the simulated portfolio price and then the simulated portfolio returns.

By repeating the procedure from step 4 to step 8 one can obtain a simulated probability density function for the one day returns that may be used to calculate the VaR of the portfolio.

Remark 5.4.1 The approach can be extended to consider a time horizon longer than one day. Indeed, $z^k_{j,T+1}$ can be used to forecast $\hat{h}_{j,T+2}$:

$$\hat{h}_{j,T+2} = \alpha_0 + \alpha_1 (z^k_{j,T+1})^2 + \alpha_2 \hat{h}_{j,T+1} \ .$$

Then, by picking randomly a second standardized residual, say e^k_2, it is possible to simulate the innovation for time $T + 2$:

$$z^k_{j,T+2} := e^k_2 \sqrt{\hat{h}_{j,T+2}} \ ,$$

that is used to simulate the risk factor value at $T+2$:

$$S^k_{j,T+2} = S^k_{j,T+1}(1 + z^k_{j,T+2}) \ .$$

The steps are iterated until the time horizon τ was reached. Hence, the simulated risk factor value in $T+\tau$ is:

$$S^k_{j,T+\tau} = S_{j,T} \prod_{i=1}^{\tau} \left(1 + e^k_i \sqrt{\hat{h}^k_{j,T+i}}\right) \ .$$

□

5.4.1 The model misspecification

In step 4 the bootstrap from the sample can be performed by assuming that all the standard residuals are equally probable. Therefore, the random date can be picked from a uniform distribution. Indeed, the use of a uniform distribution is the best one can do since one is sure that the standardized residuals are iid, or, otherwise, that model (5.9) is the true model for the market volatility.

If we consider the possibility that the model could be misspecified, we will not be sure about the iid of the standardized residuals. In such a situation, to use a uniform distribution might not be the best way to sample the random date. As in Section 5.3.1, to increase the sensibility of the VaR measure to risk changes, we can impute a higher probability of occurrence to the standard residuals obtained by the most recent returns. This can be done by extracting the random date from an exponential instead of a uniform distribution.

5.4.2 The multivariate framework

In the multivariate extension of the above method no variance-covariance matrix is used. Indeed, the bootstrap is not directly applied on the residual returns, but on the past states of the world. A state of the world is represented by the m underlying returns observed at a certain date.

In practice, one has to construct a $T \times m$ matrix where in the column there are the time series of the each underlying returns. The returns have to be standardized with the procedure described above. Note that in spite we are working in a multivariate framework the GARCH model for the returns filtration is a univariate GARCH for each time series.

The bootstrap is carried out by picking randomly a row vector from the standardized return matrix. Now $\bar{e}$ is an m-vector and it is used to generate the simulated innovation as in step 6. Since all the standardized returns relative to the same date are picked at the same time, the correlation among different underlying is preserved. The portfolio is revalued for each randomly picked state of the world so as to obtain an empirical distribution. From the empirical distribution the VaR estimate is obtained by the same techniques which are used for the historical simulation.

In order to apply the filtered historical simulation method, one does not need to estimate the variance-covariance matrix of returns. However, the iid assumption of the standardized residuals implies that their correlations are constant over time. For a long data set, Pritsker [56] shows that the differences in correlations are significant from a statistical as well as an economic point of view.

The same author shows that the reduction of the sample size can mitigate the problem. However, a too short data set could lead to a VaR underestimate due to the potential lack of extreme observations. Hence, one needs to find a proper data set size in order to manage the trade-off between stationarity property and the number of extreme observations considered.

Chapter 6

Testing VaR estimation methods

In the last few years, financial institutions have the possibility to choose among a variety of VaR estimation methods. The choice has to be consistent with the dimension and the composition of the trading book. However, in order to make this choice properly, a comparison between the different VaR estimation methods has to be done.

There is a lot of literature that tries to evaluate the accuracy of the different methods, especially for equity portfolios[1]. However, most of this literature does not consider the effects on VaR estimation of the introduction of non-linear positions in the portfolio.

The aim of this chapter is to test the VaR estimation methods introduced in Chapter 5 for portfolios containing non-linear positions. To measure the performances of the estimation methods, they will be tested on the period between June 1999 and May 2001 in order to include the crash of the NASDAQ index.

The choice of this period is not casual. Indeed, our attention is also on the VaR sensitivity to sudden changes of risk due to market crashes. It seems natural to require this kind of sensitivity for a risk measure; yet, clearly, not all estimation methods supply this feature.

To verify the accuracy of VaR estimation methods one of the most used approaches is the so-called reality check test based on the observation of the VaR performances during a period of time. Hence, the performances are tested by an unconditional coverage test and by a conditional coverage test. While the first test assesses the goodness of the coverage of VaR value on all the sample period, the second test tries to detect the presence of a relevant time dependence among VaR performances.

Without loss of generality, hereafter the time horizon will be assumed equal to one day ($\Delta t = 1$). One of the problems in using longer time intervals

[1] See among others Hendricks [39] and Pritsker [55].

is that the assumption of a constant portfolio allocation becomes unrealistic. Moreover, a longer time horizon involves a reduction of the sample size for the test and it can lead to a reduction of the power of the test.

Section 6.1 describes the four portfolios considered for the analysis. Section 6.2 is a first glance at the VaR estimation. Section 6.3 is devoted to testing the accuracy of the VaR estimates during the test period. Comments and conclusions are in Section 6.4.

6.1 The portfolios considered

For the following analysis, four portfolios are considered. They consist of the same kind of assets (equities and options on index) but with different composition rates. A certain percentage of equities is diversified inside the American biotechnological industry. The choice of this industry is motivated by the fact that we want to consider an high volatility portfolio whose stocks are quoted in a market that suffered of some strong crashes in the recent past. It enables us to observe the sensibility of the VaR estimate with respect to an increase of risk due to a market crash.

In particular, portfolio A is completely composed by the biotech equities. The percentage of biotech equities reduces to 63% in portfolio B. The residual 37% consists of other stocks having low correlation with the biotechnological industry.

In portfolio C, biotech stocks are 56% of the total portfolio while 12% of total wealth is represented by long positions in the S&P 500 Index European put options and in the NASDAQ 100 Index European put options. Both options have four months time to maturity and are in the money. Their purpose is to reduce the market risk of the equity portfolio. Indeed, the number of put options and the strike price are chosen to reduce, in the following four months, the probability of a loss of more than 95% of the equity portfolio. The aim is to show how the hedging strategy can change the VaR value.

In portfolio D the biotech stocks are 56% of the total portfolio while 7% of total wealth is represented by short positions in the S&P 500 Index European call options and in the NASDAQ 100 Index European call options. Both options have four months time to maturity and are out of the money. They are not used for hedging purposes but for a sort of speculation that increases the expected portfolio profitability. This kind of behavior strongly increases the portfolio risk. Our aim is to show how the different VaR methods are able to detect such a risk increase.

The exact composition of the portfolios is in appendix D.

Table 6.1: VaR estimated by the methods described in the previous sections for the four portfolios. The VaR is expressed in percentage with respect to the portfolio price at time t. The time horizon is one day.

	Portfolios			
$\alpha = 0.01$	A	B	C	D
1) Parametric VaR	7.50%	6.06%	3.92%	19.76%
2) Standard Historical Simulation	10.11%	7.41%	5.41%	10.10%
3) Weighted H. S. ($\lambda = 0.99$)	9.24%	7.40%	4.95%	10.77%
4) Weighted H. S. ($\lambda = 0.97$)	7.16%	6.24%	4.17%	10.09%
5) Filtered H. S. (Uniform)	9.32%	6.90%	8.57%	16.58%
6) Filtered H. S. (Exponential)	7.75%	5.93%	8.79%	16.53%
$\alpha = 0.05$	A	B	C	D
1) Parametric VaR	5.30%	4.28%	2.62%	12.55%
2) Standard Historical Simulation	5.88%	4.44%	3.39%	6.79%
3) Weighted H. S. ($\lambda = 0.99$)	6.35%	4.69%	3.51%	6.89%
4) Weighted H. S. ($\lambda = 0.97$)	4.73%	4.53%	2.23%	5.52%
5) Filtered H. S. (Uniform)	5.22%	4.34%	5.72%	10.55%
6) Filtered H. S. (Exponential)	4.34%	4.11%	4.88%	9.73%

6.2 A first glance at the empirical results

The different VaR estimates at 30.05.2001 for the following day are summarized in table 6.1. In the application of the parametric approach for the linear portfolios A and B the assumption is only that the returns are normally distributed, while for portfolio C and D a delta-gamma approximation is used.

The variance-covariance matrixes are estimated by the so called Orthogonal GARCH(1,1)[2]. The parameters of the GARCH model are estimated by means of 300 daily observations. Moreover, in order to apply the delta-gamma approach, equation (5.6) has to be solved numerically[3].

Figure 6.1 shows the conditional probability density function of the return of portfolios C and D compared with a normal distribution. The probability density function is obtained by assuming the normality of the risk factors return and by using a quadratic approximation for the options return. The non-linear positions produce an asymmetric shape of the density function. As expected, the left tail of the distribution is higher for portfolio

[2]See among others Alexander [2] and Byström [11] for an application to Nordic stock markets during the Asian financial crisis.

[3]Note that in the left hand side of the equation there is an integral that has to be solved numerically. For this purpose a Gauss-type numerical integration with Laguerre polynomials of order 14 is used.

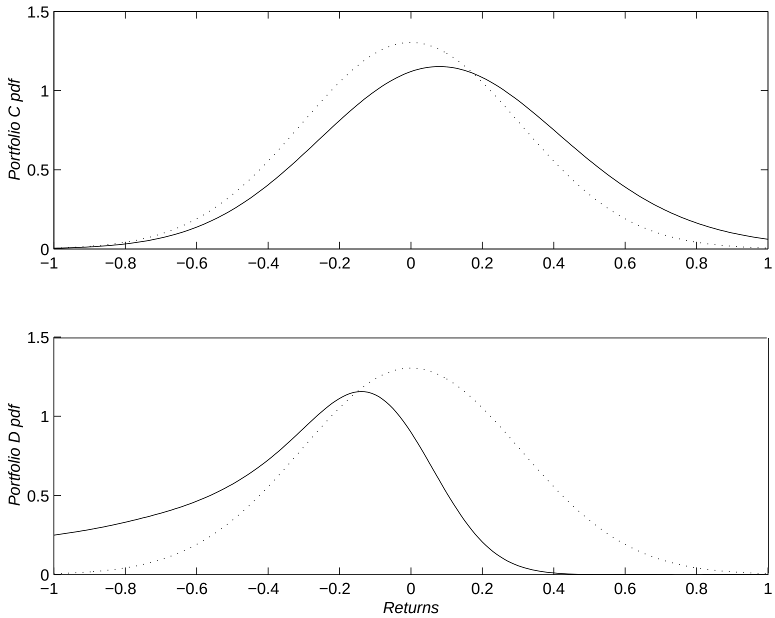


Figure 6.1: Conditional probability density function of the return of the portfolios C (above) and D (below) compared with a normal distribution (shaded line).

D than for portfolio C. This is due both to the larger risk in portfolio D and to the risk overestimation of the delta-gamma method for short option positions[4].

The historical simulations are done by using equation (5.7) with 500 past daily returns. In portfolios C and D the options are revalued by using the Black-Scholes formula with a constant volatility equal to the implied volatility at date t. The weighted historical simulation approach is performed by using two different decay factors. The first one (0.99) gives a lower decay effect than the second one (0.97). The latter yields a VaR estimate more sensitive to the changes in market risk.

In the filtered historical simulation method, the sample size is composed by 400 daily observations. The choice of this sample size is justified by the need to manage the trade-off between the covariance stationary of the standardized residuals and the possibility to include a sufficient number of extreme observations. Within the sample the standardized residuals are picked randomly 5000 times. The bootstrap is done either by picking randomly a date from a uniform distribution or by using an exponential distribution with parameter λ equal to 0.99.

[4]See Remark 5.2.1 for a description of the phenomenon.

Table 6.2: Number of failures x and their proportion $\hat{p}$ with respect to the whole sample of 500 observations (from 07.06.1999 to 30.05.2001). The time horizon is one day.

	Portfolios							
	A		B		C		D	
$\alpha = 0.01$	x	$\hat{p}$	x	$\hat{p}$	x	$\hat{p}$	x	$\hat{p}$
1) Param. VaR	6	1.2%	5	1.0%	27	5.4%	5	1.0%
2) S.H.S.	8	1.6%	9	1.8%	13	2.6%	12	2.4%
3) W.H.S. ($\lambda = 0.99$)	9	1.8%	10	2.0%	9	1.8%	6	1.2%
4) W.H.S. ($\lambda = 0.97$)	9	1.8%	10	2.0%	12	2.4%	10	2.0%
5) F.H.S. (Uniform)	7	1.4%	7	1.4%	8	1.6%	7	1.4%
6) F.H.S. (Exp.)	8	1.6%	8	1.6%	7	1.4%	8	1.6%
	A		B		C		D	
$\alpha = 0.05$	x	$\hat{p}$	x	$\hat{p}$	x	$\hat{p}$	x	$\hat{p}$
1) Param. VaR	24	4.8%	23	4.6%	41	8.2%	10	2.0%
2) S.H.S.	52	10.4%	50	10.0%	45	9.0%	33	6.6%
3) W.H.S. ($\lambda = 0.99$)	32	6.4%	32	6.4%	30	6.0%	29	5.8%
4) W.H.S. ($\lambda = 0.97$)	28	5.6%	28	5.6%	27	5.4%	28	5.6%
5) F.H.S. (Uniform)	40	8.0%	39	7.8%	30	6.0%	27	5.4%
6) F.H.S. (Exp.)	25	5.2%	24	4.8%	25	5.0%	26	5.2%

6.3 Testing the VaR estimations

To verify the accuracy of VaR estimates, one of the most used methods is the so called reality check. It is based on the observation of the VaR performances during a period of time. The observed variable is the number of VaR failures. For VaR failure we mean a loss higher than the estimated VaR. The percentage of failures with respect to the total of the considered observations, say $\hat{p}$, should be as near as possible to the defined loss probability α.

Table 6.2 shows the number of failures and their proportion with respect to the whole sample. They are obtained by estimating VaR for the different portfolios during a period of 500 days from June 1999 to May 2001. The obtained VaR is compared with the loss of the following day.

The first estimation method gives similar results for the two portfolios with linear positions only. For both portfolios this method seems to work quite well. On the other hand, the delta-gamma approach, seems to perform badly for both portfolios containing options. As expected, the quadratic approximation underestimates the risk for a long position while overestimates the risk for a short position. This is also evident from table 6.1, where the

VaR estimated by the delta-gamma approach is twice the value obtained by the historical simulations.

The standard historical simulation method performs poorly both for equity portfolios and for non-linear portfolios. This is due to the lack of sensibility of the method with respect to changes in market risk. Indeed, from figure 6.4 one can see the weak reaction of the standard historical simulation VaR to the NASDAQ crash. The performances remarkably get worse for a 5% accepted loss probability. For this accepted loss probability weighted historical simulations seems to give better performances for all the portfolios.

As expected, the filtered historical simulations seems to work very well for the accepted loss probability of 1%. Indeed, as a result of bootstrap, the extreme left tail has an accurate definition. The exponentially sampled filtered historical simulations dominates the standard filtered historical simulation method for the accepted loss probability of 5%. Moreover, due to the GARCH forecast of the variance, the filtered historical simulation methods are very sensitive to market crashes. This is evident in figure 6.4 and 6.5. Also in this case the sensitivity is higher for the modified exponential sampled approach than for the standard approach.

6.3.1 Unconditional coverage test

According to Kupiec [47], by using a likelihood ratio statistic, a test can be developed to decide whether to reject the null hypothesis that the probability of failure p is equal to α. The starting point is to assume that the sample is drawn from a Bernoulli population with two possible events: the VaR can cover the loss or the VaR is not sufficient to cover the loss. If the Bernoulli random variables are independent the probability to have x failures in a sample of size n will be given by a binomial distribution:

$$\mathbb{P}(X = x) = \begin{cases} \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x} & \text{for } x = 0, 1, \ldots, n \\ 0 & \text{elsewhere} \end{cases}$$

For the null hypothesis $p = \alpha$ and assuming that x is the observed number of failures, the relevant likelihood ratio test statistic is given by

$$W_{uc}(x, \alpha) = -2 \ln \left[\alpha^x (1-\alpha)^{n-x} \right] + 2 \ln \left[\hat{p}^x (1-\hat{p})^{n-x} \right] \ ,$$

where $\hat{p} = \frac{x}{n}$.

Under the null hypothesis, the test statistic $W(x, \alpha)$ is asymptotically distributed as a χ_1^2. The rejection region is

$$\{x \ : \ W_{uc}(x, \alpha) \geq c_\xi\} \ ,$$

where ξ is the probability of the first kind error and c_ξ is the quantile of the χ_1^2 distribution associated to the probability ξ. By assuming $\xi = 0.05$ and

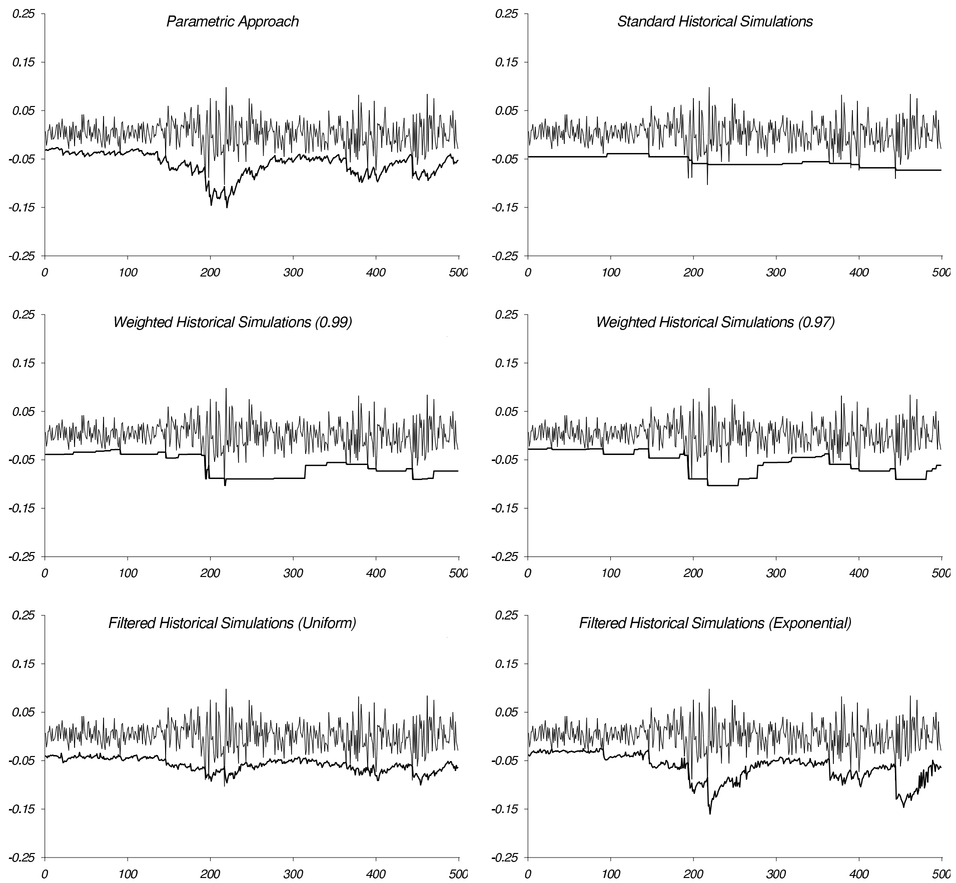


Figure 6.2: Protfolio B returns of the period from June 1999 to May 2001 compared with the 1% VaR.

a sample size of 500 observations, the non rejection region is approximately equal to

$$\begin{aligned} 1 \leq x \leq 10 \quad &\text{if} \quad \alpha = 0.01 \\ 16 \leq x \leq 35 \quad &\text{if} \quad \alpha = 0.05 \ . \end{aligned}$$

According to the Neyman-Pearson theorem, the likelihood ratio test is the uniformly most powerful test against simple alternative hypothesis. In spite of this property, the above test has poor power especially in small samples[5].

Table 6.3 shows the likelihood ratio test value and the p-value for the different estimation methods. Portfolio A is omitted since it gives the same results of portfolio B.

Some very low p-values can be noted when testing the parametric method and the standard historical simulation method applied to non-linear portfolio. This pattern is very evident especially for portfolio C. The p-values for

[5] For a close discussion on the power of the test see Kupiec [47].

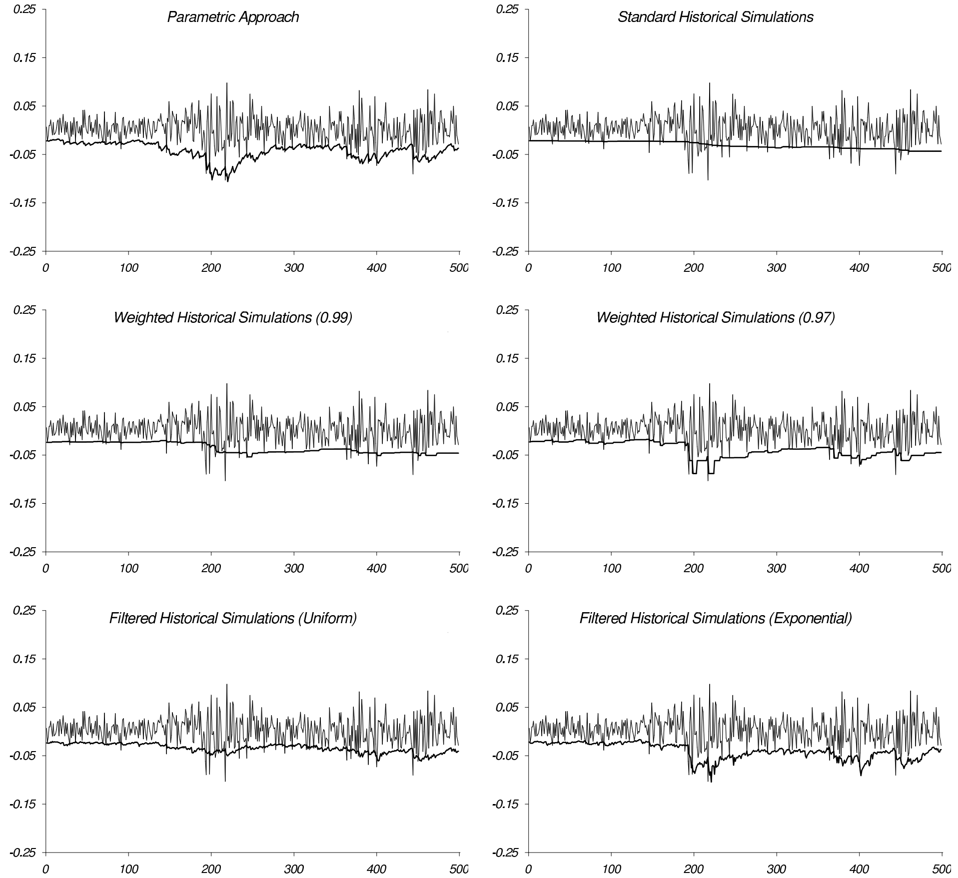


Figure 6.3: Protfolio B returns of the period from June 1999 to May 2001 compared with the 5% VaR.

portfolio D are much more uncertain. For the equity portfolio the null is never rejected for the parametric method and for the exponentially sampled historical simulations.

6.3.2 Conditional coverage test

Since the above test refers to the unconditional coverage of the VaR, it can suggest to accept the null hypothesis also when there is a clustering on VaR failures. According to Christoffensen [13], the unconditional test does not have any power against the alternative hypothesis that the failures are time dependent.

The clustering phenomenon in the failures indicates that when risk raises, the VaR is not able to quickly adapt to the new market conditions. Hence, the unconditional coverage test seems not sufficient to assess the goodness of a VaR estimation method.

Let us denote by δ_t a variable that is 1 if in t there is a failure and 0

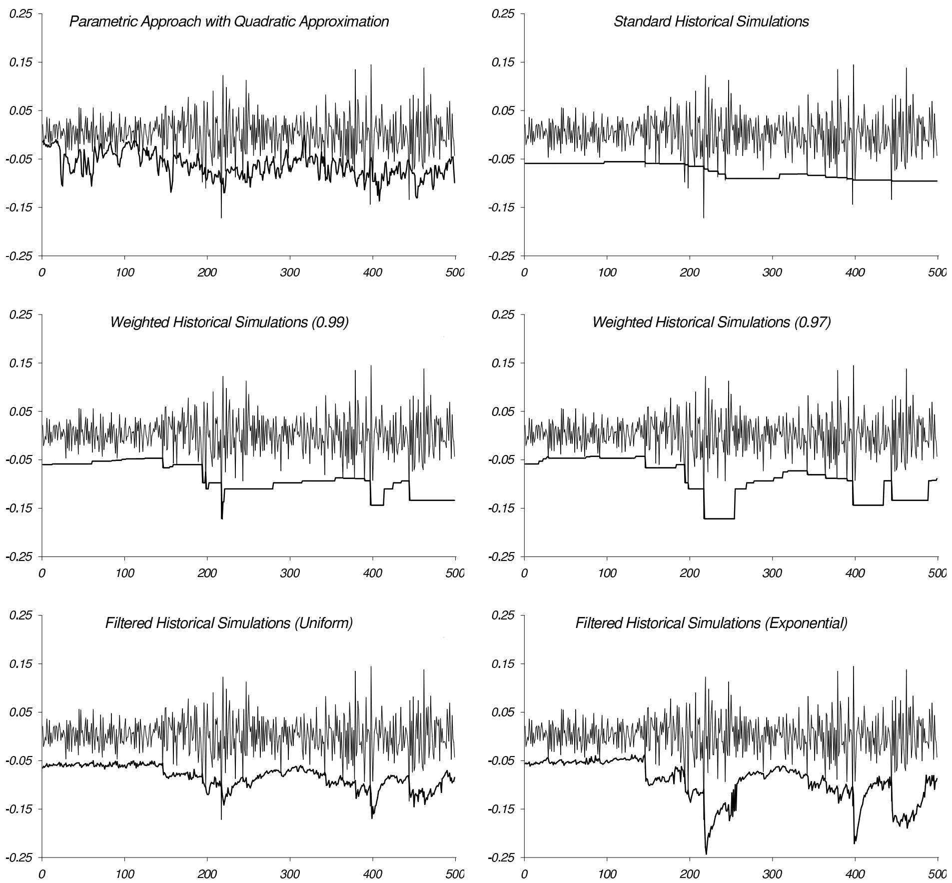


Figure 6.4: Protfolio C returns of the period from June 1999 to May 2001 compared with the 1% VaR.

elsewhere:

$$\delta_t = \mathbf{1}_{\{V_t a' \bar{R}_{t+\Delta t} < -VaR_t\}} \ .$$

Christoffensen [13] generalizes the idea of independent failures introducing the concept of VaR efficiency. The VaR at time t is efficient conditionally on the information available at $t-1$ when $E[\delta_t|\mathcal{F}_{t-1}] = \alpha$ for all t. Testing the conditional efficiency is equivalent to test that the process $\{\delta_t\}_{t=1,2,\dots,n}$ is iid as a Bernoulli with parameter α.

Remark 6.3.1 In the special case of trivial conditional information $\mathcal{F}_{t-1} = \{\emptyset, \Omega\}$, the conditional VaR efficiency becomes $E[\delta_t] = \alpha$. In this case, testing the efficiency is equivalent to the unconditional test proposed by Kupiec [47]. $\square$

Let us denote $\pi_{i,j}$ the probability that $\delta_t = j$ conditional to $\delta_{t-1} = i$. Hence, $\pi_{0,0}$ is the probability that VaR covers the loss conditional to the correct

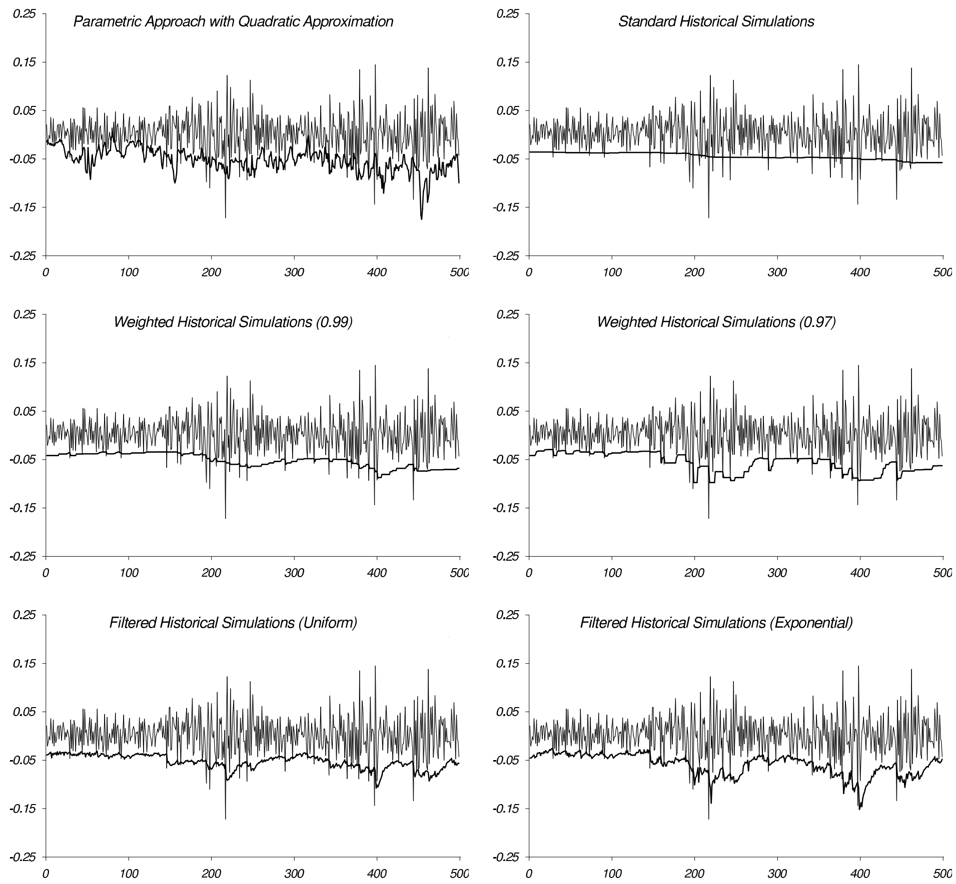


Figure 6.5: Protfolio C returns of the period from June 1999 to May 2001 compared with the 5% VaR.

coverage also in the previous period; $\pi_{1,1}$ is the probability to have a failure conditional to a failure in the previous period. In a similar way $x_{i,j}$ is the number of times that $\delta_t = j$ conditional to $\delta_{t-1} = i$.

The conditional likelihood function of δ_t is

$$L = \pi_{0,0}^{x_{0,0}}(1-\pi_{0,0})^{n-x-x_{0,0}}\pi_{1,1}^{x_{1,1}}(1-\pi_{0,0})^{x-x_{1,1}} ,$$

where $\pi_{0,0}$ and $\pi_{1,1}$ can be estimated with $\hat{\pi}_{0,0} = \frac{x_{0,0}}{n-x}$ and $\hat{\pi}_{1,1} = \frac{x_{1,1}}{x}$. The likelihood ratio statistic can be written as

$$\begin{aligned} W_{cc}(x, x_{0,0}, x_{1,1}, \alpha) &= -2\ln[\alpha^x(1-\alpha)^{n-x}] + \\ &\quad + 2\ln[\hat{\pi}_{0,0}^{x_{0,0}}(1-\hat{\pi}_{0,0})^{n-x-x_{0,0}}\hat{\pi}_{1,1}^{x_{1,1}}(1-\hat{\pi}_{0,0})^{x-x_{1,1}}] , \end{aligned}$$

and is asymptotically distributed as a χ_2^2.

The conditional coverage test was applied only for the accepted loss probability of 5%. Indeed, only in this case we have a sufficient number of

Table 6.3: Lakilewood ratio value for the unconditional test and p-value. The likelihood ratio values in the rejection region are in bold.

	Portfolios					
	B		C		D	
$\alpha = 0.01$	W_{uc}	p-value	W_{uc}	p-value	W_{uc}	p-value
Param. VaR	0.000	1.0000	**48.058**	0.0000	0.000	1.0000
S.H.S.	2.613	0.1060	**8.973**	0.0027	**7.111**	0.0077
W.H.S. (0.99)	**3.914**	0.0479	2.613	0.1060	0.190	0.6630
W.H.S. (0.97)	**3.914**	0.0479	**7.111**	0.0077	**3.914**	0.0479
F.H.S. (Unif.)	0.719	0.3966	1.538	0.2149	0.719	0.3966
F.H.S. (Exp.)	1.538	0.2149	0.719	0.3966	1.538	0.2149
	B		C		D	
$\alpha = 0.05$	W_{uc}	p-value	W_{uc}	p-value	W_{uc}	p-value
Param. VaR	0.173	0.6776	**9.110**	0.0025	**12.143**	0.0005
S.H.S.	**20.654**	0.0000	**13.755**	0.0002	2.459	0.1168
W.H.S. (0.99)	1.903	0.1678	0.992	0.3192	0.642	0.4229
W.H.S. (0.97)	0.365	0.5455	0.164	0.6852	0.365	0.5455
F.H.S. (Unif.)	**7.102**	0.0077	0.992	0.3192	0.164	0.6852
F.H.S. (Exp.)	0.043	0.8364	0.000	1.0000	0.042	0.8384

For the χ_1^2 $c_{0.05} = 3.841$ and $c_{0.1} = 2.706$.

Table 6.4: Lakilewood ratio value for the conditional test and p-value. The likelihood ratio values in the rejection region are in bold.

	Portfolios					
	B		C		D	
$\alpha = 0.05$	W_{cc}	p-value	W_{cc}	p-value	W_{cc}	p-value
Param. VaR	0.920	0.6313	0.462	0.7937	1.157	0.5608
S.H.S.	**16.532**	0.0003	**4.612**	0.0996	0.749	0.6876
W.H.S. (0.99)	0.185	0.9119	1.463	0.4812	0.420	0.8107
W.H.S. (0.97)	0.489	0.7829	0.454	0.7970	0.935	0.6265
F.H.S. (Unif.)	2.463	0.2919	0.019	0.9905	0.454	0.7970
F.H.S. (Exp.)	0.066	0.9677	0.327	0.8491	0.059	0.9711

For the χ_2^2 $c_{0.05} = 5.991$ and $c_{0.1} = 4.605$.

failures to investigate their time dependence. The test gives an idea on the sensitivity of the VaR estimation method to the rise of risk. Indeed, when a VaR estimation method is sufficiently sensitive to changes in market risk, it should react after the first failure, thus preventing a clustering of violations.

Hence, test results penalizes the poor responsiveness of the standard historical simulation method. Table 6.4 shows that for this kind of estimation method, the null is rejected both for portfolio B and C.

6.4 Conclusion

This chapter is devoted to assessing the performances of the different VaR estimation methods, by testing them on two years of daily data. The sample period considered for the test includes the NASDAQ crash. The observed variable is the number of times that the loss exceeds VaR. We investigate both the accuracy of the VaR coverage and the time dependence of the VaR failures.

The conclusions are quite different for the different portfolios considered. Indeed, for linear portfolios all the methods give quite accurate results. Unfortunately, most of the financial institutions have portfolios with strong positions on options or other non-linear instruments.

The results are visibly different for portfolios which contain option positions. Indeed, the increase of volatility due to the leverage effect of option positions, reduces the accuracy of some methods. The worst performances are obtained by parametric methods. The quadratic approximation underestimates the risk on long option positions and overestimates the risk on short option positions.

The standard historical simulation approach gives poor results both for the coverage percentage and for the failure time dependence. The latter result indicates that the standard historical simulation approach has an inadequate sensitivity to the change of market risk conditions. The generalization of the standard historical simulation strongly improves the accuracy of VaR estimate and, particularly, its sensitivity to crashes of the market.

The best performances are obtained by the semi-parametric method of the filtered historical simulations. The standardization of the returns reduces the non-iid problem while the bootstrap enables us to have a well-defined simulated distribution. At the 5% level, the filtered historical simulations with exponential sampled bootstrap dominate the uniformly sampled approach.

Appendix A

Proofs of propositions

A.1 Proof of Proposition 2.2.1

Let us define $\tilde{W}(S_t, t) := e^{-rt} W^+(S_t, t)$. Equation (2.3) becomes

$$\operatorname*{ess\ sup}_{\mathbb{Q} \in \mathcal{Q}_\theta} \; E^{\mathbb{Q}} \left[\tilde{W}(S_T, T) - \tilde{W}(S_t, t) | \mathcal{F}_t \right] = 0 \; ,$$

By applying Ito's lemma to $\tilde{W}(S_t, t)$, we obtain

$$\begin{aligned} \operatorname*{ess\ sup}_{\mathbb{Q} \in \mathcal{Q}_\theta} \; E^{\mathbb{Q}} \Bigg[\int_t^T \left(\frac{\partial \tilde{W}}{\partial u} + \mathcal{A}_u \tilde{W}(S_u, u) \right) du + \\ + \int_t^T \frac{\partial \tilde{W}}{\partial S} \sqrt{v_u} S_u \; dw^*_{1,u} + \int_t^T \frac{\partial \tilde{W}}{\partial v} \eta(v_u; \theta) \; dz_u \Bigg| \mathcal{F}_t \Bigg] = 0 \; , \end{aligned} \tag{A.1}$$

where $z_t = \rho w^*_{1,t} + \sqrt{1-\rho^2} w^*_{2,t}$ and the infinitesimal generator $\mathcal{A}_t$ is defined as

$$\begin{aligned} \mathcal{A}_t := r S_t \frac{\partial}{\partial S} + \mu_v^{\mathbb{Q}}(v_t; \theta) \frac{\partial}{\partial v} + v_t S_t^2 \frac{1}{2} \frac{\partial^2}{\partial S^2} + \\ + \frac{1}{2} \eta^2(v_t; \theta) \frac{\partial^2}{\partial v^2} + |v_t|^{1/2} S_t \rho \; \eta(v_t; \theta) \frac{\partial^2}{\partial S \partial v} \end{aligned}$$

Under Assumption 1 equation (A.1) becomes

$$\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_\theta} E^{\mathbb{Q}}\left[\int_t^T \left(\frac{\partial\tilde{W}}{\partial u} + \mathcal{A}_u\tilde{W}(S_u,u)\right)du\,\middle|\,\mathcal{F}_t\right] = 0\ ,$$

$$\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_\theta} E^{\mathbb{Q}}\left[\int_t^{t+h} \left(\frac{\partial\tilde{W}}{\partial u} + \mathcal{A}_u\tilde{W}(S_u,u)\right)du\ +\right.$$

$$\left. + \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_\theta} E^{\mathbb{Q}}\left[\tilde{W}(S_T,T) - \tilde{W}(S_{t+h},t+h)|\mathcal{F}_{t+h}\right]\,\middle|\,\mathcal{F}_t\right] = 0\ ,$$

$$\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_\theta} \frac{1}{h}E^{\mathbb{Q}}\left[\int_t^{t+h} \left(\frac{\partial\tilde{W}}{\partial u} + \mathcal{A}_u\tilde{W}(S_u,u)\right)du\right] = 0\ ,$$

By taking the limit for $h \to 0$ we have

$$\sup_\theta\left[\frac{\partial\tilde{W}}{\partial t} + \mathcal{A}_t\tilde{W}(S_t,t)\right] = 0\ .$$

Recalling the definition of $\tilde{W}(S_t,t)$ we get

$$\sup_\theta\left[\frac{\partial W^+}{\partial t} + \mathcal{A}_tW^+(S_t,t) - rW^+\right] = 0\ .$$

This is known as the Bellman differential equation for the optimal control problem (2.3)[1].

For the buyer's price the same considerations are true by substituting the supremum by the infimum.

A.2 Proof of Proposition 4.3.1

Let us consider the discounted values $\tilde{V}_t := B_t^{-1}V_t$, $\tilde{S}_t := B_t^{-1}S_t$ and $\tilde{W}_{j,t} := B_t^{-1}W_{j,t}$ for $j = 1,2$. The self-financing constraint can be written as

$$d\tilde{V}_t = a_t\ d\tilde{S}_t + c_t\ d\tilde{W}_{2,t}\ ,$$

whereas equation (2.12) becomes

$$\frac{\partial\tilde{W}_1^+}{\partial t} + \frac{\partial\tilde{W}_1^+}{\partial v}\varphi_{1,t} + \frac{1}{2}\frac{\partial^2\tilde{W}_1^+}{\partial\tilde{S}^2}v_t\tilde{S}_t^2 + \frac{1}{2}\frac{\partial^2\tilde{W}_1^+}{\partial v^2}\varphi_{2,t} + \frac{\partial^2\tilde{W}_1^+}{\partial\tilde{S}\partial v}|v_t|^{1/2}\tilde{S}_t\varphi_{3,t} = 0\ .$$

[1] See Krylov [46] among the others.

At time $s \in [t, T]$ the super hedging portfolio value is

$$\begin{aligned}
\tilde{V}_s(\phi^+) &= \tilde{W}_t^+ + \int_t^s a_u^+ \, d\tilde{S}_u + \int_t^s c_u^+ \, d\tilde{W}_{2,u} \\
&= \tilde{W}_t^+ + \int_t^s a_u^+ \, d\tilde{S}_u + \int_t^s c_u^+ \, d\tilde{W}_{2,u} + \\
&+ \int_t^s \left(\frac{\partial \tilde{W}_1^+}{\partial u} + \frac{1}{2} \frac{\partial^2 \tilde{W}_1^+}{\partial \tilde{S}^2} v_u \tilde{S}_u^2 \right) du + \\
&+ \int_t^s \left(\frac{\partial \tilde{W}_1^+}{\partial v} \varphi_{1,u} + \frac{1}{2} \frac{\partial^2 \tilde{W}_1^+}{\partial v^2} \varphi_{2,u} + \frac{\partial^2 \tilde{W}_1^+}{\partial \tilde{S} \partial v} |v_u|^{1/2} \tilde{S}_u \varphi_{3,u} \right) du
\end{aligned}$$

Moreover, by applying Ito's lemma to the discounted option price $\tilde{W}_{1,t}^+$ and using equations (2.10) and (4.5), we obtain

$$\begin{aligned}
d\tilde{W}_1^+ &= a_t^+ \, d\tilde{S} + c_t^+ \, d\tilde{W}_2 + \frac{\partial \tilde{W}_1^+}{\partial t} \, dt + \frac{1}{2} \frac{\partial^2 \tilde{W}_1^+}{\partial \tilde{S}^2} v_t \tilde{S}_t^2 dt + \\
&+ \frac{\partial \tilde{W}_1^+}{\partial v} \mu_v^{\mathbb{Q}}(v_t; \theta) \, dt + \frac{1}{2} \frac{\partial^2 \tilde{W}_1^+}{\partial v^2} \eta^2(v_t; \theta) \, dt + \\
&+ \frac{\partial^2 \tilde{W}_1^+}{\partial \tilde{S} \partial v} |v_t|^{1/2} \tilde{S}_t \rho \eta(v_t; \theta) \, dt
\end{aligned}$$

Hence, for all $s \in [t, T]$ the discounted replication error

$$\begin{aligned}
e_s(\phi^+) &= \int_t^s \frac{\partial \tilde{W}_1^+}{\partial v} [\varphi_{1,u} - \mu_v^{\mathbb{Q}}(v_u; \theta)] \, du + \\
&+ \int_t^s \frac{1}{2} \frac{\partial^2 \tilde{W}_1^+}{\partial v^2} [\varphi_{2,u} - \eta^2(v_u; \theta)] \, du + \\
&+ \int_t^s \frac{\partial^2 \tilde{W}_1^+}{\partial \tilde{S} \partial v} |v_u|^{1/2} \tilde{S}_u [\varphi_{3,u} - \rho \eta(v_u; \theta)] \, du
\end{aligned}$$

is positive if condition (2.11) is satisfied.

A.3 Proof of Proposition 5.2.1

By the properties of the positive definite matrixes we can decompose Σ as HH' such that

$$\tilde{R}_{t+\Delta t} = HZ \ ,$$

where Z is an $m \times 1$ vector of independent standard normal variables. Hence, equation (5.3) can be written as

$$R_{t+\Delta t} = \kappa + c'HZ + Z'H'BHZ \ .$$

Since B is symmetric also $H'BH$ will be symmetric. Its decomposition is

$$H'BH = G\Lambda G' \ ,$$

where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is the diagonal matrix formed from the eigenvalues of $H'BH$, while G is the matrix of the orthogonal and normalized eigenvectors of $H'BH$. Since $H'BH$ is symmetric all the eigenvalues are real. Hence, the portfolio return takes the following form:

$$\begin{aligned} R_{t+\Delta t} &= \kappa + c'HGG'Z + Z'G\Lambda G'Z \\ &= \kappa + \eta'\tilde{z} + \tilde{z}'\Lambda\tilde{z} \\ &= \kappa + \sum_{j=1}^{m} \eta_j \tilde{z}_j + \sum_{j=1}^{m} \lambda_j \tilde{z}_j^2 \ . \end{aligned}$$

where $\eta := G'H'c$ and $\tilde{z} := G'Z$ is a vector of independent standard normal variables. Let us define $h := \frac{\kappa}{m}$ and consider the random variable $x_j = h + \eta_j\tilde{z}_j + \lambda_j\tilde{z}_j^2$ for $j = 1, 2, \ldots, m$. The characteristic function of x is[2]

$$\begin{aligned} \varphi_x(u) &= E[e^{iu(h+\eta\tilde{z}+\lambda\tilde{z}^2)}] \\ &= \frac{e^{iuh}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[iu(\eta\tilde{z} + \lambda\tilde{z}^2)\right] \exp\left[-\frac{\tilde{z}^2}{2}\right] \, d\tilde{z} \\ &= \frac{e^{iuh}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\tilde{z}^2(1-2iu\lambda) - 2iu\eta\tilde{z}\right)\right] \, d\tilde{z} \\ &= \exp\left[iuh - \frac{u^2\eta^2}{2(1-2iu\lambda)}\right] \times \\ &\quad \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\tilde{z}\sqrt{1-2iu\lambda} - \frac{iu\eta}{\sqrt{1-2iu\lambda}}\right)^2\right] \, d\tilde{z} \end{aligned}$$

by defining the variable y as $\tilde{z}\sqrt{1-2iu\lambda}$, the above integral can be solved in the following way:

$$\begin{aligned} \varphi_x(u) &= \frac{1}{\sqrt{1-2iu\lambda}} \exp\left[iuh - \frac{u^2\eta^2}{2(1-2iu\lambda)}\right] \times \\ &\quad \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(y - \frac{iu\eta}{\sqrt{1-2iu\lambda}}\right)^2\right] \, dy \\ &= \frac{1}{\sqrt{1-2iu\lambda}} \exp\left[iuh - \frac{u^2\eta^2}{2(1-2iu\lambda)}\right] \ . \end{aligned}$$

[2]To simplify the notation the subscript j is left out.

Since the variable x_j for $j = 1, 2, \ldots, m$ are independent, the characteristic function of $R_{t+\Delta t}$ is

$$\begin{aligned} \varphi_R(u) &= \left(\prod_{j=1}^{m} \frac{1}{\sqrt{1-2iu\lambda_j}} \right) \exp\left(iu\kappa - \frac{1}{2} \sum_{j=1}^{m} \frac{u^2\eta_j^2}{1-2iu\lambda_j} \right) \\ &= |I - 2iu\Lambda|^{-1/2} \exp\left[iu\kappa - \frac{1}{2}u^2\eta'(I-2iu\Lambda)^{-1}\eta \right] \end{aligned} \quad \text{(A.2)}$$

where

$$\begin{aligned} |I - 2iu\Lambda| &= |G(G'IG - 2iuG'\Lambda G)G'| \\ &= |I - 2iuH'BH| \\ &= |I - 2iu\Sigma B| \ , \end{aligned}$$

and

$$\begin{aligned} \eta'(I-2iu\Lambda)^{-1}\eta &= \eta'G'[G(I-2iu\Lambda)G']^{-1}G\eta \\ &= c'H(I-2iuH'BH)^{-1}H'c \\ &= c'(\Sigma^{-1} - 2iuB)^{-1}c \\ &= c'(I-2iu\Sigma B)^{-1}\Sigma c \ , \end{aligned}$$

Hence equation (A.2) becomes

$$\varphi_R(u) = |I - 2iu\Sigma B|^{-1/2} \exp\left[iu\kappa - \frac{1}{2}u^2 c'(I-2iu\Sigma B)^{-1}\Sigma c \right] \ .$$

□

A.4 Some notes on Theorem 5.2.1

Theorem A.4.1 *Let $f(y)$ and $\varphi_Y(u)$ be Leabesgue-integrable, if the mean and variance of the random variable Y exist, then its cumulative density function $F(y)$ will be*

$$F(y) = \frac{1}{2} - \frac{1}{2\pi} \int_0^\infty \Delta_u \left[\frac{e^{-iuy}\varphi(u)}{iu} \right] du \ , \quad \text{(A.3)}$$

where $\Delta_u g(u) = g(u) + g(-u)$. □

For a proof see Shephard [58].

It remains to show that equation (A.3) and equation (5.5) are equal or equivalently that

$$\int_0^\infty Re \left[\frac{e^{-iuy}\varphi(u)}{iu} \right] du = \frac{1}{2} \int_0^\infty \left[\frac{e^{-iuy}\varphi(u) - e^{iuy}\varphi(-u)}{iu} \right] du \ . \quad \text{(A.4)}$$

Let us consider the left part of the equation (A.4). By the definition of the characteristic function we can write

$$\int_0^{\infty} Re\left[\frac{1}{iu}e^{-iuy}\int_{-\infty}^{+\infty} e^{iux} f(x)\ dx\right]\ du$$
$$\int_0^{\infty} Re\left[\frac{1}{iu}\int_{-\infty}^{+\infty} e^{i(ux-uy)} f(x)\ dx\right]\ du\ ,$$

where $f(x)$ is the probability density function. By applying Euler's rule we have

$$\int_0^{\infty} Re\left[\int_{-\infty}^{+\infty} \frac{\cos(ux-uy)f(x)}{iu}\ dx + \right.$$
$$\left. + \int_{-\infty}^{+\infty} \frac{\sin(ux-uy)f(x)}{u}\ dx\right]\ du$$
$$\int_0^{\infty}\left[\int_{-\infty}^{+\infty} \frac{\sin(ux-uy)f(x)}{u}\ dx\right]\ du\ .$$

On the other hand, the right part of the equation (A.4) can be written as

$$\frac{1}{2}\int_0^{\infty}\frac{1}{iu}\left[\int_{-\infty}^{+\infty}\left(e^{i(ux-uy)} - e^{-i(ux-uy)}\right) f(x)\ dx\right]\ du\ .$$

By applying again Euler's rule

$$\frac{1}{2}\int_0^{\infty}\left[\int_{-\infty}^{+\infty} \frac{\cos(ux-uy)+i\sin(ux-uy)}{iu} f(x) - \right.$$
$$\left. - \frac{\cos(ux-uy)-i\sin(ux-uy)}{iu} f(x)\ dx\right]\ du$$
$$\int_0^{\infty}\left[\int_{-\infty}^{+\infty}\frac{\sin(ux-uy)}{u} f(x)\ dx\right]\ du\ .$$

Appendix B

Option price under stochastic volatility

B.1 Heston option price (exact formula)

Heston's call price is given by the following formula

$$W_t = S_t P_{1,t} - K e^{-r(T-t)} P_{2,t} \ ,$$

where S_t is the underlying price in t, K is the strike price and $P_{1,t}$ and $P_{2,t}$ have the same meaning than $\Phi(d_1)$ and $\Phi(d_2)$ in Black-Scholes formula. They are obtained by an inversion theorem of the characteristic function due to Shephard [58]. Since the characteristic function, $\varphi_j(.)$ for $j = 1, 2$, is obtained analytically by Heston [40], the two probabilities can be obtained by the following one-dimension integration:

$$\begin{aligned} P_{j,t} &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{e^{-iu \ln K} \varphi_j(u)}{iu}\right] du \\ &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty e^{-u} \mathrm{Re}\left[\frac{e^{u-iu \ln K} \varphi_j(u)}{iu}\right] du \qquad \text{for } j = 1, 2 \end{aligned}$$

where the characteristic function is defined in the following exponential way:

$$\varphi_j(u) = \exp\left[C_j(T-t, u) + D_j(T-t, u) v_t + iu \ln S_t\right] \quad \text{for } j = 1, 2$$

where

$$\begin{aligned}
C_j(T-t,u) &= uir(T-t) + \frac{\zeta}{\delta^2}\left\{[\tilde{\xi} - \rho\delta(a_j + ui) + d_j] - \right. \\
&\quad \left. -2\ln\left[\frac{1 - g_j \exp(d_j(T-t))}{1-g_j}\right]\right\} \\
D_j(T-t,u) &= \frac{\tilde{\xi} - \rho\delta(a_j+ui) + d_j}{\delta^2}\left[\frac{1-\exp(d_j(T-t))}{1-g_j\exp(d_j(T-t))}\right] \\
g_j(T-t,u) &= \frac{\tilde{\xi} - \rho\delta(a_j+ui) + d_j}{\tilde{\xi} - \rho\delta(a_j+ui) - d_j} \\
d_j(T-t,u) &= \sqrt{[\rho\delta(a_j+ui) - \tilde{\xi}]^2 + \delta^2(u^2 + ui - 2a_j ui)}
\end{aligned}$$

and a_j is an indicator function equal to 1 when $j = 1$ and zero elsewhere.

The following C++ code computes the characteristic function an the second part of the integrand functions:

```
double intgrnd(double u, double S0, double K, double T, double r, double v0,
       double delta, double rho, double xi, double zeta, int intflag)
{   double b;
    complex<double> d, e, g, C, D, car;

    if (intflag == 1)
    {   b = xi - rho*delta;
        d = sqrt((rho*delta*u*i - b)*(rho*delta*u*i - b)
          - delta*delta*(u*i - u*u));
    }
    else
    {   b = xi;
        d = sqrt((rho*delta*u*i - b)*(rho*delta*u*i - b)
          + delta*delta*(u*i + u*u));
    }
    e = exp(d*T);
    g = (b - rho*delta*u*i + d)/(b - rho*delta*u*i - d);
    C = r*u*i*T + zeta/(delta*delta)*((b - rho*delta*u*i + d)*T
      - 2*log((1 - g*e)/(1 - g)));
    D = ((b - rho*delta*u*i + d)/(delta*delta))*((1 - e)/(1 - g*e));
    car = exp(C + D*v0 + i*u*log(S0));
    return (real((exp(u-i*u*log(K))*car)/(i*u)));
}
```

Since the integrand function is sufficiently regular, it can be integrated by using a Gauss type integration rule. For this particular problem, it is convenient to use Laguerre polynomial and to approximate the integral with the following sum:

$$\int_0^\infty e^{-u}\mathrm{Re}\left[\frac{e^{u-iu\ln K}\varphi_j(u)}{iu}\right]du \approx \sum_{k=1}^{n} w_k \mathrm{Re}\left[\frac{e^{u_k - iu_k\ln K}\varphi_j(u_k)}{iu_k}\right]$$

where w_k, for $k = 1, 2, \ldots, n$, are Laguerre polynomial coefficients and they depend on values of u_k[1]. A table with the value of u_k and the corresponding

[1] See equation 25.4.45 in Abramowitz and Stegun [1] for more details.

value of w_k is in Abramowitz and Stegun [1] (Table 25.9). The following C++ code computes the Heston option price according to a Laguerre polynomial of degree 12. Input parameters are `CallPutFlag`, the flag that identifies the option kind ("c" for call option and "p" for put option), the underlying price `S0`, the strike price `K`, time to maturity `T`, risk free rate `r`, the instantaneous variance `v0`, and the four parameters `delta`, `rho`, `xi` and `zeta`.

```
const int POLDEGR = 12;
const double x[POLDEGR]={0.115722117358,    0.611757484515,
                         1.512610269776,    2.833751337744,
                         4.599227639418,    6.844525453115,
                         9.621316842457,   13.006054993306,
                        17.116855187462,   22.151090379397,
                        28.487967250984,   37.099121044467};
const double W[POLDEGR]={2.64731371055e-1,  3.77759275873e-1,
                         2.44082011320e-1,  9.04492222117e-2,
                         2.01023811546e-2,  2.66397354187e-3,
                         2.03231592663e-4,  8.36505585682e-6,
                         1.66849387654e-7,  1.34239103052e-9,
                         3.06160163504e-12, 8.14807746743e-16};

double Hprice(char CallPutFlag, double S0, double K, double T, double r,
       double v0, double delta, double rho, double xi, double zeta)
{   int k;
    double P1, P2, Hcall, Hput, gl;

    if(T==0)
    {   if(CallPutFlag == 'c') return max(S0-K,0.0);
        else return max(K-S0,0.0);
    }
    P1 = 0.0; P2 = 0.0;
    for (k=0;k<POLDEGR;k++)
    {   P1 += W[k]*intgrnd(x[k],S0,K,T,r,v0,delta,rho,xi,zeta,1);
        P2 += W[k]*intgrnd(x[k],S0,K,T,r,v0,delta,rho,xi,zeta,2);
    }
    P1 = 0.5 + (1.0/Pi)*P1;
    P2 = 0.5 + (1.0/Pi)*P2;
    Hcall = S0*P1 - K*exp(-r*T)*P2;
    if (CallPutFlag == 'c') return Hcall;
    else {Hput = Hcall - S0 + K*exp(-r*T); return Hput;}
}
```

The constant for the integration are defined as global constants. In this way they can be used also for other routines that exploit a similar integration. An example is the routine that computes the first derivative of the option price according to Heston with respect to the underlying price:

```
double Hdelta(char CallPutFlag, double S0, double K, double T, double r,
       double v0, double delta, double rho, double xi, double zeta)
{   int j;
    double P1;

    if(T==0)
    {   if(CallPutFlag == 'c')
```

```
        {   if(S0-K>0) return 1.0;
            else return 0.0;
        }
        else
        {   if(K-S0>0) return 1.0;
            else return 0.0;
        }
    }
    P1 = 0.0;
    for (j=0;j<POLDEGR;j++)
    {   P1 += W[j]*intgrnd(x[j],S0,K,T,r,v0,delta,rho,xi,zeta,1);
    }
    P1 = 0.5 + P1/Pi;
    P1 = min(max(P1,0.0),1.0);
    if (CallPutFlag == 'c') return P1;
    else return (P1-1);
}
```

The computation of the first derivative with respect to the instantaneous variance is slightly more complicated. Indeed, another integrand function has to be introduced:

```
double vegaintgrnd(double u, double S0, double K, double T, double r,
       double v0, double delta, double rho, double xi, double zeta, int intflag)
{   double b;
    complex<double> d, e, g, C, D, car;

    if (intflag == 1)
    {   b = xi - rho*delta;
        d = sqrt((rho*delta*u*i - b)*(rho*delta*u*i - b)//
            - delta*delta*(u*i - u*u));
    }
    else
    {   b = xi;
        d = sqrt((rho*delta*u*i - b)*(rho*delta*u*i - b)//
            + delta*delta*(u*i + u*u));
    }
    e = exp(d*T);
    g = (b - rho*delta*u*i + d)/(b - rho*delta*u*i - d);
    C = r*u*i*T + zeta/(delta*delta)*((b - rho*delta*u*i + d)*T//
        - 2*log((1 - g*e)/(1 - g)));
    D = ((b - rho*delta*u*i + d)/(delta*delta))*((1 - e)/(1 - g*e));
    car = exp(C + D*v0 + i*u*log(S0));
    return (real((D*exp(u-i*u*log(K))*car)/(i*u)));
}
```

The integration of the above integrand function is similar to the integration for the option price:

```
double Hvega(double S0, double K, double T, double r,
       double v0, double delta, double rho, double xi, double zeta)
{   int j;
    double V1, V2, vega;

    if(T==0) return 0.0;
```

```
    V1 = 0.0; V2 = 0.0;
    for (j=0;j<POLDEGR;j++)
    {   V1 += W[j]*vegaintgrnd(x[j],S0,K,T,r,v0,delta,rho,xi,zeta,1);
        V2 += W[j]*vegaintgrnd(x[j],S0,K,T,r,v0,delta,rho,xi,zeta,2);
    }
    V1 /= Pi; V2 /= Pi;
    vega = (S0*V1 - K*exp(-r*T)*V2)*0.01;
    vega = max(vega,0.0);
    return vega;
}
```

B.2 Option pricing by perturbation approach

Considering the perturbation approach described by Lewis [48], option price according to the Heston model can be approximated with a power series expansion around $\delta = 0$. More precisely, Lewis describes a power series approximation for BSIV according to a generic stochastic volatility model where instantaneous variance dynamics has the form

$$dv_t = b(v_t)dt + \delta\eta(v_t)dw_t$$

where the $\mathbb{Q}$-Brownian motion w_t can be correlated with the $\mathbb{Q}$-Brownian motion that drives the price process, δ is constant and functions $b(v_t)$ and $\eta(v_t)$ do not depend on δ. In this framework the expansion takes the following form[2]:

$$\begin{aligned}\sigma^2 = \bar{v} + \frac{1}{T}\delta J^{(1)} R^{(1,1)} + \delta^2 \Bigg[& \frac{J^{(3)} R^{(2,0)}}{T^2} + \frac{J^{(4)} R^{(1,2)}}{T} + \\ & + \frac{1}{2}\frac{(J^{(1)})^2}{T}\left(\frac{X}{Z} + \frac{1}{8}\frac{12+Z}{Z^2} - \frac{5}{2}\frac{X^2}{Z^3}\right)\Bigg] + O(\delta^3) \ ,\end{aligned} \tag{B.1}$$

where

$$\bar{v} = \frac{\zeta}{\xi} + \left(v_0 - \frac{\zeta}{\xi}\right)\left(\frac{1-e^{-\xi T}}{\xi T}\right) \ , \quad X = \ln\left(\frac{S_0}{Ke^{-rT}}\right) \ , \quad Z = \bar{v}T \ ,$$

$$\begin{aligned}R^{(2,0)}(S_0, v_0, T) &= T\left(\frac{1}{2}\frac{X^2}{Z^2} - \frac{1}{2Z} - \frac{1}{8}\right) \ , \\ R^{(1,1)}(S_0, v_0, T) &= \frac{1}{2} - \frac{X}{Z} \ , \\ R^{(1,2)}(S_0, v_0, T) &= \frac{X^2}{Z^2} - \frac{X}{Z} - \frac{1}{4Z}(4 - Z)\end{aligned}$$

[2] To simplify the notation, hereafter we assume $t = 0$.

In order to apply this general formula, we have to define the three variables $J^{(1)}$, $J^{(3)}$ and $J^{(4)}$. This can be conveniently done only by assuming a specific model. For instance, in the Heston model the three variable are given by the following equations:

$$
\begin{aligned}
J^{(1)}(v_0,T) &= \frac{\omega\left[2+T\theta+e^{T\theta}\left(T\theta-2\right)\right]+\theta\left(e^{T\theta}-T\theta-1\right)v_0}{\theta^2 e^{T\theta}}\,,\\
J^{(3)}(v_0,T) &= \frac{\omega+e^{2T\theta}\omega\left(2T\theta-5\right)+4e^{T\theta}\left[\omega+\omega T\theta-T\theta^2 v_0+\theta v_0\sinh(T\theta)\right]}{2\theta^2 e^{2T\theta}}\,,\\
J^{(4)}(v_0,T) &= \frac{\omega\left[6+2e^{T\theta}\left(T\theta-3\right)+T\theta\left(4+T\theta\right)\right]-\theta\left[2-2e^{T\theta}+T\theta\left(2+T\theta\right)\right]v_0}{2\theta^3 e^{T\theta}}\,.
\end{aligned}
$$

The following C++ code computes BSIV according to the above approximation applied to the Heston model. Input parameters have the same meaning of those in routine `Hprice`.

```
double HApxImpVol(double S0, double K, double T, double r, double v0,
       double delta, double rho, double xi, double zeta)
{   double X, Z, vmean, J1, J3, J4, impvar, R11, R20, R12;

    X = log(S0) - log(K) + r*T;
    vmean = zeta/xi + (v0-zeta/xi)*((1-exp(-xi*T))/(xi*T));
    J1 = rho/xi*(zeta*(2+T*xi+exp(T*xi)*(T*xi-2)) + xi*(exp(T*xi)-T*xi-1)*v0)
       / (exp(T*xi)*xi*xi);
    J3 = (zeta + exp(2*T*xi)*zeta*(2*T*xi-5) +
       +  4*exp(T*xi)*(zeta + zeta*T*xi - T*xi*xi*v0 + xi*v0*sinh(T*xi)))
       / (2*exp(2*T*xi)*xi*xi)
       / (2*xi*xi);
    J4 = rho*rho/xi*(zeta*(6+2*exp(T*xi)*(T*xi-3) + T*xi*(4 + T*xi))
       - xi*(2 - 2*exp(T*xi) + T*xi*(2 + T*xi))*v0)
       / (2*exp(T*xi)*xi*xi*xi);
    Z = vmean*T;
    R11 = 0.5 - X/Z;
    R20 = T*(0.5*(X*X)/(Z*Z)-1/(2*Z)-0.125);
    R12 = (X*X)/(Z*Z)-X/Z-(4-Z)/(4*Z);
    impvar =  vmean + delta*J1*R11/T
           + delta*delta*(J3*R20/(T*T)
        + J4*R12/T + 0.5*J1*J1*(X/Z+0.125*(12+Z)/(Z*Z)-5/2*(X*X)/(Z*Z*Z))/T);
    return sqrt(impvar);
}
```

For 3/2 model, $J^{(1)}$, $J^{(3)}$ and $J^{(4)}$ take a more complicated form:

$$\begin{aligned}
J^{(1)}(v_0,T) &= \frac{1}{2\theta^3 e^{2T\theta}}\left\{\omega^2\left[1+4e^{T\theta}(1+T\theta)+e^{2T\theta}(2T\theta-5)\right]+\right.\\
&\left.+\left(e^{T\theta}-1\right)^2\theta^2v^2+4e^{T\theta}\omega\theta v\left[\sinh(T\theta)-T\theta\right]\right\}\ ,\\
J^{(3)}(v_0,T) &= \frac{1}{6\theta^4 e^{3T\theta}}\left\{2\omega^3\left[1+9e^{T\theta}(2+T\theta)+9e^{2T\theta}(2T\theta-1)+e^{3T\theta}(3T\theta-10)\right]+\right.\\
&+3\omega\theta^2\left[2+e^{T\theta}\left(3-6e^{T\theta}+e^{2T\theta}+6T\theta\right)\right]v^2+2\left(e^{T\theta}-1\right)^3\theta^3v^3+\\
&\left.+12e^{\frac{3}{2}T\theta}\omega^2\theta v\left[9\sinh\left(\frac{T\theta}{2}\right)+\sinh\left(\frac{3T\theta}{2}\right)-6T\theta\cosh\left(\frac{T\theta}{2}\right)\right]\right\}\ ,\\
J^{(4)}(v_0,T) &= \frac{1}{6\theta^5 e^{3T\theta}}\left\{\omega^3\left[1+6e^{T\theta}(1+T\theta)+\right.\right.\\
&\left.+e^{3T\theta}(6T\theta-22)+3e^{2T\theta}(5+2T\theta(3+T\theta))\right]-\\
&-3\omega^2\theta\left[1+e^{T\theta}\left(2+4T\theta+e^{T\theta}\left(2T\theta+2T^2\theta^2-1-2e^{T\theta}\right)\right)\right]v+\\
&\left.+3\left(e^{T\theta}-1\right)\omega\theta^2\left(e^{2T\theta}-2e^{T\theta}T\theta-1\right)v^2+\left(e^{T\theta}-1\right)^3\theta^3v^3\right\}
\end{aligned}$$

The C++ code for the computation BSIV approximation according to the 3/2 model is the following:

```
double M32ImpVol(double S0, double K, double T, double r,
       double v0, double delta, double rho, double xi, double zeta)
{   double X, Z, vmean, J1, J3, J4, impvar, R11, R20, R12;

    X = log(S0) - log(K) + r*T;
    vmean = zeta/xi + (v0-zeta/xi)*((1-exp(-xi*T))/(xi*T));
    J1 = rho/xi*(zeta*zeta
       * (1+4*exp(T*xi)*(1+T*xi)+exp(2*T*xi)*(2*T*xi-5))
       + (exp(T*xi)-1)*(exp(T*xi)-1)*xi*xi*v0*v0
       + 4*exp(T*xi)*zeta*xi*v0*(sinh(T*xi)-T*xi))
       / (2*exp(2*T*xi)*xi*xi*xi);
    J3 = (2*zeta*zeta*zeta*(1 + 9*exp(T*xi)*(2+T*xi)
       + 9*exp(2*T*xi)*(2*T*xi-1) + exp(3*T*xi)*(3*T*xi-10))
       + 3*zeta*xi*xi*(2
       + exp(T*xi)*(3-6*exp(T*xi)+exp(2*T*xi)+6*T*xi))*v0*v0
       + 2*(exp(T*xi)-1)*(exp(T*xi)-1)*(exp(T*xi)-1)
       * xi*xi*xi*v0*v0*v0
       + 12*exp(1.5*T*xi)*zeta*zeta*xi*v0
       * (9*sinh(0.5*T*xi)+sinh(1.5*T*xi)-6*T*xi*cosh(0.5*T*xi)))
       / (6*exp(3*T*xi)*xi*xi*xi*xi)
       / (2*xi*xi);
    J4 = 2*rho*rho/xi*(zeta*zeta*zeta*(1 + 6*exp(T*xi)*(1+T*xi)
       + exp(3*T*xi)*(6*T*xi-22)
       + 3*exp(2*T*xi)*(5+2*T*xi*(3+T*xi)))
       - 3*zeta*zeta*xi*(1 + exp(T*xi)*(2 + 4*T*xi
       + exp(T*xi)*(2*T*xi*(1 + T*xi)-1-2*exp(T*xi))))*v0
       + 3*(exp(T*xi)-1)*zeta*xi*xi*(exp(2*T*xi)
```

```
      - 2*exp(T*xi)*T*xi-1)*v0*v0
      + (exp(T*xi)-1)*(exp(T*xi)-1)*(exp(T*xi)-1)*(exp(T*xi)-1)
      * xi*xi*xi*v0*v0*v0)
      / (6*exp(3*T*xi)*xi*xi*xi*xi*xi);
   Z = vmean*T;
   R11 = 0.5 - X/Z;
   R20 = T*(0.5*(X*X)/(Z*Z)-1/(2*Z)-0.125);
   R12 = (X*X)/(Z*Z)-X/Z-(4-Z)/(4*Z);
   impvar = vmean + delta*J1*R11/T
          + delta*delta*(J3*R20/(T*T)
  + J4*R12/T + 0.5*J1*J1*(X/Z+0.125*(12+Z)/(Z*Z)-5/2*(X*X)/(Z*Z*Z))/T);
   return sqrt(impvar);
}
```

Appendix C

The discretization of the structural model

Let us assume that the dynamics of the price process $\{S_t\}_{t\in[0,T]}$ can be described by the following equations:

$$\begin{aligned} dS_t &= \mu S_t dt + |v_t|^{1/2} S_t dw_{1,t} \ , \\ dv_t &= a(v_t)dt + \eta_1(v_t)dw_{1,t} + \eta_2(v_t)dw_{2,t} \ , \end{aligned}$$

where $w_{1,t}$ and $w_{2,t}$ are two Brownian motions whereas $a(v_t)$, $\eta_1(v_t)$ and $\eta_2(v_t)$ are functions of v_t. Both the processes $\{S_t\}$ and $\{v_t\}$ are *cadlag*.

Let us consider an equidistant discretization composed by N subintervals of the period $T-t$ such that $\Delta t = T/N$. In order to discretize the volatility process, it is convenient to write it again in the following way

$$dv_t = a(v_t)dt + b'(v_t)dw_t \ ,$$

where $w_t = [w_{1,t} \ w_{2,t}]'$ and $b(v_t) = [\eta_1(v_t) \ \eta_2(v_t)]'$. In Euler scheme, the volatility process can be discretized in the following way:

$$v_{t+\Delta t} = v_t + a(v_t)\Delta t + b'(v_t)\varepsilon_t\sqrt{\Delta t} \ ,$$

where $\varepsilon_t = (\varepsilon_{1,t} \ \varepsilon_{2,t})'$ is a vector of independent normally distributed random variable with mean 0 and variance 1.

Under Lipschitz and linear growth conditions on $a(.)$ and $b(.)$, the Euler scheme strongly converges to the true process and the convergence order is 0.5. In order to improve the convergence order, the Milstein discretization schemes can be used[1]. By following the Milstein discretization scheme, the volatility dynamics becomes:

$$v_{t+\Delta t} = v_t + a(v_t)\Delta t + b'(v_t)\varepsilon_t\sqrt{\Delta t} + \frac{1}{2}b'(v_t)\left[\varepsilon_t\varepsilon_t' - I_2\right]\frac{\partial b}{\partial v}\Delta t \ ,$$

[1] See Kloeden and Platen [44] pp. 345-351.

where I_2 is a 2 by 2 identity matrix.

Under the assumption that $a(v_t)$ is once and $b(v_t)$ twice differentiable, the Milstein scheme has a strong convergence of order one[2].

Since the process $\{v_t\}$ is *cadlag*, the discretization of the price process can be written as

$$\begin{aligned} S_{t+\Delta t} &= S_t \exp\left[r\Delta t - \frac{1}{2}\int_t^{t+\Delta t} v_u du + \int_t^{t+\Delta t} |v_u|^{1/2} dw_{1,u}\right] \\ &= S_t \exp\left[\left(r - \frac{v_{t+\Delta t}}{2}\right)\Delta t + |v_{t+\Delta t}|^{1/2}\varepsilon_{1,t}\sqrt{\Delta t}\right] \ . \end{aligned}$$

For model 3.14, drift and diffusion functions are specified in the following way:

$$a(v_t) = [\zeta - \xi v_t] \ , \quad \eta_1(v_t) = \delta v_t^\gamma \rho \ , \quad \eta_2(v_t) = \delta v_t^\gamma \sqrt{(1-\rho^2)} \ ,$$

with $\xi, \zeta, \delta > 0$ and $\rho \in [-1, 1]$. In order to apply the Milstein discretization schemes we can state that

$$b(v_t) = \delta v_t^\gamma \begin{bmatrix} \rho & \sqrt{1-\rho^2} \end{bmatrix}' \ , \qquad \frac{\partial b}{\partial v} = \delta\gamma v_t^{\gamma-1} \begin{bmatrix} \rho & \sqrt{1-\rho^2} \end{bmatrix} \ .$$

Hence, the volatility process will have the following approximation:

$$\begin{aligned} v_{t+\Delta t} = v_t + [\zeta - \xi v_t]\Delta t + \delta v_t^\gamma \rho\sqrt{\Delta t}\ \varepsilon_{1,t} + \delta v_t^\gamma \sqrt{\Delta t(1-\rho^2)}\ \varepsilon_{2,t} + \\ + \frac{1}{2}\gamma\delta^2 v_t^{2\gamma-1}\left[\rho^2\varepsilon_{1,t}^2 + (1-\rho^2)\varepsilon_{2,t}^2 + 2\rho\sqrt{1-\rho^2}\varepsilon_{1,t}\varepsilon_{2,t} - 1\right]\Delta t \ , \end{aligned}$$

which for $\rho = 0$ becomes,

$$v_{t+\Delta t} = v_t + [\zeta - \xi v_t]\Delta t + \delta\sqrt{v_t \Delta t}\ \varepsilon_{2,t} + \frac{1}{2}\gamma\delta^2 v_t^{2\gamma-1}\left(\varepsilon_{2,t}^2 - 1\right)\Delta t \ .$$

[2] See again Kloeden an Platen [44] Theorem 10.3.5.

Appendix D

Portfolios composition

The composition of the portfolios is the following:

Assets	Portfolios A	B	C	D
Advanced Tissue Sciences	3.15%	2.00%	1.76%	2.30%
Alexion Pharmaceuticals	4.99%	3.16%	2.79%	3.64%
Amgen	10.32%	6.54%	5.76%	7.53%
Applera Biosystems Graup	2.82%	1.79%	1.58%	2.06%
Aradigm	2.27%	1.44%	1.27%	1.66%
Atrix Labs	3.05%	1.93%	1.70%	2.23%
Avigen	3.60%	2.28%	2.01%	2.62%
Aviron	9.99%	6.33%	5.58%	7.29%
Cell Genesys	11.17%	7.08%	6.24%	8.15%
Chiron	6.68%	4.23%	3.73%	4.87%
Cortex Pharmaceuticals	4.31%	2.73%	2.41%	3.14%
Diversa	3.63%	2.30%	2.03%	2.65%
Genome Therapeutics	2.68%	1.70%	1.50%	1.96%
Genzime-GENL Division	6.75%	4.28%	3.77%	4.93%
Human Genome Sciences	8.52%	5.40%	4.76%	6.21%
Immunex	2.01%	1.27%	1.12%	1.46%
Medimmune	5.10%	3.23%	2.85%	3.72%
Myriad Genetics	4.06%	2.57%	2.26%	2.96%
Sciclone Pharmaceuticals	2.57%	1.63%	1.43%	1.87%
Valentis	2.34%	1.48%	1.30%	1.70%
Intel	-	8.80%	7.76%	10.14%
Cisco	-	7.86%	6.93%	9.05%
Lucent Technologies	-	0.67%	0.59%	0.78%
General Electrics	-	4.06%	3.58%	4.67%
Wells Fargo & Co	-	11.59%	10.22%	13.35%
US Bancorp	-	3.66%	3.22%	4.21%
S&P 500 put option	-	-	3.87%	-
NASDAQ 100 put option	-	-	7.99%	-
S&P 500 call option	-	-	-	-9.43%
NASDAQ 100 call option	-	-	-	-5.72%

Bibliography

[1] ABRAMOWITZ, M., AND STEGUN, I. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* Dover Publications, 1974.

[2] ALEXANDER, C. "Orthogonal GARCH". In *Mastering Risk, Volume 2: Applications*, C. Alexander, Ed. Financial Times Management, 2001.

[3] ANDERSEN, T., CHUNG, H. J., AND SØRENSEN, B. E. "Efficient method of moments estimation of a stochastic volatility model: a Monte Carlo study". *Journal of Econometrics 91* (1999), pp. 61–87.

[4] ANDERSEN, T., AND LUND, J. "Estimating continuous time stochastic volatility models of the short-term interest rate". *Journal of Econometrics 77* (1997), pp. 343–377.

[5] AVELLANEDA, M., LEVY, A., AND PARÁS, A. "Pricing and Hedging Derivatives Securities in Market with Uncertain Volatility". *Applied Mathematical Finance* (1995), pp. 73–88.

[6] BAKSHI, G., CAO, C., AND CHEN, Z. "Empirical Performances of Alternative Option Pricing Models". *Journal of Finance 5* (1997), pp. 2003–2049.

[7] BALL, C. A., AND ROMA, A. "Stochastic Volatility Option Pricing". *Journal of Financial and Quantitative Analysi 29* (1994), pp. 589–607.

[8] BARONE-ADESI, G., BOURGOIN, F., AND GIANNOPOULOS, K. "Don't look back". *Risk August* (1998), pp. 100–104.

[9] BLACK, F., AND SCHOLES, M. "The Pricing of Options and Corporate Liabilities". *Journal of Political Economy 81* (1973), pp. 637–654.

[10] BOUDOUKH, J., RICHARDSON, M., AND WHITELAW, R. "The best of both worlds". *Risk May* (1998), pp. 64–67.

[11] BYSTRÖM, H. "Orthogonal GARCH and covariance matrix forecasting in a stress scenario". Lund University, 2000.

[12] Chernov, M., and Ghysels, E. "A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation". *Journal of Financial Econometrics 56* (2000), pp. 407–458.

[13] Christoffensen, P. F. "Evaluating Interval Forecasts". *International Economic Review 39* (1998), pp. 841–862.

[14] Cox, J., Ingersoll, J., and Ross, A. "A Theory of Term Structure of Interest Rates". *Econometrica 53* (1995), pp. 385–407.

[15] Cox, J., Ingersoll, J., and Ross, A. "An Intertemporal General Equilibrium Model of Asset Prices". *Econometrica 53* (1995), pp. 363–384.

[16] Cox, N. "Notes on Option Pricing I: Constant Elasticity of Variance Diffusions". Stanford University, 1975.

[17] Delbaen, F. "Representing Martingale Measures when Asset Prices are Continuous and Bounded". *Mathematical Finance 2* (1992), pp. 107–130.

[18] Duffie, D., and Pan, J. "An Overview of Value at Risk". *Journal of Derivatives Spring* (1997), pp. 7–48.

[19] El-Jahel, L., Perraudin, W., and Sellin, P. "Value at Risk for Derivatives". *The Journal of Derivatives Spring* (1999), pp. 7–26.

[20] El Karoui, N., Jeanblanc-Picqué, M., and Shreve, S. E. "Robustness of the Black and Scholes Formula". *Mathematical Finance 8* (1998), pp. 93–126.

[21] El Karoui, N., and Quenez, M. C. "Dynamic Programming and Pricing of Contingent Claim in an Incomplete Market". *SIAM Journal on Control and Optimization* (1995), pp. 29–66.

[22] Eraker, B., Johannes, M., and Polson, N. "The Impact of Jumps in Volatility and Returns". *Journal of Finance 3* (2003), pp. 1269–1300.

[23] Fama, E. F. "The Behavior of Stock-Market Process". *The Journal of Business 38* (1965), pp. 34–105.

[24] Feuerverger, A., and Wong, A. C. M. "Computation of Value-at-Risk for Nonlinear Portfolios". *The Journal of Risk 3* (2000), pp. 37–55.

[25] Figlewski, S. "Forecasting volatility using historical data". Stern School of Business, 1995.

[26] FIORENTINI, G., LEÓN, A., AND RUBIO, G. "Estimation and empirical performance of Heston's stochastic volatility model: the case of a thinly traded market". *Journal of Empirical Finance 9* (2002), pp. 225–255.

[27] FÖLMER, H., AND SCHWEIZER, M. "Hedging by Sequential Regression: an Introduction to the Mathematics of Option Trading". *ASTIN Bulletin 18* (1989), pp. 147–160.

[28] FÖLMER, H., AND SCHWEIZER, M. "Hedging of Contingent Claim under Incomplete Information". In *Applied Stochastic Analysis*, M. Davis and R. Elliot, Eds. Gordon and Breach, 1991, pp. pp. 389–414.

[29] FREY, R. "Derivative Asset Analysis in Models with Level Dependent and Stochastic Volatility". *CWI Quaterly 10* (1996), pp. 1–34.

[30] FREY, R., AND SIN, C. "Bounds on European Option Prices under Stochastic Volatility". *Mathematical Finance* (1999), pp. 97–116.

[31] GALLANT, A., AND TAUCHEN, G. "Which moments to match?". *Econometric Theory 12* (1996), pp. 657–681.

[32] GOURIEROUX, C., AND MONFORT, A. *Statistics and Econometric Models.* Cambridge University Press, 1995.

[33] GOURIEROUX, C., AND MONFORT, A. *Simulation-Based Econometric Methods.* Oxford University Press, 1996.

[34] GOURIEROUX, C., MONFORT, A., AND RENAULT, E. "Indirect inference". *Journal of Applied Econometrics 8* (1993), pp. S85–S118.

[35] HAMILTON, J. *Time Series Analysis.* Princeton University Press, 1994.

[36] HARRISON, J. M., AND KREPS, D. M. "Martingales and Arbitrage in Multiperiod Securities Markets". *Journal of Economic Theory 20* (1979), pp. 381–408.

[37] HARRISON, J. M., AND PLISKA, S. R. "Martingales and Stochastic Integrals in the Theory of Continuous Trading". *Stochastic Processes and their Applications 11* (1981), pp. 215–260.

[38] HARRISON, J. M., AND PLISKA, S. R. "A Stochastic Calculus Model of Continuous Trading: Complete Markets". *Stochastic Processes and their Applications 15* (1983), pp. 313–316.

[39] HENDRICKS, D. "Evaluation of Value-at-Risk Models Using Historical Data". *FRBNY Economic Policy Review April* (1996), pp. 39–70.

[40] HESTON, S. L. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options". *The Review of Financial Studies 6* (1993), pp. 327–343.

[41] HULL, J., AND WHITE, A. "The Pricing of Options on Assets with Sthocastic Volatility with Applications to Bond and Currency Options". *Journal of Finance 42* (1987), pp. 281–300.

[42] KARATZAS, I. "Optimizations Problems In the Theory of Continuous Trading". *SIAM Journal on Control and Optimization 27* (1989), pp. 1221–1259.

[43] KARATZAS, I., AND SHREVE, S. *Brownian Motion and Stochastic Calculus.* Springer, 1997.

[44] KLOEDEN, P., AND PLATEN, E. *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, 1999.

[45] KRAMKOV, D. O. "Optional Decomposition of Supermartingales and Hedging Contincent Claims in Incomplete Security Markets". *Probability Theory and Related Fields 105* (1996), pp. 459–479.

[46] KRYLOV, N. V. *Controlled Diffusion Processes.* Springer, 1980.

[47] KUPIEC, P. H. "Techniques for Verifying the Accuracy of Risk Measurement Models". *The Journal of Derivatives Winter* (1995), pp. 73–84.

[48] LEWIS, A. *Option Valuation under Stochastic Volatility.* Finance Press, 2000.

[49] MANDELBROT, B. "The Variation of Certain Speculative Prices". *The Journal of Business 36* (1963), pp. 394–419.

[50] MCNEIL, A., AND FREY, R. "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach". *Journal of Empirical Finance 7* (2000), pp. 271–300.

[51] MERTON, R. C. "Theory of Rational Option Pricing". *Bell Journal of Economics and Management Science 4* (1973), pp. 141–183.

[52] MINA, J., AND ULMER, A. "Delta-Gamma Four Ways". RiskMetrics, 1999.

[53] PASTORELLO, S., RENAULT, E., AND TOUZI, N. "Statistical inference for Random-Variance option pricing". *Journal of Business and Economic Statistics 18* (2000), pp. 358–367.

[54] PRESS, W., TEUKOLSKY, S., VETTERLING, W., AND FLANNERY, B. *Nemerical Recipes in C++*. Cambridge University Press, 2002.

[55] PRITSKER, M. "Evaluating Value-at-Risk Methodologies: Accuracy Versus Computational Time". *Journal of Financial Services Research 2* (1997).

[56] PRITSKER, M. "The Hidden Dangers of Historical Simulation". University of California at Berkeley, 2001.

[57] SCOTT, L. "Random-Variance Option Pricing: Empirical Tests of the Model and Delta-Sigma Hedging". *Advanced in Future and Options Research 5* (1991), pp. 113–135.

[58] SHEPHARD, N. "From Characteristic Function to Distribution Function: a Simple Framework for the Theory". *Econometric Theory 7* (1991), pp. 519–529.

[59] SIN, C. A. "Complications with Stochastic Volatility Models". *Advances in Applied Probability 30* (1998), pp. 256–268.

[60] STEIN, E. M., AND STEIN, J. C. "Stoch Price Distributions with Stochastic Volatility: an Analytic Approach". *The Review of Financial Studies 4* (1991), pp. 727–752.

[61] WHITE, H. "Maximum Likelihood Estimation and Misspecified Models". *Econometrica 50* (1982), pp. 1–25.

[62] WIGGINS, J. B. "Option Values Under Stochastic Volatility". *Journal of Financial Economics 19* (1987), pp. 351–372.

[63] WILMOTT, P., HOWISON, S., AND DEWYNNE, J. *The Mathematics of Financial Derivatives: A Student Introduction.* Cambridge University Press, 1995.

[64] ZANGARI, P. "A VaR Methodology for Portfolios that Include Options". *RiskMetrics Monitor 1Q* (1996), pp. 4–12.

[65] ZANGARI, P. "How Accurate is the Delta-Gamma Methodology?". *RiskMetrics Monitor 3Q* (1996), pp. 12–29.