Model Risk Quantification in Option Pricing

Michael Montag & Fredrik Persson

Master's thesis
2015:E19

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

CENTRUM SCIENTIARUM MATHEMATICARUM

Abstract

This thesis investigates a methodology for quantification of model risk in option pricing. A set of different pricing models is specified and each model is assigned a probability weight based on the Akaike Information Criteria. It is then possible to obtain a price distribution of an exotic derivative from these probability weights. Two measures of model risk inspired by the regulatory standards on prudent valuation are proposed based on this methodology.

The model risk measures are studied for different equity options which are priced using a set of stochastic volatility models, with and without jumps. The models are calibrated to vanilla call options from the S&P 500 index, as well as to synthetic option prices based on market data simulated using the Bates model. For comparable options, the model risk is higher for up-and-out barrier options compared to vanilla, digital and Asian options. Moreover, the model risk measure, in relative terms of option price, increases quickly with strike level for call options far out of the money, while the model risk in absolute terms is lowest when the option is deep out of the money. The model risk for up-and-out barrier options tends to be higher when the barrier is closer to the spot price, although the increase in risk does not have to be monotonic with decreasing barrier level.

The methodology is flexible and easy to implement, yielding intuitive results. However, it is sensitive to different assumptions in the structure of the pricing errors.

Keywords: Model risk, Parameter uncertainty, AIC, Exotic options, Prudent valuation

Acknowledgements

The thesis was written during the spring of 2015 at the Centre for Mathematical Sciences, Lund University, in cooperation with the model risk team at Nordea Wholesale Banking, Copenhagen.

We would like to express our gratitude towards our supervisors, Magnus Wiktorsson and Jens Nilsson, for providing excellent feedback during our thesis work. We would also like to thank Kenneth Nielsen and our examiner Erik Lindström for their valuable comments and suggestions.

Finally, we would like to convey our sincere thanks to friends and family for their strong support throughout the entire period of studies at Lund University.

Lund, June 2015

Fredrik Persson
Michael Montag

Contents

CHAPTER 1

Introduction

> *"Statisticians, like artists, have the bad habit of falling in love with their models."*
>
> — George E. P. Box

1.1 Background

In today's financial world the use of models is ubiquitous. Market participants use mathematical and statistical models to try to predict future market outcomes, price derivatives and hedge positions. In derivatives pricing, continuous time models formulated as SDEs, *stochastic differential equations*, are primarily used. The most well known such model is the Black-Scholes model, formulated by Fischer Black and Myron Scholes [7] in 1973, which models the underlying asset as a geometric Brownian motion. Since then several extensions of the Black-Scholes model have been proposed, by for example including jump processes and stochastic volatility to try to capture more realistic dynamics of the stock market. With all these models one could ask the question which one is the most suitable model? How can we quantify the risk associated with pricing a derivative with an inappropriate model?

Even when a model can calibrate perfectly, i.e. within bid-ask spreads, to vanilla options, it still carries a risk when used to price exotic options. Schoutens, Simons, and Tistaert show in [40] that several different stochastic volatility models that calibrate almost perfectly to vanilla options still can give widely different prices on exotic options.

History has shown us how poorly specified models and neglect of what could be called *model risk* can lead to disastrous results. A well known example of misused models is the now infamous Gaussian copula in credit default obligation (CDO) pricing. It was first presented in a paper by David Li in [42], and became a standard method in CDO pricing, but it severely underestimated extreme outcomes. The CDO market grew from $275 billion in 2000 to $4.7 trillion by 2006. By 2009 the market for CDOs had collapsed and had brought down much of the financial world with it [39].

To decrease the risk of future financial disasters, there has been an increasing amount of regulatory standards. One example of such regulation is the upcoming regulatory technical standards from European Banking Authority (EBA) on *prudent valuation* [3]. The purpose of this regulation is to make banks and institutions calculate *additional valuation adjustments* (AVA) for held positions. These adjustments should compensate for close-out-costs, price uncertainty and model risk. To fulfill these requirements, two types of approaches can be taken. The first is called the *simplified approach*, which is available to smaller institutions with assets and liabilities worth less than 15 billion euros. In this approach, the AVA can simply be taken as 0.1% of the aggregate absolute value of held positions.

The second approach, called the *core approach*, is mandatory for larger institutions that exceed the threshold, but can also be used by smaller institutions as an active choice. The core approach is more comprehensive and defines the AVA based on a prudent value for each asset. The prudent value of an asset is in turn calculated based on plausible values of the asset and should be taken with a certainty level of 90%.

1.2 Purpose

The main objective of the thesis is to find a statistically sound methodology for quantifying model risk as part of the core approach in the regulatory standards on prudent valuation. Since the prudent value should be based on a 90% certainty level, the model risk quantification in this thesis aims to find a distribution of prices for each asset. The studied assets are vanilla call and puts, digital calls, up-and-out barrier options and arithmetic Asian call options. The work was carried out in collaboration with the model risk team at Nordea Wholesale Banking, Copenhagen.

1.3 Methodology

To find the price distribution of a derivative, a model set $\mathcal{Q} = \cup_{i=1}^{m} \mathbb{Q}_i$ is formed. Each pricing model $\mathbb{Q}_i$ is assigned a probability weight w_i, depending on how well it can price liquid benchmark options. The method used in this thesis is based on the work of Detering and Packham in [19], where the probability weights are determined by the Akaike Information Criterion. Detering and Packham's method is expanded by utilizing a flat-top Gaussian likelihood function for the pricing errors. An algorithm used in the selection of models in $\mathcal{Q}$ is also developed. The pricing of derivatives is done with Fourier transform methods for vanilla and digital calls, and Monte Carlo methods otherwise.

1.4 Outline

The outline for the rest of the thesis is as following:

Chapter 2: In this chapter fundamental results of derivative pricing are presented as well as the models that used in the empirical testing: Black-Scholes, Heston, Bates and NIGCIR. Least squares calibration and Monte-Carlo pricing are also covered. The ill-posedness of the calibration problem in option pricing is presented, which serves as a motivation to why quantification of model risk is relevant.

Chapter 3: The concept of model risk is introduced and the regulation from the EBA is presented in this chapter. Furthermore, the basics of Bayesian statistics are covered as well as the theory behind AIC and BIC model weights. Detering and Packham's way of using AIC weights is presented along with their distributional assumptions on the differences between market and model prices. Some other, arguably more realistic, noise structures are also discussed. Finally a couple of model risk measures are introduced, inspired by the concept of additional value adjustment.

Chapter 4: The model risk methodology is evaluated for European calls, digital calls, up-and-out options and arithmetic Asian call options. Both real market data from the S&P 500 index as well as simulated data are used. The AIC-weighting methodology is compared to the Metropolis-Hastings algorithm in a special case of the Heston model.

Chapter 5: The results and methodology are discussed, as well as directions on future research in the topic of model risk quantification.

Chapter 6: Concluding remark on the results.

1.5 Resources

Financial data is provided by Nordea in form of bid and ask prices on the S&P 500 index for call and put options during the month of August 2014, as well as yield curve data. All calculations are carried out in MATLAB 2014b.

CHAPTER 2

Option pricing, models and calibration

2.1 Arbitrage theory and models

In order to understand the methodology and results in the coming chapters the main results from continuous time arbitrage theory will briefly be presented in this section. The primary reference will be [6] by Björk which gives a comprehensive treatment of the topic.

2.1.1 Stochastic differential equations

The building block of stochastic calculus is the standard *Brownian motion.*

Definition 2.1 *A* Brownian motion *is a continuous time stochastic process* $\{W_t\}_{t\geq 0}$ *that fulfills the following properties*

(i) $W_0 = 0$

(ii) $W_{t+h} - W_t$ *is independent of* W_t *for all* $t, h \geq 0$*.* $W_{t+h} - W_t$ *has a distribution that only depends on* h *for all* $t \geq 0$*.*

(iii) $W_{t+h} - W_t \sim N(0, h)$

(iv) W *has continuous sample paths.*

A stochastic differential equation (SDE) with initial condition $X_0 = x$ is written as

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t,$$

which is simply another way of writing the stochastic integral equation

$$X_t = x + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W_s.$$

The first integral is a regular Riemann integral and the second integral is an *Ito integral.* A common stochastic process in financial mathematics is the *geometric Brownian motion.*

Definition 2.2 *A* geometric Brownian motion X_t *satisfies the following SDE*

$$\begin{aligned} \mathrm{d}X_t &= \mu X_t\,\mathrm{d}t + \sigma X_t\,\mathrm{d}W_t, \\ X_0 &= x. \end{aligned}$$

2.1.2 Arbitrage theory

Let a stochastic process S be defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, where $\mathbb{P}$ is the objective, real world measure and $\{\mathcal{F}_t\}$ is the market filtration up to time t. The market is defined by S and B, where B is a risk free asset

$$\mathrm{d}B_t = r(t)B_t\,\mathrm{d}t, \tag{2.1}$$

with risk free rate $r(t)$. The concept of *contingent claim* is defined below.

Definition 2.3 *Consider a financial market with vector price process S. A* contingent claim *with* date of maturity *(exercise date) T, is any stochastic variable $\mathcal{X} \in \mathcal{F}_T^S$. A contingent claim is called a* simple *claim if it is of the form*

$$\mathcal{X} = \Phi(S_T),$$

where Φ is called the contract function*.*

Two of the most common simple claims are the *European call option* and the *European put option.*

Definition 2.4 *A* European call option *with* strike price *K, time to maturity T on an underlying asset S is a simple claim with the payoff function*

$$\Phi(S_T) = \max(S_T - K, 0). \tag{2.2}$$

Similarly, a European put option *has payoff function*

$$\Phi(S_T) = \max(K - S_T, 0). \tag{2.3}$$

The price at time t of a call and put with strike K and time to maturity T will be denoted by $C(t, S_t, K, T)$ and $P(t, S_t, K, T)$. When obvious, the dependence on t and S_t will be suppressed. There exists a simple relationship, *put-call parity*, between a put and a call option

$$C(t, S_t, K, T) - P(t, S_t, K, T) = S_t - D_t - K \cdot e^{-r(T-t)}, \tag{2.4}$$

where D_t is the total discounted value of dividends that will be paid until the maturity date. It is common to talk about the *moneyness* of an option instead of its strike, which is defined below[1].

Definition 2.5 *The* moneyness *of an option is given by K/S, where S is the spot price and K is the strike price.*

The main theorem of arbitrage theory is the risk-neutral valuation formula (RNVF), given below.

Theorem 2.6 *The arbitrage free price at time t of a simple claim $\mathcal{X} = \Phi(S_T)$ with time to maturity T is given by*

$$\Pi(t) = N_t \mathbb{E}^{\mathbb{N}} \left(\frac{\Phi(S_T)}{N_T} \,\middle|\, \mathcal{F}_t \right), \tag{2.5}$$

where $\mathbb{N}$ is the martingale measure with N_t as numeraire. In particular, if the numeraire is chosen as B_t and the risk-free rate is assumed to be constant the arbitrage free price is

$$\Pi(t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}} \left(\Phi(S_T) \mid \mathcal{F}_t \right),$$

where $\mathbb{Q}$ is the risk neutral measure.

[1] Some literature have different definitions of moneyness. The definition used in this thesis is spot moneyness, not forward moneyness.

2.1.3 Black-Scholes

In the Black-Scholes market, the stock price S_t under the objective measure $\mathbb{P}$ is assumed to follow a geometric Brownian motion with constant drift μ and diffusion σ

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t. \tag{2.6}$$

Under the risk-neutral measure $\mathbb{Q}$ the stock price instead follows

$$\mathrm{d}S_t = r S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t. \tag{2.7}$$

Solving the SDE in equation 2.7 one obtains

$$S_T = S_t \exp\left((r - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t) \right),$$

i.e. the stock price is log-normally distributed.

Definition 2.7 *The* implied volatility *of an option is defined as the volatility $\sigma_{BS}(T,K)$ that solves the equation*

$$C(t, S_t, K, T; \sigma_{BS}(T,K)) = C^*(t, S_t, K, T),$$

where C is the theoretical call price given by Black-Scholes and C^* is the market price of the call.

The implied volatility at time t can thus be viewed as a function of strike and time to maturity which is called the *volatility surface*. If the market would follow the Black-Scholes model, the volatility surface would be constant. When time to maturity is set to a fixed value, the implied volatility is a function of only moneyness, and is then called the *volatility smile*. The name suggests the empirical fact that the function is often skewed, forming a "smile". As Derman, Kani, and Zou mention in [18], before the stock market crash of 1987 the volatility smile was relatively constant, but since then almost always implied volatility increases with decreasing strike. Much of the motivation for development of models that came after Black-Scholes has been to correctly capture the skew in the volatility surface. An example of a volatility surface is shown in figure 2.1.

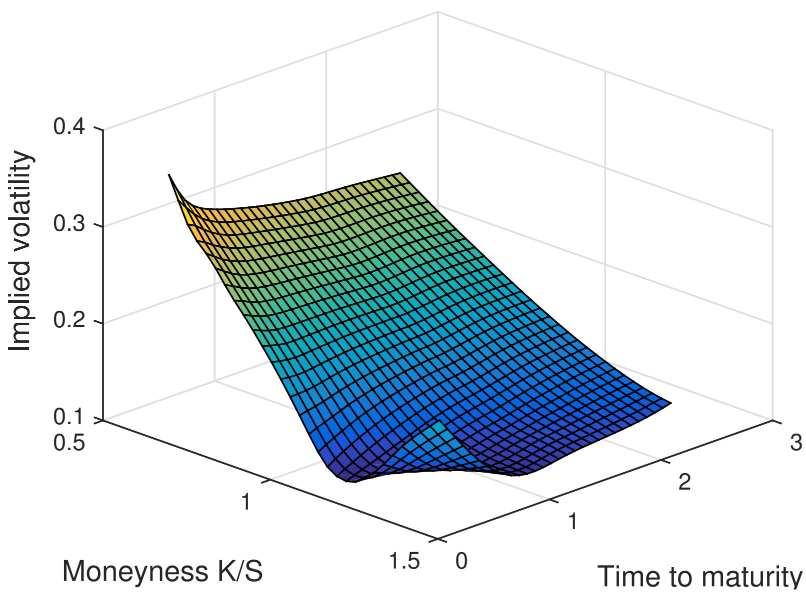


Figure 2.1: The volatility surface of the S&P 500 index on 8th of August 2014.

2.1.4 Heston

The Heston model expands the Black-Scholes model by incorporating stochastic volatility. It was introduced by Heston [25] in 1993. Under $\mathbb{Q}$ it is defined as

$$\begin{aligned} \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \sqrt{V_t}S_t\,\mathrm{d}W_t^{(S)}, \\ \mathrm{d}V_t &= \kappa(\xi - V_t)\,\mathrm{d}t + \sigma_V\sqrt{V_t}\,\mathrm{d}W_t^{(V)}, \\ \mathrm{d}W_t^{(S)}\,\mathrm{d}W_t^{(V)} &= \rho\,\mathrm{d}t. \end{aligned} \tag{2.8}$$

The volatility process V_t is unobservable and follows a Cox-Ingersoll-Ross (CIR) process [16]. It has long term variance $\xi > 0$, mean reversion rate $\kappa > 0$, volatility of volatility $\sigma_V > 0$ and correlation coefficient $-1 \leq \rho \leq 1$. The Heston model therefore relaxes the assumption of constant volatility as opposed to the Black-Scholes model. For the volatility to be strictly positive, the *Feller condition* $2\kappa\xi > \sigma_V^2$ has to be fulfilled.

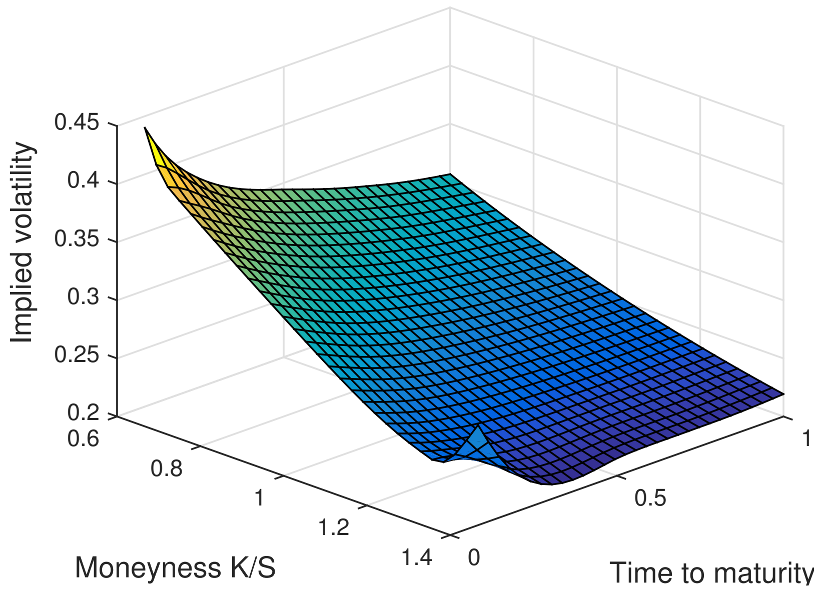


Figure 2.2: The volatility surface implied by the Heston model with $S = 1$, $V = 0.1$, $\kappa = 5$, $\xi = 0.06$, $\sigma_V = 0.5$ and $\rho = -0.7$.

2.1.5 Bates

The Bates model expands the Heston model by including jumps and it was first introduced by Bates in [4]. Using slightly different notation, the model has the following dynamics under $\mathbb{Q}$

$$\begin{aligned} \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \sqrt{V_t}S_t\,\mathrm{d}W_t^{(S)} + S_t\exp\left(J_t - 1\right)\mathrm{d}N_t - a_jS_t\,\mathrm{d}t, \\ \mathrm{d}V_t &= \kappa(\xi - V_t)\,\mathrm{d}t + \sigma_V\sqrt{V_t}\,\mathrm{d}W_t^{(V)}, \\ \mathrm{d}W_t^{(S)}\,\mathrm{d}W_t^{(V)} &= \rho\,\mathrm{d}t, \\ J_t &\sim N(\mu_j, \sigma_j), \\ N_t &\sim Poi(\lambda_j), \\ a_j &= \lambda_j\exp\left(\mu_j + \frac{\sigma_V^2}{2} - 1\right). \end{aligned} \tag{2.9}$$

The jump intensity $\lambda_j > 0$ determines the frequency of jumps, whereas $\mu_j \in \mathbb{R}$ and $\sigma_j > 0$ determines the average size and standard deviation of the jumps respectively.

2.1.6 NIGCIR

The NIGCIR model was developed by Carr et al. in [13]. It is a Normal Inverse Gaussian (NIG) process subordinated with an integrated CIR process. The model is defined under the risk-neutral measure $\mathbb{Q}$ as

$$S_t = S_0 \exp\{rt + X(I_t)\} + \mu_t,$$

where X is a NIG process which has the moment generating function at time t

$$M(z) = \exp\left\{t\delta\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+z)^2}\right)\right\},$$

and where

$$\begin{aligned}
I_t &= \int_0^t V_s \,\mathrm{d}s, \\
\mathrm{d}V_t &= \kappa(\eta - V_t)\,\mathrm{d}t + \sigma_V\sqrt{V_t}\,\mathrm{d}W_t, \\
\mu_t &= -\log \mathbb{E}\left[\exp\left\{I_t\delta\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+1)^2}\right)\right\}\right].
\end{aligned}$$

The parameter restrictions are $\kappa, \eta, \sigma_V, \delta, \alpha > 0$ and $|\beta| < \alpha$.

2.1.7 Example of model parameters

An example of parameter choices from literature for Heston, Bates and NIGCIR can be seen in appendix B.1. These particular parameters are calibrated to the Eurostoxx 50 index on October 7th, 2003 by Schoutens, Simons, and Tistaert in [40].

2.2 Model calibration

2.2.1 Least squares

Given some liquid vanilla options $C^*(t, K_1, T_1), \dots, C^*(t, K_N, T_N)$, where C^* is the mid price

$$C^* = \frac{C^{ask} + C^{bid}}{2},$$

one is interested in finding the most likely or best fitting parameter vector θ. Often this is done by minimizing the value of some loss function $L(\theta)$

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}}\, L(\theta), \tag{2.10}$$

and a common choice used in practice is the ordinary least squares loss function,

$$L^{LS}(\theta) = \sum_{i=1}^{N} (C^*(t, K_i, T_i) - C^{model}(t, K_i, T_i; \theta))^2. \tag{2.11}$$

Other common choices discussed in [34] are weighted least squares (WLS)

$$L^{WLS}(\theta) = \sum_{i=1}^{N} \frac{(C^*(t, K_i, T_i) - C^{model}(t, K_i, T_i; \theta))^2}{(C^{ask}(t, K_i, T_i) - C^{bid}(t, K_i, T_i))^2}, \tag{2.12}$$

and relative least squares (R-LS)

$$L^{R-LS}(\theta) = \sum_{i=1}^{N} \frac{(C^*(t,K_i,T_i) - C^{model}(t,K_i,T_i;\theta))^2}{C^*(t,K_i,T_i)^2}. \tag{2.13}$$

To carry out the minimization numerically, MATLAB offers several different deterministic algorithms in the following functions: `lsqnonlin`, `fminunc`, `fmincon` and `fminsearch`, which use algorithms such as trust-region-reflective least squares, Levenberg-Marquardt, BFGS and Nelder-Mead method. One could also use stochastic optimization techniques such as simulated annealing. The function primarily used in this report will be `lsqnonlin` since it's specifically designed for nonlinear least squares minimization problems and the choice of algorithm is trust-region-reflective least squares (the default choice).

To take into account some of the parameter restrictions in the optimization procedure, a transformation is first applied to the parameter such that the transformed parameters $\tilde{\theta}$, defined on $\mathbb{R}$, can be optimized without fear of them violating their restrictions. As in [34] the log-transform is applied to parameters that are strictly positive, such as $\log(\sigma_V)$ for σ_V in the Heston model and $\operatorname{atanh}(\rho)$ is applied to the correlation parameter ρ that has to be in the interval $[-1, 1]$.

2.2.2 Inverse problems and ill-posedness

The calibration problem presented above is a nonlinear inverse problem, which in general can be formulated as

$$f(\theta) = Y \text{ for } \theta \in \Theta, \tag{2.14}$$

where the inverse f^{-1} is unknown and Θ is some parameter set. As discussed in Gupta [23], an inverse problem is *well-posed* if it satisfies Hadamard's criteria:

(i) For all admissible data, a solution exists.

(ii) For all admissible data, the solution is unique.

(iii) The solution depends continuously on the data.

A problem is *ill-posed* if it violates one or more of the above criteria. When calibrating option pricing models, only the first condition can be guaranteed. For example, multiple parameters might calibrate to within bid-asks spreads and the observed prices might be perturbed by measurement noise. Figure 2.3 and 2.4 show the loss function $L^{WLS}(\theta)$ in two cases that present difficulty in parameter estimation. Both cases show valleys where $L^{WLS}(\theta)$ remains relatively constant.

One can approximate the solution to an ill-posed problem by *regularization*. In the context of calibration of option pricing models, the *penalized* least squares is one possible regularization method; see Cont and Tankov [15] and Lindström et al. [34]. One possible way to formulate this is by considering the loss function

$$L_t^{P-WLS}(\theta) = L_t^{WLS}(\theta) + \|\theta - \theta_{t-1}\|^2_{\Sigma^{-1}}. \tag{2.15}$$

Here θ_{t-1} is the previous day's least squares estimate and

$$\|\theta - \theta_{t-1}\|^2_{\Sigma^{-1}} = (\theta - \theta_{t-1})^T \Sigma^{-1} (\theta - \theta_{t-1}), \tag{2.16}$$

where Σ^{-1} is the inverse of a symmetric, positive semi-definite covariance matrix Σ. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are other regularization methods which are discussed in chapter 3 and used in this thesis. In particular, AIC and BIC adds regularization on the number of model parameters.

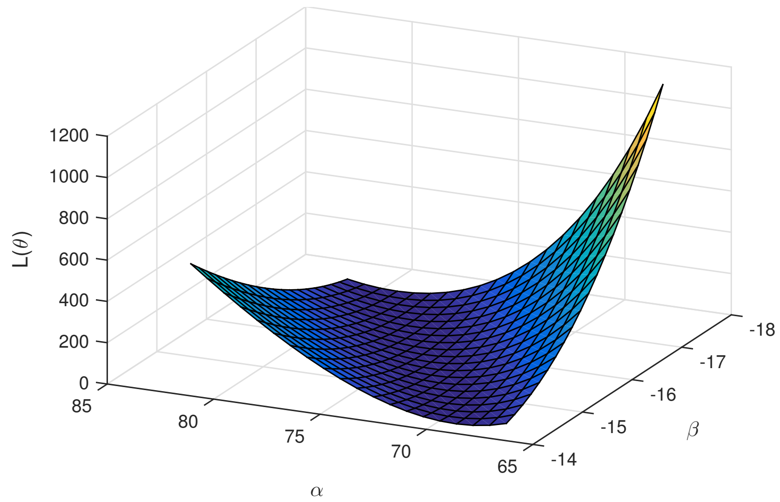


Figure 2.3: The loss function $L^{WLS}(\theta)$ for different α and β in the NIGCIR model (keeping other parameters fixed). The model is calibrated to S&P 500 data on August 11, 2014. The parameter estimates are $\hat{\kappa} = 5.3928$, $\hat{\xi} = 0.8484$, $\hat{\sigma}_V = 5.0636$, $\hat{\delta} = 1.0094$, $\hat{\alpha} = 74.1225$ and $\hat{\beta} = -15.7767$. One can see that $L^{WLS}(\theta)$ has a "valley" where the function is fairly flat.

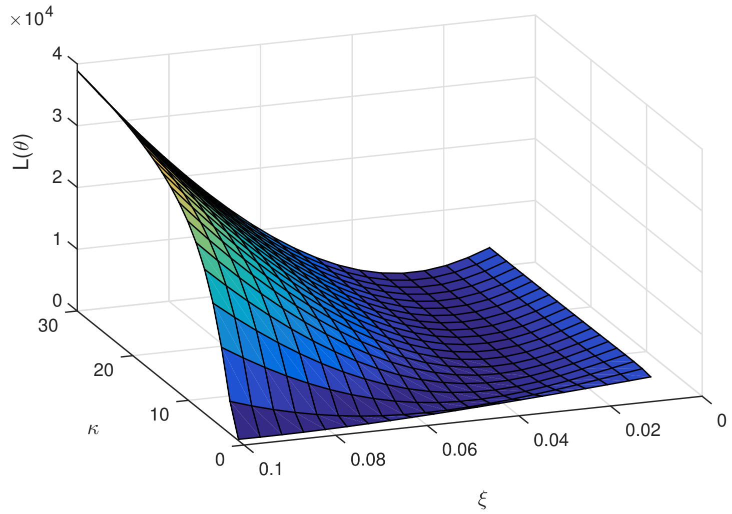


Figure 2.4: The loss function $L^{WLS}(\theta)$ for different ξ and κ in the Heston model. The model is calibrated to *one* time to maturity, $T = 2.37$, on the S&P 500 data, August 11, 2014. The parameter estimates are $\hat{V}_0 = 0.0678$, $\hat{\kappa} = 4.2498$, $\hat{\xi} = 0.0447$, $\hat{\sigma}_V = 1.4260$ and $\hat{\rho} = -0.7817$. Again, one can see that $L^{WLS}(\theta)$ has a "valley" making the parameter κ hard to estimate.

Multimodality in the calibration problem

Due to the ill-posedness of the calibration problem, there can exist more than one local minima of the loss function. Cont and Tankov [15] discuss the sensitivity of the nonlinear least squares calibration problem for exponential Lévy models, which might be sensitive to initial starting points in gradient descent optimization algorithms. As discussed in section 2.2.2 regularization can be used to obtain a well-posed calibration problem, but one could also employ a multistart methodology. The idea is to run multiple instances of local minimization solvers in parallel, starting from different points in the parameter space. All found solutions can then be compared and the ones that calibrate well enough can be kept. The implications on model risk from multimodality in the calibration problem are discussed further in section 3.7.2.

2.3 Option pricing

In addition to the European call and put, the following options will be considered in chapter 4: digital call, up-and-out call and arithmetic Asian call option. Their payoffs are defined below.

Definition 2.8 *The* digital call *has payoff*

$$\Phi(S_T) = \mathbb{1}\{S_T > K\}.$$

The up-and-out call *has payoff*

$$\Phi(S) = \max(S_T - K, 0)\mathbb{1}\left\{\sup_{0\le t\le T} S_t < B\right\}.$$

The arithmetic Asian call *has payoff*

$$\Phi(S) = \max\left\{\sum_{i=1}^{n} S_{t_i}/n - K, 0\right\}.$$

In general, when no analytic formula or approximation exists for the price of an option, there exists three main choices for numerically computing the price:

(i) PDE methods;

(ii) Fourier transform methods;

(iii) Monte Carlo methods.

The latter two will be used in this thesis.

2.3.1 Fourier transform methods

The European call as well as the digital call option can be priced fast and accurately with Fourier methods described in the paper [12] by Carr and Madan. As discussed in Lee [32] and Lindström et al. [34], to price a vanilla European call and digital call an inverse Fourier transform needs to be computed; see [34, eq. 2.1,2.2]. Numerically this can be done with the Fast Fourier Transform (FFT) as in [32] or the Gauss-Laguerre quadrature method in [34]. A Matlab function developed by Wiktorsson that uses the latter method will be used when pricing European and digital calls in chapter 4.

2.3.2 Monte Carlo methods

To calculate the expected value of a random variable X with the crude Monte Carlo method, one can simulate i.i.d. samples $X_1, \dots, X_N$ and form the estimate as

$$\widehat{\mathbb{E}(X)} = \frac{1}{N}\sum_{i=1}^{N} X_i.$$

In the context of option pricing, the risk neutral valuation formula states that the fair price of a derivative is given by an expected value under the $\mathbb{Q}$ measure. Thus, to price a derivative with payoff $\Phi(S_T)$ one can simulate N stock price processes $S^i = \{S^i_t\}_{0\le t\le T}$ for $i = 1, \dots, N$ under $\mathbb{Q}$, and then form the estimate of the price at time t by

$$\widehat{\Pi}(t) = \frac{1}{N}\sum_{i=1}^{N} e^{-r(T-t)}\Phi(S^i). \tag{2.17}$$

The estimate is unbiased and asymptotically normally distributed. The variance of the estimate can be estimated as

$$\mathbb{V}(\widehat{\Phi(S)}) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\Phi(S^i) - \frac{1}{N}\sum_{j=1}^{N}\Phi(S^j)\right). \tag{2.18}$$

Simulation of trajectories

When it is difficult to simulate the stock price process from its exact transition density one can use an Euler discretization as an approximation. Given an SDE

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t,$$

one can simulate X_{t+h} given X_t as

$$X_{t+h} \approx X_t + \mu(t, X_t)h + \sigma(t, X_t)\sqrt{h}Z,$$

where $Z \sim N(0, 1)$. The discretization error approaches zero as the step size h goes to zero.

If one considers only the volatility process V_t in the Heston model, the Euler scheme gives

$$V_{t+h} = V_t + \kappa(\xi - V_t)h + \sigma_V\sqrt{V_t h}Z,$$

with $Z \sim N(0, 1)$. The problem with the above scheme is that there is a possibility that the simulated volatility becomes negative. Instead of using the above scheme, trajectories of the underlying asset and volatility are simulated using the MATLAB code package "Monte Carlo Simulation and Derivatives Pricing" by Kienitz and Wetterau [29], which uses the Quadratic Exponential scheme of Andersen [2]. The above package also offers effective ways of simulating the Bates and NIGCIR models.

Control Variates

One method of decreasing the estimator variance is to use *control variates*, which is described by Glasserman in [21]. Suppose one wishes to estimate the expected value of a random variable Y and let X be a random variable whose expectation is *known*. Define a new random variable $Y(b)$ as

$$Y(b) = Y - b(X - \mathbb{E}(X)),$$

where $b \in \mathbb{R}$. The mean of this variable is

$$\mathbb{E}(Y(b)) = \mathbb{E}(Y - b(X - \mathbb{E}(X))) = \mathbb{E}(Y) - b\mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(Y).$$

Let $(X_1, Y_1), \dots, (X_N, Y_N)$ be i.i.d. samples from X and Y and

$$Y_i(b) = Y_i - b(X_i - \mathbb{E}(X)).$$

One can thus estimate $\mathbb{E}(Y)$ with the sample mean of $Y_1(b), \ldots, Y_N(b)$

$$\widehat{\mathbb{E}(Y)} = \frac{1}{N} \sum_{i=1}^{N} Y_i(b),$$

which is an unbiased and consistent estimator. One is interested in selecting the constant b such that the variance

$$\mathbb{V}(Y(b)) = \mathbb{V}(Y) - 2b\mathbb{C}(Y, X) + b^2\mathbb{V}(X),$$

is minimized. By differentiating with respect to b, the optimal choice can be found as

$$b^* = \frac{\mathbb{C}(X, Y)}{\mathbb{V}(X)}.$$

If the variance of X and the covariance $\mathbb{C}(X, Y)$ between X and Y are unknown, it is possible to use an estimation of b^* as

$$\hat{b}^* = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{N}(X_i - \bar{X})^2},$$

where $\bar{X}$ and $\bar{Y}$ are the sample means. The variance of $Y(b)$ given the optimal b^* is

$$\mathbb{V}(Y(b)) = \mathbb{V}(Y)(1 - \rho_{XY}^2),$$

where ρ_{XY} is the correlation between X and Y. This indicates that a high absolute value of the correlation ρ_{XY} gives a greater variance reduction.

When using control variates for pricing derivatives, a common choice of control variate is the underlying asset $X = S_T$, which has known expectation $\mathbb{E}^{\mathbb{Q}}(S_T|\mathcal{F}_t) = S_t e^{r(T-t)}$. With $Y = e^{-r(T-t)}\Phi(S)$, the estimator becomes

$$\widehat{\Pi}(t) = \frac{1}{N} \sum_{i=1}^{N} \left(e^{-r(T-t)}\Phi(S^i) - b(S_T^i - e^{r(T-t)}S_t) \right). \tag{2.19}$$

Instead of the underlying asset, a European call with strike K and time to maturity T will be used as a control variate, i.e. $X = \max(S_T - K, 0)$. The estimator is now formed as

$$\hat{\Pi}(t) = \frac{1}{N} \sum_{i=1}^{N} \left(e^{-r(T-t)}\Phi(S^i) - b(\max(S_T^i - K, 0) - e^{r(T-t)}C(t, K, T)) \right),$$

where $C(t, K, T)$ is the call price, which can be considered known since the Fourier method is so fast and accurate.

CHAPTER 3

Quantifying model risk

"Essentially, all models are wrong, but some are useful."
— George E. P. Box

3.1 What is model risk?

3.1.1 Value and price approach

In his 1995 research note [17], Derman discusses many of the risks associated with valuation of securities. Much can go wrong in the modelling process; one might assume an incorrect model, implement a model incorrectly or use a model inappropriately. In addition one must be careful when assessing numerical approximation errors, coding errors and unreliable data. Morini summarizes Derman's view of model risk in [36] as:

> Model risk is the risk that the model is not a realistic/plausible representation of the factors affecting the derivative's value.

This view of model risk is called the *value approach*. In contrast to this is the *price approach* by Rebonato, which Morini summarizes as:

> Model risk is the risk of a significant difference between the mark-to-model value of an instrument, and the price at which the same instrument is revealed to have traded in the market.

In other words: the only thing that matters is that the model gives prices that are consistent with the market. To highlight this further, assume that one has access to the following from the market at time t:

- The history of the underlying asset price, $\{S_u\}_{u\in[0,t]}$.
- Prices of N liquid benchmark instruments $\{C^*(u, K_1, T_1), \dots, C^*(u, K_N, T_N)\}_{u\in[0,t]}$.

Using the value approach one would incorporate how well the model is able to explain the history $\{S_u\}_{u\in[0,t]}$ into the risk analysis, while in the price approach only the benchmark instruments would be included. Which viewpoint of model risk one adopts will need to depend on the goal of the analysis.

3.1.2 Model risk AVA

The European Banking Association's draft on Regulatory Technical Standards (RTS, [3]) discusses model risk additional value adjustment in article 11. In paragraph 1 it states:

RTS Article 11 paragraph 1

> Institutions shall estimate a model risk AVA for each valuation model ("individual model risk AVA") by considering valuation model risk which arises due to the potential existence of a range of different models or model calibrations, which are used by market participants, and the lack of a firm exit price for the specific product being valued.

Article 11 paragraph 3 covers the method in which model risk AVA should be calculated:

RTS Article 11 paragraph 3

> Where possible, institutions shall calculate the model risk AVA by determining a range of plausible valuations produced from alternative appropriate modelling and calibration approaches. In this case, institutions shall estimate a point within the resulting range of valuations where they are 90% confident they could exit the valuation exposure at that price or better.

Throughout the thesis, without loss of generality, a *long* position is assumed to be held in the contract under consideration. This means that the point of interest is the 10% quantile of the price distribution of the contract. If instead a short position would be assumed, then the relevant point in the price distribution would be the 90% quantile. In light of the above excerpt from the EBA, the price approach will be taken when quantifying the price distribution. The AVA for a contract when using a particular model for pricing can then be calculated as the difference between the model price and the quantile.

3.2 Academic research on model risk

As discussed in section 2.2.2 one knows how to compute prices given the parameters, but several different parameters might fit the market data, which gives rise to parameter uncertainty. In reality one does not even know from which model class the market data was generated from.

In literature there exists a few different approaches to model risk quantification in option pricing. In [28], Kerkhof, Melenberg, and Schumacher make a distinction between parameter uncertainty and model misspecification, while Cont argues in [14] that no such distinction is necessary. Cont identifies two different schools of thought that have evolved regarding how model uncertainty[1] should be assessed:

1. Worst case approach.
2. Bayesian model averaging.

The development of an axiomatic framework for model risk and uncertainty has been strongly inspired by that of market risk. In [14], Cont presents four axioms for a measure of model uncertainty, available in appendix A. He considers both benchmark instruments where the bid and ask prices are known and illiquid products, where market prices are not known. For a benchmark instrument, a product with a liquid market price, the model risk is in some sense constrained to the bid-ask spread, while one requires more sophistication to quantify it for an illiquid product.

[1] In this thesis no distinction is made between *model risk* and *model uncertainty*. For more discussion on the difference between risk and uncertainty see Cont [14].

Let as before $C_1^*, \ldots, C_N^*$ be observed market prices of benchmark instruments with payoffs $H_1, \ldots H_N$ and $\mathcal{Q}$ be a model set containing arbitrage-free pricing models $\mathbb{Q}$ that satisfy

$$\mathbb{E}^{\mathbb{Q}}(H_i) \in [C_i^{bid}, C_i^{ask}], \qquad \mathbb{E}^{\mathbb{Q}}(|H_i|) < \infty \qquad \text{for all } i \in \{1, \cdots N\}. \tag{3.1}$$

For an exotic derivative $\mathcal{X}$, Cont then proposes a risk measure as

$$\mu_{\mathcal{Q}}(\mathcal{X}) = \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}(\mathcal{X}) - \inf_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}(\mathcal{X}), \tag{3.2}$$

which satisfies the axioms in appendix A. This measure is in some sense a min-max or worst-case risk measure.

Many models are not able to fully capture the market dynamics, giving model prices of benchmark instruments that are not always within bid-ask spreads. In order to be able to include such models in the model set $\mathcal{Q}$, Cont expands the risk measure to allow discrepancies from benchmark option prices by instead adding a punishment for them as

$$\mu_{\mathcal{Q}}(\mathcal{X}) = \sup_{\mathbb{Q} \in \mathcal{Q}} \{\mathbb{E}^{\mathbb{Q}}(\mathcal{X}) - \|C^* - \mathbb{E}^{\mathbb{Q}}(H)\|\} - \inf_{\mathbb{Q} \in \mathcal{Q}} \{\mathbb{E}^{\mathbb{Q}}(\mathcal{X}) + \|C^* - \mathbb{E}^{\mathbb{Q}}(H)\|\}. \tag{3.3}$$

The norm $\|C^* - \mathbb{E}^{\mathbb{Q}}(H)\|$ can here be chosen as any p-norm

$$\|C^* - \mathbb{E}^{\mathbb{Q}}(H)\| = \left(\sum_{i=1}^{N} |C_i^* - \mathbb{E}^{\mathbb{Q}}(H_i)|^p \right)^{1/p}.$$

An alternative axiomatic framework that extends that of Cont is presented by Gupta in his PhD thesis [23]. Gupta also studies the uncertainty in the local volatility model using a Bayesian approach. Other interesting work on model uncertainty is Lindström [33], where a revised RNVF that takes parameter uncertainty into account is presented. In this thesis a method based on Detering and Packham [19] will be implemented and evaluated and it is presented in sections 3.4 and 3.7.

The concept of complete and incomplete markets is important in option pricing. A market is complete if every contingent claim can be replicated. As discussed in Branger and Schlag [8] it is not true that incompleteness of the market is equivalent to existence of model risk. There can still exist model risk even in complete markets. One example would be in the Black-Scholes model (which is complete), where there might exist uncertainty in the volatility σ.

3.3 Bayesian statistics and model uncertainty

Bayesian statistics is based on Bayes' theorem which states, given two events A and B,

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

In the inference setting, where one assumes y is data generated by a model with parameters θ, Bayes' theorem says

$$f(\theta \mid y) = \frac{f(y \mid \theta) f(\theta)}{\int f(y \mid \tilde{\theta}) f(\tilde{\theta})\, \mathrm{d}\tilde{\theta}} \propto f(y \mid \theta) f(\theta).$$

In other words, the posterior density $f(\theta \mid y)$ is proportional to the likelihood $f(y \mid \theta)$ times the prior $f(\theta)$. The difference between frequentist inference and Bayesian inference, is that frequentists use the likelihood while Bayesians use the posterior. The prior density reflects the statisticians subjective knowledge about the parameters and the choice of prior will affect the posterior. A flat improper prior $f(\theta) \propto 1$ can be used when one has no additional knowledge about the parameters.

The method of Bayesian model averaging is presented in [26]. In general consider the problem when one is interested in estimating a model dependent quantity X. The probability of X given the data D is

$$\mathbb{P}(X \mid D) = \sum_{i=1}^{K} \mathbb{P}(X \mid M_i, D)\mathbb{P}(M_i \mid D),$$

where M_i, $i = 1, \ldots, K$ are the models under consideration. The posterior probability for model i is given by

$$\mathbb{P}(M_i \mid D) = \frac{\mathbb{P}(D \mid M_i)\mathbb{P}(M_i)}{\sum_k \mathbb{P}(D \mid M_k)\mathbb{P}(M_k)},$$

where

$$\mathbb{P}(D \mid M_i) = \int \mathbb{P}(D \mid \theta_i, M_i)\mathbb{P}(\theta_i \mid M_i)\,\mathrm{d}\theta_i$$

is the integrated likelihood of model M_i with θ_i being the parameter vector to model i. The priors are thus both on the model parameters $\mathbb{P}(\theta_i \mid M_i)$ and on the models $\mathbb{P}(M_i)$.

In the context of option pricing, Bunnin, Guo, and Ren investigates in [9] how Bayesian model averaging can be used to improve valuation. Consider a set of models $\mathcal{M}$ where each model $M_i \in \mathcal{M}, i = 1, \ldots, m$ gives a fair value

$$\Pi_i(t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}(\Phi(S) \mid \mathcal{F}_t, M_i)$$

of a contingent claim $\mathcal{X} = \Phi(S_T)$. Averaging over the models, one obtains the weighted average price

$$\Pi_{\mathcal{M}}(t) = \sum_{i=1}^{m} e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}(\Phi(S) \mid \mathcal{F}_t, M_i)\mathbb{P}(M_i \mid \mathcal{F}_t). \tag{3.4}$$

The probability $\mathbb{P}(M_i \mid \mathcal{F}_t)$ can here be interpreted as the weight for model M_i. It is worth noting that Bunnin, Guo, and Ren do their averaging under the objective measure $\mathbb{P}$ and not under $\mathbb{Q}$.[2]

As mentioned the goal is to find the 10% quantile of the price distribution and from this infer the model risk AVA. Practically, two interpretations of the EBA guide lines to estimating the quantile of the price distribution of an exotic option would be:

(i) Specify a finite, predetermined set of pricing models $\mathcal{Q} = \cup_{i=1}^{m}\mathbb{Q}_i$ where each pricing model is assigned a probability weight w_i depending on how well it can reproduce the benchmark instruments C^*. The weights w_i are interpreted (approximately) as posterior probabilities.

(ii) Sampling from the posterior $p_{\Pi|C^*}(z)$. The quantile α is then found as

$$\mathbb{P}(\Pi < \alpha \mid C^*) = \int_{-\infty}^{\alpha} p_{\Pi|C^*}(z)dz = 0.1. \tag{3.5}$$

[2] From now on when discussing models and model sets in general the notation M_i and $\mathcal{M}$ will be used and when specifically option pricing models are discussed the notation $\mathbb{Q}_i$ and $\mathcal{Q}$ are used.

These are two different viewpoints on how one should approach the problem and the focus of this thesis will be on (i). If one is interested in using the second approach, sampling from the posterior of the price distribution, one would have to resort to, for example, Markov Chain Monte Carlo methods. However, these methods have drawbacks:

(i) It might be difficult to determine how one should specify the priors of the model classes as well as the parameters. This requires statistical sophistication by the user. One way of choosing a prior density on the parameters θ is to consider $p(\theta) \propto \exp\left\{-\frac{1}{2}\|\theta - \theta^{(t-1)}\|^2_{\Sigma^{-1}}\right\}$ where θ_{t-1} could be the previous day's least squares parameters.[3]

(ii) MCMC methods might require long burn in periods before they reach the invariant distribution. In the option pricing context in each step new prices must be computed to evaluate the likelihood, which is computationally intensive, although the Fourier transform methods are generally very fast.

(iii) The most common MCMC algorithms such as Metropolis-Hastings suffer from "curse of dimensionality" and they might get trapped a long time in a local mode.

(iv) Good mixing often relies on careful choice of tuning parameters.

Nevertheless, MCMC methods have been widely applied in litterature to calibration of market models, such as the Heston model. In most articles the calibration is done to historic prices of the underlying, although some articles take into account both the moves of the underlying and option prices (see for example [37]). As previously mentioned, one especially interesting use of MCMC is [23] where the uncertainty in the local volatility model is analyzed.

One difficulty when using MCMC in the general model risk setting is that risk comes from both misspecification of the parameters and the model class. To price a derivative one needs to know the pair $x = (\theta^{(k)}, \mathbb{Q}_k)$ where $\theta^{(k)}$ is some parameter vector that belongs to the parameter space of model $\mathbb{Q}_k$. Thus one would need to simulate one chain for each model class or a chain $x_1, \dots, x_n$ which include jumps between model classes and where the model classes can have different number of dimensions. An example of such *trans-dimensional* sampling is the *reversible jump* MCMC, presented in [22]. Trans-dimensional sampling methods have however not seen much application in the literature of financial economics.

While MCMC is a powerful method that allows for sampling from the posterior distribution, it might not be the most suitable method for determining model risk AVA. First, there are the implementation difficulties discussed above, but one should also take into consideration what the EBA actually calls for in their technical standards. One should take into account "a range of plausible valuations produced from alternative modelling and calibration approaches", but for example the standard Metropolis-Hastings algorithm does not incorporate any least squares calibration. A simpler and more appropriate methodology for model risk AVA is presented in the following sections 3.4 and 3.7 and tested in chapter 4. Despite these considerations, the Metropolis-Hastings algorithm will be used to analyse the special case of the Heston model under perfect calibration. It is presented in section 4.3.1.

3.4 Model weights with AIC and BIC

The *Aikaike information criterion* (AIC) and *Bayesian information criterion* (BIC) are two information criteria that can be used for model order selection (see [1, 41]). These criteria

[3]This is the Bayesian interpretation of the regularization discussed in section 2.2.2.

allow for comparison between non-nested models, which cannot be compared with ordinary statistical tests such as likelihood ratio tests. In both criteria, the likelihood of each model is penalized with the number of parameters used in the model. This penalization is added due to the fact that models with a larger number of parameters are expected to have a higher likelihood. A result of the increasing penalty with number of parameters is that overfitting can be prevented.

Definition 3.1 *The* Akaike information criterion *is defined as*

$$AIC = -2\log(\mathcal{L}) + 2k,$$

where k is the number of independent model parameters and $\mathcal{L}$ is the maximum likelihood of the model.

The AIC is an unbiased estimate of the expected Kullbeck-Leibler distance

$$I(\theta_k) = \mathbb{E}\left(\log\left(\frac{g(y)}{f(y|\theta_k)}\right)\right), \tag{3.6}$$

where g is the density of the true (but unknown) data generating process, and f is the density of the model under consideration.

Definition 3.2 *The* Bayesian information criterion *is defined as*

$$BIC = -2\log(\mathcal{L}) + k\log(N),$$

where N is the sample size.

A detailed derivation of the BIC is given in [5], starting from an approximation of the likelihood

$$\mathbb{P}(D \mid M_i) = \int \exp\left(\log(p(D \mid \theta_i)p(\theta_i))\right) \mathrm{d}\theta_i. \tag{3.7}$$

When using AIC or BIC for model selection, the model that has the lowest value is selected as the most suitable model. The first thing that can be noted about the difference between the two is that BIC penalizes based on the sample size n, whereas AIC does not. Based on the definition of AIC and BIC, it can also be seen that BIC will, in general, penalize an increasing number of parameters harder than AIC would.

Given a set of models $\{M_1, \ldots, M_m\}$ with equal prior probabilities, the posterior of model i can be approximated by

$$\mathbb{P}(M_i|D) \approx \frac{\exp\left\{-\frac{1}{2}BIC_i\right\}}{\sum_{k=1}^{m}\exp\left\{-\frac{1}{2}BIC_k\right\}}, \tag{3.8}$$

where BIC_i is the BIC score of model i (see [38] for more discussion on BIC). Note that there is an implicit assumption that the true model belongs to the model set. In a similar way, AIC scores can also be transformed into model weights

$$w_i = \frac{\exp\left\{-\frac{1}{2}AIC_i\right\}}{\sum_{k=1}^{m}\exp\left\{-\frac{1}{2}AIC_k\right\}}. \tag{3.9}$$

As discussed in Burnham and Anderson [10] the AIC weights can also be interpreted as approximations of posterior probabilities but not under the assumption of equal prior probabilities of the models. An obvious drawback of AIC and BIC is that it is not clear how one should handle models with an infinite number of parameters, or models where the number of parameters increases with the data. In practice, one can often either find a parametrization of such a model, or compute the "effective" number of parameters in a semiparametric model.

3.5 Likelihood function

Some assumptions regarding the relationship between the prices given by the model and the observed market prices has to be made in order to set up a likelihood function. In some sense the choice of likelihood function is subjective and different choices will effect the resulting inference. Several choices are presented in this section, and the Gaussian likelihood from section 3.5.1, as well as the "flat-top Gaussian" likelihood from section 3.5.3 are later tested in chapter 4.

3.5.1 Likelihood assuming Gaussian error terms

One needs to assume a structure of the discrepancies between a model's prices and the market prices

$$\varepsilon_j = C^*(K_j, T_j) - C^{model}(K_j, T_j) \text{ for } j = 1 \dots N. \tag{3.10}$$

In [19], Detering and Packham assume that ε_j are i.i.d. realizations of a zero-mean normal distributed variable with constant variance σ^2. The likelihood for the realization $\varepsilon_1, \dots, \varepsilon_N$ under the normal distribution is then given by

$$\mathcal{L}(\sigma \mid \varepsilon_1, \dots, \varepsilon_N) = \prod_{j=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} e^{-\varepsilon_j^2/2\sigma^2}. \tag{3.11}$$

Taking the logarithm of $\mathcal{L}$, one gets the log-likelihood

$$\ell = -\frac{N}{2}\left(\log(2\pi) + \log\sigma^2 + \frac{\sum_{j=1}^{N}\varepsilon_j^2}{N\sigma^2}\right). \tag{3.12}$$

The mean square error (MSE) for model $\mathbb{Q}_i$ is

$$MSE_i = \frac{1}{N}\sum_{j=1}^{N}\varepsilon_j^2, \tag{3.13}$$

which can be used as the maximum likelihood estimator for the unknown noise variance σ^2 under model $\mathbb{Q}_i$. The log-likelihood (3.12) can then be formulated as

$$\begin{aligned} \ell_i &= -\frac{N}{2}\left(\log(2\pi) + \log MSE_i + \frac{N \cdot MSE_i}{N \cdot MSE_i}\right) \\ &= -\frac{N}{2}\left(\log(2\pi) + \log MSE_i + 1\right). \end{aligned} \tag{3.14}$$

It can be noted that the relation between model prices and market prices (3.10) as described by Detering and Packham assumes that the order of magnitude of model errors is the same for all options. It might be reasonable to take into account the bid-ask spread of a quoted option, since a wide bid-ask spread indicates a larger uncertainty in the price. To take this into consideration, the errors are normalized with the bid-ask spread as

$$\varepsilon_j = \frac{C^*(K_j, T_j) - C^{model}(K_j, T_j)}{C^{ask}(K_j, T_j) - C^{bid}(K_j, T_j)}. \tag{3.15}$$

An equivalent way of writing the above is

$$C^*(K_j, T_j) \sim N(C^{model}(K_j, T_j), \sigma^2[C^{ask}(K_j, T_j) - C^{bid}(K_j, T_j)]), \tag{3.16}$$

i.e. the observed market price is the true price disturbed by some Gaussian noise proportional to the bid-ask spread.

There is a close connection between the assumptions of the distribution of the pricing errors, and the loss functions discussed in section 2.2.1. The WLS estimate works under the assumption of i.i.d. normal ε_j as defined in equation 3.15.

3.5.2 Likelihood using implied volatility

Instead of defining the errors as the difference in model and market price, one can instead look at the difference in model implied volatility and market implied volatility

$$\varepsilon_j = \sigma^*_{\mathrm{BS}}(K_j, T_j) - \sigma^{model}_{\mathrm{BS}}(K_j, T_j) \text{ for } j = 1 \dots N, \tag{3.17}$$

with ε_j being i.i.d. normal distributed with variance σ^2. One could also define ε_j as

$$\varepsilon_j = \frac{\sigma^*_{\mathrm{BS}}(K_j, T_j) - \sigma^{model}_{\mathrm{BS}}(K_j, T_j)}{\sigma^{ask}_{\mathrm{BS}}(K_j, T_j) - \sigma^{bid}_{\mathrm{BS}}(K_j, T_j)} \text{ for } j = 1 \dots N. \tag{3.18}$$

One advantage of using implied volatility over prices is that the errors in implied volatility are similar in order of magnitude for all options.

3.5.3 Likelihood assuming flat-top Gaussian distribution

One alternative way to deal with option prices that are known only within the interval $[C^{bid}, C^{ask}]$ is to assume that the error terms

$$\varepsilon_j = \frac{C^*(K_j, T_j) - C^{model}(K_j, T_j)}{C^{ask}(K_j, T_j) - C^{bid}(K_j, T_j)},$$

are realizations of a distribution with pdf

$$p_j(\varepsilon) = \frac{\Delta_j}{\sqrt{2\pi}\sigma + \Delta_j} \exp\left\{-\frac{\Delta_j^2}{2\sigma^2}\left[\mathbb{1}\left\{\varepsilon < -\frac{1}{2}\right\}\left(\varepsilon + \frac{1}{2}\right)^2 + \mathbb{1}\left\{\varepsilon > \frac{1}{2}\right\}\left(\varepsilon - \frac{1}{2}\right)^2\right]\right\},$$

for $\sigma > 0$. Here $\Delta_j = C^{ask}(K_j, T_j) - C^{bid}(K_j, T_j)$ is the bid-ask spread. As $\sigma \to 0$ this distribution will approach the uniform distribution $U(-1/2, 1/2)$, motivating a slight addition to the definition above as

$$p_j(\varepsilon) = \begin{cases} \frac{\Delta_j}{\sqrt{2\pi}\sigma + \Delta_j} \exp\left\{-\frac{\Delta_j^2}{2\sigma^2}\left[\mathbb{1}\left\{\varepsilon < -\frac{1}{2}\right\}\left(\varepsilon + \frac{1}{2}\right)^2 + \mathbb{1}\left\{\varepsilon > \frac{1}{2}\right\}\left(\varepsilon - \frac{1}{2}\right)^2\right]\right\}, & \text{for } \sigma > 0 \\ \mathbb{1}\{-\frac{1}{2} \leq \varepsilon \leq \frac{1}{2}\}, & \text{for } \sigma = 0. \end{cases}$$

The idea behind this distribution builds on the assumption that two models that both calibrate within the bid-ask spread of all observed options should have the same likelihood. The assumption is fulfilled since $\varepsilon \in [-\frac{1}{2}, \frac{1}{2}]$ is equivalent to $C^{model} \in [C^{bid}, C^{ask}]$. The pdf is shown in figure 3.1. The probability of the model price being inside a bid-ask spread is simply

$$\mathbb{P}\left(C^{model}(K_j, T_j) \in [C^{bid}(K_j, T_j), C^{ask}(K_j, T_j)]\right) = \frac{\Delta_j}{\sqrt{2\pi}\sigma + \Delta_j}. \tag{3.19}$$

The log-likelihood under this noise structure is

$$\begin{aligned} \ell_i =& N\left(\log(\Delta_j) - \log(\sqrt{2\pi}\sigma_i + \Delta_j)\right) \\ &- \sum_{j=1}^{N} \Delta_j^2 \left(\frac{\mathbb{1}\left\{\varepsilon_j < -\frac{1}{2}\right\}\left(\varepsilon_j + \frac{1}{2}\right)^2 + \mathbb{1}\left\{\varepsilon_j > \frac{1}{2}\right\}\left(\varepsilon_j - \frac{1}{2}\right)^2}{2\sigma_i^2}\right), \end{aligned}$$

which can be maximized numerically to find the estimate of σ_i. One instance where a flat-top Gaussian kernel has been used is in [31, pp. 29], however we have not yet seen any financial applications of the distribution at the time of writing this thesis.

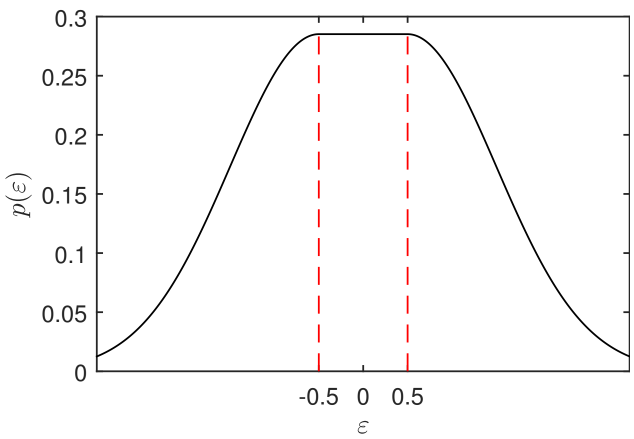


Figure 3.1: Example of the flat-top pdf of the measurement error with $\sigma = 1$ and $\Delta = 1$.

One should note that the different likelihoods presented are not entirely realistic, since in them there is no restriction that the price of the benchmark instrument or the implied volatility has to be positive. To take this issue into account one could instead look at the errors defined as

$$\varepsilon_j = \frac{\log(C^*(K_j, T_j)) - \log(C^{model}(K_j, T_j))}{C^{ask}(K_j, T_j) - C^{bid}(K_j, T_j)}.$$

However, in practice the probability of obtaining a negative call price (or implied volatility) assuming the structure in equation 3.15 is negligible.

3.5.4 Likelihood using call option price surface integral

A more sophisticated noise structure is given by Duembgen and Rogers in [20]. Assume that the models call option surface is disturbed by a Gaussian noise in the following way

$$p(C^* \mid C^{model}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{\frac{-Q(C^*, C^{model})}{2\sigma^2}\right\},$$

where $Q(C^*, C^{model})$ is some quadratic penalty function. One example of a possible quadratic loss function is

$$Q(C^*, C^{model}) = \int_0^T \int_0^\infty \{c_K(v, K) - c_K^j(v, K)\}^2 dK dv. \tag{3.20}$$

Here, $c_K(v, K)$ denotes

$$c_K(v, K) = \frac{\partial}{\partial K} C(v, KS_0)/S_0.$$

In practice only a finite number of options are observed, which means that the double integral in equation 3.20 needs to be approximated numerically, using for example the trapezoidal rule.

The pricing errors for two options which have very similar strike and time to maturity should be highly correlated. The other choices of likelihood assume an independence between individual pricing errors, which will be unreasonable if one keeps adding more options to the option surface. In contrast, the surface integral will not suffer from this drawback, since a finer option surface will only result in a smaller discretization error.

3.6 Risk measure

The weighted average price $\Pi_{\mathcal{Q}}(t)$ and the 10% quantile α defined in equations 3.4 and 3.5 respectively can be used to define a new risk measure.

Definition 3.3 *The* absolute model risk measure *is defined as*

$$\tilde{\mu}(\mathcal{X}) = \Pi_{\mathcal{Q}}(t) - \alpha. \tag{3.21}$$

It is clear that derivatives with a large price uncertainty will have a price distribution with fatter tails, giving a greater distance between the weighted average and the percentile, and thus a greater value of $\tilde{\mu}(\mathcal{X})$.

To compare the risk measure for different derivatives one can also consider the relative risk measure.

Definition 3.4 *The* relative model risk measure *is defined as*

$$\mu(\mathcal{X}) = \frac{\Pi_{\mathcal{Q}}(t) - \alpha}{\Pi_{\mathcal{Q}}(t)}. \tag{3.22}$$

If we know that we will use a certain model $\mathbb{Q}$ in $\mathcal{Q}$ to price the derivative when trading it might be more relevant to look at the difference between this models price and the quantile of the price distribution.

Definition 3.5 *The* absolute model risk measure *with respect to* $\mathbb{Q} \in \mathcal{Q}$ *is defined as*

$$\tilde{\mu}_{\mathbb{Q}}(\mathcal{X}) = \Pi_{\mathbb{Q}}(t) - \alpha. \tag{3.23}$$

Definition 3.6 *The* relative model risk measure *with respect to* $\mathbb{Q} \in \mathcal{Q}$ *is defined as*

$$\mu_{\mathbb{Q}}(\mathcal{X}) = \frac{\Pi_{\mathbb{Q}}(t) - \alpha}{\Pi_{\mathbb{Q}}(t)}. \tag{3.24}$$

In most applications the model $\mathbb{Q}$ would be the model in $\mathcal{Q}$ with the highest probability weight or the model used by the trader when pricing the derivative.

The risk measures above are strongly influenced by the definition of model risk AVA. One can note that they are conceptually closely related to the market risk concept of value-at-risk (VaR), since VaR is essentially a quantile in the loss distribution. In [19] a VaR-like model risk measure is introduced, and it is shown that it fails to fulfil Cont's axiom of convexity.

Similar to Gupta in [23] it is also possible to consider a measure of the absolute deviation[4] with the following definition.

Definition 3.7 *The* model risk measure of absolute deviation *is defined as*

$$\bar{\mu}(\mathcal{X}) = E^{Q}[|\mathbb{E}^{\mathbb{Q}}(\mathcal{X}) - E^{Q}[\mathbb{E}^{\mathbb{Q}}(\mathcal{X})]|], \tag{3.25}$$

where Q is the probability measure of the models in $\mathcal{Q}$.

Note here that $E^{Q}[\cdot]$ is the expected value of *prices* over the model set $\mathcal{Q}$ and $\mathbb{E}^{\mathbb{Q}}[\mathcal{X}]$ is the expected value of an option payoff $\mathcal{X}$ under the risk neutral measure $\mathbb{Q}$. Given model weights w_i and prices $\Pi_i(t)$ for the contingent claim $\mathcal{X}$, the risk measure of absolute deviation is calculated as

$$\bar{\mu}(\mathcal{X}) = \sum_i w_i \Big|\Pi_i(t) - \sum_k w_k \Pi_k(t)\Big|. \tag{3.26}$$

[4] Gupta eventually worked with a measure slightly different than this.

3.6.1 Sample quantile

The quantile α is calculated in this thesis numerically using a MATLAB package inspired by Hyndman and Fan [27]. Assume all model prices $\Pi_i(t)$, $i = 1 \dots m$ are sorted in increasing order of price and that the sum of all weights is 1. For each model weight w_i, the plotting position can be defined as

$$p_i = \sum_{k=1}^{i} w_k - \frac{w_i}{2}, \tag{3.27}$$

which can be looked upon as the midpoint between each step in the cumulative sum of weights. This particular choice of plotting position is common in hydrology and satisfies all desirable properties for a sample quantile defined by Hyndman and Fan. The 10% sample quantile can then be estimated through linear interpolation between the model pairs $\{p_l, \Pi_l(t)\}$ and $\{p_{l+1}, \Pi_{l+1}(t)\}$, where $p_l < 0.1 < p_{l+1}$. The interpolated price $\Pi(t)$ at $p = 0.1$ is then the sample quantile α.

3.7 Model risk using AIC or BIC

Detering and Packham apply the idea of AIC weights to quantify the model risk with respect to hedging. In [30], Kivilo and Olofsson apply AIC model weight idea of Detering to calculation of model risk AVA for exotic option based on data from the article [40]. As discussed in section 3.4 model weights can be formed as

$$w_i = \frac{\exp\left\{-\frac{1}{2} IC_i\right\}}{\sum_k \exp\left\{-\frac{1}{2} IC_k\right\}}, \tag{3.28}$$

where IC_i is AIC_i or BIC_i for model i. There exists different choices of the model log-likelihood ℓ_i, where several choices are presented in section 3.5. The general algorithm for calculating the model risk for an exotic option using a finite model set can then be found in algorithm 1. In chapter 4 this algorithm is tested on real and simulated data and practical considerations are discussed.

Algorithm 1 Model risk with finite and fixed model population

Require: Current stock price S_t, risk-free rate r, market prices of benchmark instruments $C_1^*, \dots, C_N^*$, a model set $\mathcal{Q} = \cup_{i=1}^m \mathbb{Q}_i$, payoff function Φ of an exotic derivative.
  for $i = 1 : m$ **do**
    Calculate weights w_i for model $\mathbb{Q}_i$ from benchmark instruments using expression 3.28.
    Calculate the price $\Pi_i(t)$ under $\mathbb{Q}_i$ using the RNVF, equation 2.5.
  end for
  return 10% quantile of the price distribution and risk measure μ from section 3.6.

3.7.1 Choosing model set

The question of how to choose the model set $\mathcal{Q}$ is not trivial. The directions given by the EBA are fairly vague, but one should consider different models and calibration approaches. One could, for each model class, use different calibration procedures (for example different loss functions in least squares), and thus include a model $\mathbb{Q}_i$ in $\mathcal{Q}$ for each calibration. If there are only a few model classes, and one model receives the highest weight, then it is likely that the measure of model risk will only give crude and sensitive answers. For example, if one has 20 models in $\mathcal{Q}$ and $w_1 = 0.91$, $w_2 = 0.08$ and $\sum_{i=3}^{20} w_i = 0.01$, then an interpolation method used to calculate the 10% quantile might itself introduce large uncertainty.

Instead of doing the above, one could try to include many models from regions of high likelihood in the parameter space. One way of doing this is to (for each model class) first do a least squares calibration, yielding a parameter vector $\hat{\theta}$. The estimate $\hat{\theta}$ serves as a reference point to a region of high likelihood from which one can place points in the vicinity of $\hat{\theta}$. For each parameter, an interval can be spanned, thus yielding a d-dimensional hyperrectangle for a model with d parameters. One way of finding the size of the hyperrectangle, which is used in this thesis, is described in algorithm 2. This is inspired by Detering and Packham in [19], where the included parameter ranges are spanned such that the complementary parameter space has negligible weight. Parameter vectors could be chosen from this hyperrectangle deterministically by using a grid, or they could be chosen by uniform Monte Carlo sampling. In chapter 4 uniform random sampling will be used when selecting models from the hyperrectangle.

Algorithm 2 Finding hyperrectangle

Require: Least squares estimate $\hat{\theta}$, weight threshold ω, market prices of benchmark instruments.
  Calculate model prices, likelihood (section 3.5) and information criterion value $IC(\hat{\theta})$ for the least squares model.
  for each model parameter $\hat{\theta}_i$ **do**
    $c \leftarrow 10^{-3}$
    while upper search has not converged **do**
      start from $(1+c)\cdot\hat{\theta}_i$ and numerically search for the $\theta_i^{upper} > \hat{\theta}_i$ that minimizes $\left|\exp\left\{\left(IC(\theta) - IC(\hat{\theta})\right)/2\right\} - \frac{1}{\omega}\right|$, where $IC(\theta)$ is the information criterion value for the model with parameter set $\theta = [\hat{\theta}_1, \cdots, \hat{\theta}_{i-1}, \theta_i^{upper}, \hat{\theta}_{i+1}, \cdots, \hat{\theta}_N]$.
      $c \leftarrow x_{start} \cdot 2$
    end while
    $c \leftarrow 10^{-3}$
    while lower search has not converged **do**
      start from $(1-c)\cdot\hat{\theta}_i$ and numerically search for the $\theta_i^{lower} < \hat{\theta}_i$ that minimizes $\left|\exp\left\{\left(IC(\theta) - IC(\hat{\theta})\right)/2\right\} - \frac{1}{\omega}\right|$, where $IC(\theta)$ is the information criterion value for the model with parameter set $\theta = [\hat{\theta}_1, \cdots, \hat{\theta}_{i-1}, \theta_i^{lower}, \hat{\theta}_{i+1}, \cdots, \hat{\theta}_N]$.
      $c \leftarrow x_{start} \cdot 2$
    end while
  end for
  return upper and lower parameter bound vectors θ^{upper} and θ^{lower} respectively.

The reason for repeatedly trying to find the bounds for different starting points by varying c in algorithm 2 is because the weight ratio $\exp\left\{\left(IC(\theta) - IC(\hat{\theta})\right)/2\right\}$ is constant and equal to 1 in an area around the least squares estimate for models that calibrate perfectly when the flat-top likelihood is used. This causes problems for the numerical routine `fminsearch`, which is eased by starting from parameter values farther and farther away from the least squares estimate until `fminsearch` is able to converge. Note in particular that

$$\left|\exp\left\{\left(IC(\theta) - IC(\hat{\theta})\right)/2\right\} - \frac{1}{\omega}\right|$$

is zero when the ratio in weight between the least squares estimate and the model with parameter vector θ is equal to the weight threshold ω.

Unlike the method presented above, Kivilo and Olofsson only include one model, the least squares estimate, from each model class. They also investigate the use of interpolation to find the quantile of the price distribution, which is more important when dealing with few models. In this thesis, interpolation does not make as great a difference since a higher total number of models is present, causing the cumulative sum of weights to be fairly smooth.

3.7.2 Multimodality

As discussed in section 2.2.2, it is not guaranteed that the solution to the calibration problem is unique and several parameter sets found in different local minimas might calibrate well to market data. When multiple solutions are found, the model set can be spanned over each minima, again using algorithm 2.

CHAPTER 4

Results

4.1 The S&P 500 dataset

The dataset at our disposal consists of daily option data from August, 2014 on the S&P 500 index. For each option, the following information is available:

- Underlying spot price
- Maturity date
- Strike level
- Bid price
- Ask price

The traded options on this index are very liquid and for other indices there may be wider bid-ask spreads. For many pairs of maturity date and strike level, there exists bid and ask prices for both put and call options. Only these pairs will be used since both put and call prices are needed to calculate a dividend adjusted spot price of the underlying index as

$$S_t - D_t = C(t, S_t, T, K) - P(t, S_t, T, K) + K \cdot e^{-r(T-t)},$$

which comes from the put-call parity in equation 2.4. The dividend adjusted spot price is used as the underlying spot price for all option pricing in order to take the dividends into account. The risk-free rates are extracted from yield curve data, based on overnight indexed swaps. The yield curve of August 8^{th} can be seen in figure 4.1.

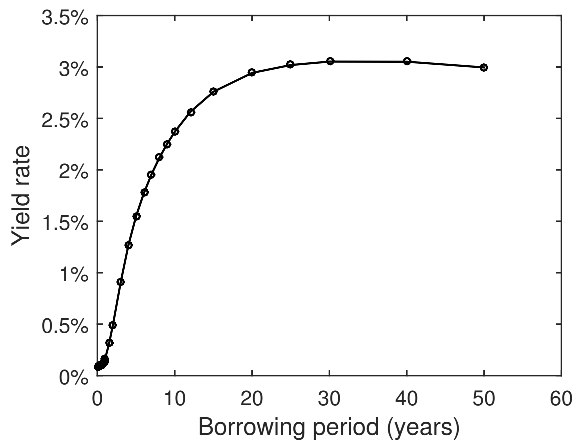


Figure 4.1: Shows the yield curve from August 8, 2014.

With the intention of reducing the large number of options available and focus on the most liquid options, the following rules are defined for including an option in the calibration set:

- Moneyness in the range of $[0.6, 1.4]$;
- Time to maturity between 0.25 and 2.5 years.

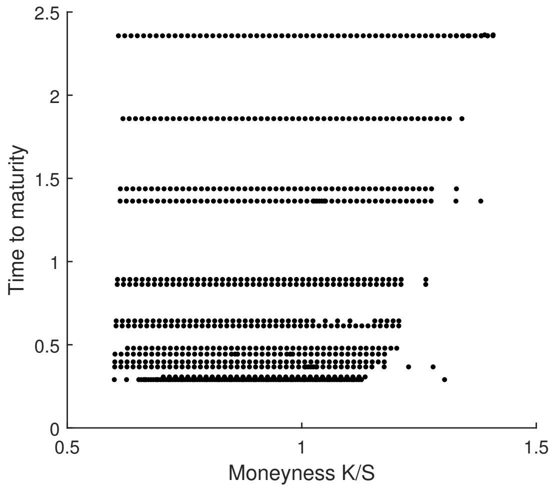
The reason for setting the lower limit of time to maturity at 3 months is that a higher level of liquidity for the studied options is probable by excluding short maturities. Another reason is that the density of options is much higher for short maturities in the dataset, which would cause the calibration to become too heavily affected by short maturities without some systematical exclusion.

In order to test the model risk framework under different market conditions, a couple of days have been selected for which results will be presented:

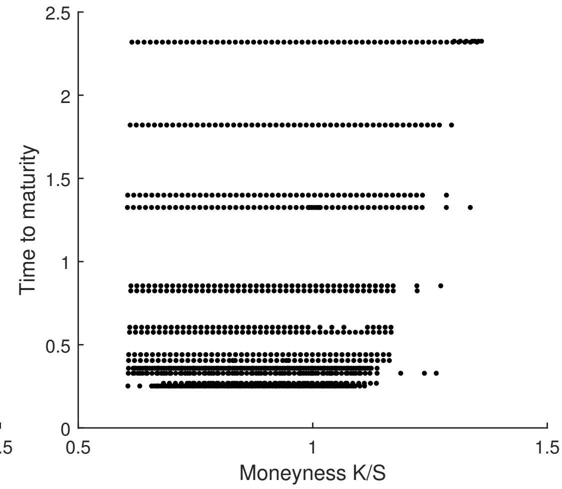
(i) August 8, 2014 when Heston, Bates and NIGCIR all calibrate perfectly;

(ii) August 22, 2014 when one model class attains almost all model weight.

The chosen data set then contains 750 and 880 call options for the given days respectively and the grids of call option prices can be seen in figure 4.2.



(a) 8 August 2014



(b) 22 August 2014

Figure 4.2: Call option grids for the S&P 500 index.

In order to speed up the calculations, subsets of the data are created where every 10th option is selected. This gives two sets of 75 and 88 options respectively. The data set from August 8$^{\text{th}}$ will be studied in more detail. One million trajectories are simulated for each model when pricing barrier options and Asian options in order to keep the variance low. For each trajectory, the number of steps taken in time equals the number of days to maturity for the option priced. Because of memory constraints, pricing is done for 10^4 trajectories 100 times, where the call option is used as a control variate. The final price is then taken as the average of the 100 prices.

4.1.1 Models in the set

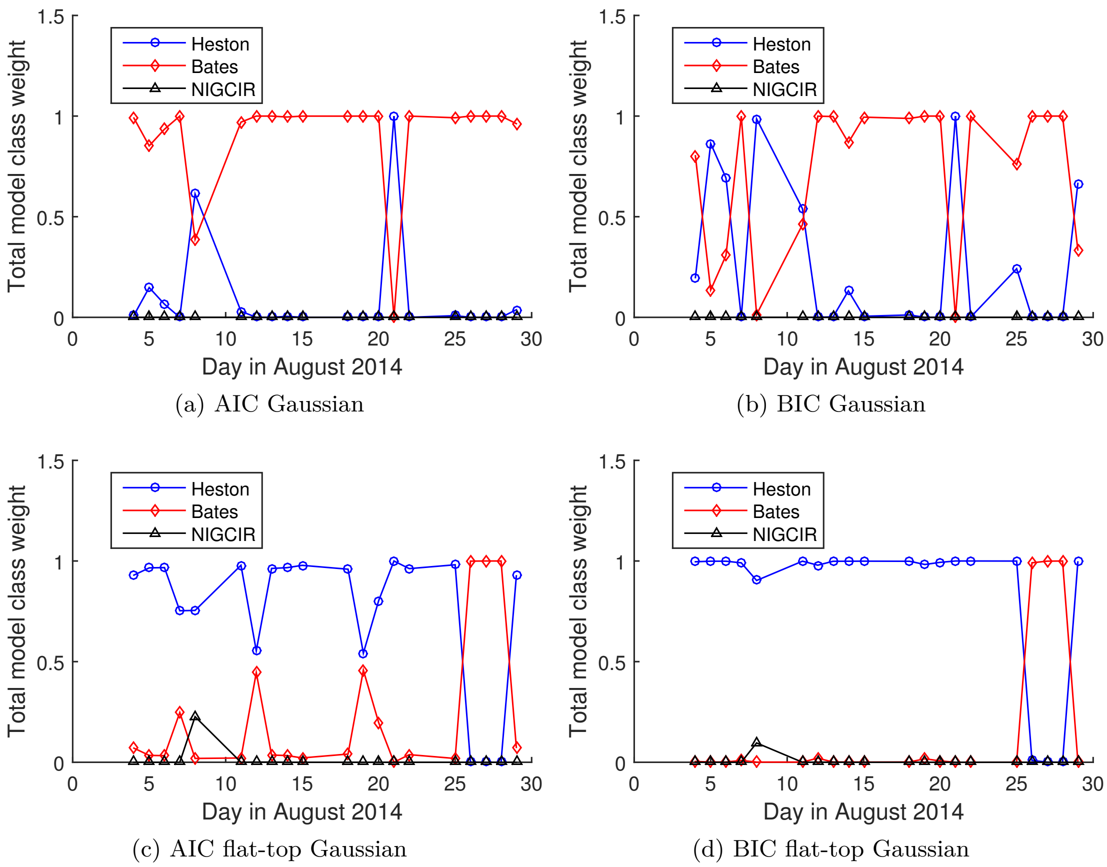
The model classes used in the empirical study are Black-Scholes, Heston, Bates and NIG-CIR. For each model class, the weighted least squares estimate is found, starting from some indiscriminate initial point. Models to include in the set are then sampled uniformly in an interval around the least squares estimate for each parameter dimension. Note that the process of sampling is simply for convenience and one could instead create an equidistant grid over the parameter space.

A question to consider is how wide the intervals around the least squares estimates should be. Algorithm 2 is used to study the parameters one by one, letting the others be fixed and search for the parameter interval where the model weights are above a given threshold ($\omega = 1/1000$) relative to the least squares model weight. A set of 1000 models from each model class is then sampled uniformly over the intervals.

Even when one follows the process above, many models will have very small weight. To reduce the computational burden, the models are sorted by weight and the L worst models which do not have a cumulative weight sum above 0.001 are filtered out, and the weights are re-normalized. This "filtering" reduces the computational burden by a significant amount, but will not have a large effect on the end result.

4.1.2 Choice of likelihood function and information criteria

In order to find a weight for each model, a choice of information criteria and likelihood function must be made. Different combinations of model weights over time for AIC/BIC and flat-top/standard Gaussian likelihood are plotted in figure 4.3. The weights are plotted here as the sum of all model weights within each model class. Black-Scholes is not included since the total weight is negligible for all tested days and likelihoods.

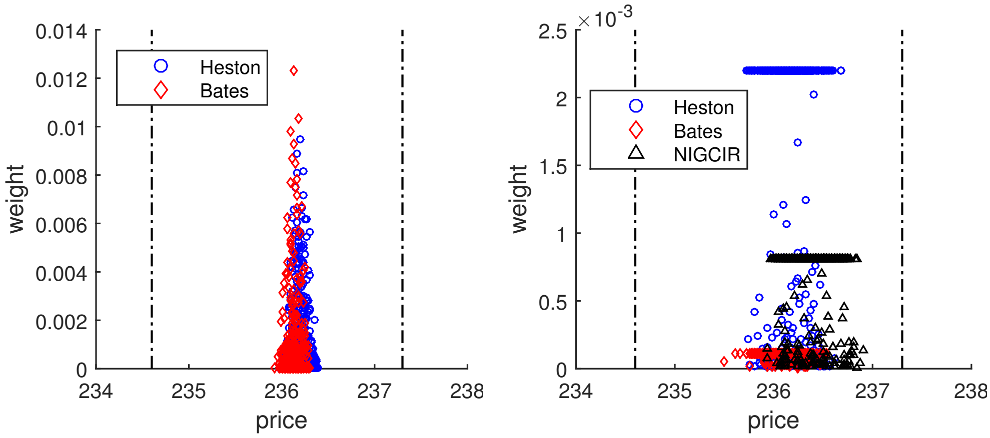


(a) AIC Gaussian (b) BIC Gaussian

(c) AIC flat-top Gaussian (d) BIC flat-top Gaussian

Figure 4.3: Model weights are shown over time for different combinations of information criteria and likelihood functions.

One can see that there is a large difference in the choice of method and likelihood. Also, the weights do fluctuate quite a bit over time. The price distribution of a vanilla call option is calculated using both the standard Gaussian and flat-top Gaussian likelihoods combined with AIC. The option is priced on August 8, 2014, has moneyness $K/S = 0.89$ and time to maturity $T = 0.397$. The price given by each model in the model set is plotted against its weight in figure 4.4.



(a) Gaussian

(b) Flat-top Gaussian

Figure 4.4: Model price is shown plotted against model weight for different choices of likelihood. The dotted lines display the bid-ask spread since the price is known for this option.

Since the weights are more stable over time and the price distribution is the widest for the AIC flat-top Gaussian methodology, it is selected for further analysis. Because the price distribution is the widest for this choice, it will also be a conservative method when it comes to calculating the model risk on August 8^{th}. Under this choice of information criteria and likelihood, the parameter intervals of models included in the set are presented in table 4.1 for August 8^{th} and similar results for August 22^{nd} in appendix B.2.

4.1.3 Price distribution and quantile

Once the models are sampled over the parameter intervals and the weights have been calculated for each model, the prices of out-of-sample options can be studied using the same model set. In particular, the prices of a vanilla call option, a digital call option, an up-and-out barrier option and an arithmetical Asian call will be calculated. For each option, three plots are presented. In the first plot, model weights are plotted against model price and each dot in this figure represents a model instance that gives a price for the option. The second plot shows a weighted histogram of the prices, with 30 bins between the minimum and maximum price. The third plot displays a kernel smoothing function estimate using the MATLAB `ksdensity` function. Note that all models with negligible weights have been excluded for better visualization and lower computational burden.

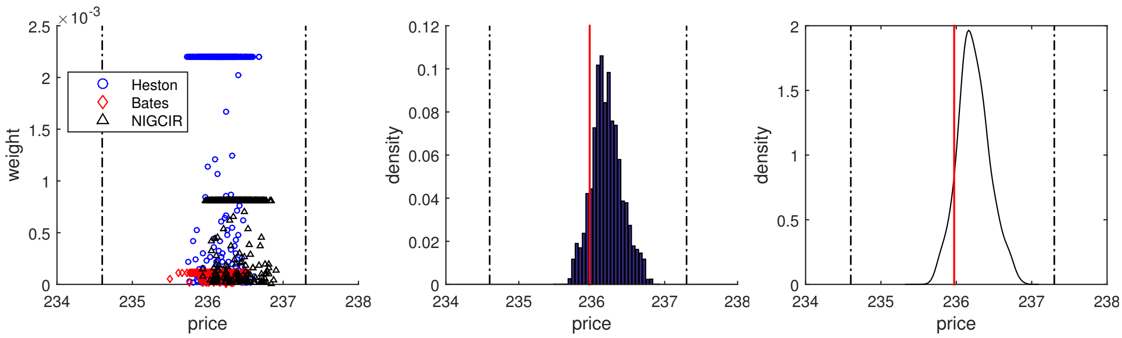
Vanilla call option

A vanilla call option with moneyness $K/S = 0.89$ and time to maturity $T = 0.397$ years is priced with each model for August 8^{th}, which can be seen in figure 4.5. The relative risk measure for this option is $\mu = 0.11\%$.

Black-Scholes		Heston	
Parameter	Interval	Parameter	Interval
σ	$[0.1322, 0.2012]$	V_0	$[0.0201, 0.0211]$
		κ	$[1.7800, 1.8487]$
		ξ	$[0.0516, 0.0526]$
		σ_V	$[0.5835, 0.6074]$
		ρ	$[-0.8153, -0.7836]$

NIGCIR		Bates	
Parameter	Interval	Parameter	Interval
κ	$[4.7305, 4.9252]$	V_0	$[0.0184, 0.0195]$
ξ	$[0.7503, 0.7744]$	κ	$[1.7984, 1.8745]$
σ_V	$[4.6340, 4.8092]$	ξ	$[0.0503, 0.0515]$
δ	$[0.9648, 0.9800]$	σ_V	$[0.5786, 0.6042]$
α	$[59.7850, 60.6579]$	ρ	$[-0.8747, -0.8400]$
β	$[-14.1998, -13.9118]$	λ_j	$[0.8646, 1.1845]$
		μ_j	$[-0.0125, 0.0103]$
		σ_j	$[0.0345, 0.0407]$

Table 4.1: Parameter intervals for Black-Scholes, Heston, NIGCIR and Bates on August 8^{th}.

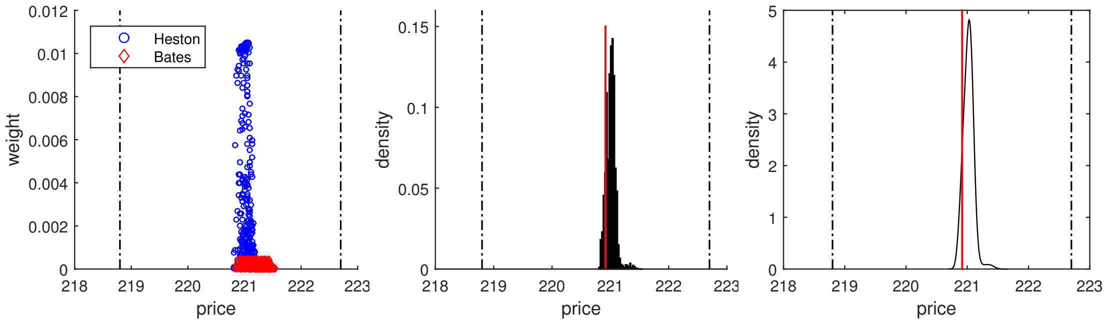


(a) Model weight vs price (b) Weighted histogram (c) Kernel density

Figure 4.5: Distribution of prices for a vanilla call option on August 8^{th}. The red line is the 10% quantile. The dotted lines display the bid-ask spread since the price is known for this option.

The reason for the horizontal lines of weights in figure 4.5a is that all models that are within the bid-ask spread of all calibration options will get the same weight due to the flat-top Gaussian likelihood. As can be seen by the high number of points generating the horizontal lines, there are many models that all calibrate perfectly. Note also that the horizontal line for Heston is on a higher level of weight than the other models since it has fewer parameters and therefore gets less punishment for complexity by AIC.

In order to display the differences between pricing with models generated by the data from August 8^{th} (when several model classes have significant weight) and the models generated by the data from August 22^{nd} (when only one model class has weight), the corresponding result for August 22^{nd} has been included in figure 4.6. The option priced here is a vanilla call option with moneyness $K/S = 0.90$ and time to maturity $T = 0.405$ years. The option is therefore comparable to the corresponding option on August 8^{th}. The relative risk in this case is calculated to $\mu = 0.047\%$, which is lower than on August 8th. Note that on August 22nd perfect calibration is not achieved in any of the model classes, meaning that no horizontal lines are present in figure 4.6a.

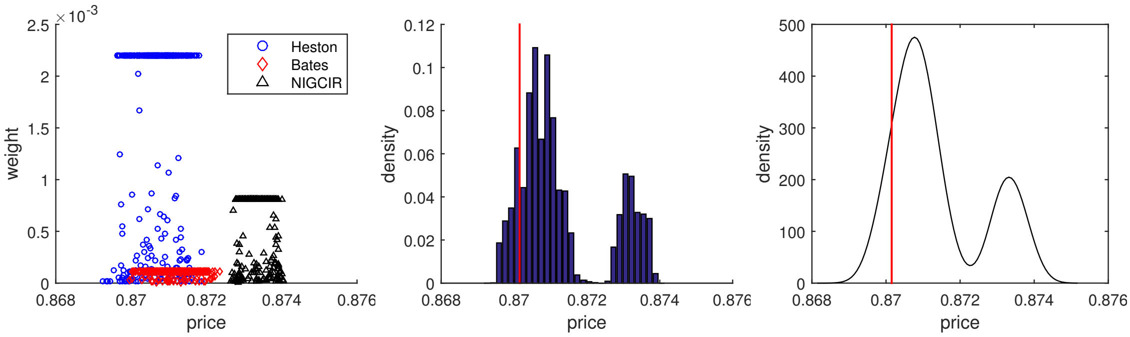


(a) Model weight vs price (b) Weighted histogram (c) Kernel density

Figure 4.6: Distribution of prices for a vanilla call option on August 22nd. Red line is the 10% quantile. The dotted lines display the bid-ask spread since the price is known for this option.

Digital call option

A digital call option with moneyness $K/S = 0.89$ and time to maturity $T = 0.397$ years is priced with each model from August 8th, giving a relative risk measure of $\mu = 0.14\%$, somewhat higher than for a vanilla call with the same strike and time to maturity. The price distributions can be seen in figure 4.7.



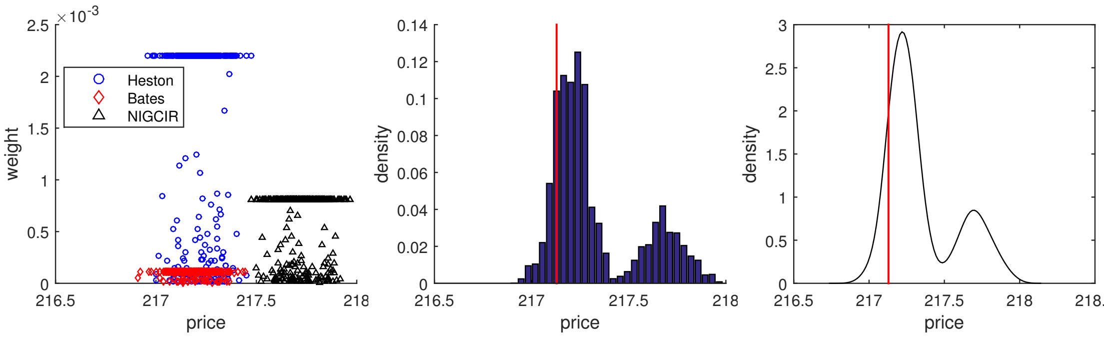
(a) Model weight vs price (b) Weighted histogram (c) Kernel density

Figure 4.7: Distribution of prices for a digital call option on August 8th. Red line is the 10% quantile.

The price distribution for the digital option consists of two modes, where the left mode is generated by Heston and Bates, and the right mode is generated by NIGCIR. This shows that two models can give approximately the same prices for vanilla options as seen in figure 4.5, but are still able give differing prices for other types of contracts, such as digital options.

Asian call option

An Asian call option with moneyness $K/S = 0.89$ and time to maturity $T = 0.397$ years is priced with each model from August 8th, giving a relative risk measure of $\mu = 0.10\%$. The price distributions can be seen in figure 4.8.

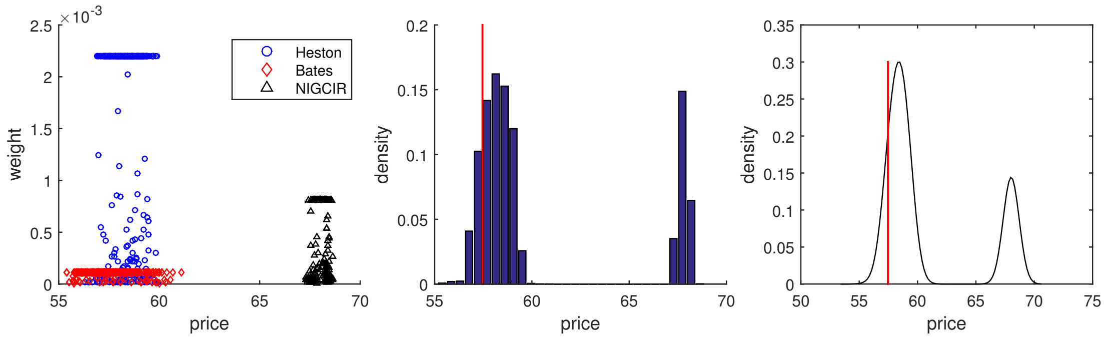


(a) Model weight vs price (b) Weighted histogram (c) Kernel density

Figure 4.8: Distribution of prices for an Asian call option on August 8$^{\text{th}}$. The red line is the 10% quantile.

Up-and-out barrier option

An up-and-out barrier option with moneyness $K/S = 0.89$, barrier level $B/S = 1.05$ and time to maturity $T = 0.397$ years is priced with each model from August 8$^{\text{th}}$, giving a a relative risk measure of $\mu = 5.4\%$.



(a) Model weight vs price (b) Weighted histogram (c) Kernel density

Figure 4.9: Distribution of prices for an up-and-out barrier option on August 8$^{\text{th}}$. The red line is the 10% quantile.

The difference in prices between the model classes is even greater here than for the Asian and digital call options, which gives rise to a high relative risk measure. This clearly demonstrates that multiple models which calibrate well to vanilla options can give very different prices to exotic derivatives.

4.1.4 Study of risk measure

The relative and absolute risk measures μ and $\hat{\mu}$ from section 3.6 can be used to study how the model risk depends on which option that's being priced. The study gives an indication of which changes in properties of options are more sensitive to model risk.

Varying strike level for vanilla options

In order to display the difference between vanilla options, the risk measures are calculated for different values of moneyness by varying the strike level, keeping the time to maturity fixed at $T = 0.397$. The results can be seen in figure 4.10. One can see that for options far in the money, the relative risk measure is low and for options far out of the money the relative risk increases. For low moneyness the payoff is approximately linear and thus many different models will agree on the price, while for a high moneyness the payoff will depend more on the tail of the distribution, which will make the choice of model more sensitive. By studying the absolute risk measure instead, it is clear that there is a quick drop for moneyness above $K/S = 1.1$. This is due to very low prices of options far out of the money, causing a low risk in absolute terms, even though it's high in relative terms.

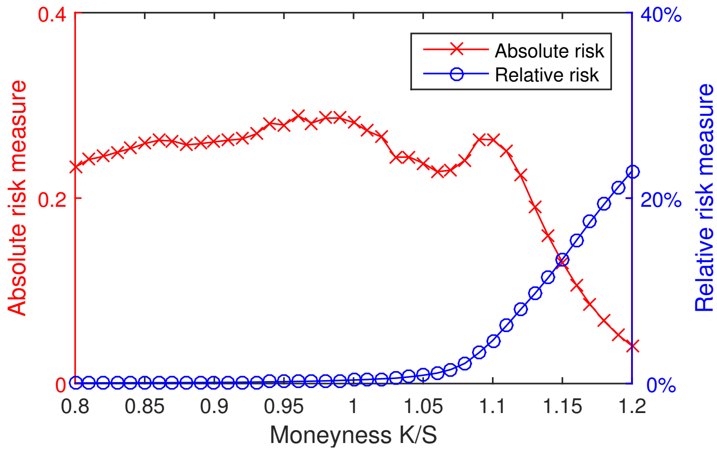


Figure 4.10: The model risk measures μ and $\tilde{\mu}$ are shown as a function of the moneyness for a vanilla call option.

Varying properties of barrier options

The risk measures are also studied for a set of up-and-out barrier options with varying barrier and strike levels. In figure 4.11, the relative risk measure is displayed as a 3D plot for time to maturity fixed at $T = 0.397$. The same data is also presented as a 2D line plot for in the money and at the money strike levels in figure 4.13 for better readability. The absolute risk measure $\tilde{\mu}$ is shown in figure 4.12. One can see that the behaviour of the absolute risk measure is quite different from the relative risk measure.

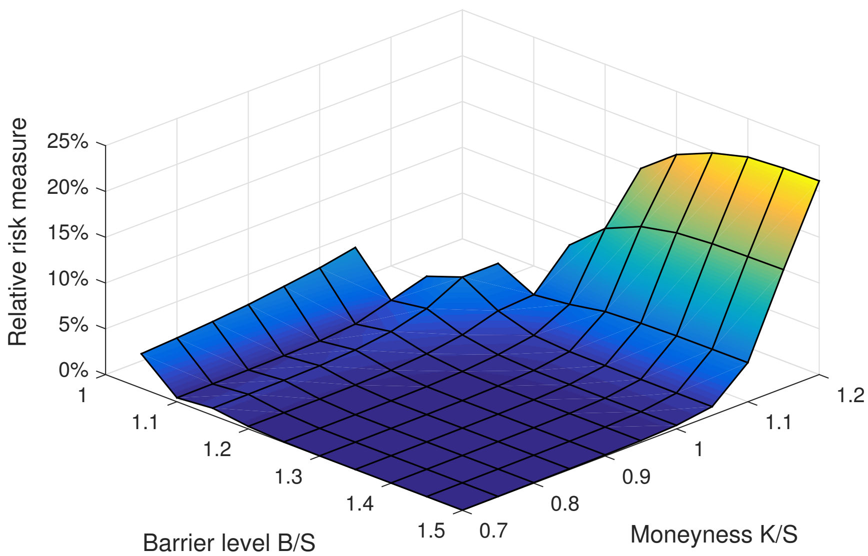


Figure 4.11: The relative model risk measure μ is shown as a function of barrier level and moneyness for an up-and-out option.

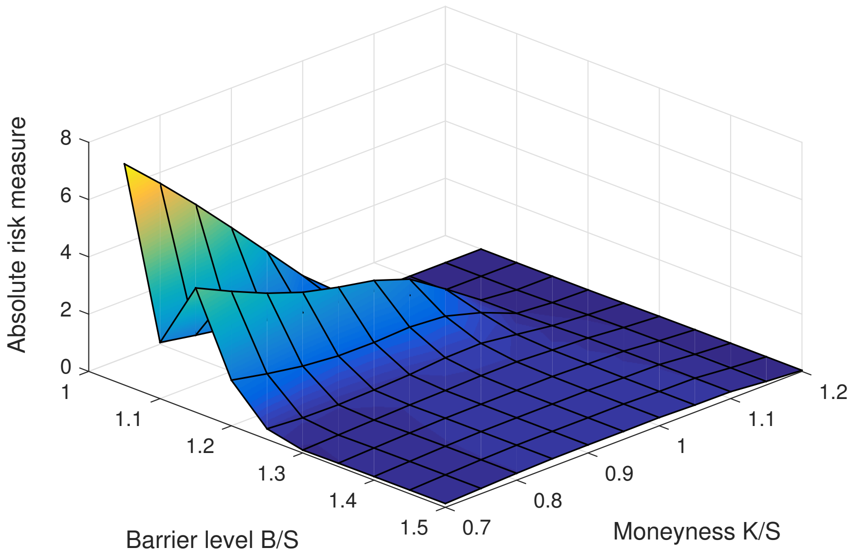


Figure 4.12: The absolute model risk measure $\tilde{\mu}$ is shown as a function of barrier level and moneyness for an up-and-out option.

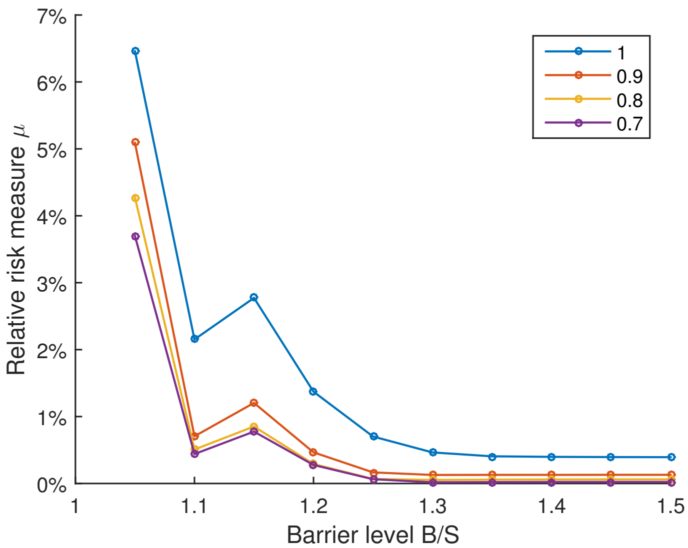
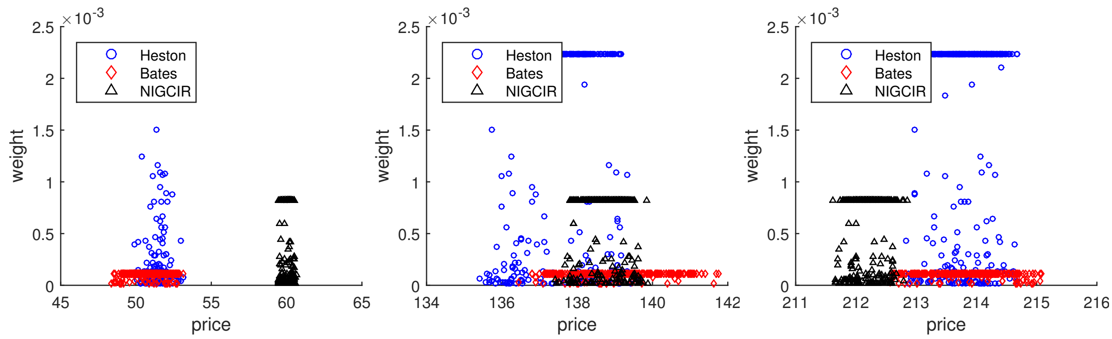


Figure 4.13: Shows the relative model risk measure as a function of barrier level for barrier options with 4 different values of moneyness K/S. Each line represents one value of moneyness.

The results indicate that lower barrier levels in general give higher risk measure, which is reasonable since a lower barrier will more often lead to the option being knocked out. The tail of the model return distributions are more important for closer barriers, where two different models might show very different tail characteristics and then also more different prices, leading to a higher risk measure.

For high barrier levels, the barrier option starts behaving similar to a vanilla option since a knock-out is unlikely. This gives rise to a much lower risk measure, which is in line with the risk for the corresponding vanilla call options. Interestingly, there is a small bump in the risk measures around barrier level $B/S = 1.1$. A close look at how the distribution changes around this bump is shown in figure 4.14. One can see that the NIGCIR models crosses over the Heston and Bates models from higher prices for a low barrier level to lower prices for a high barrier level. The risk measures are lower during the point of crossover since the model classes temporarily agree on the price, causing a tighter distribution.



(a) $B/S = 1.05$ (b) $B/S = 1.09$ (c) $B/S = 1.19$

Figure 4.14: Distribution for the up-and-out option for $B/S = 1.05, 1.09, 1.19$.

Temporal development

It is desirable that the risk measure is reasonably smooth over time, assuming the market characteristics don't change rapidly. In order to investigate this property, the model weighing procedure is repeated for each day in the dataset and three option types are priced with each model. These types consist of one up-and-out barrier option with relative barrier level $B/S = 1.10$, one Asian call option and one digital call option, all with time to maturity $T = 0.397$ and moneyness $K/S = 0.95$. The relative risk measure μ for each day and option type is shown in figure 4.15. It can be seen that the risk is highest for the up-and-out and Asian options on August 8$^{\text{th}}$ and this is also the day when Heston, Bates and NIGCIR all attain perfect calibration.

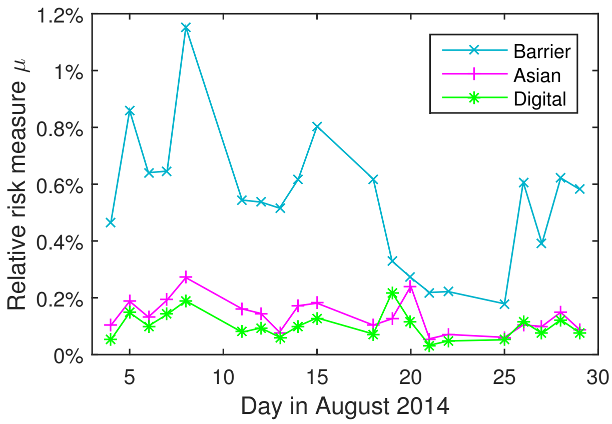


Figure 4.15: The relative risk measure is shown as a function of calibration date for three option types. The irregularities along the time axis are due to the fact that the markets are closed on weekends.

4.1.5 Multiple local minima

In order to investigate how common multiple solutions for the least squares minimization problem are, the multistart principle as described in section 2.2.2 is tested for the S&P 500 dataset. For each model, the starting points for the minimization are sampled uniformly

Models	Occurrences of local minima	
	1	2
Heston	20	0
Bates	20	0
NIGCIR	14	6

Table 4.2: Occurrences of local minima.

across the large intervals as seen in appendix B.4. The number of starting points is set to 100 for each model and the estimation is run for each day in the dataset. When the calibration is finished, the best fit of all parallel minimizations is kept and selected as the base model. The other local minima are then iterated and kept if the following conditions are fulfilled:

- The model weight using the flat-top AIC weighing should be at least a thousand of the model weight for the best model.
- At least one parameter in the model should differ 5% or more from the corresponding parameters of the other kept models.

These conditions guarantee that the kept models are assigned non-negligible weight, but also that there are no duplicates. For the Bates model, all solutions with very low jump frequency ($\lambda_j < 0.01$) are removed as well. The reason for this is that the Bates model coincides with the Heston model for jump frequency zero, and it's not interesting to see how many local minima that can be found with varying jump parameters μ_j and σ_j when jumps never occur anyway. For the S&P 500 dataset in August 2014, the frequency of multiple minima can be seen in table 4.2.

It is clear that more than one local minima satisfying the conditions is never found for Heston and Bates. For NIGCIR, there are some occurrences of multiple minima, in particular 30% of the tested days. There are no multiple local minima for August 8$^{\text{th}}$ and August 22$^{\text{nd}}$ tested in the previous sections.

One example of a day where the procedure finds 2 distinct local minima for NIGCIR is on the 29th of August. The two parameter vectors, $\theta_1 = [4.28, 1.11, 5.41, 1.90, 170, -13.6]$ and $\theta_2 = [4.27, 1.11, 5.40, 1.75, 156, -13.6]$, both price within bid-ask of 90% of the benchmark options.

4.2 Simulation study

In this section the market will be simulated under the Bates model, and the model classes used will be Black-Scholes, Heston and NIGCIR. The simulated options will have a spot moneyness K/S between 0.8 and 1.2 (in increments of 0.02) and time to maturity 1, 6, 12 and 24 months. The risk-free rate r is set to 1%, and $S_0 = 100$. In total there will be 84 options. The model risk measure will be studied when changing parameters in the Bates simulation. First, varying jump intensity λ_j will be studied, and then changes in the mean reversion rate κ will be studied as well.

4.2.1 Varying jump intensity in simulation

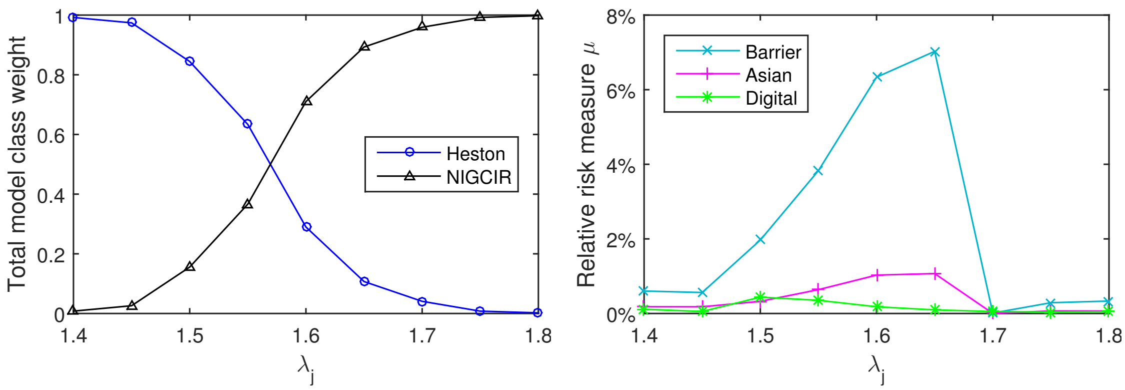
The parameters used to simulate the Bates model are found in table 4.3. The jump intensity parameter is varied between 1.4 and 2 to see how the model risk measure is affected. For each model class, 200 models are included. Since there are no simulated bid-ask spreads, calibration is done with ordinary least squares, and the likelihood used is the standard Gaussian in section 3.5.1. The least squares estimates are found in appendix B.3.1.

Parameter	Value
V_0	0.006
κ	1.6
ξ	0.05
σ	0.6
ρ	-0.8
λ_j	1.4 : 0.05 : 1.8
μ_j	-0.07
σ_j	0.04

Table 4.3: Parameters used for Bates simulation.

The total weight distributed between Heston and NIGCIR varies for different λ_j as seen in figure 4.16a. An increase in jump frequency will lead to a higher weight of NIGCIR, which is reasonable since the NIGCIR model includes jumps, while Heston does not. Equal probability between the Heston models and NIGCIR models will in this case be achieved at roughly $\lambda_j \approx 1.58$.

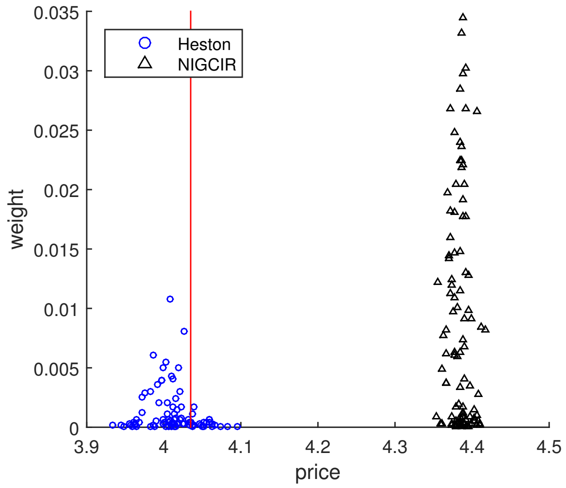
To study the effect on the risk measure μ when λ_j changes, one up-and-out barrier option, one digital call and one Asian call are priced. The barrier B is chosen as 110 and the strike as 95 for both the up-and-out and Asian option. The results are shown in figure 4.16b.



(a) Weights (b) Risk measures

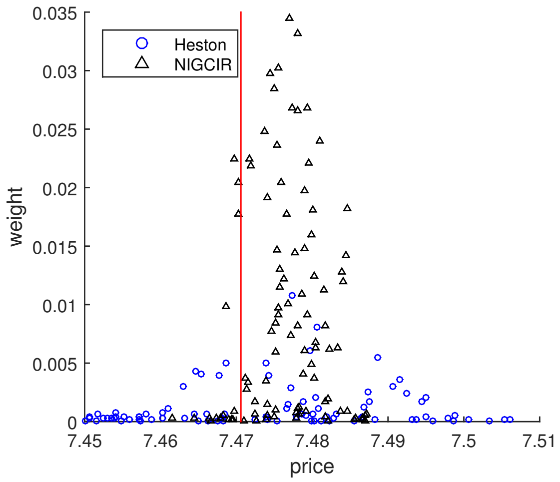
Figure 4.16: This figure shows how the summed weight for each model class varies for different values of the jump intensity λ_j in the simulated market, as well as how the risk measure changes.

In figure 4.16b, it is clear that the model risk increases as λ_j increases until $\lambda_j = 1.70$, where there is a quick decrease. This is linked to the model weights, as Heston goes below 10% in weight for the same value of λ_j, causing the 10% quantile to be closer to the weighted average price and thus leading to a lower value of the risk measure. The phenomenon can be seen in figure 4.17. The opposite case would be that of a vanilla call which can be seen in figure 4.18 and in this case the distribution is not multimodal. Moneyness and time to maturity for this option are the same as for the barrier option.



(a) $\lambda_j = 1.65$ (b) $\lambda_j = 1.70$

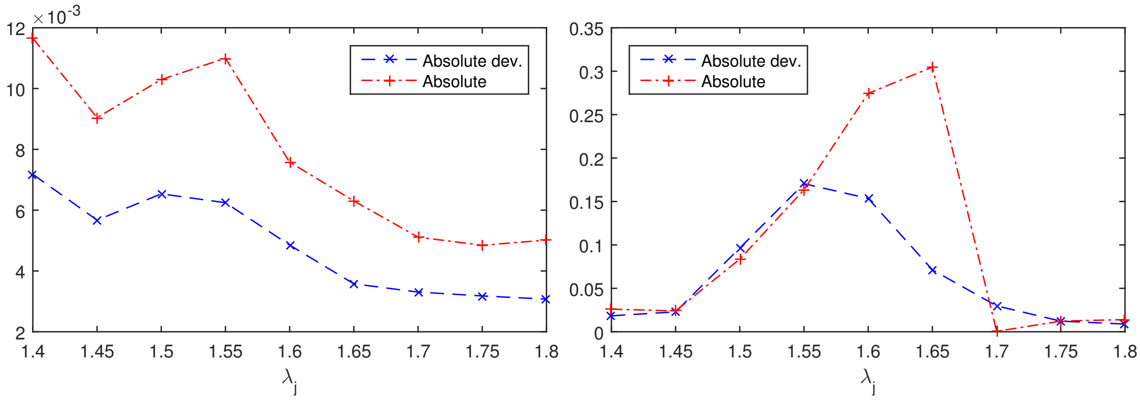
Figure 4.17: The figure shows each model weight plotted against its corresponding barrier option price before and after Heston gets less than 10% total weight. The red line is the 10% percentile.



(a) $\lambda_j = 1.65$ (b) $\lambda_j = 1.70$

Figure 4.18: The figure shows each model weight plotted against its corresponding vanilla call option price.

One could also consider the risk measure of absolute deviation, as described in definition 3.7. A comparison between the absolute risk measure and the risk measure of absolute deviation can be seen in figure 4.19. The options under consideration are one vanilla call and one up-and-out barrier. As seen in figure 4.17 the price distribution is multimodal for the barrier option, but it will not be multimodal for the vanilla call as seen in figure 4.18. In the case of no multimodality the two risk measures $\tilde{\mu}$ and $\bar{\mu}$ both behave in the same manner, aside from some amplitude scaling factor. On the other hand, in the case of multimodality as in the case of the up-and-out barrier option, $\bar{\mu}$ will be much more stable than $\tilde{\mu}$. Looking back at figure 4.16a one saw that the weight between the Heston and NIGCIR model classes was most equal when the market was simulated under a Bates model with $\lambda_j = 1.55$. Under this market, the model risk measured with $\bar{\mu}$ is also the highest.



(a) Vanilla (b) Barrier

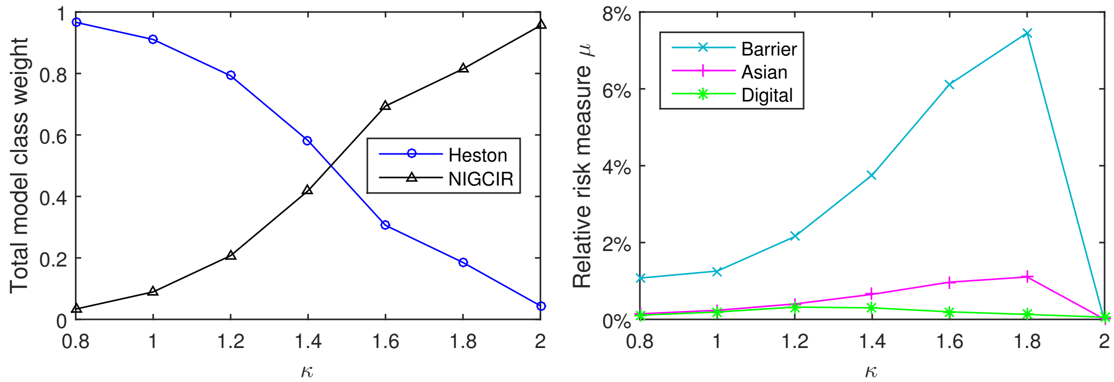
Figure 4.19: The absolute risk measure $\tilde{\mu}$ (red dashed and dotted line) and measure of absolute deviation $\bar{\mu}$ (dashed blue line). The left figure shows the risk measures for vanilla call and the right figure risk measures for a barrier option. Again, the results are for different λ_j in the Bates simulation. The spot moneyness for both options is $K/S = 0.95$ and time to maturity equals $T = 100/250$. The barrier level is again chosen as $B/S = 1.1$.

4.2.2 Varying mean reversion rate in simulation

In a similar way to the previous section, a market based on the Bates model is simulated with varying mean reversion κ. The parameters used for the simulation are given in table C.1.

Parameter	Value
V_0	0.006
κ	0.8 : 0.2 : 2
ξ	0.05
σ_V	0.6
ρ	-0.8
λ_j	1.6
μ_j	-0.07
σ_j	0.04

Table 4.4: Parameters used for Bates simulation.



(a) Weights (b) Risk measures

Figure 4.20: This figure shows how the summed weight for each model class varies for different values of the mean reversion rate κ in the simulated market, as well as how the risk measure changes.

It can be seen that the model risk increases with κ until there once again is a quick drop at the point where the model weights are focused to only one of the model classes.

Unidentifiable parameters

The mean reversion rate κ in the volatility process is often difficult to estimate and as discussed by Detering and Packham in [19] it cannot be uniquely identified by options with the same maturity. To demonstrate this, consider again the Bates simulation with varying κ, but now the market only consists of options with time to maturity 1 year. As seen now in figure 4.21, the behaviour of the model weights is strange and the method is no longer reliable. When only using one time to maturity in the analysis the results will be widely different compared to when using more than one time to maturity; compare figure 4.20a and 4.21. The least squares estimate of κ in the Heston model can be seen in figure 4.22.

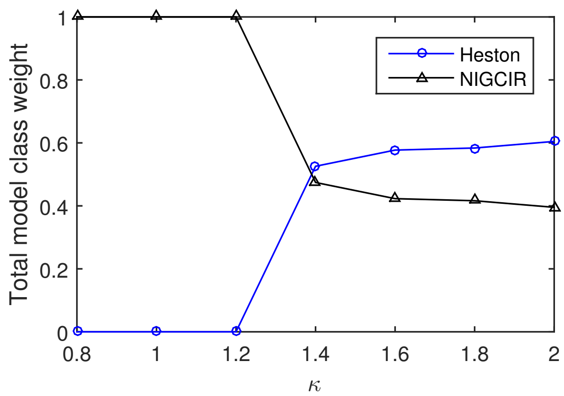


Figure 4.21: This figure shows how the summed weight for each model class varies for different values of the mean reversion rate κ, when only options with time to maturity 1 year are considered.

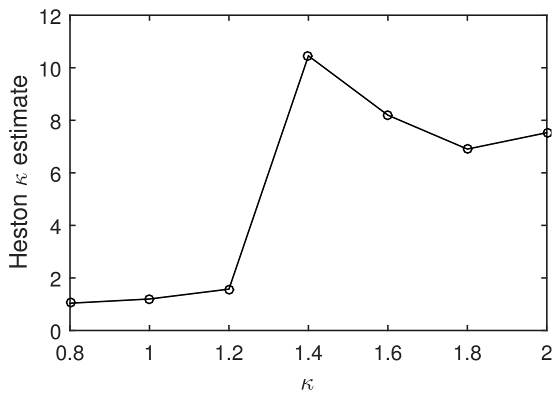


Figure 4.22: The least squares estimate of κ for different κ used in the Bates simulation.

When reliable prices of exotic options are available on the market, the mean reversion rate κ might still be identifiable, since exotic options such as barrier options and Asian options are path dependent. The difference between model and market prices for exotic options could then be incorporated into the model likelihood as error terms, together with the ordinary errors to vanilla options.

4.3 Calibration risk in the Heston model

It was noted that the Heston, Bates and NIGCIR models all achieved perfect calibration on August 8^{th}, on the subset of options described in the beginning of the chapter. This raises two questions:

- How different are the models that can achieve perfect calibration?
- How wide is the range of prices generated by models that can calibrate perfectly to benchmark instruments?

Here, specifically, the Heston model class will be studied, using the MCMC Metropolis-Hastings algorithm.

4.3.1 Metropolis-Hastings

The Metropolis-Hastings (MH) algorithm was developed by Metropolis et al. [35] and Hastings [24], in 1953 and 1970. Assuming a *symmetric jump proposal*, the algorithm is given by algorithm 3 (with a symmetric proposal the algorithm is also simply called the Metropolis algorithm, which is a special case of the Metropolis-Hastings algorithm).

Algorithm 3 Metropolis-Hastings with symmetric proposal

Require: Benchmark prices C^*, prior $p_0(\theta \mid C^*)$, likelihood function $p(C^* \mid \theta)$, symmetric jump proposal $J(\theta \mid \theta_{k-1})$.

Draw $\theta_0 \sim p_0(\theta \mid C^*)$.

for $k = 1 : n$ **do**

Draw $\theta^{\#} \sim J(\theta \mid \theta_{k-1})$.

Let $\theta_k = \begin{cases} \theta_k^{\#} & \text{with probability } \min\left\{1, \frac{p(\theta^{\#}|C^*)}{p(\theta_{k-1}|C^*)}\right\}, \\ \theta_{k-1} & \text{otherwise.} \end{cases}$

end for

return the sequence $\theta_1, \dots, \theta_n$.

The chain $\{\theta_k\}_{k\geq 0}$ has $p(\theta \mid C^*)$ as a stationary distribution.

Since the goal is to find only parameters in the Heston model that calibrate perfectly, i.e. parameters that belong to the set

$$\{\theta\colon C(\theta) \in [C^{bid}, C^{ask}]\} \subset \mathbb{R}^4_+ \times [-1, 1], \tag{4.1}$$

the likelihood is assumed to be uniform over the bid-ask spreads

$$p(C^* \mid \theta) \propto \prod_{j=1}^{N} \mathbb{1}\{C^{model}(K_j, T_j; \theta) \in [C^{bid}(K_j, T_j), C^{ask}(K_j, T_j)]\}.$$

This is a special case of the flat-top distribution with $\sigma = 0$. Clearly, the proposal can be accepted only if all model prices are within bid-ask spreads, and the chain is therefore restricted to cover only models with perfect fit. If one would be interested in a full Bayesian treatment, the σ parameter would instead enter as a nuisance parameter.[1] The prior is set as a multivariate Gaussian around the previous day's (August 7^{th}) weighted least squares estimate $\theta^{(t-1)}$,

$$p(\theta) \propto \exp\left\{-\frac{1}{2}(\theta - \theta^{(t-1)})^T \Sigma^{-1} (\theta - \theta^{(t-1)})\right\}.$$

The covariance matrix Σ, as well as the least squares estimate $\theta^{(t-1)}$ are found in appendix C.1. The posterior is then given by

$$\begin{aligned} p(\theta \mid C^*) \propto & \prod_{j=1}^{N} \mathbb{1}\{C^{model}(K_j, T_j; \theta) \in [C^{bid}(K_j, T_j), C^{ask}(K_j, T_j)]\} \\ & \cdot \exp\left\{-\frac{1}{2}(\theta - \theta^{(t-1)})^T \Sigma^{-1} (\theta - \theta^{(t-1)})\right\}. \end{aligned}$$

The jump distribution is chosen as

$$J(\theta \mid \theta_{k-1}) \sim N(\theta_{k-1}, \Sigma_J),$$

where the covariance Σ_J is found in appendix C.2, and chosen in such a way such that the acceptance rate of the proposal $\theta_k^{\#}$ is in the range of $20 - 30\%$. Initial starting values for the chains are chosen randomly around the current least squares estimate.

The algorithm is run with a chain length of $n = 2 \cdot 10^5$, with 4 parallel chains. The first 20% of the chains are thrown away as a burn-in period and thinning is applied by keeping every 10th sample. Each sample constitutes an instance of the Heston model and can thus be used for pricing an option. In order to allow for comparison with the previous results, the option to price is once again chosen as the vanilla call option with $K/S = 0.89$ and $T = 0.397$. This gives a histogram of prices as seen in figure 4.23. Despite the high chain lengths, it is difficult to be certain that the chain has converged. However, the minimum and maximum price in the histogram gives an idea of which prices that models that are not restricted to the hyperrectangle can reproduce.

[1] Actually, this was initially attempted, but it was difficult to obtain good mixing.

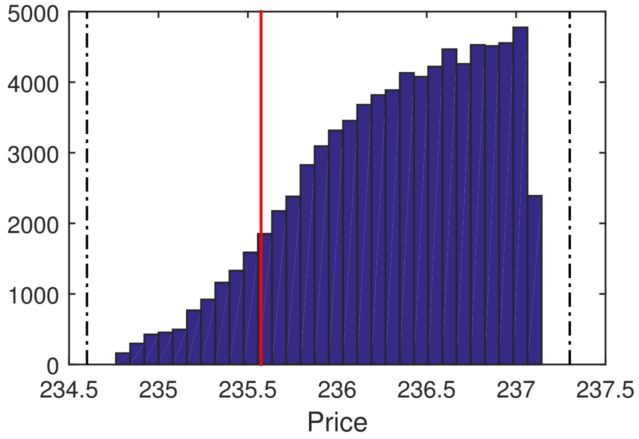


Figure 4.23: The posterior price distribution from the MH algorithm is shown. The dotted lines display the bid-ask spread since the price is known for this option.

Comparison with previous results indicates that the distribution of prices is wider with the Metropolis-Hastings algorithm. The setup has thus found parameter combinations outside the previously spanned hyperrectangle that was still able to give perfect fit within bid-ask spreads of all benchmark options. Since the hyperrectangle is in five dimensions, it is hard to visualize its extension in all dimensions at once. In figure 4.24, it is instead plotted as a projection for all pairwise combinations of dimensions.

The fact that the distribution of prices is wider is not surprising, since the MH algorithm will sample from the entire set $\{\theta\colon C(\theta) \in [C^{bid}, C^{ask}]\}$ (assuming the space of perfect calibration is a connected subset of the parameter space, otherwise the MH might not efficiently explore the whole space). The hyperrectangle's width in one dimension was chosen when only changing one parameter, keeping others fixed. There might exist complex interactions between the parameters that this method will not take into account (see figure 2.4).

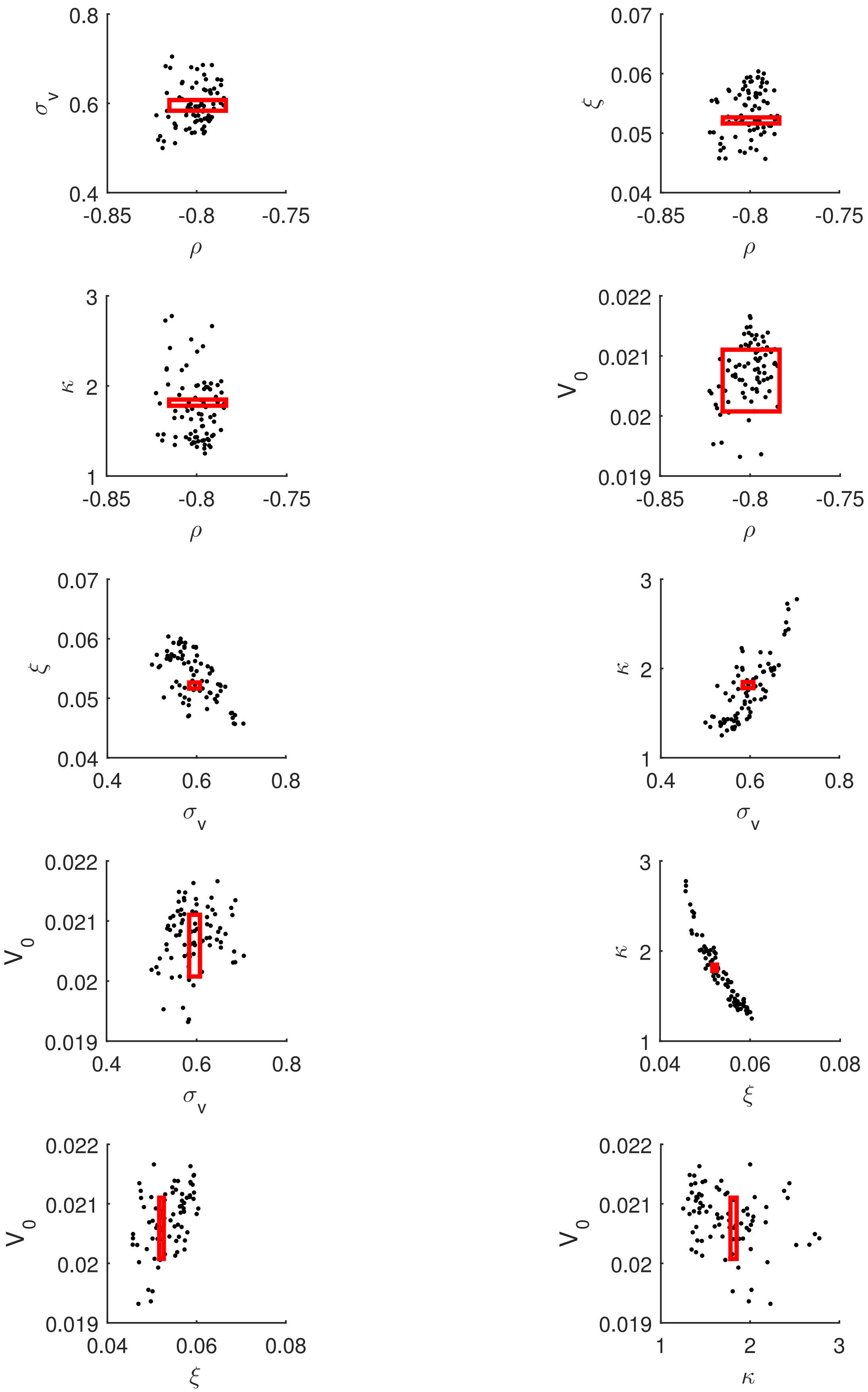


Figure 4.24: The figure shows the extension of the hyperrectangle in red lines and roughly 80 models from the sample as black dots. Note that each dot represents a model with perfect calibration.

CHAPTER 5

Discussion

5.1 Assumptions made in the study

The task of quantifying model risk is highly non-trivial. Approaches like the worst-case risk measure of Cont which do not rely on a distribution over a model and/or parameter set are easy to obtain. On the other hand, if one wants to obtain a *price distribution* of an exotic option, one suddenly needs to make assumptions on pricing errors and prior probabilities. This makes the task of quantifying model risk somewhat paradoxal; in order to quantify the risk, one needs to assume a model! It is therefore important to have a variety of alternative model validation approaches.

One assumption made to produce the results in chapter 4 was that the modeling errors are mutually independent. This might be a strong assumption and it could be reasonable to believe that there is some correlation of the modeling errors for options with similar strike and time to maturity. One could set up a correlation structure between similar strikes and time to maturities, but there would then be parameter choices regarding the correlation subject to the modeller. Another option would be to use the likelihood based on the option price surface integral, which was presented in section 3.5.4.

Furthermore, a flat-top Gaussian distribution was assumed for the pricing error. This choice was motivated by the property that two models, both calibrating perfectly within the bid-ask spread of all options, should have the same likelihood. Of course, a counter argument would be that a model that calibrates to mid prices on all options should be better than a model that calibrates on the edges of the bid-ask spread on all options. On a practical level, it could be seen in figure 4.3, where the model weights were calculated over time, that the flat-top likelihood gave slightly more stable weight distribution compared to the standard Gaussian likelihood. When using the flat-top Gaussian likelihood one should make sure that the data is reliable and that the bid-ask spreads are not large enough to introduce arbitrage opportunities.

5.2 The risk measures

The risk measures $\hat{\mu}$ and μ defined in section 3.6 were proven useful for comparing model risk between different options. The definition of these measures is motivated by the concept of model risk AVA. The relative risk measure was introduced to make options with differing prices comparable, but what must be kept in mind is that it will not work for contracts that have a value of zero. One such example is a swap contract at inception, where two parties have agreed to exchange future cash flows with the equal present values.

Another possible shortcoming with the risk measures is that the risk measure can change rapidly when the price distribution is multimodal, as illustrated in figure 4.17. The problem is present since the risk measures are based on a specific quantile, which will always be located within one of the modes of the distribution and will thus be able to jump when the quantile location moves from one mode to another. The model risk measure of absolute deviation $\bar{\mu}$ does not suffer from this shortcoming since it is not based on the quantile. However, the measure of absolute deviation does not have a direct connection to prudent valuation, as it is not calculated with respect to a certainty level.

When the relative risk measure was calculated for similar options over several days during August 2014 in section 4.1.4, it could be seen that the risk varied over time. This result is related to the varying distribution of weights among model classes, as models from different model classes tend to price exotics more differently than models from the same class, which is shown in figure 4.9. Note however that the risk measures did not vary as much as the distribution of weights among model classes over time. In particular, the risk measures are fairly robust to switches in weight between Heston and Bates because of their similarities in pricing for low jump frequencies in Bates. The risk measure was the greatest for most option types on the day when not only Heston and Bates, but also NIGCIR attained weight. This is reasonable since there is a greater uncertainty related to which model that accurately describes the market when a higher number of model classes perform well.

5.3 The choice of models to include

The choice of models that enter the model set $\mathcal{Q}$ should be based on the range of models used by market participants. One could include only the least squares estimate from each model class, or some other models using other calibration procedures. Alternatively, one could use the method investigated in this thesis, where a hyperrectangle is spanned around the least squares estimate and models are chosen uniformly from this hyperrectangle. This latter method is closest in spirit to what was presented in Detering and Packham's examples in [19]. Another possibility is to only consider models that calibrate within bid-ask spreads to the benchmark instruments as valid in $\mathcal{Q}$, but this might be a too conservative restriction and one might end up with the case that $\mathcal{Q}$ is empty if no models achieve perfect calibration.

The hyperrectangle was spanned for each model class such that the model weights for parameter vectors outside the hyperrectangle were supposed to be negligible and thus unable to affect the result if excluded. It must be noted here that the non-negligble parameter space might not be approximated well by a hyperrectangle for models with strong dependencies between parameters. For example, there might exist models where the likelihood of the model is almost constant along some contour line in the parameter space. It would then be more suitable to span a hyperellipsoid or even an arbitrary hyperspace of models. In particular, when the Metropolis-Hastings algorithm was used to find instances of the Heston model that all calibrated perfectly on August 8, it could be seen that the hyperrectangle did not cover all the samples. This means that the risk associated with complex relationships among the model parameters is not accounted for in the main results, and this risk would likely increase the width of the distribution further. To accomodate a greater inclusion of models it is tempting to decrease the weight threshold ω in algorithm 2, where the hyperrectangle is defined. However, a larger hyperrectangle would make sampling troublesome as most of the parameter space in the hyperrectangle would yield models with zero weight when the models with good fit are focused to a thin subspace.

5.4 Pros and cons of the methodology

Pros and cons of the methodology in this thesis are summarized below.

Pros:

- The methodology can be integrated with existing systems for pricing and calibration. In essence, the methodology only needs market prices for calibration and a set of models with corresponding pricing routines to assign weights to the models.
- It is easy to implement and transparent conceptually. Unlike MCMC there are very few tuning parameters needed to implement it.
- It is flexible enough to accommodate the main requirements of the regulatory technical standards, since it can account for model class uncertainty, parameter uncertainty and calibration risk.

Cons:

- It can take a long time to calculate the model risk for exotic derivatives since a pricing is needed for each model in the model set $\mathcal{Q}$.
- It is not clear how one could use AIC or BIC weights in a model with infinite or very large parameter space, such as the local volatility model. Another problem arises when the number of parameters increases when the data increases. Either one would have to assume a parametric form or develop a different methodology.
- The model weights are not very stable over time, although the risk measures themselves show a higher grade of stability.
- It's a non-trivial task to select which model classes and parameter sets to include. In particular, a hyperrectangle of parameter sets around the least squares estimate does not necessarily cover all models with good fit.
- The methodology is relatively sensitive to the choice of likelihood function.

5.5 Future research

Markov Chain Monte Carlo methods where discussed as a possible way of determining model risk, but it is not trivial to combine model class uncertainty with parameter uncertainty. This area could be researched further.

In the methodology presented in this thesis, yesterday's information does not effect the model weights today. One way to account for previous information is to cast the problem of model risk quantification as a filtering problem. Assuming only one model class, the filtering problem would look like

$$\theta^{(t)} = \theta^{(t-1)} + e_t,$$

where $\theta^{(t)}$ are the unknown parameters at time t and e_t is a white noise process. This would then constitute the state equation. The measurement equation would look like

$$C_t^* = C^{model}(\theta^{(t)}) + \varepsilon_t.$$

Since $C^{model}(\cdot)$ is a non-linear function, some non-linear filter, such as the particle filter, could be used. The above state space can also easily be extended to take the underlying asset moves into account. Sequential calibration of the sort above is covered in [34], but it would be interesting to apply it more in the context of model risk.

One possible improvement to the choice of models to include could be to consider models from a hyperellipsoid based on an approximation of the covariance matrix of estimated parameters. As mentioned in [11], maximum likelihood estimation has asymptotic normality

$$\sqrt{n}(\hat{\theta} - \theta^*) \xrightarrow{d} N(0, \Sigma)$$

for some covariance matrix Σ, where n is the number of observations, θ^* is the true parameter, $\hat{\theta}$ is the ML estimate and $\xrightarrow{d}$ denotes convergence in distribution. The covariance matrix can for reasonably large n be approximated by

$$\Sigma \approx H(\hat{\theta})^{-1}/n,$$

where $H(\hat{\theta})^{-1}$ is the inverse Hessian of the negative log likelihood function, evaluated in the maximum likelihood estimate. If the modeling errors of option prices were assumed to be Gaussian and mutually independent, the least squares estimate of the model parameters would coincide with the maximum likelihood estimate. The least squares estimate could thus be used as the centre for a hyperellipsoid spanned to some confidence level of the multivariate normal distribution. If the covariance matrix Σ would be non-diagonal, it means that dependencies between model parameters in the loss function would be somewhat accounted for. Another approach seeking to solve the same problem could also be to find transformations in the parameter space such that parameter dependencies are minimized, making the hyperrectangle more suitable.

There are also possible extensions to the method when dealing with parameters that are difficult to estimate, such as mean reversion shown in section 4.2.2. One may have a best-guess of such parameters based on some other estimation techniques, which could be used instead of the ordinary least squares estimation. An interval could then be spanned manually around this best-guess based on the uncertainty of the guess, allowing for sampling of models with the ordinary method. If the parameter would be unidentifiable from vanilla prices such that the prices are truly independent of the parameter, then all models would get equal weights when varying only the unidentifiable parameter. This would lead to a high model risk for exotic options that are dependent of the unidentifiable parameter. However, if there are reliable exotic option prices available on the market, they could be incorporated as benchmark instruments when calculating the weights. These option prices might contain information about the unidentifiable parameter, and would therefore decrease the parameter uncertainty. Taking the moves of underlying asset into account could perhaps also improve the inference.

CHAPTER 6

Conclusion

The purpose of this thesis was to find a statistical method for quantifying model risk. In order to fulfill the requirements of the regulatory technical standards on prudent valuation, the keystone of the methodology was to find a distribution of prices for traded options. The evaluated methodology consisted of setting up a set of parametric models such as Black-Scholes, Heston, Bates and NIGCIR with various parameter vectors. These parameter vectors were chosen uniformly from a hyperrectangle around the respective least squares estimate.

The likelihood for each parameter combination was calculated by assuming that the errors between model prices and market prices of options followed a flat-top Gaussian distribution. Using the calculated likelihoods, a weight was assigned to each model using the Akaike Information Criterion. The models could then be used to price various options, which combined with the model weights gave rise to price distributions of these options. An interesting result was that although multiple models gave approximately the same prices for vanilla options, they were still able to give differing prices for exotic options.

Several risk measures were introduced to quantify the model risk with respect to the price distribution, with the regulatory technical standards in mind. Both empirical and simulated studies showed that up-and-out barrier options were more risky from a model risk perspective than arithmetic Asian call options and digital call options, at the studied levels of moneyness and barrier levels. It was also shown that the model risk measure in relative terms of option price increased quickly with moneyness K/S for call options far out of the money. In absolute terms the risk measure was lower for options out of the money compared to options in the money. The distribution of model weights over the different model classes seemed not to be very stable over time, however the risk measures themselves did not vary as heavily.

Although the calibration problem did not have multiple local minima for the Heston model, the study in the Heston model with the Metropolis-Hastings algorithm showed that the spanned hyperrectangle around the least squares estimate was not sufficient to include all models with perfect calibration. To capture the risk associated with complex dependencies of model parameters, some other method than the hyperrectangle might have to be used, if one would want to guarantee that all models with perfect calibration are included in the model set.

It could be seen that the proportion of total weight assigned to the NIGCIR model increased as the jump intensity in the simulation was increased. This might be reasonable since the NIGCIR model includes jumps, while Heston and Black-Scholes do not. The Black-Scholes models never attained any weight, neither on real data or in the simulation study.

In summary, the evaluated methodology was able to take into account both model class uncertainty and parameter uncertainty in accordance with the guidelines from the EBA. However, it was found to be sensitive to distributional assumptions on the pricing errors.

Appendices

APPENDIX A

Cont's axioms

Define a set $\mathcal{C}$ containing all contingent claims with well defined price

$$\mathcal{C} = \left\{ \mathcal{X} \in \mathcal{F}_T \colon \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}\left(|\mathcal{X}|\right) < \infty \right\}.$$

Let a self-financing trading strategy be denoted by $(\phi)_{t\in[0,T]}$ and $\mathcal{S}$ the set of such trading strategies with the requirement that for any $\phi \in \mathcal{S}$ the integral $\int_0^T \phi \, \mathrm{d}S$ is well defined.

Cont introduces four axioms that a model risk measure should satisfy:

(i) For liquid instruments, model uncertainty reduces to the uncertainty on market value:

$$\forall i, \quad \mu(H_i) \leq |C_i^{ask} - C_i^{bid}|. \tag{A.1}$$

(ii) Effect on model risk when hedging with the underlying:

$$\forall \phi \in \mathcal{S}, \quad \mu\left(\mathcal{X} + \int_0^T \phi_t \, \mathrm{d}S_t\right) = \mu(\mathcal{X}). \tag{A.2}$$

(iii) Convexity:

$$\forall \mathcal{X}_1, \mathcal{X}_2 \in \mathcal{C}, \forall \lambda \in [0,1] \quad \mu(\lambda \mathcal{X}_1 + (1-\lambda)\mathcal{X}_2) \leq \lambda \mu(\mathcal{X}_1) + (1-\lambda)\mu(\mathcal{X}_2). \tag{A.3}$$

(iv) Static hedging with traded options:

$$\forall \mathcal{X} \in \mathcal{C}, \forall u \in \mathbb{R}^K, \quad \mu\left(\mathcal{X} + \sum_{i=1}^K u_i H_i\right) \leq \mu(\mathcal{X}) + \sum_{i=1}^K |u_i(C_i^{ask} - C_i^{bid})|. \tag{A.4}$$

APPENDIX B

Parameter estimates and intervals

B.1 Parameter estimates from literature

The following parameter estimates are presented by Schoutens, Simons, and Tistaert in [40] and is included for comparison. The models are calibrated to the Eurostoxx 50 index on October 7th, 2003.

Heston		Bates		NIGCIR	
Parameter	Interval	Parameter	Interval	Parameter	Interval
V_0	0.0654	V_0	0.0576	κ	1.2101
κ	0.6067	κ	0.4963	ξ	0.5507
ξ	0.0707	ξ	0.0650	σ_V	1.7864
σ_V	0.2928	σ_V	0.2286	δ	1.0867
ρ	−0.7571	ρ	−0.9900	α	16.1975
		λ_j	0.1382	β	−3.1804
		μ_j	0.1791		
		σ_j	0.1346		

Table B.1: Parameter estimates for Heston, Bates and NIGCIR from literature.

B.2 Parameter intervals on August 22, 2014

Black-Scholes		Heston	
Parameter	Interval	Parameter	Interval
σ	$[0.1121, 0.1899]$	V_0	$[0.0101, 0.0105]$
		κ	$[1.6495, 1.6772]$
		ξ	$[0.0489, 0.0494]$
		σ	$[0.5655, 0.5771]$
		ρ	$[-0.7873, -0.7750]$

NIGCIR		Bates	
Parameter	Interval	Parameter	Interval
κ	$[4.3958, 4.8633]$	V_0	$[0.0074, 0.0083]$
ξ	$[1.2365, 1.3021]$	κ	$[1.7733, 1.8109]$
σ	$[6.0154, 6.4982]$	ξ	$[0.0474, 0.0480]$
δ	$[0.6553, 0.6784]$	σ	$[0.6237, 0.6409]$
α	$[69.3617, 71.6636]$	ρ	$[-0.8582, -0.8372]$
β	$[-14.8736, -14.2560]$	λ_j	$[1.9443, 2.2120]$
		μ_j	$[-0.0099, -0.0030]$
		σ_j	$[0.0278, 0.0298]$

Table B.2: Parameter intervals for Black-Scholes, Heston, NIGCIR and Bates on August 22^{nd}.

B.3 Parameter estimates in simulation study

B.3.1 Varying jump intensity in Bates simulation

Black-Scholes									
Parameter	λ_j								
estimate	1.4	1.45	1.50	1.55	1.60	1.65	1.70	1.75	1.80
$\hat{\sigma}$	0.1818	0.1828	0.1837	0.1846	0.1855	0.1864	0.1872	0.1881	0.1890

Heston									
Parameter	λ_j								
estimate	1.4	1.45	1.50	1.55	1.60	1.65	1.70	1.75	1.80
$\hat{V}_0$	0.0130	0.0133	0.0135	0.0138	0.0141	0.0143	0.0146	0.0148	0.0151
$\hat{\kappa}$	2.1808	2.1968	2.2127	2.2283	2.2438	2.2590	2.2741	2.2891	2.3038
$\hat{\xi}$	0.0521	0.0523	0.0525	0.0528	0.0530	0.0532	0.0534	0.0536	0.0539
$\hat{\sigma}$	0.5006	0.4987	0.4968	0.4950	0.4932	0.4914	0.4897	0.4880	0.4864
$\hat{\rho}$	−0.7762	−0.7756	−0.7750	−0.7745	−0.7741	−0.7736	−0.7732	−0.7729	−0.7725

NIGCIR									
Parameter	λ_j								
estimate	1.4	1.45	1.50	1.55	1.60	1.65	1.70	1.75	1.80
$\hat{\kappa}$	2.5728	2.5379	2.5004	2.4667	2.4354	2.4024	2.3694	2.3363	2.3064
$\hat{\xi}$	1.0466	1.0289	1.0107	0.9937	0.9774	0.9614	0.9453	0.9292	0.9146
$\hat{\sigma_V}$	2.3884	2.3071	2.2257	2.1515	2.0822	2.0146	1.9488	1.8850	1.8274
$\hat{\delta}$	0.3247	0.3267	0.3286	0.3307	0.3331	0.3353	0.3374	0.3396	0.3420
$\hat{\alpha}$	26.9888	27.0883	27.2244	27.3880	27.5872	27.7906	28.0363	28.3252	28.5985
$\hat{\beta}$	−14.5322	−14.8369	−15.1765	−15.5154	−15.8686	−16.2256	−16.6164	−17.0400	−17.4330

Table B.3: Parameter estimates for Black-Scholes, Heston and NIGCIR in the simulation study when varying λ_j in the Bates simulation.

B.3.2 Varying mean reversion rate in Bates simulation

Black-Scholes							
Parameter	κ						
estimate	0.8	1	1.2	1.4	1.6	1.8	2
$\hat{\sigma}$	0.1599	0.1679	0.1746	0.1804	0.1855	0.1899	0.1937

Heston							
Parameter	κ						
estimate	0.8	1	1.2	1.4	1.6	1.8	2
$\hat{V_0}$	0.0136	0.0139	0.0140	0.0140	0.0141	0.0140	0.0140
$\hat{\kappa}$	1.9232	1.9012	1.9738	2.0948	2.2438	2.4099	2.5872
$\hat{\xi}$	0.0408	0.0454	0.0488	0.0512	0.0530	0.0543	0.0553
$\hat{\sigma_V}$	0.4580	0.4644	0.4741	0.4841	0.4932	0.5012	0.5082
$\hat{\rho}$	−0.7621	−0.7638	−0.7668	−0.7702	−0.7741	−0.7780	−0.7820

NIGCIR							
Parameter	κ						
estimate	0.8	1	1.2	1.4	1.6	1.8	2
$\hat{\kappa}$	1.7081	1.8495	2.0172	2.2149	2.4356	2.6701	2.9200
$\hat{\xi}$	0.8650	0.9106	0.9400	0.9615	0.9776	0.9894	0.9987
$\hat{\sigma_V}$	1.7220	1.8431	1.9318	2.0108	2.0826	2.1455	2.2048
$\hat{\delta}$	0.2407	0.2649	0.2882	0.3110	0.3330	0.3542	0.3748
$\hat{\alpha}$	22.4330	23.2241	24.4370	25.9065	27.5800	29.4113	31.4092
$\hat{\beta}$	−13.4896	−13.5974	−14.1361	−14.9063	−15.8619	−16.9727	−18.2236

Table B.4: Parameter estimates for Black-Scholes, Heston and NIGCIR in the simulation study when varying κ in the Bates simulation.

B.4 Multistart intervals

Heston		Bates		NIGCIR	
Parameter	Interval	Parameter	Interval	Parameter	Interval
V_0	$[10^{-6}, 0.3]$	V_0	$[10^{-6}, 0.3]$	κ	$[10^{-2}, 100]$
κ	$[10^{-6}, 500]$	κ	$[10^{-6}, 500]$	ξ	$[10^{-2}, 100]$
ξ	$[10^{-6}, 0.5]$	ξ	$[10^{-6}, 0.5]$	σ_V	$[10^{-2}, 100]$
σ_V	$[10^{-6}, 0.1]$	σ_V	$[10^{-6}, 0.5]$	δ	$[10^{-2}, 100]$
ρ	$[-0.999, 0.999]$	ρ	$[-0.999, 0.999]$	α	$[10^{-2}, 500]$
		λ_j	$[10^{-6}, 10]$	β	$[-100, 100]$
		μ_j	$[-1, 1]$		
		σ_j	$[10^{-2}, 1]$		

Table B.5: Intervals for the multistart calibration for Heston, Bates and NIGCIR.

APPENDIX C

Metropolis-Hastings specifications

C.1 Prior covariance

The prior is a multivariate Gaussian centered around the previous days WLS estimate.

Parameter	Value
V_0	0.0221
κ	1.7941
ξ	0.0535
σ_V	0.5891
ρ	-0.8116

Table C.1: Weighted least squares estimate on August 7^{th}.

The covariance matrix for the prior distribution is chosen as

$$\Sigma = \begin{pmatrix} 0.05 & 0 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 & 0 \\ 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0.1 \end{pmatrix}. \tag{C.1}$$

C.2 Proposal covariance

The covariance matrix for the proposal distribution is chosen as

$$\Sigma_J = 500000^{-1} \begin{pmatrix} 1/150 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1/2 \end{pmatrix}. \tag{C.2}$$

Bibliography

[1] H. Akaike. “A New Look at the Statistical Model Identification”. In: *IEEE Transactions on automatic control* AC-19.6 (1974).

[2] Leif BG Andersen. “Efficient simulation of the Heston stochastic volatility model”. In: *Available at SSRN 946405* (2007).

[3] European Banking Authority. *EBA FINAL draft Regulatory Technical Standards*. 2014.

[4] Ds Bates. “Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options”. In: *Rev. Financ. Stud.* 9.1 (1996), pp. 69–107.

[5] HS Bhat and N Kumar. “On the derivation of the Bayesian Information Criterion”. In: *School of Natural Sciences, University of California* (2010).

[6] T. Björk. *Arbitrage Theory in Continuous Time*. Oxford scholarship online. Oxford University Press, 1998. ISBN: 9780191525100.

[7] Fischer Black and Myron Scholes. “The Pricing of Options and Corporate Liabilities”. English. In: *Journal of Political Economy* 81.3 (1973), pp. 637–654. ISSN: 00223808.

[8] Nicole Branger and Christian Schlag. “Model risk: A conceptual framework for risk measurement and hedging”. In: *EFMA 2004 Basel Meetings Paper*. 2004.

[9] FO Bunnin, Y Guo, and Y Ren. “Option pricing under model and parameter uncertainty using predictive densities”. In: *Statistics and Computing* 12 (2002), pp. 37–44.

[10] K.P. Burnham and D.R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, 2002. ISBN: 9780387953649.

[11] Xumeng Cao. “Relative performance of expected and observed Fisher information in covariance estimation for maximum likelihood estimates”. In: (2013).

[12] Peter Carr and Dilip B. Madan. “Option Valuation Using the Fast Fourier Transform”. In: *Journal of computational finance* 2 (1999), pp. 61–73.

[13] Peter Carr et al. “Stochastic Volatility for Lévy Processes”. In: *Mathematical Finance* 13.3 (2003), pp. 345–382.

[14] Rama Cont. “Model uncertainty and its impact on the pricing of derivative instruments”. In: *Mathematical finance* 16.3 (2006), pp. 519–547.

[15] Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC Financial Mathematics Series, 2003.

[16] John C Cox, Jr Ingersoll Jonathan E, and Stephen A Ross. “A Theory of the Term Structure of Interest Rates”. In: *Econometrica* 53.2 (1985), pp. 385–407.

[17] Emanuel Derman. “Model risk”. In: *Quantitative Strategies Research Notes, Goldman Sachs* (1995).

[18] Emanuel Derman, Iraj Kani, and Joseph Zou. "The Local Volatility Surface, Unlocking the Information in Index Option Prices". In: *Quantitative Strategies Research Notes, Goldman Sachs* (1995).

[19] Detering and Packham. "Model risk of contingent claims". In: *Working paper* (2014).

[20] Moritz Duembgen and Leonard Rogers. "Estimate nothing". In: *Quantitative Finance* 14.12 (2014), pp. 2065–2072.

[21] Paul Glasserman. *Monte Carlo methods in financial engineering*. Vol. 53. Springer Science & Business Media, 2003.

[22] Peter J. Green. "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination". In: *Biometrika* 82 (1995), pp. 711–732.

[23] Alok Gupta. "A Bayesian approach to financial model calibration, uncertainty measures and optimal hedging". PhD thesis. Citeseer, 2010.

[24] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (Apr. 1970), pp. 97–109.

[25] Steven L. Heston. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options". In: *Review of Financial Studies* 6 (1993), pp. 327–343.

[26] Jennifer A Hoeting et al. "Bayesian model averaging: a tutorial". In: *Statistical science* (1999), pp. 382–401.

[27] Rob J. Hyndman and Yanan Fan. "Sample Quantiles in Statistical Packages". In: *The American Statistician* 50.4 (1996).

[28] F.L.J. Kerkhof, Bertrand Melenberg, and Johannes M. Schumacher. *Model Risk and Regulatory Capital*. Discussion Paper 2002-27. Tilburg University, Center for Economic Research, 2002.

[29] J. Kienitz and D. Wetterau. *Financial Modelling: Theory, Implementation and Practice with MATLAB Source*. The Wiley Finance Series. Wiley, 2013. ISBN: 9780470744895.

[30] Erik Kivilo and Carl Olofsson. "Prudent Valuation & Model Risk Quantification". MA thesis. Lund University, 2014.

[31] Christiaaan Nicolaas Klijn. "Finding cancer genes in copy number data and insertional mutagenesis data". PhD thesis. TU Delft, Delft University of Technology, 2011.

[32] Roger W. Lee. "Option Pricing by Transform Methods: Extensions, Unification, and Error Control". In: *Journal of Computational Finance* (2004).

[33] Erik Lindström. "Implications of parameter uncertainty on option prices". In: *Advances in Decision Sciences* 2010 (2010).

[34] Erik Lindström et al. "Sequential calibration of options." In: *Computational Statistics Data Analysis* 52.6 (Mar. 10, 2008), pp. 2877–2891.

[35] Nicholas Metropolis et al. "Equation of state calculations by fast computing machines". In: *The journal of chemical physics* 21.6 (1953), pp. 1087–1092.

[36] M. Morini. *Understanding and Managing Model Risk: A Practical Guide for Quants, Traders and Validators*. The Wiley Finance Series. Wiley, 2011. ISBN: 9781119960850.

[37] Nicholas G Polson and Jonathan R Stroud. "Bayesian inference for derivative prices". In: *Bayesian Statistics* 7 (2003).

[38] Adrian Raftery. *Bayesian Model Selection in Social Research (with Discussion by Andrew Gelman Donald B. Rubin, and Robert M. Hauser, and a Rejoinder)*. 1995.

[39] Felix Salmon. *Recipe for Disaster: The Formula That Killed Wall Street.* [Online; accessed 22-January-2015]. 2010. URL: `http://archive.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all`.

[40] W. Schoutens, E. Simons, and J. Tistaert. "A perfect calibration! Now what?" In: *Wilmott* (2004), pp. 66–78.

[41] Gideon Schwarz et al. "Estimating the dimension of a model". In: *The annals of statistics* 6.2 (1978), pp. 461–464.

[42] Wall St, David X. Li, and David X. Li. "On Default Correlation: A Copula Function Approach". In: *The Journal of Fixed Income* (2000).

Master's Theses in Mathematical Sciences 2015:E19
ISSN 1404-6342

LUTFMS-3277-2015

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/