

Robust Calibration of Financial Models Using Bayesian Estimators

Alok Gupta* Christoph Reisinger†

February 7, 2012

Abstract

We consider a general calibration problem for derivative pricing models, which we reformulate into a Bayesian framework to attain posterior distributions for model parameters. It is then shown how the posterior distribution can be used to estimate prices for exotic options. We apply the procedure to a discrete local volatility model and work in great detail through numerical examples to clarify the construction of Bayesian estimators and their robustness to the model specification, number of calibration products, noisy data and misspecification of the prior.

1 Introduction

Since the model proposed by Black and Scholes in their seminal 1973 paper [10], the variety and complexity of financial models has grown dramatically. Typically, agents will want to use a model to price or hedge an instrument in the market. But before they can do this, they must first mark the model to market — that is, *calibrate* the model to observable prices. Most commonly, vanilla instruments such as European calls or puts are used. This calibration is necessary to avoid introducing arbitrage into the market by making the agent vulnerable to other agents creating riskless profits from the first agent's incorrect prices.

In the original Black-Scholes model, there is one scalar volatility parameter to be estimated. In contrast, in some of the commonly used models today, entire functions have to be calibrated, which raises questions not just of numerical complexity but of identifiability of the model from a restricted set of observations (market data) and their robust and stable estimation.

Calibration can be classed as an *inverse problem*: a parametrised model has been specified, we then observe market prices and try to find the model parameter which gives those prices. Put abstractly, the calibration problem often fails one or more of Hadamard's criteria for well-posedness (see for example [17]), which are:

*Mathematical Institute, University of Oxford, alok.gupta@maths.ox.ac.uk

†Mathematical Institute, University of Oxford, christoph.reisinger@maths.ox.ac.uk

A. Gupta acknowledges financial support from the UK Engineering and Physical Sciences Research Council (EPSRC) and Nomura International plc under CASE and PhD Plus Awards. The work was completed while C. Reisinger was a Visiting Fellow at the Isaac Newton Institute for Mathematical Sciences, University of Cambridge, during the programme *Inverse Problems*.

- 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.

We assume the first criteria is true, i.e. there exists a parameter for which the model reproduces market prices, for otherwise our model is poorly designed and introduces arbitrage into the market. In reality, the calibration instruments, to which we try to mark our model, are only observable in the market upto some *bid-ask spread*, which is the interval of values between what an agent is willing to pay for the instrument (at the lower end) and what an agent is willing to sell the instrument for (at the upper end).

The latter two conditions are not certain to be satisfied by the calibration problem. It is clear that picking the wrong solution (by condition ii.) or choosing a solution that is not stable (by condition iii.) can have disastrous effects on the pricing and hedging of an instrument.

For illustration, we consider here the local volatility model, in which the underlying asset price S is assumed to follow

$$dS_t/S_t = \mu dt + \sigma(S_t, t) dZ_t, \quad (1)$$

where the drift μ is the expected growth rate and the volatility σ a function of both the asset price and time, and Z is a standard Brownian motion. The function $\sigma(\cdot, \cdot)$ is a priori unknown and must be inferred from observed prices. Dupire [15] derives an explicit formula, expressing this function in terms of European call prices and their sensitivities with respect to strike and maturity. Hindering the direct practical application of the formula, it requires prices for a continuum of strikes and maturities, which are not quoted in reality, making additional assumptions necessary, in practice by interpolation and discretisation. The formula illustrates two fundamental facts: that a discrete set of observation prices is not sufficient to pin down the functional parameter, and even if a continuum of prices was available, the solution of the inverse problem is unstable.

To address the difficulties of calibrating the local volatility model, authors such as Jackson et al. [26], Lagnado & Osher [28] and Coleman et al. [12] have developed minimisation techniques and penalty functions for finding the ‘best-fit’ local volatility surface with a certain regularity. Further analysis of these methods and their improvement has been detailed by Chiarella et al. [11]. Crepey [14] and Egger & Engl [16] show that a carefully (Tikhonov) regularised problem is well-posed in the above sense, and rates for the convergence of the regularised solution can be derived for vanishing data noise and regularisation parameter. This approach has been extended to American calibration options by Achdou and Pironneau [1, 2].

Choosing a regularising functional restricts the solution to a more well-behaved class, but the resulting solution does not contain information on its uncertainty.

Given that financial models typically make very specific assumptions on the processes they are describing, we would like to add *robustness* as signifying

- iv. insensitivity to small deviations from the assumptions

to our list of desiderata. In the context of this study on financial model calibration, the main assumptions made are: the assumed class of models which

are being calibrated; approximation of these models, e.g. via discretisation of a local volatility function; any regularising assumptions, e.g. smoothing penalty terms; the choice of calibration instruments. As any financial market model will only be an approximation to the true data generating process, the main requirement of a robust method is that the predictions made from the estimated model chosen out of the ‘wrong’ model class are a sufficiently accurate approximation. We will assess this property by numerical tests in synthetic (i.e. the calibration data are generated by a model) and real markets later on.

A different viewpoint is taken by Avellaneda et al. [5, 6] and Lyons [29]. Rather than imposing a detailed description of the term-structure and leverage of instantaneous variance, only an upper and lower bound is assumed. Upper and lower price bounds for derivatives (the sub- and superhedging prices) are obtained by solving a stochastic control problem, where quoted market prices form constraints and can narrow the no-arbitrage price bands for other contracts considerably. Using relative entropy regularisation introduced by Avellaneda et al. [4], Samperi [32] shows that the infimum of a regularised error function is continuous (in fact, differentiable) with respect to calibration prices. Extending this to uncertain volatility function bounds, He et al. [25] obtain more realistic bid and ask prices than for constant bounds.

The Bayesian approach of this paper demonstrates a shift in philosophy of the aforementioned approaches. Acknowledging that the calibration problem is ill-posed, we no longer focus on finding a *best-fit* solution, but we are interested in finding a *distribution* of solutions. The essential idea behind the Bayesian approach is to begin with some *prior* distribution for the unknown parameter and update this distribution using the observable market prices to give a *posterior* distribution for the model parameter. So instead of finding *a* model which, in some measure, *best* replicates prices, we seek all models which *sufficiently replicate* prices to within a pre-decided tolerable level of error. This is not dissimilar from uncertain parameter models so far, and related to an approach that more recently Hamida & Cont [24] have adopted as part of their investigation into model risk, to obtain a spread of possible prices of exotic options which are all consistent with those of calibration options. Where this paper differs from [24] is by recasting the problem into a Bayesian framework, while [24] use a prior distribution only to generate initial populations for an evolutionary optimisation algorithm. The Bayesian approach to calibration has been used before by authors such as Jacquier & Jarrow [27] in Black-Scholes models, Bhar et al. [9] for the calibration of instantaneous spot and forward interest rates, and Monoyios [30] in the context of drift parameter uncertainty. The specification of a prior corresponds to the regularising penalty in Tikhonov regularisation and opens the possibility of incorporating prior information in a rigorous framework, although this does mean that the impact of prior assumptions has to be assessed critically.

In this paper we concentrate on providing a practical method for constructing prior and likelihood functions and on exploring the robustness of the Bayesian posteriors, especially with the view towards pricing exotic options using Bayesian estimates. This is to be seen as “proof of concept” for a challenging example (of a high-dimensional parameter), and improvements to the computational strategy would aid the practical application in this setting. A marked advantage of the Bayesian approach, which this work highlights, is that the posterior distribution can be translated into price spreads for derivatives, in the spirit of the model

uncertainty measures in [13, 24]. We also demonstrate by case studies that the Bayesian mean yields reliable predictions of exotic derivative prices, which are more robust than those based on parameters obtained by Tikhonov regularisation, and much more accurate also than suggested by model uncertainty bounds not taking into account information of the posterior distribution.

The paper is divided as follows. In Section 2 we formalise the calibration problem and review relevant results from Bayesian theory. In Section 3, we discuss the construction of the prior and likelihood function as applied to the local volatility model. Section 4 provides details of its discretisation and on the tailoring of Metropolis sampling to this application, and give calibration examples in Section 5. Finally in Section 6 we present some case studies which demonstrate the robustness of the proposed calibration method for pricing other contracts. Section 7 concludes.

2 Set-up of the Calibration Problem

We consider financial derivatives on an asset whose price process is modelled by $S(\theta) = (S_t(\theta))_{t \geq 0}$ over time t , where the possible models are labeled by a parameter $\theta \in \Theta$, chosen out of a set Θ of parameters (models). The perhaps simplest example is the constant volatility parameter in the Black-Scholes model. In the example of local volatility, which will be studied extensively in the following sections, $\theta = \sigma(\cdot, \cdot)$ is a function of two arguments (the stock price S and time t), chosen out of a suitable class of functions to be specified later.

Now consider an option over a finite time horizon $[0, T]$ written on S and with payoff function h . We write the time t value of this option under the above model for S as $f_t(\theta)$. The price building mechanism which leads to this functional form is of course important (e.g., a replication argument in a complete arbitrage free market), but we will leave this open for chosen applications. We include the argument θ in f to emphasise the dependence of this price on the model (parameter).

Suppose at time $t \in \Upsilon_n([0, T]) = \{t_1, \dots, t_n : 0 = t_1 < t_2 < \dots < t_n \leq T\}$ we observe market quotes $V_t^{(i)}$ for these options, where $i \in I_t$ an index set. These could be European call options with different strikes and maturities. The common approach to calibration is to find a value of θ consistent with the observed prices $V = \{V_t^{(i)} : i \in I_t, t \in \Upsilon_n\}$.

In practice, instead of a single $V_t^{(i)}$ one usually has quotes available for a bid price $V_t^{(i)bid}$, and an ask price $V_t^{(i)ask}$, the best prices for which agents are willing to buy and sell the option, respectively. A calibration is therefore only arbitrage-free if the model price lies in the interval $[V_t^{(i)bid}, V_t^{(i)ask}]$, but without the input of additional information one cannot distinguish between models which calibrate within the bid-ask spread.

To this end, we cast this problem in a Bayesian framework. Assume we have some prior information for θ (for example that it belongs to a particular subset of the original parameter space, e.g. represents a positive constant or smooth function), summarised by a *prior* density $p(\theta)$ for θ . We then observe data $V = \{V_t^{(i)} : i \in I_t, t \in \Upsilon_n\}$, for instance

$$V_t^{(i)} = \frac{1}{2} \left(V_t^{(i)bid} + V_t^{(i)ask} \right),$$

and write their relation to the ‘true’ prices (identical to the model prices with ‘true’ parameter θ^*) by

$$V_t^{(i)} = f_t^{(i)}(\theta^*) + e_t^{(i)} \quad (2)$$

with additive noise $\{e_t^{(i)} : i \in I_t\}$. A possible interpretation of (2) is that there is an underlying true model, unknown to the observer, under which the market is complete and arbitrage-free (i.e. derivatives can be hedged perfectly knowing the model); the bid-ask spread reflects the model uncertainty which causes the buyer/seller to demand a risk premium. The existence of a true model is not necessary for the definition of the calibration procedure in this paper as such, as long as the observed market prices are attainable within the class of assumed models subject to the assumed noise. It would clearly become relevant if we were to address questions of consistency of Bayesian estimators and hedging based on those parameters.

Then $p(V|\theta)$, the probability of observing the data V given θ , is determined by the distribution of the noise e and is called the *likelihood* function. We will discuss a complete specification of the noise in 3.2.

An application of Bayes rule gives that the *posterior* density of θ is given by

$$p(\theta | V) = \frac{p(V|\theta)p(\theta)}{p(V)} = \frac{p(V|\theta)p(\theta)}{\int p(V|\theta)p(\theta) d\theta}.$$

If the noise is modelled such that observations only have positive likelihood if the model price lies within the bid-ask spread, we can turn this around to say that any parameter with positive posterior density gives model prices for the calibration options within the bid-ask spread.

The estimator

$$\theta_{MAP}(V) = \arg \max_{\theta \in \Theta} \{p(\theta|V)\},$$

the *maximum a posteriori* (MAP) estimator, is the value which maximises the posterior density. A family of estimators $\theta_L(V)$ can be defined as

$$\theta_L(V) = \arg \min_{\theta'} \left\{ \int_{\Theta} L(\theta, \theta') p(\theta|V) d\theta \right\},$$

where $L : \mathbb{R}^{2M} \rightarrow \mathbb{R}$ is a *loss function* with the property

$$\begin{cases} L(\theta, \theta') = 0 & \text{if } \theta' = \theta, \\ L(\theta, \theta') > 0 & \text{if } \theta' \neq \theta. \end{cases}$$

The minimiser $\theta_L(V)$ is not necessarily unique. $L_1(\theta, \theta') = \|\theta - \theta'\|_2^2$ gives the Bayes estimator $\theta_{L_1}(Y) = \mathbb{E}[\theta|V]$, which is the *mean* value of θ with respect to the Bayesian posterior density $p(\theta|V)$. The MAP estimator does not correspond to a non-negative bounded loss function.

We discuss possible interpretations of the result of such a calibration further.

In an ideal world, the following is given: even if we do not *a priori* know the true model (data generating process), we know that the model belongs to a certain class of models, parametrised by θ , and the observations allow us to differentiate between the true model and any other model from this class.

In the reality of financial markets, a class of candidate models can only be assumed. Most certainly, the observed data are generated by a process outside

the assumed model class. It can therefore not be expected that the estimator reproduces the values of all financial instruments outside the class of calibration instruments exactly (or within bid and ask). A robust estimator will have the property that if the true model is in some sense close to the assumed class of models, predictions made from the estimated parameter, say values of exotic derivatives or hedge parameters, will be close to the value of those derivatives under the true model and can be hedged accurately by a trading strategy based on the estimated parameter.

In the example of the local volatility model, θ is an infinite dimensional (functional) parameter. This means on the one hand that any finite number of observations will be insufficient to identify the parameter exactly. Moreover, one will only be able to compute a finite-dimensional (discretised) approximation. So even if the market is governed by a local volatility model, this function will almost certainly lie outside, albeit close to, the assumed (computable) class of local volatility models, which is necessarily part of a finite-dimensional space.

The relation of the number and type of parameters to the number and type of calibration instruments then becomes relevant. If the number of data is smaller than the number of parameters, the parameter will generally be under-determined and regularisation, here via the Bayesian prior, favours particular parameters over others. As Wasserman [33] remarks, we should desire that the Bayesian posterior is not dominated and led astray by the priors, so care has to be taken with its construction.

There are two routes to increasing the available information in this context: to observe the value of different financial derivatives, e.g. vanilla options with different strikes and maturities, and/or to observe the prices of the same financial derivatives at different times, in which case past calibrations provide prior information for re-calibrations. We will later give examples for both. As the number of calibration products, and/or the re-calibration frequency, increase, a relevant property of Bayesian estimators is *consistency*. Ghosal [21] points out that consistency is crucially important for parameter estimation since the violation of consistency puts serious doubts against inferences based on the inconsistent posterior distribution. We note that there is a vast body of literature on Bayesian consistency in a more general and more advanced setting, see e.g. [22] and the references therein. It ensures that the estimator converges to the ‘true’ parameter, if such a ‘true’ parameter exists, i.e. on the assumption that the data are indeed generated by the assumed model. This cannot be assumed to be the case in financial applications.

We discuss these points by numerical illustrations in Sections 5 and 6, after introducing the construction of Bayesian posteriors and their numerical realisation in the following two sections.

3 Bayesian Estimation of Local Volatility

In this section, we discuss the main ingredients of the Bayesian set-up. Recall θ is the unknown parameter in the model for the evolution of the asset price $S(\theta)$, which we take as local volatility function in our running example. A contention of this paper is that certain features such as positivity or certain asymptotic behaviours might be expected of θ and that these prior views should be incorporated into the calibration procedure. To this end the Bayesian theory

previously introduced serves as a useful rigorous framework.

3.1 Prior Distribution (Regularisation)

We identify key characteristics expected of the local volatility surface that can be recast into a Bayesian prior. We encapsulate the information available of σ at the outset in the following three properties:

Positivity: Since the squared price variation $\sigma^2 > 0$ we adopt the convention $\sigma > 0$.

Asymptotics: For small values of t especially, σ should be close to today's at-the-money volatility σ_{atm} ; at-the-money volatility is usually equated with the Black-Scholes implied volatility for a European call option with strike equal to the current price of the underlying and for a short maturity. Today's at-the-money volatility will roughly determine the position of the local volatility surface in \mathbb{R}^3 . A more refined asymptotic shape of the local volatility for small times and in the wings can be deduced from the implied volatility skew for short maturities and far in- and out-of-the-money options, as detailed in [7].

Smoothness: We exclude sharp spikes or troughs in the surface; we do not argue that sudden changes could not happen, but follow Jackson et al. [26] to point out that a small set of currently observed prices will not be capable of accurately predicting abrupt changes in future volatility.

The following approach is to reformulate prior beliefs of θ into a norm *cost functional* $\|\cdot\|$ of θ so that parameters which better satisfy the prior beliefs have smaller norm. Then the natural Gaussian prior is

$$p(\theta) \propto \exp \left\{ -\frac{1}{2} \tilde{\lambda} \|\theta - \theta_0\|^2 \right\}, \quad (3)$$

where $\tilde{\lambda}$ is a constant which quantifies how strong our prior assumptions are: a higher value of $\tilde{\lambda}$ indicating greater confidence in our assumptions. $1/\tilde{\lambda}$ can be thought of as the prior variance of θ . From (3) we see that those θ which better satisfy prior beliefs have greater density.

To illustrate a possible choice of norm $\|\cdot\|$, we continue our example of the local volatility model. In light of the assumptions presented earlier, we choose volatility functions from the set

$$\{\sigma \in H^1(\mathbb{R}_+^2 \cap K), \sigma > 0 \text{ a.e.}, \|\log \sigma\|_{1,K} < \infty \text{ for all compactly contained } K\},$$

where H^1 is the Sobolev space of weakly differentiable functions and $\|\cdot\|_{1,K}$ the standard H^1 norm on K . The restriction to compact sets allows different behaviour of the volatility in the tails at this stage.

Following the regularisation functional used, for example, by Fitzpatrick [18], we choose for the prior

$$p_{lv}(\sigma) \propto \exp \left\{ -\frac{1}{2} \lambda_p \|\log(\sigma) - \log(\sigma_{atm})\|_{\kappa}^2 \right\}, \quad (4)$$

where

$$\|u\|_{\kappa}^2 = (1 - \kappa) \|u\|_0^2 + \kappa \|\nabla u\|_0^2, \quad (5)$$

and here $\nabla = (\frac{\partial}{\partial t}, \frac{\partial}{\partial S})$ is the grad operator, $\|\cdot\|_0$ is the standard L_2 norm on a suitable domain, and $\kappa \in (0, 1)$ is a pre-specified constant.

Using the logarithm penalises σ approaching zero. The first part of the norm is to ensure greater prior density is attached to σ that are closer to the ATM volatility. The second part ensures that the volatility is locally linear in its arguments — once discretised on a grid, this will ensure that the prior covariance of volatilities at neighbouring grid nodes approaches one if the grid points are close.

3.2 Likelihood Function (Calibration Error)

Recall from Section 2 that $V_t^{(i)}$ is a market-observed price at time t of a calibration option, written on underlying S taking value S_t at time t , and $f_t^{(i)}(\theta)$ is the model price for the same derivative when the model parameter is θ . Typically, one will choose liquidly traded options for the calibration, such as European calls with a range of strikes and maturities.

The definition of a likelihood relies on assumptions on the data noise. We shall use the observations $V_0^{(i)} = \frac{1}{2}(V_0^{(i)bid} + V_0^{(i)ask})$, and define $\delta_i = \frac{10^4}{S_0} |V_0^{(i)ask} - V_0^{(i)bid}|$ as the basis point bid-ask spread for the i^{th} option at time 0. In the first instance, we model the basis point error for the i^{th} option by a normal distribution

$$\frac{10^4}{S_0} (f_0^{(i)}(\theta^*) - V_0^{(i)}) \sim N(0, \delta_i^2), \quad (6)$$

which we assume independent across data. To avoid introducing arbitrage for calibration options, one may truncate the density, e.g. attach zero density to all observations for which the model values lie outside the bid-ask spread.

It is important that inferences are not sensitive to the noise model, which influences the shape of the posterior distribution for finite sample size and has implications for asymptotic properties such as consistency. Assumption of the normal distribution (6) implies that the likelihood of the bid-ask mid-point is largest close to the ‘true’ value, and that the bid-ask spread is a measure for the width of the distribution. As we truncate the distribution outside an interval around the mid-price, the tails of the assumed (un-truncated) distribution become irrelevant. We shall see in the numerical examples that the impact of the (assumed) standard deviation of the price error on the calibration is small within a relevant range.

The assumption of independent noises is made for simplicity, and may have to be revised especially for observations for very similar strikes or maturities. We comment on re-calibration at different times t_n later.

For computational purposes, we will modify the above definition slightly. Using the terminology of Jackson et al. [26], we define the basis point square-error function for a calibration at time $t = 0$ as

$$G(\theta) = \frac{10^8}{S_0^2} \sum_{i \in I} w_i |f_0^{(i)}(\theta) - V_0^{(i)}|^2, \quad (7)$$

where the w_i are pre-specified weights summing to one, and are chosen depending on relative volumes likely to be traded (so, for instance, at-the-money calls or puts are weighted more heavily).

Then we shall only attach positive Bayesian posterior weight to parameters θ which on average reproduce prices to within the average basis point bid-ask spread. In other words, we will attach positive likelihood under θ only if

$$G(\theta) \leq \delta^2, \quad (8)$$

where $\delta^2 = \sum_{i \in I} w_i \delta_i^2$ is the pre-specified average basis point square-error tolerance. Additionally, the smaller the value of $G(\theta)$, the more *likely* the observation. Hence, for the Bayesian likelihood we will take

$$p(V|\theta) \propto 1_{G(\theta) \leq \delta^2} \exp \left\{ -\frac{1}{2\delta^2} G(\theta) \right\}. \quad (9)$$

So those parameters (surfaces) σ which reproduce prices closest to the market observed prices V (the bid-ask mid-point) give the greatest likelihood values.

Note that we could modify the likelihood so that every price is calibrated to within its tolerance, i.e. $(10^4/S_0)|f_0^{(i)}(\theta) - V_0^{(i)}| \leq \delta_i$ for all $i \in I$. This would be computationally more intense because it requires greater exploration of the parameter space to find sufficiently many acceptable candidate solutions. Satisfaction of (8) is the approach also taken by Hamida & Cont [24], however, they consider all surfaces σ satisfying the constraint (8) equally well calibrated, whereas in this study we will differentiate between different degrees of calibration, which subsequently leads to different weight in the posterior distribution.

3.3 Posterior Distribution

Combining prior and likelihood functions, we get the posterior explicitly as

$$p(\theta|V) \propto 1_{G(\theta) \leq \delta^2} \exp \left\{ -\frac{1}{2\delta^2} [\lambda \|\theta - \theta_0\|_{\kappa}^2 + G(\theta)] \right\}, \quad (10)$$

where $\lambda = \delta^2 \tilde{\lambda}$. From this, estimates for θ and other predictions can be derived as discussed in Section 2.

Observe that maximising the posterior (10) is equivalent to minimising the expression

$$\lambda \|\theta - \theta_0\|_{\kappa}^2 + G(\theta) \quad (11)$$

over the set $\{\theta : G(\theta) < \delta^2\}$, and (11) is precisely the form of function authors such as Lagnado & Osher [28] and Jackson et al. [26] seek to minimise to find their optimal calibration parameter. This shows how the Bayesian approach reformats and generalises traditional Tikhonov regularisation methods into a unified framework, as is already noted by Fitzpatrick [18].

The posterior density, however, contains more information than the MAP estimator and we will use this in Section 6 for the robust pricing of further options.

4 Numerical Method

This section outlines the numerical approach leading to samples from the posterior distribution.

4.1 Parameter Discretisation and the Value Function

We first restrict θ to a finite-dimensional space, and represent the local volatility surface $\sigma(S, t)$ by a grid of nodes whose positions are given by: $S_{min} = s_1 < \dots < s_j < \dots < s_J = S_{max}$ in the spatial direction and $0 = t_1 < \dots < t_l < \dots < t_L = t_{max}$ in the temporal direction. Following the ordering convention $\sigma_{j+(l-1)J} = \sigma(s_{j+1}, t_l)$, the discrete representation of $\sigma(S, t)$ is defined by the parameter vector

$$\theta = (\log \sigma_1, \dots, \log \sigma_m, \dots, \log \sigma_M),$$

and a spine interpolant $\Theta(\cdot, \cdot)$ of θ . We emphasise this dependence by writing $\sigma(\cdot, \cdot; \theta)$.

For each time t_l we construct the unique natural cubic spline through the nodes $(s_1, t_l), \dots, (s_J, t_l)$ to give all values $\Theta(S, t_l)$. Then for $(S, t) \in [s_j, s_{j+1}] \times [t_l, t_{l+1}]$ the value of $\Theta(S, t)$ is found by linear interpolation of the two values $\Theta(S, t_l)$ and $\Theta(S, t_{l+1})$. Then $\sigma(S, t) = \exp(\Theta(S, t))$. By interpolating the logarithm of the volatility and then exponentiating we ensure that $\sigma > 0$.

With this discretisation, the norm function in (4) can be written as

$$\|\Theta - \theta_0\|_{\kappa}^2 = (\theta - \theta_0)^T C (\theta - \theta_0),$$

where $\theta_0 = \log(\sigma_{atm})$ and C is the inverse covariance matrix induced by the norm. This follows because the spline basis coefficients are linear in the nodal values θ , and the squared Sobolev norms of the splines are quadratic in the splines. With the norm given by (5), C is non-singular so for the sake of convention write $A^{-1} \equiv C$.

This is similar to the approach taken in [26].

Finally, to calculate the likelihood value (9) for each θ , using (7), we must price all calibration options, say European call options, $f_t^{(i)}(\theta)$ for $i \in I_t$, using the model parameter θ . For the local volatility model, we follow the method of [24] of solving the Dupire PDE [15] with appropriate boundary conditions:

$$\begin{aligned} \frac{\partial f}{\partial T} + K(r - d) \frac{\partial f}{\partial K} - \frac{K^2 \sigma^2(K, T; \theta)}{2} \frac{\partial^2 f}{\partial K^2} &= 0 \quad \forall K, T \geq 0, \\ f(S, K, 0) &= (S - K)^+ \quad \forall K \geq 0. \end{aligned}$$

To solve this PDE numerically, we use a Crank-Nicolson finite difference scheme to give all the prices for the range of K and T simultaneously. This is computationally more efficient than solving Black-Scholes PDEs (in S and t) for all combinations of K and T separately.

4.2 Metropolis Sampling

Given the price data V , we are interested in the posterior distribution $p(\theta|V)$ given by (10). However, because of the high dimensionality of θ it is unfeasible to analytically find $p(\theta|V)$ or a numerical, e.g. grid-based, representation. Instead, the best we can do is try to draw samples from this distribution and draw conclusions based on these samples.

First of all observe that the form of the posterior given by (10) makes direct sampling slightly difficult. It will be computationally more efficient to first generate a set of sample from the non-truncated (normal) density

$$g(\theta|V) \propto \exp \left\{ -\frac{1}{2\delta^2} [\lambda \|\theta - \theta_0\|_{\kappa}^2 + G(\theta)] \right\}, \quad (12)$$

say $\{\theta_1, \dots, \theta_n\}$. Then $\{\theta_i : G(\theta_i) \leq \delta^2\}$ is a set of samples from $p(\theta|V)$ given by (10).

We now concentrate on generating samples from $g(\theta|V)$. To do this we will use the Markov Chain Monte-Carlo (MCMC) Metropolis algorithm which proceeds as follows (see [19] for further detail):

1. Select a starting point θ_0 for which $g(\theta_0|V) > 0$.
2. For $r = 1, \dots, n$, sample a proposal $\theta^\#$ from a *symmetric jumping distribution* $J(\theta^\#|\theta_{r-1})$ and set

$$\theta_r = \begin{cases} \theta^\# & \text{with probability } \min \left\{ \frac{g(\theta^\#|V)}{g(\theta_{r-1}|V)}, 1 \right\}, \\ \theta_{r-1} & \text{otherwise.} \end{cases}$$

Then the sequence of iterations $\theta_1, \dots, \theta_n$ converges to the target distribution $g(\theta|V)$.

To optimise the routine we run m parallel chains, each starting from a different point $\theta_0^{(j)}$ such that the set of starting points $\{\theta_0^{(1)}, \dots, \theta_0^{(m)}\}$ is an *overdispersed* sample of the target distribution. By overdispersed we mean that the samples are more widely distributed than the target distribution (see [20]). We also discard the first b iterations of the run (known as the *burn-in*) since it takes some exploratory time for the algorithm to settle on the target distribution. And we only keep every k th draw from the remaining iterations (known as *thinning*) of the sequence to reduce the correlation between samples.

Authors such as Beskos & Stuart [8] and Hairer et al. [3] have recommended associating the jump function with a random walk for which the transition kernel is associated with the prior density (3). Let A^{-1} be the inverse non-singular covariance matrix introduced in the previous section. By Cholesky decomposition we can find a matrix B such that $A = BB^T$. Then the jump function $J(\theta'|\theta)$ is given by

$$\theta' = \theta + \sqrt{2du}B\xi,$$

where $\xi \sim N(0, I_M)$ and du is the *step size* of our random walk. The value of du is chosen so that the acceptance rate of jumps is close to the optimum value of 23% found by Gelman et al. [19].

The values for n , m , b , k , du for each numerical example are given in Appendix A.

4.3 Monitoring Convergence of Metropolis Sampling

Convergence can be assured by checking that the *potential scale reduction factor* (PSRF) [19] of estimands of interest, the calibration prices for example, is close to 1 — and at least less than 1.1 in particular. Full details of the calculation and explanation of the procedure are set out below.

Suppose we have found m chains of length n (after discarding the burn-in and using thinning). For each calibration price v we notate by v_{ij} the i^{th} draw from the j^{th} simulated chain, for $i = 1, \dots, n$ and $j = 1, \dots, m$. Then we compute the between-sequence variances, B , and within-sequence variances,

W , as follows:

$$B = \frac{n}{m-1} \sum_{j=1}^m (\bar{v}_{.j} - \bar{v}_{..})^2 \quad \text{and} \quad W = \frac{1}{m} \sum_{j=1}^m s_j^2,$$

where

$$\bar{v}_{.j} = \frac{1}{n} \sum_{i=1}^n v_{ij} \quad \bar{v}_{..} = \frac{1}{m} \sum_{j=1}^m \bar{v}_{.j} \quad s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (v_{ij} - \bar{v}_{.j})^2.$$

Convergence is then assessed by estimating the factor by which the scale of the current distribution for v might be reduced if the length of the chain was allowed to continue in the limit $n \rightarrow \infty$. The potential scale reduction factor (PSRF) is estimated by

$$PSRF(v) = \sqrt{1 - \frac{1}{n} \left(1 - \frac{B}{W}\right)}, \quad (13)$$

which tends to 1 as $n \rightarrow \infty$. If the PSRF is high, a lot greater than 1.1 for example, then it is likely that continuing simulation will improve inferences based on the target distribution.

The estimate works because B usually overestimates posterior variance assuming the starting distribution is overdispersed. Whereas W usually underestimates the posterior variance because within chain samples have not had sufficient time to range over all the target distribution. But the longer we run the chain, the closer B gets to W and so the closer the ratio B/W gets to 1 and the PSRF estimate for v given by (13) goes to 1 also.

See [20] or [19] for further references on PSRF values.

5 Calibration Examples

Using raw data cited by other papers, we attempt to calibrate the local volatility model. We use a Markov Chain Monte-Carlo (MCMC) Metropolis algorithm to sample the posterior distribution of calibrated parameters.

5.1 Two Datasets

We first calibrate to prices generated using the local volatility surface used in [26]. Hence, it is a synthetic market where the true model is known and of the same class as the model we try to calibrate. Then we repeat the procedure using S&P 500 data taken from [12]. The datasets are as follows.

1. We price 66 European call options on the local volatility surface given in [26] (4.4) with 11 strikes and 6 maturities. Similar to [26], we take $S_0 = 5000$, interest rate $r = 0.05$, dividend rate $d = 0.03$. To each of the prices we add Gaussian noise with mean zero and standard deviation 0.1% of the original price [24] and treat these as the observed prices, similar to the approach in Coleman et al. [12]. We take the calibration error acceptance level to be $\delta = 3$ basis points following the results of [26]. The prices are given in Appendix A.

2. We take real S&P 500 implied volatility data with 10 strikes and 7 maturities used in [12] to determine the prices of 70 corresponding European call options. The spot price of the underlying at time 0 is $S_0 = 590$, the interest rate is $r = 0.060$ and dividend rate $d = 0.026$. The prices are given in Appendix A.

In both examples, the local volatility is discretised as explained in 4.1.

For the first example, we take nodes positioned on the grid given by

$$\begin{aligned} s &= 2500, 4000, 4500, 4750, 5000, 5250, 5500, 7000, 10000, \\ t &= 0.0, 0.5, 1.0, \end{aligned} \quad (14)$$

so $J = 9$, $L = 3$ and there are a total of $M = J \times L = 27$ free parameters (cf. 66 calibration prices).

For the second example, we take nodes positioned on the grid given by

$$\begin{aligned} s &= 300, 500, 560, 590, 620, 670, 800, 1200, \\ t &= 0.0, 0.5, 1.0, 2.0, \end{aligned}$$

so $J = 8$, $L = 4$ and there are a total of $M = J \times L = 32$ free parameters (cf. 70 calibration prices).

Clearly, discretisation provides another form of regularisation, and specifically the number of options used here is greater than the number of calibration parameters. However, tests in Section 6 will show that the methodology is robust with respect to an increase in model parameters, and the present choice is a compromise between resolution of the surface and computational tractability.

5.2 Calibration Results for a Simulated Dataset

Using the MCMC Metropolis algorithm, 16 chains of 10000 surfaces each were generated, which, after burn-in removal and thinning (see Appendix A), resulted in 90 surfaces per chain. Out of these 1440 surfaces, 479 were accepted as samples for the local volatility surface, i.e. gave $G(\theta) \leq \delta^2$ (cf. the 50 found by Hamida & Cont [24]). The chains were run in parallel using Matlab on four Dual-Core 2.8GHz AMDs with 32 GB RAM, where the computational time of a single chain on a single processor was about 5 minutes.

To check the convergence of Metropolis Sampling, we compute the PSRF numbers for the European call prices to which we are trying to calibrate. From Table 1 we see that almost all prices have PSRF value 1.0 which indicates excellent sampling. Only some short dated far out-of-the-money options have values slightly greater than 1.1; but these options had smaller weights w_i in our algorithm so we would expect slower convergence.

The resulting surfaces for prior confidence parameter $\lambda = 1$ are plotted in Figure 1. Figure 1 clearly demonstrates the variety of surfaces which can be calibrated to the same set of prices. We see that, especially for stock values far from S_0 the volatility becomes very uncertain and varied.

Using this distribution of surfaces we can construct confidence intervals (or ‘credible sets’) of the value of the local volatility surface $\sigma(S, t)$ at any point (S, t) . Figure 2 shows the 95% and 68% pointwise confidence intervals. Close to the spot price S_0 , the bounds are very tight, but the further away-from-the-money we go, the looser these bounds become. However, we see that the true

Maturity	Strike (units of S_0)								
	0.80	0.90	0.94	0.98	1.00	1.02	1.06	1.10	1.20
0.083	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.039	1.140
0.167	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.022	1.156
0.250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.012	1.160
0.500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.071
0.750	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.009
1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001

Table 1: For the simulated dataset: PSRF values for the calibration call prices (using [26]).

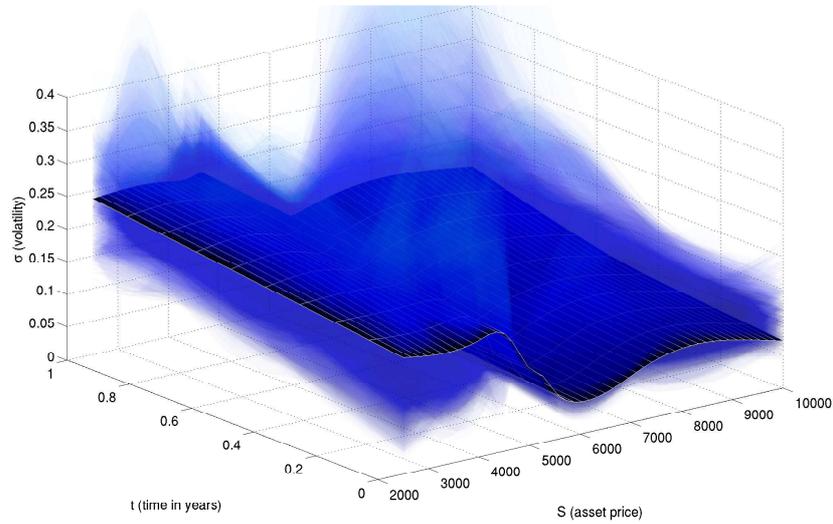


Figure 1: For the simulated dataset, using Metropolis Sampling, 479 surfaces from the posterior distribution were sampled and are plotted with the same degree of transparency. The true surface is plotted in opaque black.

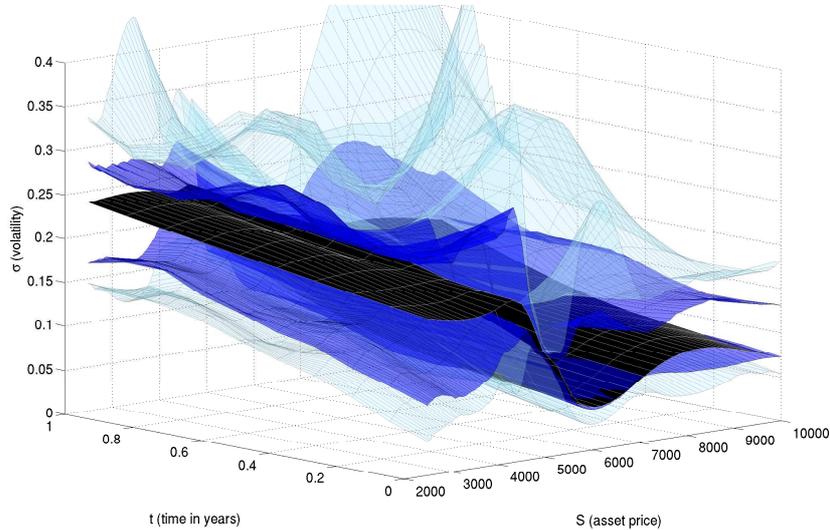


Figure 2: For the simulated dataset: using the 479 sampled surfaces (where $\lambda = 1$), 95% and 68% confidence intervals are found pointwise. The true surface is plotted in opaque black.

surface is almost captured within the 68% confidence interval and completely captured by the 95% confidence interval. This result could, for example, be used to find lower and upper volatility function bounds for implementation of the uncertain volatility model studied by He et al. [25]. Similarly, one could deduce confidence intervals for the integrated variance and use conservative hedging as proposed by Mykland [31]. Using only marginal distributions of the volatility values at specific points in time and asset values, however the dependence structure of the values on the surface, encoded in the posterior distribution, is lost.

5.3 Recalibration and Consistency

We can also look at verifying convergence of the posterior by checking that for an increasing number of calibration data the posterior concentrates around the true underlying model. To this end, we simulate a stock price path over 12 weeks on the assumed surface and price the same 66 European call options at the start of each week. We recalculate the error functional $G(\theta)$ and update the Bayesian posterior to incorporate the new observed prices every week. To update the Bayesian posterior we implement importance sampling by using the new error function to update the posterior density of the surfaces we have already sampled via the Metropolis algorithm. Instead of updating the Bayesian weights of the surfaces we found at time 0, we could resample the Bayesian posterior each week or each day using the Metropolis algorithm. Then for the first re-sampling, we are calibrating to $66 + 66 = 132$ prices and 198 prices on the second re-sampling which makes sampling much harder and takes longer to converge.

The results of the change in Bayesian posterior are illustrated in Figure 3. In

the Bayesian mean calculation (15) at the first calibration time, the ‘Bayesian weights’ are $y_0^{(i)} = 1/N$ for each surface θ_i for $i = 1, \dots, N$ — hence in Figure 1 all plotted surfaces have the same degree of transparency. However, after the first recalibration, the new Bayesian mean calculation for a function f will be

$$\sum_{i=1}^N y_1^{(i)} f(\theta_i)$$

for some Bayesian weights $y_1^{(1)}, \dots, y_1^{(N)}$ summing to 1. However, the weights are no longer equal. And to reflect this in Figure 3 we have varied the transparency of the plotted surfaces to reflect the weight. A surface with greater Bayesian weight will be more opaque.

Figure 3 shows that after about 5 weeks, the Bayesian posterior has settled and only a handful of surfaces have significant weight. Moreover, these surfaces are close to the true surface (plotted in opaque black). At recalibration time t_k , say, the section of the local volatility surface corresponding to $\{\sigma(S, t) : 0 \leq t < t_k\}$ no longer contributes to the observed calibration prices so the section $\{\sigma(S, t) : 0 \leq t < t_k\}$, for small t_k , is very different to the true surface for some of the heavily weighted surfaces. This is especially noticeable in the wings for very small S and very large S . We must remember that we have only sampled the Bayesian posterior and hence if none of our samples is the true surface (which it is not) then we will never settle on this true surface, but settle on the closest few, as Figure 3 shows. Nevertheless, using the proxy updating procedure, we still see a clear sense of convergence to the true surface.

5.4 Calibration to S&P 500 Dataset

For the second test case, we try to calibrate to 70 S&P 500 European call prices. We give the PSRF values in Table 2 below. Again we see that most options have PSRF value 1.0 indicating that the Metropolis sampling routine has been allowed to run for a sufficient time. Only some short dated far out-of-the-money calls have values just above 1.1 which is again attributable to their smaller weights w_i in the error function (7).

Maturity	Strike (units of S_0)									
	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.30	1.40
0.175	1.000	1.000	1.000	1.000	1.002	1.022	1.105	1.177	1.166	1.119
0.425	1.000	1.000	1.000	1.000	1.000	1.001	1.025	1.130	1.218	1.169
0.695	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.015	1.123	1.188
0.940	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.018	1.106
1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.011	1.088
1.500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.022
2.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.002

Table 2: For S&P 500 dataset: PSRF values for the calibration call prices (using [12]).

Figure 4 gives a plot of 600 samples from the posterior (so satisfied $G(\theta) \leq \delta^2$, this time for $\delta = 4.5$ basis points). Again we see that, especially in the wings and for short times, the spread of volatilities is enormous.

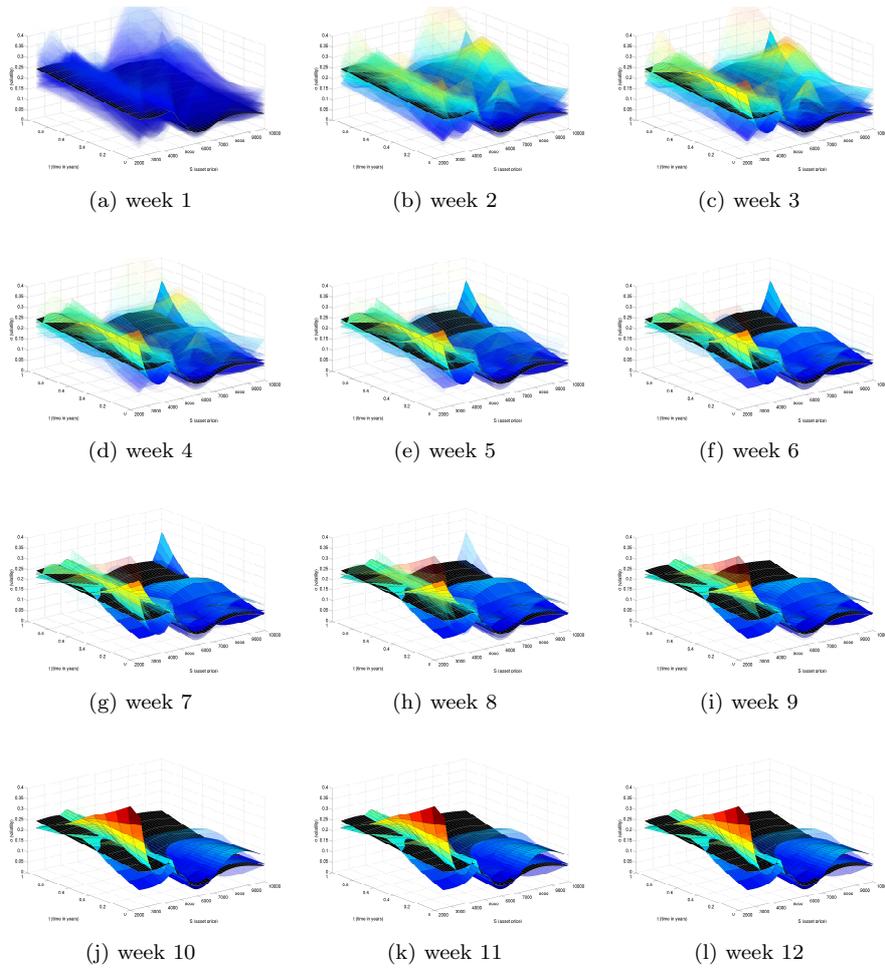


Figure 3: For the simulated dataset: a path is simulated on the true local volatility surface and the Bayesian posterior is updated using the newly observed prices each week for 12 weeks. The transparency of each surface reflects the Bayesian weight (see main text) of the surface. The true surface is plotted in opaque black.

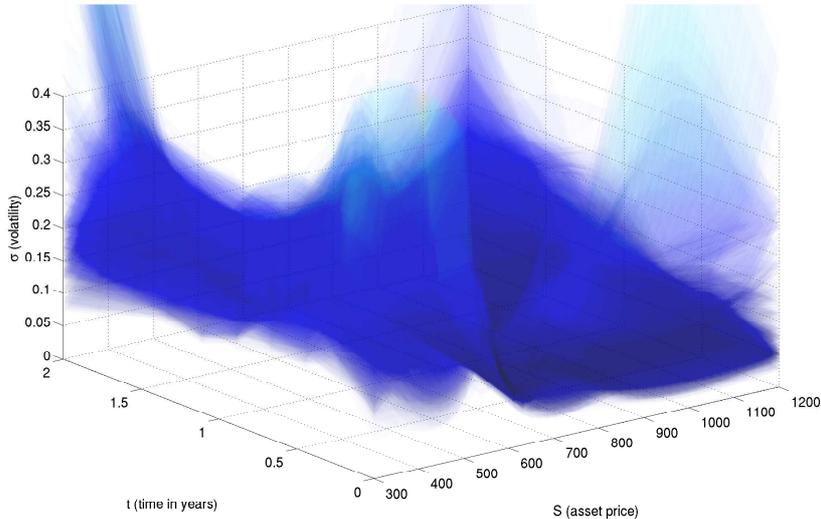


Figure 4: For S&P 500 dataset: using Metropolis Sampling, 600 surfaces from the posterior distribution were found and are plotted (where $\lambda = 1$).

6 Robust Pricing of Other Options

The posterior distribution of local volatility surfaces obtained in 5.2 and 5.3 can be used to price other contracts and assess the spread of possible prices.

6.1 Estimating Option Values

For a new, say, exotic option, each local volatility surface with positive posterior gives a possible value of this option. The Bayesian posterior distribution of local volatility surfaces therefore translates into a probability density of prices.

The Bayesian model average price of such a new contract is given by

$$\frac{1}{N} \sum_{i=1}^N f(\theta_i) \approx \int_{\theta} f(\theta) p(\theta|V) d\theta, \quad (15)$$

where f is the pricing function and $\theta_1, \dots, \theta_N$ are the surfaces found by Metropolis sampling. Note that, because the parameters $\theta_1, \dots, \theta_N$ are *samples*, the Bayesian weighting of each in the sum (15) is $1/N$ rather than $p(\theta_i|V)$.

We assess this price estimate, where possible, against the true value $f(\theta^*)$ as priced on the correct (assumed) surface θ^* , where the bid-ask spread is estimated by $[f(\theta^*) - \delta S_0/10^4, f(\theta^*) + \delta S_0/10^4]$.

A short remark on arbitrage is due here. If the market under the model with parameter θ^* is assumed complete, any price of the new contract different from the model price under θ^* could potentially be arbitrated (i.e. someone could make a risk-free profit by dynamically trading the underlying and a bond) by someone who knew the true parameter. Based on the information contained in the (noisy) calibration prices alone (bid-ask spreads), however, prices with positive posterior weight cannot normally be arbitrated.

The MAP price is taken to be $f(\theta_{MAP})$, where θ_{MAP} is the sample that has greatest posterior density, i.e. $p(\theta_{MAP}|V) \geq p(\theta_i|V)$ for all $i = 1, \dots, N$. Note that the MAP price does not correspond to the maximum density of $f(\theta_i)$, but that of θ_i . We noted earlier that the MAP parameter corresponds to the value calculated on the surface which gives the smallest regularised calibration error and is therefore identical to the classical Tikhonov solution.

The MAP estimator we use will not maximise the posterior density (10) precisely, due to the finite sample size. However, we observe for the example of Section 5.2 that our MAP estimator gives a weighted average basis point calibration error (7) of 1.84 for 66 prices, compared with the 2.65 for 10 prices achieved by Jackson et al. [26]. Hence, our density sampling has found a MAP surface which gives optimisation comparable to that found by papers using different optimisation routines.

In Figure 5 we highlight this pricing method on the example of an up-and-out barrier call option. The barrier option is path dependent and, as such, much more sensitive to changes in the local volatility surface. In the graph we plot the Bayesian posterior probability density of prices, true, MAP, Bayes prices and estimated bid-ask spread of the true price.

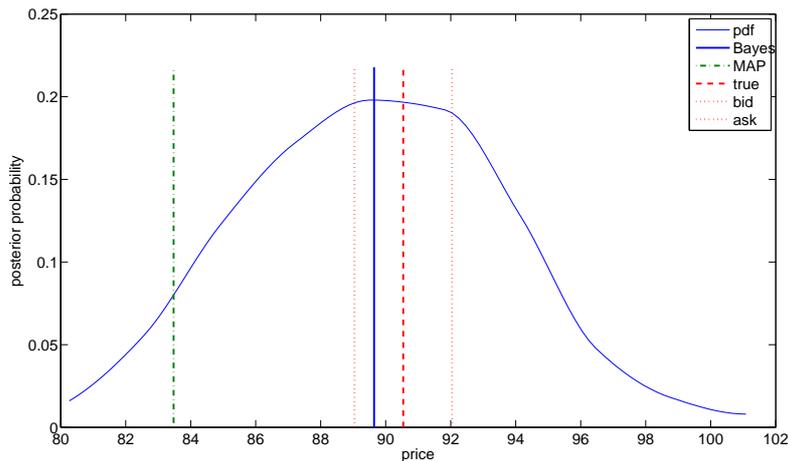


Figure 5: For simulated dataset: prices for up-and-out barrier call option strike 5000 ($S_0 = 5000$), barrier 5500 and maturity 3 months. Included are the true price (found on the true surface) with an assumed bid-ask spread of 6 basis points as per 5.1, the MAP/Tikhonov price, and the Bayes price with its associated posterior pdf of prices.

Figure 5 shows that the barrier price obtained with the surface with the smallest regularised calibration error for European options can lie many basis points away from the true price. The Bayesian price on the other hand reflects the entire distribution and the incorporated prior information (i.e. regularisation) to give a much closer price, which lies well within the bid-ask spread.

Recall that the Bayesian posterior we calculate depends on, besides the data and model used, the form of the (subjective) prior and the level of the data noise. In the following subsections, we conduct some robustness tests on the

datasets studied in the previous sections to quantify how much the solution varies with respect to model assumptions and changes in both the form of the prior and the observed prices.

6.2 Robustness with Respect to Discretisation

As discussed at the end of Section 2 and detailed in Section 4.1, we are computing an approximation to a two-dimensional function, and it is hoped that the choice of representation, e.g. the type of splines, number and placement of spline knots, does not influence inferences from this estimate significantly.

We first analyse the robustness with respect to the placement of the spline knots, by repeating the computation leading to Fig. 5, but for a re-arrangement of the spline knots from (14). Fig. 6 shows the result for one example, (a), where the spline knots are denser around the spot than in the original choice (14), and one where the knots are more evenly spread out, (b). Although the shape of the posterior distribution of barrier prices is to some extent affected by the spline knots, the standard deviation and most notably the Bayesian mean price are very robust. We conducted further tests with a variety of placements, including e.g. some where the knots were shifted by ± 100 and ± 300 , and the results were very similar.

Secondly, in Fig. 7, we investigate the robustness of estimators with respect to the number of spline knots. For three different sets of knots, we plot the estimated barrier option value for an increasing number of calibration options. The prices were generated by starting with noisy realisations of the model prices on the assumed surface, for 16 maturities and 64 strikes (total 1024 prices) and then reducing by a factor 2 (alternating) the maturities and strikes to get 512, 256, 128, 64, 32 data. The knots picked were 27 (3 times x 9 spatial, the original ones from earlier), 56 (4 times x 14 spatial) and 108 (6 times x 18 spatial).

The plot does not indicate that the estimators converge to the true value. This is explained by the fact that computations were carried out with a finite sample of calibrated surfaces for each set of knots, which in particular does not contain the true surface. As a consequence, the posterior settles around a small number of calibrated surfaces in the vicinity of the true surface, similar to what is seen in Fig. 3. These surfaces determine the estimated price. The main limitation at present is the computation time. As the dimension of the parameter increases with the number of knots, exploration of the parameter space requires longer and more MCMC chains and becomes computationally costly. The number of calibrated surfaces is further reduced as the number of prices to match is increased. So for 108 knots, 2000 calibrated surfaces were found for 32 calibration options, but only 15 for 512 options and 4 for 1024 (this compares, e.g., to 534 surfaces for 27 knots and 1024 options). For applications where calibration of a high-dimensional parameter to a very large number of prices is needed, further work is needed to improve the efficiency of the MCMC algorithm.

The positive result in terms of robustness is that the magnitude of the error depends only weakly on the number of spline knots for both estimators, while the Bayesian average gave more robust results than the MAP for all examples considered. We repeated the tests for different parameters (strikes, maturities, barriers), and the results were similar. Bayesian averaging appears to have an additional regularisation effect on top of the smoothing of surfaces dictated by

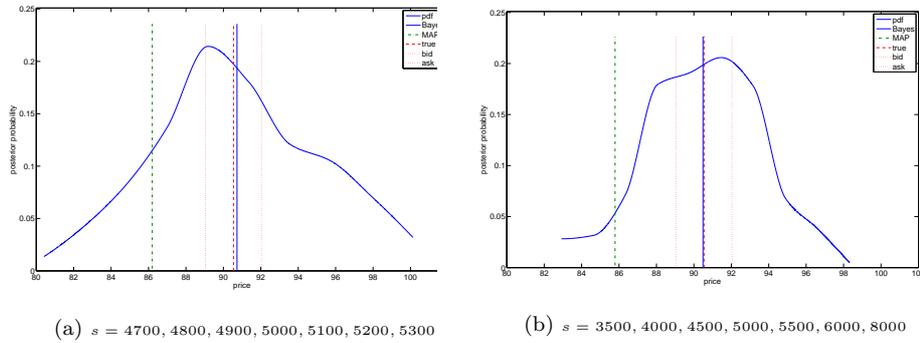


Figure 6: Same plots as Fig. 5, for different placement of the spline knots, with the lowest ($s = 2500$) and highest ($s = 10000$) as in (14), but with higher (a) and lower (b) density around the spot.

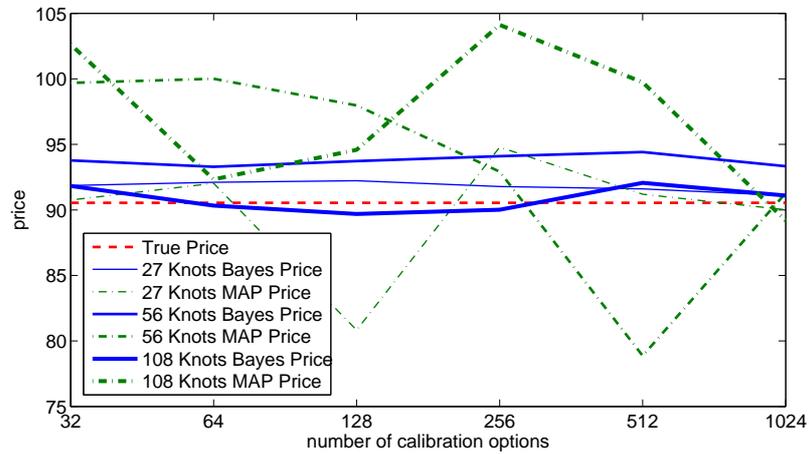


Figure 7: For the simulated data set: Prices for an up-and-out barrier call option with strike 5000 ($S_0 = 5000$), barrier 5500 and maturity 3 months for different number of calibration options and different number of spline knots.

the prior.

6.3 Robustness with Respect to the Prior

The choice of prior contains a degree of subjectivity, paralleling the arbitrariness of the regularising functional in Tikhonov regularisation. For the latter, the literature on local volatility estimation proposes various combinations of smoothness norms (see, e.g., [1, 2, 11, 14, 16, 26, 28]), and the choice of one over the other cannot be made (solely) based on objective market observations.

To test robustness to these assumptions, we adjust the form of the prior by changing the value of κ defined in (5). In each case, we hold all other parameters fixed, in particular the confidence parameter at $\lambda = 1$, and recalculate the distribution of prices for the same up-and-out barrier call option. In Figure 8 we plot the graph equivalent to Figure 5 of the previous section, but for different values of κ .

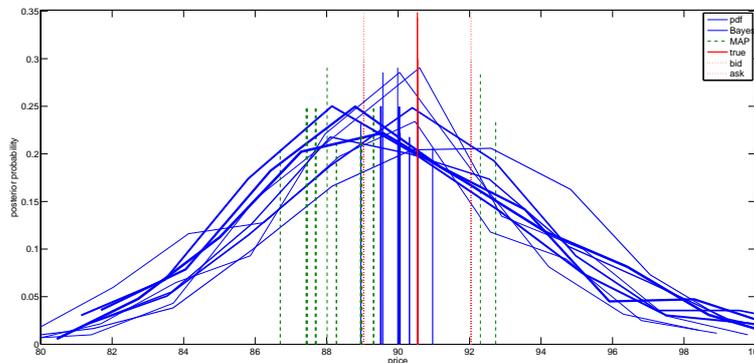


Figure 8: Prices for an up-and-out barrier call option with strike 5000 ($S_0 = 5000$), barrier 5500 and maturity 3 months for different parameters $\kappa \in \{10^{-2.00}, 10^{-1.75}, 10^{-1.50}, 10^{-1.25}, 10^{-1.00}, 10^{-0.75}, 10^{-0.50}, 10^{-0.25}, 10^{-0.1}\}$, corresponding to different priors. Included is the true price with its bid-ask spread, MAP prices, and Bayes prices with Bayesian posterior pdfs of prices. The thickness of the lines increases with κ , their length corresponds to the maximum of the corresponding posterior, so they can be optically identified.

Figure 8 shows that the distribution of prices is fairly robust to changes in κ since the peaks are roughly at the same prices and the tails exhibit similar behaviour. More encouraging even is the robustness of the Bayes estimate which lies in the bid-ask spread for all values of κ . This is in contrast to the MAP/Tikhonov estimate, which tends to fall outside of the bid-ask spread for several values of κ . These graphs provide satisfactory evidence for the robustness of Bayes price estimators to changes in the prior.

6.4 Robustness with Respect to the Noise Level

The regularising functional (prior) is designed – and shown, see e.g. [14, 16] – to give continuous dependence of the parameters on the data. To put this to the

test quantitatively, next we plot the graphs for the same experiments but for the case where we hold $\kappa = 10^{-1}$ fixed and instead add Gaussian noise e with different standard deviation ε to the observed market prices quoted in Appendix A. We consider noise levels of

$$\varepsilon \in \{10^{-4.0}, 10^{-3.5}, 10^{-3.0}, 10^{-2.5}\}$$

and run the calibration procedure for 100 independent noise additions for each ε . These levels compare to assumed bid-ask data of $\pm\delta = \pm 3 \cdot 10^{-4} \approx \pm 10^{-3.5}$ as per Section 5.1. For each value of ε we plot in Figure 9 the 100 MAP and Bayes estimates of the price and posteriors for the same barrier options previously used. First of all we see that as the noise is increased the closeness of the distributions of prices deteriorates and for $\varepsilon = 10^{-2.5}$ few surfaces have been calibrated so the distributions become non-smooth and irregular. The MAP prices are even more sensitive to the noise and can miscalculate the price by up to 10-15%. In contrast, the Bayes prices prove to be very robust. For the barrier option for $\varepsilon = 10^{-3.5}$ only one out of a hundred Bayes estimates is slightly outside the bid-ask spread and for $\varepsilon = 10^{-2.5}$ only a handful of Bayes estimates fall slightly below the bid price.

We also conducted tests varying the weights of calibration options in the error functional, and again the Bayes average was very insensitive to these changes while the MAP estimator moved by an amount equalling several bid-ask spreads.

6.5 Pricing an American Option on the S&P 500

We conducted a similar robustness test for the second test case of real market data, and report here a summary of the results.

Figure 10 shows the distribution of prices for an American put option written on the S&P 500 corresponding to the Bayesian posterior distribution for different numbers of spline knots. Again the MAP and Bayes prices are marked on the graph. However, in this case, we have used real data so we do not know the real local volatility surface (or even if one exists) so cannot compare with true results. The Bayes prices are (by construction) still closer to the centre of the distribution than the MAP prices, but the spread of prices is similar between the two estimates.

This impression is reinforced by Fig. 11, the equivalent of Fig. 7 in the previous example, which shows the estimates of American option prices over the number of calibration options, for different spline configurations. The Bayes average price shows a similar variation to the previous example. Different from before, the MAP has now also a similar variation. A possible interpretation of this is that the American option value depends on similar properties of the volatility surface as the European (calibration) options, and therefore the MAP estimates are more benign than in the earlier barrier option example, which exhibits stronger tail dependence.

7 Conclusions and Extensions

In this paper we introduced the Bayesian framework for calibrating the parameters of financial models to market prices. We then gave a practical method

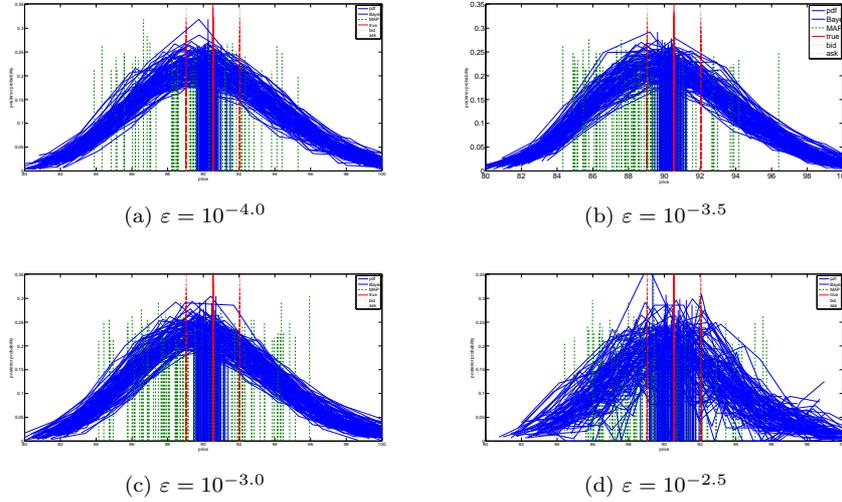


Figure 9: Prices for up-and-out barrier call option with strike 5000 ($S_0 = 5000$), barrier 5500 and maturity 3 months. Each graph corresponds to a different value of ε and shows the estimators for 100 different noise additions. Included is the true price (found on the true surface) with its bid-ask spread, the MAP price, and the Bayes price with its associated posterior pdf of prices.

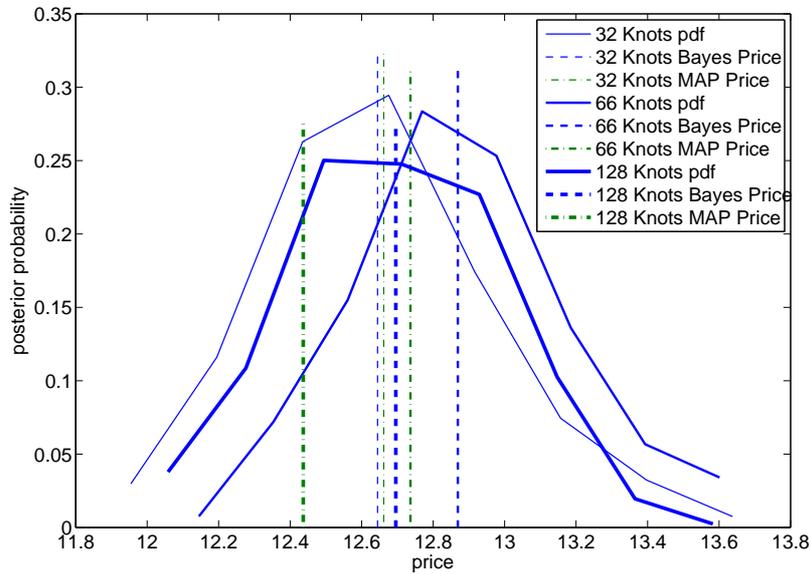


Figure 10: For S&P 500 dataset: prices for American put option with strike \$590 ($S_0 = \590) and maturity 1 year. Included are the MAP prices and the Bayes average prices with the associated posterior pdfs of prices, for three different spline bases for the local volatility with 32, 66 and 128 knots, respectively.

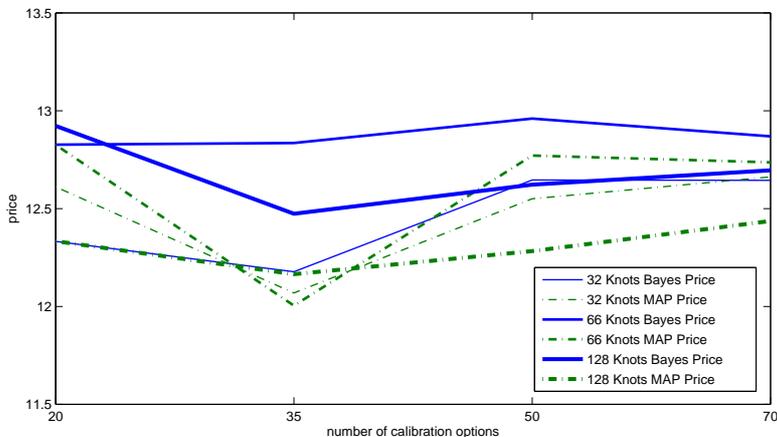


Figure 11: For S&P 500 dataset: prices for American put option with strike \$590 ($S_0 = \590) and maturity 1 year for different number of calibration options and different number of spline knots.

for formulating the prior and likelihood functions necessary for the Bayes procedure and applied it to the case of the local volatility model. Numerical examples were presented and showed the improvement in pricing of the Bayesian procedure over common maximum a posteriori methods. Moreover, we highlighted the robustness of the pricing method to inaccuracies in the model and prior, and mispricings in observed market data.

The Bayes average is more robust than the MAP in the examples considered, often significantly so. This comes at the expense of high computational cost. While the MAP estimator is equivalent to a Tikhonov regularised solution and could be found by any of a number of efficient deterministic optimisation algorithms (see, e.g., [12]), the Bayes average requires information of the whole high-dimensional posterior distribution.

The flip side of this is that the Bayesian posterior density $p(\theta|V)$ can be used as the basis for a variety of further useful analysis. A natural thing to do would be to use the posterior to derive a measure for the model uncertainty of a contract. For any payoff, a distribution of prices can be found (as we showed in Section 6 for American and barrier options), and this distribution can be used to assign a model uncertainty value to the contract in the spirit of [13] and [23]. Such measures would be important for a risk manager and for an agent trying to decide between different products.

A second, perhaps more important, use of the Bayesian posterior would be to use it to develop better hedging strategies. This is more fundamental than pricing as typically a trader will be more interested in the hedging strategy (the price of which will then correspond to the trader's price for the contract) than a stand-alone price. The technique described in this article gives accurate and robust prices, but it is not immediately clear which model (parameter) should be used for hedging. One possible way to use the posterior density for hedging would be to deduce prediction sets for the spot volatility or integrated variance, and then hedge conservatively within these sets. This corresponds to the

approaches proposed in [5] and [31] respectively. In particular, Mykland finds conservative bid and ask prices as the superreplication cost under the assumption that the prediction set is realised. In an alternative approach, motivated by the analysis of this paper, the Bayesian loss functions introduced in Section 2 could be designed to correspond to hedging losses so that the Bayes estimator is that parameter θ which minimises the expected hedging loss.

References

- [1] Y. Achdou. An inverse problem for parabolic variational inequalities in the calibration of American options. *SIAM Journal of Control*, 43(5):1583–1615, 2005.
- [2] Y. Achdou and O. Pironneau. Numerical procedure for calibration of volatility with American options. *Applied Mathematical Finance*, 12(3):201–241, 2005.
- [3] A. Apte, M. Hairer, A.M. Stuart, and J. Voss. Sampling the posterior: An approach to non-Gaussian data assimilation. *Physica D: Nonlinear Phenomena*, 230(1-2):50–64, 2007.
- [4] M. Avellaneda, C. Friedman, R. Holmes, and D. Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Applied Mathematical Finance*, 4:37–64, 1997.
- [5] M. Avellaneda, A. Lévy, and A. Paras. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2:73–88, 1995.
- [6] M. Avellaneda and A. Paras. Managing the volatility risk of portfolios of derivative securities: the Lagrangian uncertain volatility model. *Applied Mathematical Finance*, 3(1):21–52, 1996.
- [7] H. Berestycki, J. Busca, and I. Florent. Asymptotics and calibration of local volatility models. *Quantitative Finance*, 2(1):61–69, 2002.
- [8] A. Beskos and A. Stuart. MCMC methods for sampling function space. In *Sixth International Congress on Industrial and Applied Mathematics: Zurich, Switzerland, July 16-20, 2007*, page 337. European Mathematical Society, 2009.
- [9] R. Bhar, C. Chiarella, H. Hung, and W.J. Runggaldier. The volatility of the instantaneous spot interest rate implied by arbitrage pricing: A dynamic Bayesian approach. *Automatica*, 42(8):1381–1393, 2006.
- [10] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637, 1973.
- [11] C. Chiarella, M. Craddock, and N. El-Hassan. The calibration of stock option pricing models using inverse problem methodology. *QFRQ Research Papers, UTS Sydney*, 2000.

- [12] T.F. Coleman, Y. Li, and A. Verma. Reconstructing the unknown local volatility function. In *Quantitative analysis in financial markets: collected papers of the New York University Mathematical Finance Seminar*, page 192, 2001.
- [13] R. Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3):519–547, 2006.
- [14] S. Crepey. Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization. *SIAM Journal on Mathematical Analysis*, 34(5):1183–1206, 2003.
- [15] B. Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.
- [16] H. Egger and H.W. Engl. Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Problems*, 21(3):1027–1045, 2005.
- [17] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*. Kluwer Academic Pub, 1996.
- [18] B.G. Fitzpatrick. Bayesian analysis in inverse problems. *Inverse Problems*, 7(5):675–702, 1991.
- [19] A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, 2nd edition, 2004.
- [20] A. Gelman and D.B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, pages 457–472, 1992.
- [21] S. Ghosal. A review of consistency and convergence rates of posterior distributions. In *Proc. Varanasi Symp. on Bayesian Inference*, 1998.
- [22] S. Ghosal and A. Van Der Vaart. Convergence rates of posterior distributions for noniid observations. *Annals of Statistics*, 35(1):192, 2007.
- [23] A. Gupta, C. Reisinger, and A. Whitely. Model uncertainty and its impact on derivative pricing. In K. Böcker, editor, *Rethinking Risk Management and Reporting: Uncertainty, Bayesian Analysis and Expert Judgment*. Risk Books, 2010.
- [24] S.B. Hamida and R. Cont. Recovering volatility from option prices by evolutionary optimization. *Journal of Computational Finance*, 8(4), 2005.
- [25] C. He, T.F. Coleman, and Y. Li. Calibrating volatility function bounds for an uncertain volatility model. *Journal of Computational Finance*, 13(4), 2010.
- [26] N. Jackson, E. Süli, and S. Howison. Computation of deterministic volatility surfaces. *Journal of Computational Finance*, 2(2):5–32, 1999.
- [27] E. Jacquier and R. Jarrow. Bayesian analysis of contingent claim model error. *Journal of Econometrics*, 94(1–2):145–180, 2000.

- [28] R. Lagnado and S. Osher. A technique for calibrating derivative security pricing models: numerical solution of an inverse problem. *Journal of Computational Finance*, 1(1):13–25, 1997.
- [29] T.J. Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance*, 2:117–133, 1995.
- [30] M. Monoyios. Optimal hedging and parameter uncertainty. *IMA Journal of Management Mathematics*, 18(4):331, 2007.
- [31] P.A. Mykland. Financial options and statistical prediction intervals. *Ann. Statist.*, 31:1413–1438, 2003.
- [32] D. Samperi. Calibrating a diffusion pricing model with uncertain volatility: regularization and stability. *Mathematical Finance*, 12(1):71–87, 2002.
- [33] L. Wasserman. Asymptotic properties of nonparametric Bayesian procedures. *Practical nonparametric and semiparametric Bayesian statistics*, pages 293–304, 1998.

A Constants & Datasets

Constant	Description	Test 1	Test 2
S_0	time 0 asset price	5000	590
r	rate of interest	0.05	0.06
d	dividend yield	0.03	0.0262
σ_{atm}	time 0 ATM volatility	0.15	0.14
I	# of calibrating options	66	70
δ	calibration tolerance (b.p.)	3	4.5
M	# of nodes ($J \times L$)	27	32
λ_p	strength of prior (for sampling)	1	1
κ	Sobolev norm (LV model)	0.1	0.1
n	# of iterations in each MCMC chain	25000	25000
m	# of MCMC Metropolis chains	16	16
b	length of burn-in	1000	1000
k	frequency of thinning	100	100
du	jump function step size	0.00015	0.00013

Table 3: Numerical examples constants for calibration process.

Maturity	Strike (units of S_0)										
	0.80	0.90	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.10	1.20
0.083	1.003	0.507	0.312	0.219	0.136	0.070	0.030	0.010	0.003	0.000	0.000
0.167	1.010	0.518	0.332	0.246	0.168	0.104	0.058	0.029	0.012	0.001	0.000
0.250	1.012	0.531	0.352	0.270	0.196	0.132	0.083	0.048	0.025	0.004	0.000
0.500	1.029	0.577	0.414	0.337	0.265	0.200	0.146	0.102	0.068	0.024	0.000
0.750	1.052	0.623	0.469	0.396	0.327	0.264	0.208	0.160	0.119	0.059	0.004
1.000	1.079	0.671	0.525	0.457	0.390	0.329	0.274	0.224	0.180	0.110	0.021

Table 4: For the simulated dataset: European call prices (units of 10^3) (using [26]).

Maturity	Strike (units of S_0)										
	0.80	0.90	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.10	1.20
0.083	0.002	0.009	0.012	0.014	0.015	0.017	0.015	0.014	0.012	0.008	0.002
0.167	0.003	0.010	0.014	0.015	0.017	0.019	0.017	0.015	0.014	0.010	0.003
0.250	0.005	0.012	0.015	0.017	0.019	0.021	0.019	0.017	0.015	0.012	0.005
0.500	0.007	0.014	0.017	0.019	0.021	0.022	0.021	0.019	0.017	0.014	0.007
0.750	0.009	0.015	0.019	0.021	0.022	0.023	0.022	0.021	0.019	0.015	0.009
1.000	0.010	0.017	0.021	0.022	0.024	0.026	0.024	0.022	0.021	0.017	0.010

Table 5: For the simulated dataset: Weights for calibration options.

Maturity	Strike (units of S_0)									
	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.30	1.40
0.175	91.3	62.8	35.2	12.9	2.1	0.1	0.0	0.0	0.0	0.0
0.425	96.3	69.0	44.0	23.3	8.5	2.3	0.4	0.2	0.0	0.0
0.695	101.8	76.1	52.6	32.6	16.4	5.9	1.9	0.6	0.1	0.0
0.940	106.8	82.2	59.9	39.9	23.8	11.3	4.7	1.8	0.2	0.0
1.000	108.0	83.6	61.6	41.6	25.4	12.8	5.5	2.1	0.2	0.1
1.500	117.2	94.4	73.1	54.0	37.3	23.7	14.3	7.7	1.9	0.3
2.000	125.7	104.0	83.6	64.9	48.2	34.2	23.6	14.7	5.6	1.8

Table 6: For the S&P 500 dataset: European call prices (\$) (using [12]).