Calibration of the Heston Stochastic Local Volatility Model: A Finite Volume Scheme[1]

Bernd Engelmann[2], Frank Koster[3], Daniel Oeltz[4]

This version: February 2, 2012

ABSTRACT

The two most popular equity derivatives pricing models among practitioners are the local volatility model and the Heston model. While the former has the appealing property that it can be calibrated exactly to any given set of arbitrage free European vanilla option prices, the latter delivers a more realistic smile dynamics. In this article we combine both modeling approaches to the Heston stochastic local volatility model. The theoretical framework for this modeling approach was already developed in Ren, Madan, and Qian (2007). We focus on the numerical model calibration which requires special care in the treatment of mixed derivatives and in cases where the Feller condition is not met in the Heston model which leads to a singular transition density at zero volatility. We propose a finite volume scheme to calibrate the model after a suitable transformation of the model equation and demonstrate its accuracy in numerical test cases using real market data.

JEL Classification: G13
Keywords: Heston Model, Local Volatility Model, Finite Volume Scheme

[1]The opinions expressed in this article are those of the authors and do not necessarily reflect views of Macquarie Capital (Europe) Ltd.

[2]Quantsolutions, Gaertnerweg 7, D-60322 Frankfurt, e-mail: bernd.engelmann@quantsolutions.de

[3]e-mail: frank.koster@hedgeit.de

[4]Senior Manager, Macquarie Capital (Europe) Ltd, Untermainanlage 1, D-60329 Frankfurt, e-mail: daniel.oeltz@hedgeit.de

On the equity derivatives trading desks of investment banks the two most widely applied models for pricing and risk management of exotic products are the local volatility model, developed by Dupire (1994) and Derman and Kani (1994), and the Heston model that was derived in Heston (1993). The local volatility model has the appealing property that it can be calibrated to any arbitrage free set of European vanilla options with arbitrary precision. However, it was criticized by many authors for the smile dynamics it implies. Rebonato (1999) outlines that the predictions of future smile surfaces by the local volatility model are almost flat and therefore in contradiction with empirical observations. Hagan, Kumar, Lesniewski, and Woodward (2002) show by applying perturbation theory that the smile dynamics of the local volatility model is counterintuitive as it implies that the volatility smile shifts in the opposite direction of the stock price. They argue that for this reason the local volatility model is not suitable for pricing and risk managing exotic option books. In contrast to these theoretical observations, Engelmann, Fengler, and Schwendner (2009) show in a recent empirical analysis that the hedging performance of the local volatility model for reverse barrier options is satisfactory. Therefore, from a practical perspective, the model seems to work at least for the simpler types of exotic products.

Although rigorous empirical analyses are to our knowledge not yet published, there is a consensus among market participants that the satisfactory hedging performance of the local volatility model for barrier products does not carry over to products that are more sensitive to the forward volatility skew. Examples of forward skew sensitive products are forward starting options and cliquet options. In these cases a stochastic volatility model is used for pricing and risk management instead of local volatility. Several extensions of the Black-Scholes model by making the volatility stochastic have been developed by numerous authors. The most prominent examples are the models of Hull and White (1987), Heston (1993), Schöbel and Zhu (1999), and Stein and Stein (1991). As already mentioned, the most popular stochastic volatility model applied in practice is the Heston model. In comparison to the local volatility model, the Heston model delivers much more realistic descriptions of future smile dynamics but does not have the ability to calibrate exactly to any arbitrary set of arbitrage free prices of European vanilla options. Especially for options with short maturities pricing errors are typically large for the Heston model when it is applied in equity markets.

One possibility to improve the calibration for short maturities is by adding jumps to the model. This was done, for instance, by Bates (1996). Increasing the number of model parameters improves the calibration of the model to the overall volatility surface but makes it more difficult to keep the calibration stable. Furthermore, the pricing of exotic products becomes more involved by the introduction of jumps since then a partial integro-differential equation has to be solved instead of a partial differential equation. Another possibility of improving the model calibration is making one or more parameters of the Heston model time-dependent.

In this work, we consider an alternative approach by combining the Heston model with the local volatility model which results in the Heston stochastic local volatility model. Here the dynamics of the Heston model is adjusted by a local volatility factor which ensures that the

model can be calibrated with arbitrary precision to a given set of arbitrage free prices of European vanilla options. The resulting model contains both the Heston and the local volatility model as special cases. If the Heston model would calibrate perfectly to a given set of European vanilla option prices, the local volatility parameter would be equal to one and the model would reduce to the pure Heston model. On the other hand, if the stochastic component of the volatility in the Heston model is set to zero, the model would degenerate to a pure local volatility model.

The theory of this model class was already developed in an article by Ren, Madan, and Qian (2007). However, the numerical algorithm for model calibration that was presented in this article will not work in the case of combining the local volatility model with the Heston model. First, Ren, Madan, and Qian (2007) consider the case of uncorrelated processes for the stock and the volatility evolution only. Furthermore, for certain parameter combinations in the Heston model it is possible that the volatility process reaches zero with positive probability. In this case, the algorithm presented in Ren, Madan, and Qian (2007) would collapse because it is not clear how to define a reasonable boundary condition at volatility = 0 when solving the partial differential equation for the transition probability density.

In a recent article Henry-Labordére (2009) shows how to calibrate general multi-factor stochastic volatility models that are extended by a local volatility factor. The algorithm he proposes is based on a Monte-Carlo simulation approach. This has the advantage that the resulting calibration routine is very generic but the disadvantage that the routine is slow compared to a grid based method from a computational point of view. For this reason we prefer a solution that is more efficient for our specific choice of stochastic volatility model, the Heston model, but less generic.

In this article we present a finite volume scheme to calibrate the Heston stochastic local volatility model. We assume that the parameters of the Heston process have already been determined. This could be done either by calibrating the Heston model to a set of European vanilla option prices or even by determining one or more parameters of the Heston model from historical data using time series analysis possibly modified to reflect a trader's market view. We show how to compute the correcting local volatility factor to ensure a perfect calibration to a given set of arbitrage free European vanilla option prices using a finite volume scheme.

This article is organized as follows. In the next section, we will shortly explain the theory behind the Heston stochastic local volatility model along the lines of Ren, Madan, and Qian (2007) to keep the article self-contained. In Section 2 we will propose a finite volume scheme to solve the partial differential equation necessary for calibrating the model. In Section 3, we will demonstrate the accuracy of our numerical scheme for some test cases using real market data. The final section concludes.

1. The Heston Stochastic Local Volatility Model

We start with the classical Heston (1993) model. The dynamics of this two-factor model is given by

$$dS_t = (r_t - d_t)S_t dt + \sqrt{V_t} S_t dW_t, \tag{1}$$
$$dV_t = \lambda(\theta - V_t) + \eta\sqrt{V_t} dZ_t, \tag{2}$$

where S_t is the stock price process starting in S_0 at time $t = 0$, V_t the process of the stock price's instantaneous variance starting in V_0, r_t is the riskless interest rate, d_t the dividend yield, λ is the speed of mean reversion, θ the long run variance, η the volatility of variance and W_t and Z_t are correlated Wiener processes with correlation coefficient ρ. The dynamics of the local volatility model is given by

$$dS_t = (r_t - d_t)S_t dt + \sigma_{LV}(t, S_t) S_t dW_t, \tag{3}$$

where $\sigma_{LV}(t, S_t)$ is the local volatility function. All other notation has the same meaning as in (1).

The Heston stochastic local volatility model is constructed by introducing a local volatility factor $\sigma_H(t, S_t)$ in (1). This results in the model dynamics

$$dS_t = (r_t - d_t)S_t dt + \sigma_H(t, S_t)\sqrt{V_t} S_t dW_t, \tag{4}$$
$$dV_t = \lambda(\theta - V_t) + \eta\sqrt{V_t} dZ_t. \tag{5}$$

Both local volatility functions $\sigma_{LV}(t, S_t)$ and $\sigma_H(t, S_t)$ are computed in a way to ensure that the resulting model matches a given set of European vanilla option prices exactly. While this is quite easy for the local volatility model, it is more involved for the Heston stochastic local volatility model.

Ren, Madan, and Qian (2007) use a result of Gyöngy (1986) to relate the local volatility functions $\sigma_{LV}(t, S_t)$ and $\sigma_H(t, S_t)$. The marginal distributions of both the local volatility model and the Heston stochastic local volatility model coincide if the condition

$$\sigma_{LV}^2(t, s) = \sigma_H^2(t, s) E[V_t | S_t = s] \tag{6}$$

is fulfilled. Since the local volatility $\sigma_{LV}(t, S_t)$ can be computed independently by well-established calibration algorithms (see, e.g., Coleman, Li, and Verma (1999) or Crépey (2003)), the function $\sigma_H(t, S_t)$ can be computed from $E[V_t | S_t = s]$ using (6). The conditional expectation $E[V_t | S_t = s]$ can be computed as

$$E[V_t | S_t = s] = \frac{\int_0^\infty v p(t, \log(s), v) dv}{\int_0^\infty p(t, \log(s), v) dv}, \tag{7}$$

where p is the density function of the transition probability distribution of $(\log(S_t), V_t)$. This distribution gives the probability that $(\log(S_t), V_t)$ reaches the value $(\log(s), v)$ at time t when starting in $(\log(S_0), V_0)$ at time $t = 0$. The density function $p(t, \log(s), v)$ can be computed by solving the Kolmogorov forward equation

$$\begin{aligned}\partial_t p = &-\partial_x \left(\left(r - d - \frac{1}{2}\sigma_H^2(t, e^x)v \right) p \right) - \partial_v \left(\lambda(\theta - v)p \right) \\ &+ \frac{1}{2}\partial_{xx}\left(\sigma_H^2(t, e^x)vp \right) + \frac{1}{2}\eta^2 \partial_{vv}(vp) + \eta\rho\partial_{xv}\left(\sigma_H(e^x, t)vp \right),\end{aligned} \tag{8}$$

where $x = \log(s)$. Partial derivatives are abbreviated by indices of the differential operators and the arguments of p have been omitted to keep the notation simple. Because σ_H and p are coupled by (6) and (7) this is a non-linear convection diffusion equation. The initial condition for (8) is

$$p(0, \log(s), v) = \delta(\log(s) - log(S_0))\delta(v - V_0) \tag{9}$$

where δ denotes the Dirac delta function. Since p is a density we must have

$$\int_{\mathbf{R}\times\mathbf{R}_+} p(x, v, t)dxdv = 1, \ \forall t \geq 0. \tag{10}$$

This is essential for a density and motivates the use of a numerical scheme which preserves this property. For a numerical solution we have to restrict the infinite domain to a finite domain $\Omega = [-x_{low}, x_{high}] \times [0, v_{high}]$, where x_{low}, x_{high} and v_{high} are positive constants, and suitable boundary conditions must be imposed on $\partial\Omega$. Along the boundaries in $x = x_{low}, x_{high}$ and $v = v_{high}$ we have a number of options as p is very small there. For example, we might use $p = 0$ or some extrapolation conditions. In any case we must be aware that ignoring p outside Ω affects the computation of the conditional expectation in (7) which means that the domain must not be too small even if we use a boundary condition which does not much affect the interior of Ω. More complicated is the boundary condition for $v = 0$. We know that the marginal distribution $P(v,t) := \int_{\mathbf{R}} p(x, v, t)dx$ of the variance process is noncentral chi-square (v. Haastrecht and Pelsser 2008)

$$P(v,t) = \frac{1}{C_0} e^{-(z+\mu)/2} \sum_{i=0}^{\infty} \frac{(\mu/2)^i z^{d/2+i-1}}{i! 2^{d/2+i} \Gamma(d/2 + i)} \tag{11}$$

with parameters $C_0 := \eta^2(1 - e^{\lambda t})/4\lambda$, $d := 4\lambda\theta/\eta^2$, $\mu := 4\lambda e^{\lambda t} V_0/\eta^2(1 - e^{\lambda t})$, $z = v/C_0$. From this we obtain

$$P(v,t) \sim v^{2\lambda\theta/\eta^2 - 1}$$

for small v. For non-degenerate cases with non-vanishing diffusion in x-direction $p(x, v, t)$ has the same asymptotics for small v

$$p(x, v, t) \sim v^{2\lambda\theta/\eta^2 - 1} \ . \tag{12}$$

From (12) we can infer the well-known Feller condition that states if $\lambda\theta - \frac{1}{2}\eta^2 > 0$ then the variance process never can reach zero. In this case $p = 0$ would be a suitable boundary condition. However, if the Feller condition is not satisfied which is often the case in practical applications we need something else. Since (10) must hold for all t no probability can disappear from the domain Ω nor can the total probability in the domain become larger. From this condition we will derive a boundary condition on $v = 0$ that ensures that (10) is fulfilled at all times.

Having described the governing equations of the Heston stochastic local volatility model and the behavior of the solution at the boundaries of the computational domain, we present the algorithmic details of calculating the functions $\sigma_H(t,S_t)$ and $p(t,\log(S_t),V_t)$ in the next section. We propose a finite volume scheme for this calculation mainly because it is best suited in dealing with the problems in $v = 0$ when the Feller condition is violated and it automatically ensures that (10) is fulfilled.

2. A Finite Volume Scheme for Model Calibration

Before going into numerical details we outline the basic strategy of decoupling (6) and (8) for the simultaneous computation of σ_H and p. We solve (8) on a discrete time grid $0 = t_0, t_1, \ldots, t_l$ starting in $t = 0$. In $t = 0$ we can compute $\sigma_H(0,S_0)$ from (7) because we know the transition density in $t = 0$ from (9). We find

$$\sigma_H^2(0,s) = \sigma_{LV}^2(0,s)/V_0 \; .$$

From this we can compute p in t_1 solving (8). Suppose we have solved (8) until $t = t_i$. From the solution, we can compute the conditional expectation $E[V_{t_i}|S_{t_i} = s]$ in t_i using (7). This result is used to compute $\sigma_H(t_i,s)$ from (6). The function $\sigma_H(t_i,s)$ is then used to compute p in t_{i+1} by solving the Kolmogorov forward equation (8) which finishes one iteration.

The most complicated part in one iteration is the solution of the Kolmogorov forward equation. Since we will adopt techniques from fluid dynamics we will rewrite (8) into an equivalent formulation which is more suitable for the application of these techniques. This leads to

$$\partial_t p + \nabla \cdot (\mathbf{b}p \; + \; \mathbf{A} \cdot \nabla p) = 0 \; . \tag{13}$$

where

$$\mathbf{b} = \begin{pmatrix} (r - d - \frac{1}{2}\eta\rho\sigma_H(e^x,t)) \; - \; \frac{1}{2}v(\sigma_H^2(e^x,t) + \partial_x\sigma_H^2(e^x,t)) \\ \lambda\theta - \frac{1}{2}\eta^2 \; - \; v(\lambda + \frac{1}{2}\eta\rho\partial_x\sigma_H(e^x,t)) \end{pmatrix} ,$$

$$\mathbf{A} = -\frac{1}{2}v \begin{pmatrix} \sigma_H^2(e^x,t) & \eta\rho\sigma_H(e^x,t) \\ \eta\rho\sigma_H(e^x,t) & \eta^2 \end{pmatrix} ,$$

and $\nabla = (\partial_x, \partial_v)$. Depending on its arguments, the "$\cdot$" operator in (13) denotes the scalar product of two vectors or matrix-vector multiplication, respectively.

2.1. Transformation of the Forward PDE

In order to reduce the mixed derivatives in (13) we introduce a new coordinate system (y,v) by applying the coordinate transform $(y,v) := (x+\alpha v, v)$ for some constant $\alpha \in \mathbb{R}$. For a function $f(x,v)$ we define the transformed function $\hat{f}(y,v)$ as

$$\hat{f}(y,v) = f(x,v) = \hat{f}(x+\alpha v, v) \ .$$

The forward PDE (13) transforms into a PDE for $\hat{p}$

$$\partial_t \hat{p} + \nabla \cdot (\hat{\mathbf{b}}\hat{p} + \hat{\mathbf{A}} \cdot \nabla \hat{p}) = 0 \tag{14}$$

where

$$\hat{\mathbf{b}} = \begin{pmatrix} r - d - \frac{1}{2}\hat{\sigma}_H^2 v + \alpha\lambda(\theta - v) - \frac{1}{2}v\partial_y\hat{\sigma}_H^2 - \frac{1}{2}\eta\rho\hat{\sigma}_H - \frac{1}{2}\eta^2\alpha - \frac{1}{2}\eta\rho\alpha v\partial_y\hat{\sigma}_H \\ \lambda(\theta - v) - \frac{1}{2}\eta^2 - \frac{1}{2}\eta\rho v\partial_y\hat{\sigma}_H \end{pmatrix} =: \begin{pmatrix} \hat{b}_y \\ \hat{b}_v \end{pmatrix} ,$$

$$\hat{\mathbf{A}} = -\frac{1}{2}v \begin{pmatrix} \hat{\sigma}_H^2 + 2\eta\rho\alpha\hat{\sigma}_H + \eta^2\alpha^2 & \eta\rho\hat{\sigma}_H + \eta^2\alpha \\ \eta^2\alpha + \eta\rho\hat{\sigma}_H & \eta^2 \end{pmatrix} =: \begin{pmatrix} \hat{A}_{yy} & \hat{A}_{yv} \\ \hat{A}_{vy} & \hat{A}_{vv} \end{pmatrix} ,$$

and $\nabla = (\partial_y, \partial_v)$ in this context. Note that $\hat{\sigma}_H(t,y,v)$ is a function of the new coordinates (y,v) and is obtained from σ_H by

$$\hat{\sigma}_H(t,y,v) = \sigma_H(t, e^{y-\alpha v}) \ .$$

The initial condition (9) transforms to

$$\hat{p}(y,v,0) = \delta(y - \log(S_0) - \alpha V_0)\delta(v - V_0) \ .$$

The aim of the coordinate transformation was the elimination of the mixed derivatives. This would be the case for a choice of α solving the equation

$$\eta^2\alpha + \eta\rho\hat{\sigma}_H(t,y,v) = 0 \ .$$

For a constant α this is only possible if $\hat{\sigma}_H$ is also a constant, i.e. for the pure Heston case. Yet, if we choose

$$\alpha = -\frac{\rho}{\eta}\hat{\sigma}_H(0, \log(S_0) + \alpha v_0, v_0) = -\frac{\rho}{\eta}\sigma_H(0, S_0) \ ,$$

the mixed derivatives disappear at $t = 0$ and along the line $(\log(S_0) + \alpha v, v)$ which is sufficient to avoid many of the problems caused by the irregular initial condition. Even for $t > 0$ we benefit as the mixed derivative terms are damped and $\hat{p}$ is better aligned to a rectangular grid which allows us to reduce the size of the computational domain and is better suited to locally refined meshes.

In the next subsection, we will present a discretization scheme for the transformed PDE (14).

2.2. Discretization of the Transformed Forward PDE

To ease the notation we will skip the $\hat{}$ over each transformed function. However, we will stick to the new coordinate system (y, v, t) to emphasize that we work with the transformed Kolmogorov forward equation.

One possibility to discretize (14) are time marching schemes like Crank-Nicolson or implicit Euler together with a suitable discretization for the mixed derivative terms like Samarskii and Mozolevski (2002). These schemes require the solution of large linear systems of equations in each time step which makes them computationally expensive. ADI schemes offer a remedy but the standard ADI scheme is not suited to equations with mixed derivatives. A simple scheme to deal with such equations was developed in Yanenko (1971). Like an ADI scheme it requires the solution of tridiagonal systems of linear equations. The mixed derivatives terms appear on the right hand sides only. The scheme is given as

$$\frac{p^{k+1/2} - p^k}{\Delta t} + \partial_y \left(b_y^k p^{k+1/2} + A_{yy}^k \partial_y p^{k+1/2} + A_{yv}^k \partial_v p^k \right) = 0 \,, \tag{15}$$

$$\frac{p^{k+1} - p^{k+1/2}}{\Delta t} + \partial_v \left(b_v^k p^{k+1} + A_{vy}^k \partial_y p^{k+1/2} + A_{vv}^k \partial_v p^{k+1} \right) = 0 \,. \tag{16}$$

Here k denotes the time index and Δt the time step of a homogeneous time grid. Alternative schemes were developed, e.g., in Craig and Sneyd (1988), in't Hout and Welfert (2009) or Hundsdorfer and Verwer (2003). A nice overview is given in in't Hout and Foulon (2010). All these schemes are similar to the Yanenko scheme in that they treat the mixed derivatives explicitly and other terms implicitly in the spatial discretization but offer more advanced time discretization schemes. We have chosen the Yanenko scheme because the focus of this article is on an accurate spatial discretization of the Kolmogorov forward PDE (8) and we therefore decided to keep the time discretization simple. However, we remark that the theoretical convergence properties of the time discretization schemes in the mentioned references are superior to the Yanenko scheme and that the use of one of these schemes might lead to an improvement over the Yanenko scheme.

For the spatial discretization we choose a finite volume scheme. As explained in Samarskii, Lazarov, and Makarov (1987) finite volume schemes converge under weaker conditions than regular finite difference methods. The basic idea behind finite volume schemes can be mo-

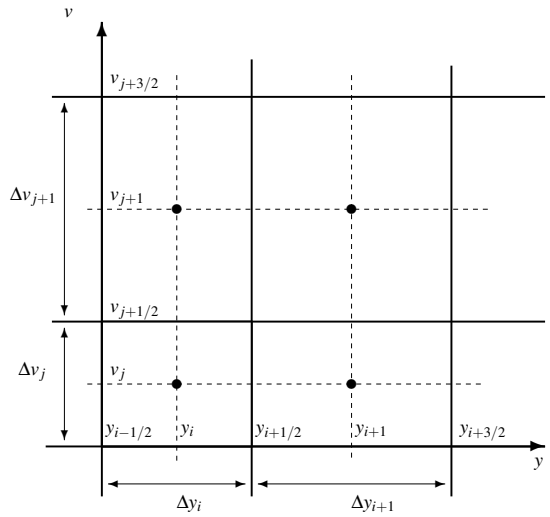


Figure 1. *Computational grid of the finite volume scheme*

tivated by the treatment of discontinuous payoffs in finite difference approximations. Here, one usually does not evaluate the payoff at each grid point but uses an average value of the payoff where the averaging is done inside a cell that is centered around a grid point. This leads to an improvement in the convergence of finite difference schemes as explained in Randall and Tavella (2000), Chapter 4. Here, we apply this kind of reasoning on the full spatial discretization scheme.

We define a computational grid $\{y_{-1/2},\ldots,y_{m+1/2}\} \times \{0 = v_{-1/2},\ldots,v_{n+1/2}\}$ which defines the volumes

$$\Omega_{i,j} := [y_{i-1/2}\ ,\ y_{i+1/2}] \times [v_{j-1/2}\ ,\ v_{j+1/2}]\ .$$

Each volume has a midpoint (y_i, v_j) with $y_i = (y_{i-1/2} + y_{i+1/2})/2\ ,\ \ v_j = (v_{j-1/2} + v_{j+1/2})/2$. We define mesh sizes $\Delta y_i := (y_{i+1/2} - y_{i-1/2})/2$ and $\Delta v_j := (v_{j+1/2} - v_{j-1/2})/2$. The grid is illustrated in Figure 1. The y-grid is concentrated around $\log(S_0) + \alpha V_0$ using the sinh-transformation that is suggested in Randall and Tavella (2000), Chapter 5. The grid spans a range of roughly $10\sigma_{ref}\sqrt{t_l}$ where σ_{ref} is the ATM implied volatility corresponding to the implied volatility curve in t_l. The v-grid is defined by

$$v_{-1/2} = 0 \quad \text{and} \quad v_{j-1/2} = c\sum_{l=0}^{j-1} d^l \ \text{ for } j > 0.$$

where c and d are such that $v_{n/2} = V_0$ and $v_{n+1/2}$ equals the pre-specified upper boundary v_{high} of the volatility grid. The primary unknowns in a finite volume scheme are the cell averages

$$\bar{p}^k_{i,j} := \frac{1}{|\Omega_{i,j}|}\int_{\Omega_{i,j}} p(t_k, y, v)\,dy\,dv \quad \text{for} \ \ 0 \le i \le m,\ 0 \le j \le n\ . \tag{17}$$

To derive a discretization scheme in the cell averages, we take cell averages of (15) and integrate with respect to y

$$\frac{\bar{p}_{ij}^{k+1/2} - \bar{p}_{ij}^{k}}{\Delta t} + \frac{1}{\Delta y_i}(f_{i+1/2,j}^{k} - f_{i-1/2,j}^{k}) = 0\,, \tag{18}$$

$$f_{l,j}^{k} = \frac{1}{\Delta v_j}\int_{v_{j-1/2}}^{v_{j+1/2}} \left(b_y^k p^{k+1/2} + A_{yy}^k \partial_y p^{k+1/2} + A_{yv}^k \partial_v p^k \right)\big|_{y_l} dv\,, l = i \pm 1/2. \tag{19}$$

Expression (19) represents an average along the edge $y_l \times [v_{j-1/2}, v_{j+1/2}]$. In the next step we explain the discretization of each of the three terms in (19). We start with the second expression because it is the easiest to discretize. We get

$$\begin{aligned} f_{i+1/2,j}^{k,y} &:= \frac{1}{\Delta v_j}\int_{v_{j-1/2}}^{v_{j+1/2}} A_{yy}^k \partial_y p^{k+1/2}\big|_{y_{i+1/2}} dv \\ &\approx -\frac{1}{2} v_j \left(\left(\sigma_{i+1/2,j}^k\right)^2 + 2\eta\rho\alpha\sigma_{i+1/2,j}^k + \eta^2\alpha^2 \right) \frac{\bar{p}_{i+1,j}^{k+1/2} - \bar{p}_{i,j}^{k+1/2}}{\Delta y_{i+1/2}} \end{aligned} \tag{20}$$

where we have used the abbreviation $\sigma_{i+1/2,j}^k = \sigma_H(t_k, y_{i+1/2}, v_j)$. To motivate this discretization, we insert (20) into the second term of (18). This leads to

$$\frac{1}{\Delta y_i}\left((A_{yy}^k)_{i+1/2,j} \frac{\bar{p}_{i+1,j}^{k+1/2} - \bar{p}_{i,j}^{k+1/2}}{\Delta y_{i+1/2}} - (A_{yy}^k)_{i-1/2,j} \frac{\bar{p}_{i,j}^{k+1/2} - \bar{p}_{i-1,j}^{k+1/2}}{\Delta y_{i-1/2}} \right)$$

which is a common second order finite difference stencil for the discretization of $\partial_y(A_{yy}\partial_y p)$. The only difference is that now we use cell averages instead of function values in the stencil.

For the third term, we find

$$\begin{aligned} f_{i+1/2,j}^{k,v} &:= \frac{1}{\Delta v_j}\int_{v_{j-1/2}}^{v_{j+1/2}} A_{yv}^k \partial_v p^k\big|_{y_{i+1/2}} dv \\ &\approx -\frac{1}{2}\left(\eta\rho\sigma_{i+1/2,j}^k + \eta^2\alpha\right)\frac{1}{\Delta v_j}\int_{v_{j-1/2}}^{v_{j+1/2}} v\partial_v p(t_k, y_{i+1/2}, v) dv \\ &= -\frac{1}{2}\left(\eta\rho\sigma_{i+1/2,j}^k + \eta^2\alpha\right)\left(v_{j+1/2} p_{i+1/2,j+1/2}^k - v_{j-1/2} p_{i+1/2,j-1/2}^k - \int_{v_{j-1/2}}^{v_{j+1/2}} p^k dv \right)/\Delta v_j \\ &\approx -\frac{1}{2}\left(\eta\rho\sigma_{i+1/2,j}^k + \eta^2\alpha\right)\left(\frac{v_{j+1/2} p_{i+1/2,j+1/2}^k - v_{j-1/2} p_{i+1/2,j-1/2}^k}{\Delta v_j} - p_{i+1/2,j}^k \right). \end{aligned} \tag{21}$$

Note that (21) does not contain cell averages $\bar{p}$, but the values of p on mid points and corners of cell boundaries which we compute by simple linear interpolation

$$p^k_{i+1/2,j} \approx a_{i+1/2}\bar{p}^k_{i+1,j} + \left(1-a_{i+1/2}\right)\bar{p}^k_{i,j}$$
$$p^k_{i+1/2,j+1/2} \approx c_{j+1/2}\bar{p}^k_{i+1/2,j+1} + \left(1-c_{j+1/2}\right)\bar{p}^k_{i+1/2,j}$$

where $a_{i+1/2} = \Delta y_i/(\Delta y_{i+1}+\Delta y_i)$ and $c_{i+1/2} = \Delta v_i/(\Delta v_{i+1}+\Delta v_i)$. The case $j=0$ is somewhat special in (21) as $p_{i+1/2,-1/2}$ is a value on the boundary $v=0$ where p is not well defined if the Feller condition is not fulfilled. Yet, given the asymptotics (12) we can safely assume that $v_{-1/2}p_{i+1/2,-1/2} = 0$.

Finally, we discretize the first term by assuming that the integrand is constant and by replacing the value of p on the cell boundary using a weighted average of the neighboring cell averages by means of an upwind scheme. This results in

$$\begin{aligned} f^{k,b}_{i+1/2,j} &:= \frac{1}{\Delta v_j}\int_{v_{j-1/2}}^{v_{j+1/2}} b^k_y p^{k+1/2}\big|_{y_{i+1/2}} dv \\ &\approx \left(r-d-\frac{1}{2}\sigma^k_{i+1/2,j}v_j + \alpha\lambda(\theta-v_j) - \frac{1}{2}v_j\partial_y\sigma^k_{i+1/2,j} - \frac{1}{2}\eta\rho\sigma^k_{i+1/2,j}\right. \\ &\qquad \left. -\frac{1}{2}\eta^2\alpha - \frac{1}{2}\eta\rho\alpha v_j\partial_y\sigma^k_{i+1/2,j}\right)\left(w_{i,j}\bar{p}^{k+1/2}_{i+1,j} + (1-w_{i,j})\bar{p}^{k+1/2}_{i,j}\right) \\ &=: b^{k,y}_{i+1/2,j}\left(w_{i,j}\bar{p}^{k+1/2}_{i+1,j} + (1-w_{i,j})\bar{p}^{k+1/2}_{i,j}\right). \end{aligned} \tag{22}$$

It remains to specify the upwind weights $w_{i,j}$. The calculation of $\bar{p}^{k+1/2}$ is done by solving tridiagonal systems of linear equations because in (18) $\bar{p}^{k+1/2}_{i,j}$ depends only on itself, $\bar{p}^{k+1/2}_{i+1,j}$ from $f^{k,y/b}_{i+1/2,j}$ and $\bar{p}^{k+1/2}_{i-1,j}$ from $f^{k,y/b}_{i-1/2,j}$. All quantities in $f^{k,v}_{i\pm1/2,j}$ refer to time t_k.

Let Z be the matrix for one of the linear systems multiplied by Δt for convenience. For interior nodes the entries of the ith row read

$$\begin{aligned} z_{i,i-1} &= \frac{\Delta t}{\Delta y_i}\left(\frac{(A^k_{yy})_{i-1/2,j}}{\Delta y_{i-1/2}} - b^{k,y}_{i-1/2,j}(1-w_{i-1,j})\right) \\ z_{i,i} &= 1-\frac{\Delta t}{\Delta y_i}\left(\frac{(A^k_{yy})_{i-1/2,j}}{\Delta y_{i-1/2}} + \frac{(A^k_{yy})_{i+1/2,j}}{\Delta y_{i+1/2}} + b^{k,y}_{i-1/2,j}w_{i-1,j} - b^{k,y}_{i+1/2,j}(1-w_{i,j})\right) \\ z_{i,i+1} &= \frac{\Delta t}{\Delta y_i}\left(\frac{(A^k_{yy})_{i+1/2,j}}{\Delta y_{i+1/2}} + b^{k,y}_{i+1/2,j}w_{i,j}\right). \end{aligned}$$

To avoid numerical problems like spurious oscillations the matrix Z should be an M-matrix (Hackbusch 1992) which is defined as a matrix Z with $z_{i,i} > 0$, $z_{i,j} \le 0$ for $i \ne j$, and Z^{-1} exists and has positive entries only. For homogeneous Dirichlet boundary conditions we have

$z_{0,0} = z_{m,m} = 1$ and $z_{0,1} = z_{m-1,m} = 0$. The requirement of negative off-diagonal entries lead to the conditions

$$w_{i,j} \leq -\frac{(A^k_{yy})_{i+1/2,j}}{b^{k,y}_{i+1/2,j}\Delta y_{i+1/2}} \quad \text{if } b^y_{i+1/2,j} > 0 \quad \text{and} \quad w_{i,j} \geq \frac{(A^k_{yy})_{i+1/2,j}}{|b^{k,y}_{i+1/2,j}|\Delta y_{i+1/2}} + 1 \quad \text{if } b^y_{i+1/2,j} < 0\,.$$

In our scheme the weights are chosen to fulfill these constraints and to be as close as possible to $a_{i+1/2}$ to keep the consistency errors small. Unfortunately, the above conditions are not sufficient to guarantee the M-matrix property. As an additional sufficient condition the diagonal dominance of Z is required, i.e. $z_{i,i} \geq |z_{i,i-1}| + |z_{i,i+1}|$, which simplifies to the condition

$$\frac{\Delta t}{\Delta y_i}|b^{k,y}_{i+1/2,j} - b^{k,y}_{i-1/2,j}| \;\leq\; 1.$$

This condition is fulfilled for a sufficiently large number of time steps only. The largest terms in $b^{k,y}$ are those including the first derivative of σ_H which might not be smooth. Thus, the condition is not met by our scheme in general. In order not to violate the constraint too much we apply additional smoothing to σ_H which will be explained in the next subsection. Also increasing the number of time steps stabilizes the scheme.

As we have already explained in Section 1 homogeneous Dirichlet boundary values are used at y_{low} and y_{high}

$$\bar{p}_{0,j} = \bar{p}_{m,j} = 0 \quad \forall j\,,$$

because the computational domain is chosen sufficiently large to ensure that p is very small near the boundaries. From (20), (21) and (22) we see that $f_{-1/2,j}$ and $f_{m+1/2,j}$ are very small as well. This means that probability mass is approximately conserved in step (15)

$$\begin{aligned}\sum_{i,j}|\Omega_{i,j}|\bar{p}^{k+1/2}_{i,j} &= \sum_{i,j}|\Omega_{i,j}|\bar{p}^k_{i,j} \;+\; \Delta t\sum_j \Delta v_j(f^k_{m+1/2,j} - f^k_{-1/2,j}) \\ &\approx \sum_{i,j}|\Omega_{i,j}|\bar{p}^k_{i,j}\,. \end{aligned} \tag{23}$$

Perfect fulfillment of condition (10) could be obtained by enforcing the boundary conditions $f^k_{m+1/2,j} = f^k_{-1/2,j} = 0$, but this would not lead to a significant improvement of the algorithm's accuracy because the effects are too small near the spot boundaries.

The algorithm for solving (16) is derived along the same lines as the algorithm presented above for approximating (15). For this reason, we only state the results for completeness and do not repeat every single step of the derivation. Computing cell averages by integrating (16) leads to

$$\frac{\bar{p}^{k+1}_{ij} - \bar{p}^{k+1/2}_{ij}}{\Delta t} + \frac{1}{\Delta v_j}(g^k_{i,j+1/2} - g^k_{i,j-1/2}) = 0\,, \tag{24}$$

$$g^k_{i,l} = \frac{1}{\Delta y_j}\int_{y_{i-1/2}}^{y_{i+1/2}} \left(b^k_v p^{k+1} + A^k_{vy}\partial_y p^{k+1/2} + A^k_{vv}\partial_v p^{k+1} \right)|_{v_l} dy \ , l = j \pm 1/2. \quad (25)$$

The above definition of $g^k_{i,l}$ only holds for interior edges $l > -1/2$. For $l = -1/2$, i.e at $v = 0$, we use the simple boundary condition

$$g^k_{i,-1/2} = 0 \ , \quad (26)$$

which enforces probability mass conservation in (16). This can be shown by a similar partial summation as in (23) together with (26) and using $g^k_{i,n+1/2} = 0$ which is justified because of the negligible p at the upper volatility boundary v_{high}. For $j \geq 0$ we discretize (25) analogously to (20), (21) and (22)

$$\begin{aligned}
g^{k,v}_{i,j+1/2} &:= \frac{1}{\Delta y_i}\int_{y_{i-1/2}}^{y_{i+1/2}} A^k_{vv}\partial_v p^{k+1}|_{v_{j+1/2}} dy \\
&\approx -\frac{1}{2} v_{j+1/2}\eta^2 \frac{\bar{p}^{k+1}_{i,j+1} - \bar{p}^{k+1}_{i,j}}{\Delta v_{j+1/2}} \ , \quad (27) \\
g^{k,y}_{i,j+1/2} &:= \frac{1}{\Delta y_i}\int_{y_{i-1/2}}^{y_{i+1/2}} A^k_{vy}\partial_y p^{k+1/2}|_{v_{j+1/2}} dy \\
&\approx -\frac{1}{2} v_{j+1/2}\left(\eta^2\alpha + \eta\rho\sigma^k_{i,j+1/2}\right) \frac{p^{k+1/2}_{i+1/2,j+1/2} - p^{k+1/2}_{i-1/2,j+1/2}}{\Delta y_i} \ , \quad (28) \\
g^{k,b}_{i,j+1/2} &:= \frac{1}{\Delta y_i}\int_{y_{i-1/2}}^{y_{i+1/2}} b^k_v p^{k+1}|_{v_{j+1/2}} dy \\
&\approx \left(\lambda(\theta - v_{j+1/2}) - \frac{1}{2}\eta^2 - \frac{1}{2}\eta\rho v_{j+1/2}\partial_y\sigma^k_{i,j+1/2}\right)\left(q_{i,j}\bar{p}^{k+1}_{i,j+1} + \left(1 - q_{i,j}\right)\bar{p}^{k+1}_{i,j}\right) \\
&=: b^{k,v}_{i,j+1/2}\left(q_{i,j}\bar{p}^{k+1}_{i,j+1} + \left(1 - q_{i,j}\right)\bar{p}^{k+1}_{i,j}\right) \ . \quad (29)
\end{aligned}$$

The values $p^{k+1/2}_{i+1/2,j+1/2}$ and $p^{k+1/2}_{i-1/2,j+1/2}$ are computed by linear interpolation as described in the context of (21). The factors $q_{i,j}$ are derived by a similar reasoning as the factors $w_{i,j}$

$$q_{i,j} = \begin{cases} \min\left(c_{j+1/2}, -\frac{(A^k_{vv})_{i,j+1/2}}{b^{k,v}_{i,j+1/2}\Delta v_{j+1/2}}\right) & \text{if } b^v_{i,j+1/2} > 0 \\ \max\left(c_{j+1/2}, \frac{(A^k_{vv})_{i,j+1/2}}{|b^{k,v}_{i,j+1/2}|\Delta v_{j+1/2}} + 1\right) & \text{otherwise} \end{cases} .$$

This finishes the discretization of (16).

2.3. Calculation of the Heston Local Volatility

As already outlined at the beginning of this chapter, the calibration of the Heston local volatility is done separately for each time slice. From the solution of the PDE (14) in one time step the Heston local volatility is computed using (6). This Heston local volatility serves then as an input to the calculation of the transition density in the succeeding time step. The solution of (14) is given in terms of cell averages $\bar{p}$ while in the calculation of $\sigma_H(t_k,s)$ using (6) and (7) we need the transition density function p. For this reason we have to extract the original function p from the cell averages $\bar{p}$. We do this using the method of monotone convex splines that is described in Hagan and West (2006). This method is well suited in this context because calculating p from $\bar{p}$ as described in Hagan and West (2006) preserves both the monotonicity and the positivity of cell averages, and most importantly the cell average values itself. As the nominator and denominator in (7) might become quite small σ_H is computed by a stabilized variant

$$\sigma_H^2 = \sigma_{LV}^2 \frac{\int_0^\infty p(t,\log(s),v)dv \; + \; \varepsilon}{\int_0^\infty v p(t,\log(s),v)dv \; + \; \theta\varepsilon} \; , \tag{30}$$

where ε is a small number, for instance, $\varepsilon = 1.e-6$. In addition σ_H^2 is smoothed by iterating the scheme

$$(\sigma_{H,i}^2)^{\gamma+1} = \left((\sigma_{H,i+1}^2)^\gamma \; + \; 2(\sigma_{H,i}^2)^\gamma \; + \; (\sigma_{H,i-1}^2)^\gamma\right) \; / \; 4, \quad \gamma = 1,\ldots,5$$

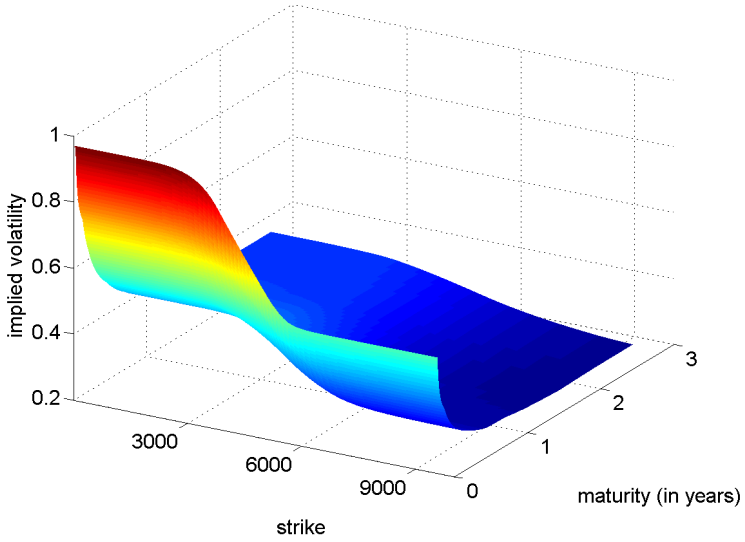
to stabilize the calculation of derivatives of σ_H. Alternatively one could specify a parameterization for σ_H and determine its parameters from a least squares fit to (30).

At this stage we also employ simple but quite effective technique to improve convergence. At each time step the forward is computed from density p and the defect to the true forward is used to adjust $\bar{p}$ along the y-coordinate.

3. Numerical Examples

We first demonstrate the convergence of the finite volume scheme by a single test case with DAX data as of 2008/10/09. The implied volatility surface is illustrated in Figure 2. The second convergence test covers 350 DAX data sets from 2007/12/28 to 2008/12/28 with roughly one data set per day except for October and November where the frequency is higher. The focus of the third part of our tests are the properties of the stochastic local volatility model in terms of the forward implied volatility and the impact on prices for some common products.

Figure 2. Implied volatility surface of the DAX on 2008/10/09



Convergence for DAX asof 2008/10/09
The starting point is the calibration of the Heston model alone. Table 1 reports the respective model parameters. We see that an extremely high value of the volatility of variance is required which leads to a strong violation of the Feller condition.

Table 1. Calibrated Heston parameters for STOXX50E 2008/7/16 and DAX 2008/10/9

	V_0	θ	λ	η	ρ	$\lambda\theta - \eta^2/2$
DAX	0.571	0.075	6.664	1.399	-0.70	-0.480

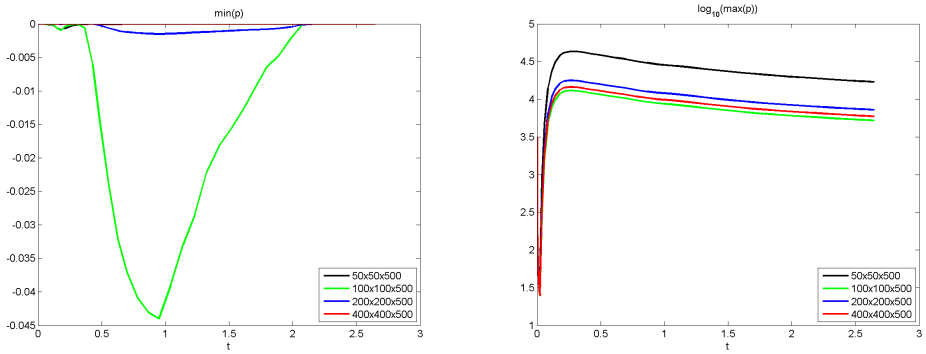
Table 2 shows errors in call prices on some of the first expiries. The errors are scaled by $10000/S_0$. We report the maximum absolute error for a strike range of $0.5S_0$ to $1.5S_0$ for the plain Heston model and the stochastic local volatility model with different numbers of grid points. The stochastic local volatility model has significantly lower calibration errors for a sufficiently large number of grid points. The CPU time for the $400 \times 400 \times 500$ grid is roughly 90 seconds on a recent entry level x86 CPU with 2 cores and OpenMP parallelization. As we have pointed out in Section 2.2 the numerical scheme is not necessarily monotone or stable if the number of time steps is not large enough. Spurious oscillations can develop and affect the solution quality. A very good indicator for the magnitude of these oscillations is the minimum value of the transition density p which is logged for the first 30 time steps and then every 15th step. Values below zero indicate oscillations. Figure 3 shows these values for the DAX example and the different numerical grids. In this figure only two graphs are visible.

For the finest grid no values below zero were observed and the graph lies on the x-axis. The graph for the coarsest grid also lies almost on the x-axis which illustrates the non-monotonous behavior of this discretization scheme. For a comparison we also plot the maximum of p. As these values are very large $log_{10}(max(p))$ is shown. In relation to the very large maximum of p the oscillations are very small. This does not hold if the transformation of the forward PDE of Section 2.1 to reduce mixed derivatives is switched off. In this case $min(p)$ took values of order -10^3 to -10^4 in the first few time steps. This observation confirms the importance of this transformation for the quality of the numerical solution.

Table 2. Dax 2008/10/09: Maximum scaled price errors for strikes between $0.5S_0$ and $1.5S_0$ corresponding to the option expiries displayed in the first row

Grid($S \times V \times T$)	17/10/08	21/11/08	20/3/09	18/9/09	18/6/10	17/6/11
Heston	-51	-52	-25	-42	-48	-95
$50 \times 50 \times 500$	12	9	31	71	135	188
$100 \times 100 \times 500$	8	8	6	6	12	20
$200 \times 200 \times 500$	3	3	3	4	7	11
$400 \times 400 \times 500$	2	3	5	6	6	9
$400 \times 400 \times 1000$	1	2	3	4	6	7

Figure 3. $\min(p)$ (left) and $\log_{10}(max(p))$ (right) for calibration of DAX as of 2008/10/9.



Fitting quality for DAX 2008

For the massive test we just report scaled fitting errors for call prices of the first and 10th expiry for a strike range of $0.5S_0$ to $1.5S_0$. We used the 400x400x500 grid. Figure 4 compares the fitting errors for the local volatility model, the Heston model and the stochastic local volatility

model. The stochastic local volatility model clearly outperforms the simple Heston model. The fitting errors are similar to the local volatility model, except for some days in November where the front expiry was closer than < 4 days to the valuation date. To achieve better convergence also under such circumstances requires some fine tuning of the model which is not the scope of the present paper.

Forward Implied Volatilities
The previous tests have shown that the stochastic local volatility model achieves a very good fit to European vanilla prices. The question arises if this spoils the dynamics of the simulated processes. To answer this question we computed the foward implied volatility σ_{fwd} defined by

$$C_{model}(t,T,K) = e^{-rt} C_{BS}(1,K,T-t,\sigma_{fwd}(t,T,K))$$

where C_{model} are the simulated call prices with expiry T, fixing date $t < T$ and relative strike K. C_{BS} is the usual Black-Scholes call price function. Our market data is DBKGn as of 2008/05/16. Figure 6 shows that the shape of the forward implied volatility is nearly preserved by our scheme and does not become flat as for the local volatility model. This is exactly what we wanted to achieve: inherit the dynamics of the Heston model and the accuracy of the local volatility model.
In the next experiment we went one step further and used Heston parameters without any relation to the market data and even a postive correlation: $V_0 = \theta = 0.16, \lambda = 2, \eta = 1.0, \rho = +0.5$. The stochastic local volatility model was fitted again to DBKGn as of 2008/05/16. Despite that the forward implied volatility of the stochastic local volatility model still looks similar to the Heston model, at least on the long run. Of coures the volatility level needs to be adjusted (moved down in our example) to fit the market data. But the smile or skew is very similar to the Heston model. This is also a quite good result which meets our expectations.

Impact on product prices
Table 3 shows prices of some typical products on DBKGn asof 16/5/2008 with barriers. Such products may also dependend on the forward dynamics. For the products with one year to expiry the price differences between the local volatility and the stochastic local volatility are just a couple of basis points and the Heston model is an outlier which in our opion this is due to the rather weak fitting quality. The model dynamics does not seem to be that important here. For the longer maturities the picture changes and now Heston and the stochastic local volatility model are closer to each other. The difference of 20 .. 30 Bps is within the range of the Heston fitting error and likely due to this error.

4. Summary

In this paper we have presented a finite volume scheme for the calibration of the Heston stochastic local volatility model. We have shown in detail how the original problem has to be transformed to avoid numerical problems and how the transformed equations have to be dis-

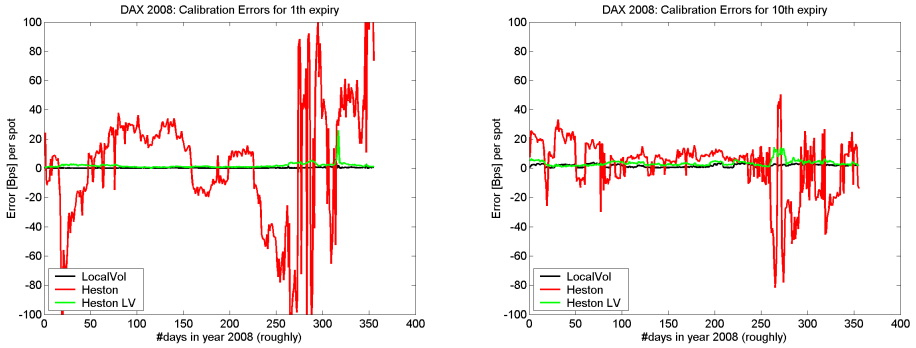


Figure 4. Maximum scaled price errors for strikes between $0.5S_0$ and $1.5S_0$ for DAX data from 2008

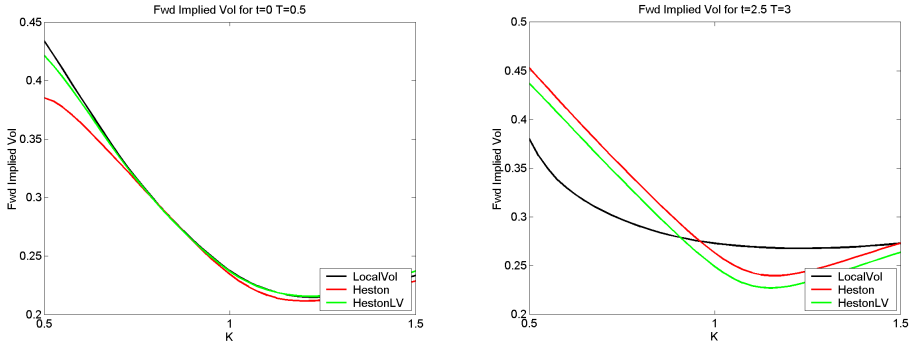


Figure 5. (Forward) implied volatility for DBKGn asof 16/05/2008 for fixing in $t = 0$ and $T = 0.5$ years (left) and $t = 2.5, T = 3$ years (right).

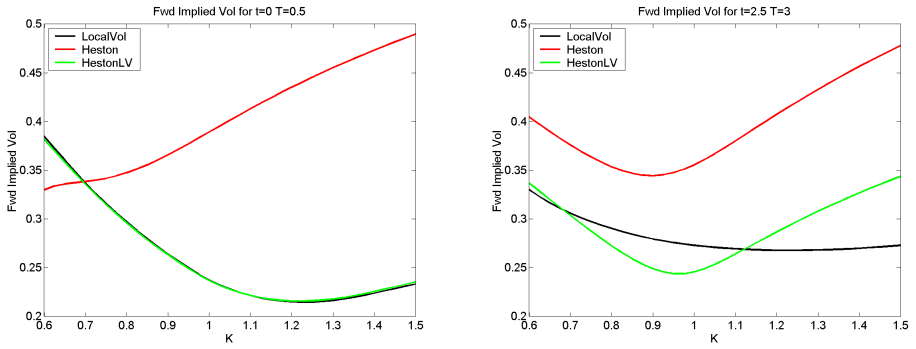


Figure 6. (Forward) implied volatility for DBKGn asof 16/05/2008 for fixing in $t = 0$ and $T = 0.5$ years (left) and $t = 2.5, T = 3$ years (right).

Product		local vol	Heston	stoch local vol
Up&Out Call	K=95%, B=118%, T=1y	1.78%	2.52%	1.85 %
	K=87%, B=135%, T=3y	3.21%	3.89%	4.22%
Down&Out Put	K=132%, B=81%, T=1y	8.19%	8.28%	8.16%
	K=132%, B=81%, T=3y	2.27%	1.62%	1.83%

Table 3. Product prices for DBKGn as of 16/05/2008 relative to the initial spot.

cretized. The resulting numerical scheme for the forward PDE requires the iterative solution of tridiagonal systems of linear equations and is therefore computationally efficient and easy to implement.

Numerical examples for the STOXX50E and the DAX have demonstrated that this method leads to a good fitting of European call option prices even for cases where the Feller condition is not fulfilled. Despite the fact that scheme is monotone and stable only for a relatively large numbers of time grid points we have observed small instabilities only when we have carried out our numerical test examples. In particular for the first time steps this is because of the coordinate transform effectively damping the mixed derivative terms. Taking into account that the calibration only needs to be performed once per underlying the CPU times are acceptable for a reasonable numbers of grid points from a practical perspective.

References

Bates, D. S., 1996, Dollar jump fears, 1984:1992, Distributional anomalies implicit in currency futures options, *Journal of International Money and Finance* 15, 65–93.

Coleman, T. F., Y. Li, and A. Verma, 1999, Reconstructing the unknown local volatility function, *Journal of Computational Finance* 2, 77–102.

Craig, I. J. D., and D. Sneyd, 1988, An Alternating-Direction Implicit Scheme for Parabolic Equations with Mixed Derivatives, *Computers and Mathematics with Applications* 16, 341–350.

Crépey, S., 2003, Calibration of the Local Volatility in a Generalized Black-Scholes Model Using Tikhonov Regularization, *SIAM Journal on Mathematical Analysis* 34, 1183–1206.

Derman, E., and I. Kani, 1994, Riding on a smile, *Risk* 7, 32–39.

Dupire, B., 1994, Skewness and kurtosis in S&P500 index returns implied by option prices, *The Journal of Financial Research* 19, 175–192.

Engelmann, B., M. R. Fengler, and P. Schwendner, 2009, Hedging under Alternative Stickyness Assumptions: An Empirical Analysis for Barrier Options, *Journal of Risk* 12, 53–77.

Gyöngy, I., 1986, Mimicking the One-dimensional Marginal Distributions of Processes Having an Ito Differential, *Probability Theory and Related Fields* 71, 501–516.

Hackbusch, W., 1992, *Elliptic Differential Equations: Theory and Numerical Treatment*. (Springer Berlin Heidelberg New York).

Hagan, P. S., D. Kumar, A. Lesniewski, and D. Woodward, 2002, Managing Smile Risk, *Wilmott Magazine* 1, 84–108.

Hagan, P. S., and G. West, 2006, Interpolation Methods for Curve Construction, *Applied Mathematical Finance* 13, 89–129.

Henry-Labordére, P., 2009, Calibration of Local Stochastic Volatility Models to Market Smiles: A Monte-Carlo Approach, Working paper.

Heston, S., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327–343.

Hull, J., and A. White, 1987, The Pricing of Options on Assets with Stochastic Volatilities, *Journal of Finance* 42, 281–300.

Hundsdorfer, W., and J. G. Verwer, 2003, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. (Springer Berlin).

in't Hout, K. J., and S. Foulon, 2010, ADI Finite Difference Schemes for Option Pricing in the Heston Model with Correlation, *International Journal of Numerical Analysis and Modeling* 7, 303–320.

in't Hout, K. J., and B. Welfert, 2009, Unconditional Stability of Second-Order ADI Schemes Applied to Multi-Dimensional Diffusion Equations with Mixed Derivative Terms, *Applied Numerical Mathematics* 59, 2593–2609.

Randall, C., and D. Tavella, 2000, *Pricing Financial Instruments: The Finite Difference Method*. (John Wiley & Sons New York).

Rebonato, R., 1999, *Volatility and Correlation*. (John Wiley & Son New York).

Ren, Y., D. Madan, and M. Q. Qian, 2007, Calibrating and Pricing with Embedded Local Volatility Models, *Risk* 20, 138–143.

Samarskii, A. A., R. D. Lazarov, and V. L. Makarov, 1987, *Difference Schemes for Differential Equations with Weak Solutions (in Russian)*. (Vysshaya Shkola Moscow).

Samarskii, A. A., Matus P. P. Mazhukin V. I., and I. E. Mozolevski, 2002, Monotone Difference Schemes for Equations with Mixed Derivatives, *Computers and Mathematics with Applications* 44, 501–510.

Schöbel, R., and J. Zhu, 1999, Stochastic Volatility with an Ornstein-Uhlenbeck Process: An Extension, *European Finance Review* 3, 23–46.

Stein, E. M., and J. C. Stein, 1991, Stock Price Distributions with Stochastic Volatility: An Analytic Approach, *Review of Financial Studies* 4, 727–752.

v. Haastrecht, A., and A. A. J. Pelsser, 2008, Efficient, Almost Exact Simulation of the Heston Stochastic Volatility Model, Working paper.

Yanenko, N. N., 1971, *The Method of Fractional Steps: The Solution of Problems of Mathematical Physics in Several Variables*. (Springer Berlin Heidelberg New York).