Calibration of the Hobson&Rogers model: empirical tests.

Paolo Foschi and Andrea Pascucci
Dipartimento di Matematica, Università di Bologna *

Abstract

The path-dependent volatility model by Hobson and Rogers is considered. It is known that this model can potentially reproduce the observed smile and skew patterns of different directions, while preserving the completeness of the market. In order to quantitatively investigate the pricing performance of the model a calibration procedure is here derived. Numerical results based on S&P500 option prices give evidence of the effectiveness of the model.

1 Introduction

The aim of this paper is to propose a flexible calibration procedure of the Hobson&Rogers model [10] and to investigate the performance by testing it on a set of S&P500 option data. Among non-constant volatility models in complete markets, the Hobson&Rogers model seems to be one of the more appealing. In this model the volatility σ is supposed to depend on the trend of the underlying asset, defined as the difference of the spot price S and a weighted average of past prices. This feature seems to be more realistic and natural compared with the usual assumption $\sigma = \sigma(t, S)$ of the widespread level-dependent models: for instance, it is known that the volatility increases after a market reversal and this is difficultly captured by a model which only takes into account of the present price of the underlying.

In the Hobson&Rogers setting no exogenous source of risk is added so that the market completeness is preserved and the arbitrage pricing theory applies. Moreover this model is potentially capable to reproduce the observed smile and volatility term structure patterns. Despite of its fine features, so far little has been done in the empirical analysis of the model. Figà-Talamanca and Guerra [8] examined the problem of the estimation of the parameters of the model and a generalization was proposed by Hubalek, Teichmann and Tompkins [11]. The Hobson&Rogers model has also been considered by Hallulli and Vargiolu [2] and an extension to the framework of term-structure modeling was given by Chiarella and Kwon [3].

In order to present our results, we first recall the main features of (a simplified version of) the Hobson&Rogers model. In a Wiener space with one-dimensional Brownian motion (W_t), we denote by S_t the stock price and by D_t the deviation of prices from the trend, defined by

$$D_t = Z_t - \int_0^{+\infty} \lambda e^{-\lambda\tau} Z_{t-\tau} d\tau, \qquad \lambda > 0, \tag{1.1}$$

*Piazza di Porta S. Donato 5, 40126 Bologna (Italy). E-mail: foschip@csr.unibo.it, pascucci@dm.unibo.it

where $Z_t = \log(e^{-rt}S_t)$ is the discounted log-price. In (1.1), the parameter λ amounts to the rate at which past prices are weighted. Hobson and Rogers assume that S_t is an Itô process, solution to the stochastic differential equation

$$dS_t = \mu(D_t)S_t dt + \sigma(D_t)S_t dW_t. \tag{1.2}$$

In (1.2), μ and $\sigma > 0$ are deterministic functions satisfying usual hypotheses in order to guarantee that the system of SDEs (1.1)-(1.2) is uniquely solvable. Finally, we denote by U_{T-t} the price at time t of an European contingent claim with exercise date T.

A key feature of the model is that the process (S_t, D_t) is Markovian (cf. Lemma 3.1 in [10]). Then if we consider the time t, the price S_t and the mean $M_t = \log(e^{rt}S_t) - D_t$ as state variables and assume that

$$U_{T-t} = e^{-rt}f(S_t, M_t, t)$$

for some smooth function f, by the Feynman-Kac formula the function f satisfies the PDE in $\mathbb{R}^3$:

$$\frac{\sigma^2(Z-M)S^2}{2}\,\partial_{SS}f + rS\partial_S f + \lambda(\log S + rt - M)\partial_M f - \partial_t f = 0. \tag{1.3}$$

As in the Black&Scholes framework, the drift term in (1.2) does not enter in the valuation PDE while a key role is played by the volatility function σ which is an input parameter of the model and has to be estimated in order to fit market observations. Aiming to motivate the model, Hobson and Rogers consider in [10] a volatility function of the form

$$\sigma(D) = \min\left\{\eta\sqrt{1+\varepsilon D^2}, N\right\} \tag{1.4}$$

for some large constant N and positive parameters ε, η: then they show that the model can indeed exhibit smiles and skews of different directions. However Hubalek, Teichmann, and Tompkins in [11] remark that the smiles obtained by (1.4) are essentially flat if compared with the ones in real markets. Hence the a priori choice of σ in (1.4) seems unsuitable to describe real option dynamics. In this note we aim to select σ without imposing a priori assumptions on its shape but simply calibrating it to market option prices. In order to maintain the approach as much flexible as possible, we only assume that σ is approximated in a space of splines.

At first glance, this calibration problem is similar to that in the framework of Dupire's implied diffusion theory [7] where the asset price S_t solves a SDE of the form

$$dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t.$$

Dupire model is consistent with the market implied volatility smile provided that the function σ is *continuously* calibrated to the market by the Dupire's local volatility formula. Several major derivatives houses have this model implemented.

On the other hand the Hobson&Rogers model seems to have two main advantages. Firstly, in a path-dependent model the volatility incorporates information on the past and, in particular, on the preceding behavior of the investors. Then, in some sense, the model "knows" how investors behave in different market circumstances and can also keep into account of the (positive or negative) trend of the asset. For this reason it seems that the Hobson&Rogers model does not

need to be continuously recalibrated: for practical use, in many cases it should be sufficiently reliable as soon as it is calibrated once a day.

Secondly, due to some invariance property of (1.3), a simple change of variables allows to evaluate all European option prices corresponding to different strikes and different time-to-maturities in a single run (see problem (2.2)-(2.5)). This considerably speed up the calibration procedure by PDEs' techniques. Actually the PDE approach also has the natural advantage of allowing to compute the derivatives with respect to the parameters (or Greeks) of the solution which will be useful in the procedure.

The paper is organized as follows. In Section 1 we recall some numerical result for the Hobson&Rogers model in the framework of PDEs of Kolmogorov type. Then, in Section 2 the inverse problem arising in the calibration is stated as a simple nonlinear least squares problem. In the last part of the paper, the results of the calibration are tested on a set of S&P500 index options prices and experimental results regarding the fitting of the model to observed prices are presented.

2 Numerical preliminaries

In this section we provide some preliminaries: we briefly recall the numerical results in [4, 5] for the Hobson&Rogers PDE (2.2) and then formulate the discretization of the equation as a block bidiagonal linear system.

2.1 Finite difference schemes for the pricing PDE

In case of an European call option with strike K, equation (2.2) is coupled with the following initial conditions

$$f(S, M, 0) = (S - K)^+. \tag{2.1}$$

We rewrite equation (1.3) as

$$\mathcal{L}u \equiv a(\partial_{xx}u - \partial_x u) + (x - y)\partial_y u - \partial_t u = 0, \tag{2.2}$$

where $u = u(x, y, t)$ is determined by the transformation

$$f(S, M, t) = Ku\left(\log(S/K) + rt, M, \lambda t\right) \tag{2.3}$$

and it has been set

$$a(x, y) = \frac{\sigma^2(x - y)}{2\lambda}. \tag{2.4}$$

By this change of variables, problem (1.3)-(2.1) is equivalent to the Cauchy problem for (2.2) in the strip $\mathbb{R}^2 \times [0, \lambda T]$ with initial condition

$$u(x, y, 0) = (e^x - 1)^+ \qquad \text{for } (x, y) \in \mathbb{R}^2. \tag{2.5}$$

Note that (2.2)-(2.5) are independent of K. Due to the additional state variable M on which the option price depends, equation (2.2) is of degenerate type since the quadratic form associated to the second order part of $\mathcal{L}$ is singular. However (2.2) belongs to the noteworthy subclass

of Hörmander PDEs today called of Kolmogorov or Ornstein-Uhlenbeck type. For this class a very satisfactory theory has been developed and many sharp analytical results are available even under weak regularity assumptions (see [12] for an exhaustive survey on this topic). In particular, in [6] it is proved that if the coefficient a is a bounded Hölder continuous function then problem (2.2)-(2.5) has a unique classical solution.

The natural framework for the study of the properties of equation (2.2) is the analysis on Lie groups. Also in the numerical approximation the best results are obtained in a non-Euclidean setting: it is known (cf., for instance, [5]) that the differential operators ∂_x and

$$Yu = (x-y)\partial_y u - \partial_t u \tag{2.6}$$

are the main (in some intrinsic sense) directional derivatives of the degenerate equation (2.2). In the numerical solution of the option pricing equation by finite-difference methods, it is natural and more efficient to approximate the main directional derivatives rather than the usual Euclidean ones. Then, on the uniform grid

$$G = \{(i\Delta_x, j\Delta_y, n\Delta_t) \ \mid \ i,j,n \in \mathbb{Z},\ n \geq 0\}, \tag{2.7}$$

we approximate as usual the derivatives $\partial_x u$ and $\partial_{xx} u$ by the centered differences and the three-point schemes, respectively:

$$\partial_x u(x,y,t) \quad \sim \quad D_{\Delta_x} u(x,y,t) = \frac{u(x+\Delta_x,y,t) - u(x-\Delta_x,y,t)}{2\Delta_x}, \tag{2.8}$$

and

$$\partial_{xx} u(x,y,t) \quad \sim \quad D^2_{\Delta_x} u(x,y,t) = \frac{u(x+\Delta_x,y,t) - 2u(x,y,t) + u(x-\Delta_x,y,t)}{\Delta_x^2}, \tag{2.9}$$

Thus, the approximation

$$\begin{aligned} \partial_{xx} u(x,y,t) - \partial_x u(x,y,t) \quad \sim \quad & D^2_{\Delta_x} u(x,y,t) - D_{\Delta_x} u(x,y,t) \\ & = d_1 u(x-\Delta_x,y,t) + d_2 u(x,y,t) + d_3 u(x+\Delta_x,y,t), \end{aligned} \tag{2.10}$$

with $d_1 = 1/\Delta_x^2 + 1/(2\Delta_x)$, $d_2 = -2/\Delta_x^2$ and $d_3 = 1/\Delta_x^2 - 1/(2\Delta_x)$, is of order Δ_x^2.

The second main derivative Y is approximated either by

$$Yu(x,y,t) \ \sim \ Y^+_{\Delta_t} u(x,y,t) = \frac{\widetilde{u}(x,y,t) - \widetilde{u}(x,y-(x-y)\Delta_t, t+\Delta_t)}{\Delta_t}, \tag{2.11}$$

or by

$$Yu(x,y,t) \ \sim \ Y^-_{\Delta_t} u(x,y,t) = \frac{\widetilde{u}(x,y+(x-y)\Delta_t, t-\Delta_t) - \widetilde{u}(x,y,t)}{\Delta_t}, \tag{2.12}$$

where $\widetilde{u}(x,y,t)$ denotes the linear interpolation of u at the point (x,y,t) based on the two nearest grid points. Specifically,

$$\widetilde{u}(x,y,t) = (1-\gamma) u(x,\widetilde{y},t) + \gamma u(x,\widetilde{y}+\Delta_y,t), \tag{2.13}$$

where $\gamma = (y - \widetilde{y})/\Delta_y$ and $\widetilde{y} = [y/\Delta_y]\,\Delta_y$ denoting by $[\cdot\,]$ the integer part. Since $\widetilde{u}(x,y,t)$ approximates $u(x,y,t)$ with an error of the order of Δ_y, then the approximations (2.11) and (2.12) are of the order of $\Delta_t + \Delta_y$. We remark that interpolation (2.13) is necessary because (x,y,t) and $(x, y-(x-y)\Delta_t, t-\Delta_t)$ cannot both belong to the same uniform grid. In [4] a different change of variables has been proposed in place of (2.3). That approach allowed for both the points to belong to the grid, but at the cost of imposing the grid size condition $\Delta_y = \Delta_x\Delta_t$.

The discrete operators $\mathcal{L}_G^+$ and $\mathcal{L}_G^-$ are defined by

$$\mathcal{L}_G^{\pm} u = a(D^2_{\Delta_x} u - D_{\Delta_x} u) + Y^{\pm}_{\Delta_x} u \tag{2.14}$$

and approximate $\mathcal{L}$ in the sense that

$$\|\mathcal{L}u - \mathcal{L}_G^{\pm} u\|_{L^\infty} \leq C\left(\Delta_x^2 + \Delta_t + \frac{\Delta_y^2}{\Delta_t}\right), \tag{2.15}$$

for some positive constant C depending on the L^∞-norms of a, $\partial_{xxx}u$, $\partial_y u$, $\partial_x^4 u$, $Y^2 u$, $\partial_{xx} Y u$, and $\partial_{xxy}u$ on the domain.

Hereafter, we refer to $\mathcal{L}_G^+$ and $\mathcal{L}_G^-$ respectively as explicit and implicit schemes for the discretization of $\mathcal{L}$. The implicit scheme is unconditionally stable, while the stability condition for the explicit method is given by $\Delta_t \leq \frac{\Delta_x^2}{2\sup a}$ and $\Delta_x < 2$ (cf. [4])

2.2 Boundary conditions

The numerical solution of (2.2) by finite-difference methods requires the discretization of the equation in a bounded region and the specification of some initial-boundary conditions. More precisely, we approximate the Cauchy problem (2.2)-(2.5) in the cylinder

$$Q = \{(x,y,t) \;\mid\; |x| < \mu,\ |y| < \nu \ \text{ and } \ 0 < \tau < \lambda T\}, \tag{2.16}$$

for some suitably large μ, ν. By transformation (2.3), this corresponds to the initial-boundary value problem for (1.3) in the domain

$$\{(S,M,t) \;\mid\; Ke^{-\mu-rt} < S < Ke^{\mu-rt},\ |M| < \nu \ \text{ and } \ 0 < t < T\}.$$

The conditions on the parabolic boundary of Q, defined by

$$\partial_P Q = \partial Q \cap \{(x,y,t) \;\mid\; \tau < \lambda T\},$$

are set as follows:

$$u(x,y,0) = (e^x - 1)^+, \qquad \text{for } \ x \in [-\mu,\mu],\ y \in [-\nu,\nu]; \tag{2.17}$$

moreover, we set

$$(\partial_{xx}u - \partial_x u)(\pm\mu, y, t) = 0, \qquad \text{for } \ y \in \,]-\nu,\nu[,\ t \in \,]0,\lambda T[. \tag{2.18}$$

We note explicitly that (2.18) corresponds to condition $\partial_{SS} f = 0$ in the original variables, which is somehow standard in the Black&Scholes framework.

It is remarkable that the approximation of $\mathcal{L}$ by its main derivatives allows to avoid imposing conditions on the lateral boundary $\{y = \pm\nu\}$, provided that ν is suitably large. To be more specific, let us first introduce some notation. Fixed $i_0, j_{0,n} \in \mathbb{N}$ for $n \in \mathbb{N} \cup \{0\}$, we denote

$$u^n_{i,j} = u(i\Delta_x, j\Delta_y, n\Delta_t), \qquad i, j \in \mathbb{Z}, \quad |i| \le i_0, \quad |j| \le j_{0,n}.$$

Applying the discrete operator in (2.10) to $u^n_{i,j}$ for $|i| \le i_0 - 1$ gives

$$D_{\Delta_x^2} u^n_{i,j} - D_{\Delta_x} u^n_{i,j} = (d_1 u^n_{i-1,j} + d_2 u^n_{i,j} + d_3 u^n_{i+1,j}). \tag{2.19}$$

Consider now the discretization (2.12) and assume that $(x, y, t) = (i\Delta_x, j\Delta_y, n\Delta_t)$ belongs to the grid. Then we have

$$\widetilde{u}(x, y + (x - y)\Delta_t, t - \Delta_t) = (1 - \gamma)\, u^{n-1}_{i,j+k} + \gamma\, u^{n-1}_{i,j+k+1}$$

and

$$Y^-_{\Delta_t} u^n_{i,j} = \frac{1}{\Delta_t} \left(u^n_{i,j} - (1 - \gamma) u^{n-1}_{i,j+k} - \gamma u^{n-1}_{i,j+k+1} \right), \tag{2.20}$$

where k and γ are, respectively, the integer and fractional part of $(x - y)\Delta_t/\Delta_y$, that is

$$k = \left[\frac{(x - y)\Delta_t}{\Delta_y} \right] = \left[\left(i \frac{\Delta_x}{\Delta_y} - j \right) \Delta_t \right] \qquad \text{and} \qquad \gamma = \left| \frac{(x - y)\Delta_t}{\Delta_y} - k \right|. \tag{2.21}$$

Applying the discrete operator $\mathcal{L}^-_G$ to $u^n_{i,j}$ reads

$$a_{ij} \left(D_{\Delta_x^2} u^n_{i,j} - D_{\Delta_x} u^n_{i,j} \right) + Y^-_{\Delta_t} u^n_{i,j} = 0, \qquad |i| \le i_0 - 1,\ |j| \le j_{0,n}, \tag{2.22}$$

where $a_{ij} = a(i\Delta_x, j\Delta_y)$. On the other hand, assuming $\mu = i_0\Delta_x$, condition (2.18) is equivalent to

$$Y^-_{\Delta_t} u^n_{i,j} = 0, \qquad i = \pm i_0,\ |j| \le j_{0,n}. \tag{2.23}$$

Next we fix $j_{0,N} \in \mathbb{N}$ and examine the domain of dependence of the set of values

$$U_N = \left\{ u^N_{i,j} \mid \quad |i| \le i_0,\ |j| \le j^N_0 \right\};$$

more precisely, for $0 \le n \le N - 1$, we specify $j_{0,n}$ as the maximum of the set of the indexes j's such that U_N depends on $u^n_{i,j}$ through conditions (2.22)-(2.23). Moreover we set $\nu_n = j_{0,n}\Delta_y$. Since, by (2.20) and (2.21), it holds

$$j_{0,n-1} = j_{0,n} + \left[j_{0,n}\Delta_t + i_0 \frac{\Delta_x\Delta_t}{\Delta_y} \right] + 1 \le j_{0,n}(1 + \Delta_t) + i_0 \frac{\Delta_x\Delta_t}{\Delta_y} + 1,$$

it follows that

$$\nu_{n-1} \le \nu_n(1 + \Delta_t) + \mu\Delta_t + \Delta_y$$

and thus $\nu_{N-n} \le z_n$, where z_n is defined by the difference equation

$$z_{n+1} = (1 + \Delta_t) z_n + \mu\Delta_t + \Delta_y, \qquad\qquad z_0 = \nu_N,$$

which has solution[1] $z_n = (1+\Delta_t)^n(y_0+\mu+\Delta_y/\Delta_t) - \mu - \Delta_y/\Delta_t$. Finally, we deduce

$$\nu_n = (1+\Delta_t)^{(N-n)}\,(\nu_N+\mu+\Delta_y/\Delta_t) - \mu - \Delta_y/\Delta_t$$

and

$$\begin{aligned}\nu_0 &= (1+\Delta_t)^{\lambda T/\Delta_t}\,(\nu_N+\mu+\Delta_y/\Delta_t) - \mu - \Delta_y/\Delta_t \\ &\le e^{\lambda T}\nu_N + (e^{\lambda T}-1)(\mu+\Delta_y/\Delta_t).\end{aligned}$$

Hence, if we assume that

$$\frac{\Delta_y}{\Delta_t} = C_0 \tag{2.24}$$

for some constant C_0, then the width of the initial region can be chosen independently of the refinement of the grid. Thus we have proved the following

Claim: *in order to approximate the solution $u(x,y,\lambda T)$ for $|x| \le \mu$ and $|y| \le \nu_N$, conditions on the lateral boundary $\{y = \pm\widetilde{\nu}\}$, where $\widetilde{\nu} = e^{\lambda T}\nu_N + (e^{\lambda T}-1)(\mu + C_0)$, are superfluous.*

Alternatively, one can solve (2.2) on the prism

$$\widetilde{Q} \equiv \{(x,y,t) \mid \quad |x| < \mu,\ |y| < e^{\lambda T - t}\nu_N + (e^{\lambda T - t}-1)(\mu + C_0),\ 0 < t < \lambda T\},$$

rather than on the whole cylinder Q. Also notice that, under condition (2.24) the approximation error of $\mathcal{L}_G^{\pm}$ in (2.15) reduces to an order of $\Delta_x^2 + \Delta_t$.

2.3 Discretization and linear systems

In order to formulate the discretization of the equation as a block bidiagonal linear system let define $I = 2i_0 + 1$, $J_n = 2j_{0,n} + 1$ and denote by $u^n \in \mathbb{R}^{IJ_n}$ the vector containing the values $u^n_{i,j}$ for $|i| \le i_0$ and $|j| \le j_{0,n}$: those values are sorted by the couple of indices (j,i) in lexicografic order.

Let consider now, the application of the discrete operator $Y^-_{\Delta_t}$ in (2.20) to the vector u^n. The generic element $Y^-_{\Delta_t}u^n_{i,j}$ is the linear combination of the corresponding element in u^n and two elements, $u^{n-1}_{i,j+k}$ and $u^{n-1}_{i,j+k+1}$ of u^{n-1}. Thus, applying $Y^-_{\Delta_t}$ to u_n is equivalent to the difference of two linear operators, $\Delta_t^{-1}\mathcal{I}_\backslash$ and $\Delta_t^{-1}Z_n$, applied respectively to u^n and u^{n-1}, where $\mathcal{I}_\backslash$ denotes the identity operator in $\mathbb{R}^{IJ_n}$. Specifically, the vector with elements $Y^-_{\Delta_t}u^n_{i,j}$ is given by

$$\frac{1}{\Delta_t}(u^n - Z_n u^{n-1}),$$

where $Z_n \in \mathbb{R}^{IJ_n \times IJ_{n-1}}$ is the matrix such that the entry corresponding to the index i,j of Zu^{n-1} is given by

$$(1-\gamma)u^{n-1}_{i,j+k} + \gamma u^{n-1}_{i,j+k+1}.$$

[1]The solution of the difference equation $z_{n+1} = \alpha z_n + \beta$, with initial value z_0 and $\alpha \ne 1$, is given by $z_n = \alpha^n(z_0 - z_*) + z_*$, where $z_* = \beta/(1-\alpha)$ the equilibrium value.

Then it turns out that the linear system (2.22) can be rewritten in matrix form

$$(\mathcal{I}_n + \Delta_t A_n D_n)u^n - Z_n u^{n-1} = 0, \tag{2.25}$$

for $1 \leq n \leq N$, where $A \in \mathbb{R}^{IJ_n \times IJ_n}$ is the diagonal matrix with elements a_{ij} and

$$D_n = \begin{pmatrix} \check{D} & 0 & \cdots & 0 \\ 0 & \check{D} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \check{D} \end{pmatrix} \qquad \text{and} \qquad \check{D} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ d_1 & d_2 & d_3 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & d_1 & d_2 & d_3 \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix}, \tag{2.26}$$

are tridiagonal matrices of order IJ_n and I, respectively.

Similarly, combining the forward and backward schemes allows to derive the θ-method:

$$\theta \Delta_t A_n D_n u^n + (1-\theta)\Delta_t Z_n A_{n-1} D_{n-1} u^{n-1} - Z_n u^{n-1} + u^n = 0,$$

or

$$\bar{A}_1^n u^n = \bar{A}_2^n u^{n-1}, \qquad 1 \leq n \leq N, \tag{2.27}$$

with $\bar{A}_1^n = (\mathcal{I}_n + \theta \Delta_t A_n D_n)$ and $\bar{A}_2^n = Z_n(\mathcal{I}_n - (1-\theta)\Delta_t A_{n-1} D_{n-1})$. As usual, the θ-method reduces to the explicit, implicit or Crank-Nicholson schemes when $\theta = 0, 1$ or 0.5, respectively. Notice that the θ-method is unconditionally stable for $0.5 \leq \theta \leq 1$. The matrices $\bar{A}_1^n$ and D_n have an identical structure, specifically they are block diagonal with tridiagonal blocks. Thus, the computational cost required to solve (2.27) is of the order of IJ_n. Furthermore, the structure of the matrices can be exploited to design computationally efficient and/or parallel algorithms for the solution of the PDE (2.2).

Finally, by setting $N = \lambda T/\Delta_t$ it turns out that considering (2.27) for $n = 1, \ldots, N$ and imposing the initial conditions $u_0 = v_0$ is equivalent to the linear system

$$\begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ -\bar{A}_2^1 & \bar{A}_1^1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & -\bar{A}_2^{N-1} & \bar{A}_1^{N-1} & 0 \\ 0 & \cdots & 0 & -\bar{A}_2^N & \bar{A}_1^N \end{pmatrix} \begin{pmatrix} u^0 \\ u^1 \\ \vdots \\ u^{N-1} \\ u^N \end{pmatrix} = \begin{pmatrix} v^0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

or, with the appropriate substitution,

$$\bar{A}\bar{u} = \bar{v}, \tag{2.28}$$

where $\bar{A}$ is block bidiagonal and $\bar{u} \in \mathbb{R}^q$ where

$$q = \sum_{n=0}^{N} IJ_n. \tag{2.29}$$

Existence and uniqueness of the solution are related to the non-singularity of $\bar{A}_1^n$ and, in view of the expressions of d_1, d_2, d_3, are clearly guaranteed if $\frac{\Delta_t}{\Delta_{x^2}}$ is suitably small. Stability and numerical stability are driven by the properties of the matrices $\bar{A}_1^n$ and $\bar{A}_2^n$.

In the calibration procedure, we consider a volatility σ (smoothly) depending on a parameter vector $\alpha \in \mathbb{R}^p_+$ for some $p \in \mathbb{N}$. Thus we rewrite the dynamics (1.2) of the price as

$$dS_t = \mu(D_t)S_t dt + \sigma(D_t;\alpha)S_t dW_t, \tag{2.30}$$

and denote by $u(\cdot;\alpha)$ the solution to (2.2)-(2.5) with $a(x,y;\alpha) = \frac{\sigma^2(x-y;\alpha)}{2\lambda}$. For what follows, it will be useful to compute the derivatives of u w.r.t the parameters α. These play the role of the Vega in the standard Black&Scholes approach. For $k = 1,\dots,p$, the derivative $\partial_{\alpha_k} u$ satisfies the PDE

$$\mathcal{L}(\partial_{\alpha_k} u) = -(\partial_{\alpha_k} a)(\partial_{xx} u - \partial_x u), \tag{2.31}$$

in $\mathbb{R}^2 \times]0, \lambda T[$, with initial conditions

$$\partial_{\alpha_k} u(x,y,0) = 0, \qquad \text{for } (x,y) \in \mathbb{R}^2. \tag{2.32}$$

A linear system, similar to (2.28), derives from the discretization of the PDE (2.31). In that case, the term $\bar{v}$ in (2.28) should include the effects of the right hand side of (2.31).

3 Calibration

We consider the calibration of the Hobson-Rogers model. The problem consists in determining the volatility function σ correspondent to a given observed solution function u. Actually, we only look for the function σ in (2.4) for which the PDE (2.2) best approximates the observations. Indeed the presence of pricing errors, inconsistencies and/or inefficiency in the market may not allow to fit exactly the data. Moreover u is observed only at a finite number of points, thus specific restrictions should be imposed on σ in order to obtain a well posed problem.

In what follows, we assume that σ belongs to a space of splines, that is

$$\sigma = \sigma(\cdot\,;\alpha) = \sum_{1=1}^{p} \alpha_i s_i(\cdot) \tag{3.1}$$

where $\alpha_i \in \mathbb{R}_+$ and s_i represents the i-th basis function for $i = 1,\dots,p$. Let $u(x,y,t;\alpha)$ denote the solution to the Cauchy problem for (2.2)-(2.5) corresponding to $\sigma(\cdot\,;\alpha)$. This defines a mapping from $\mathbb{R}^p_+$ to $C^\infty(\mathbb{R}^2 \times]0,\lambda T[)$. The scope of this section is to develop and test a numerical procedure to "invert" that function. Let further notice that, by construction, in the discrete problem (2.28) the coefficient matrix $\bar{A}$, and thus the corresponding solution $\bar{u}$, are functions of α. This legitimates the writing $\bar{A}(\alpha)$ and $\bar{u}(\alpha)$, used hereafter.

Let b_i be the observed value of u at the point $z_i \equiv (x_i, y_i, t_i)$, for $i = 1,2,\dots,M$. Furthermore, let $u(x_i,y_i,t_i;\alpha)$ be the solution of the PDE for a given α at the observation point z_i. Now, since the point z_i may not belong to the grid G, it cannot be directly approximated by using the discretizations introduced in the previous section. Instead, the value of u will be approximated from the solution of (2.28) by using a bilinear interpolation of the eight nearest points of the grid. Let us denote by $\widetilde{u}_i(\alpha)$ that approximation, then

$$\widetilde{u}_i(\alpha) = \sum_{k=1}^{8} P_{i,j_{ik}} \bar{u}_{j_{ik}}(\alpha), \tag{3.2}$$

with $P_{i,j_{ik}} \geq 0$ and

$$\sum_{k=1}^{8} P_{i,j_{ik}} = 1,$$

for $i = 1, 2, \ldots, M$, $1 \leq j_{ik} \leq M$.

Let now denote by $\widetilde{u}(\alpha)$ and $b \in \mathbb{R}^M$ the vector with elements $\widetilde{u}_i(\alpha)$ and b_i, respectively. It follows that (3.2) can be written in the compact form $\widetilde{u}(\alpha) = P\bar{u}(\alpha)$ where $P \in \mathbb{R}^{M \times q}$, with q as in (2.29), is a matrix with eight non-zero elements per row. Thus, the error made in fitting the observed values for a given α is given by

$$\varepsilon(\alpha) = P\bar{u}(\alpha) - b, \tag{3.3}$$

where $\varepsilon \in \mathbb{R}^M$.

The quality of the approximation can be measured using the Residual Sum of Squares (RSS) $|\varepsilon|^2$ while the regularity of the coefficient function can be measured by the norm $|\alpha|^2$. The scope of the calibration is to find the nicest σ that best fits the data. These considerations allows to formulate the nonlinear least squares (NLLS) problem

$$\min_{\alpha \in \mathbb{R}^p_+} \varphi(\alpha), \qquad \text{where} \qquad \varphi(\alpha) = |P\bar{u}(\alpha) - b|^2 + \rho|\alpha|^2, \tag{3.4}$$

and $\bar{u}(\alpha)$ is the solution to (2.28) correspondent to $\bar{A} = \bar{A}(\alpha)$ and $\rho \geq 0$.

The NLLS problem (3.4) is solved using a Gauss-Newton method (see, for instance, [9]). The Hessian of φ is approximated by using only first order derivatives of $\bar{u}$, that is

$$\partial_{\alpha\alpha}\varphi(\alpha) = \sum_{i=1}^{M} \varepsilon_i H_i + \mathcal{J}^T P^T P \mathcal{J} + \rho I \simeq \mathcal{J}^T P^T P \mathcal{J} + \rho I,$$

where ε_i is the i-th element of ε, $H_i = [\partial_{\alpha_j \alpha_k} \widetilde{u}_i(\alpha)]_{jk}$ is the Hessian of $\widetilde{u}_i(\alpha)$ and $\mathcal{J} = (\partial_{\alpha_1}\bar{u} \ \cdots \partial_{\alpha_p}\bar{u})$. Two equivalent approaches can be used to derive $\mathcal{J}$. In the first the PDE (2.31) is discretized and then solved. Alternatively, from (2.28), it follows that

$$\bar{A}(\alpha)\partial_{\alpha_k}\bar{u}(\alpha) = -\big(\partial_{\alpha_k}\bar{A}(\alpha)\big)\bar{u}(\alpha). \tag{3.5}$$

Notice that, when computing the gradient or the approximated Hessian of φ, it is necessary to recompute the Jacobian J and this require the solution of a set of $p+1$ discretized PDE (2.28). This represents a huge computational cost especially when the number of parameters is high.

The calibration procedure here described is applied to closing day European Option prices on S&P 500 observed in the period from February to November 1993. This dataset have been firstly used by Aït-Sahalia and Lo [1]. All the available strikes have been considered. In order to avoid inefficiencies when near to maturity, options with time-to-maturity smaller than a week have been discarded. Furthermore, those with a time-to-maturity greater than three months have also been discarded because they does not contain too much information and because in this way the computational complexity of the procedure is reduced. The resulting dataset contains 8201 observations.

For the i-th observation, the dataset contains the trading date t_i, the time to expiration T_i, the strike price K_i and the implied volatility σ_i of the corresponding option price. Then, from

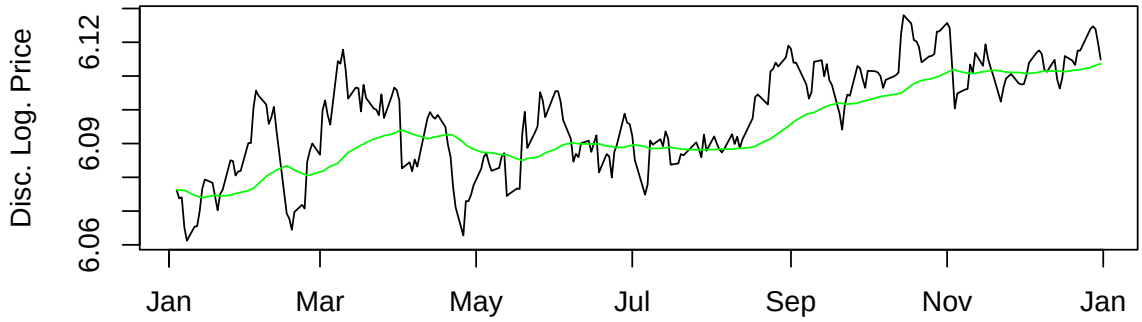

Figure 1: Discounted log prices Z_t and corresponding trend $Z_t - D_t$ with $\lambda = 7.5$ of the S&P 500 index for the year 1993.

the trading date t_i and the interest rate r_i, the closing price of the index S_i is retrieved and the deviation from the trend D_i is computed by a discretization of (1.1). A value of $\lambda = 7.5$ has been empirically chosen for that computation. Figure 1 shows the discounted log prices Z_t and the corresponding trend $Z_t - D_t$.

The market volatility smiles are represented in Figure 2, where in each box the implied volatility is plotted against the moneyness $e^{rT_i}S_i/K_i$ for a given interval of the deviation D_t and of the time-to-maturity T. These intervals are shown in the bars on the top and on the right of the figure. This figure shows how the smiles flatten as the time-to-maturity grows. On the other side, since the underlying is an index, the direction of the smile does not extremely change for different deviations D_t. Figures 3 and 4 show the volatility surface as a function of moneyness and deviation form the trend. These figures show the dependence of prices on the deviation which the Hobson&Rogers model tries to capture.

In the calibration procedure, the observations (x_i, y_i, t_i) and b_i are computed as

$$(x_i, y_i, t_i) = \big(\log(S_i/K_i) + r_iT_i,\, \log(S_i/K_i) + r_iT_i - D_i,\, \lambda T_i\big)$$

and $b_i = e^{r_iT_i}f_i/K_i$. The NLLS problem (3.4) is solved with respect to the values of the corresponding straddle obtained by using the standard Black&Scholes formula, that is $f_i = BS(S_i, T_i; K_i, \sigma_i, r_i)$. In the following tests, the calibration procedure has been performed over three months, precisely August, September and October. Figures 5, 6 and 7 show the market prices and those of the calibrated Hobson-Rogers model, with circles and lines, respectively. Three different days with different deviation from the mean have been chosen. Figure 5 show the result of an in-sample test, while in Figures 6 and 7 the calibrated model demonstrates fine predictability properties. Figure 8 shows the distribution of relative errors of prices obtained with the calibrated Hobson&Rogers model. Finally, relative errors for the whole period considered are shown in Figure 9, where the red line distinguishes in-sample from out-of sample errors.

The calibrated volatility function σ is plotted in Figure 10, where the nodes of the spline are represented by circles on the abscissas other two nodes exist at ± 100. Figure 11 shows the calibrated volatility function zoomed near the origin.

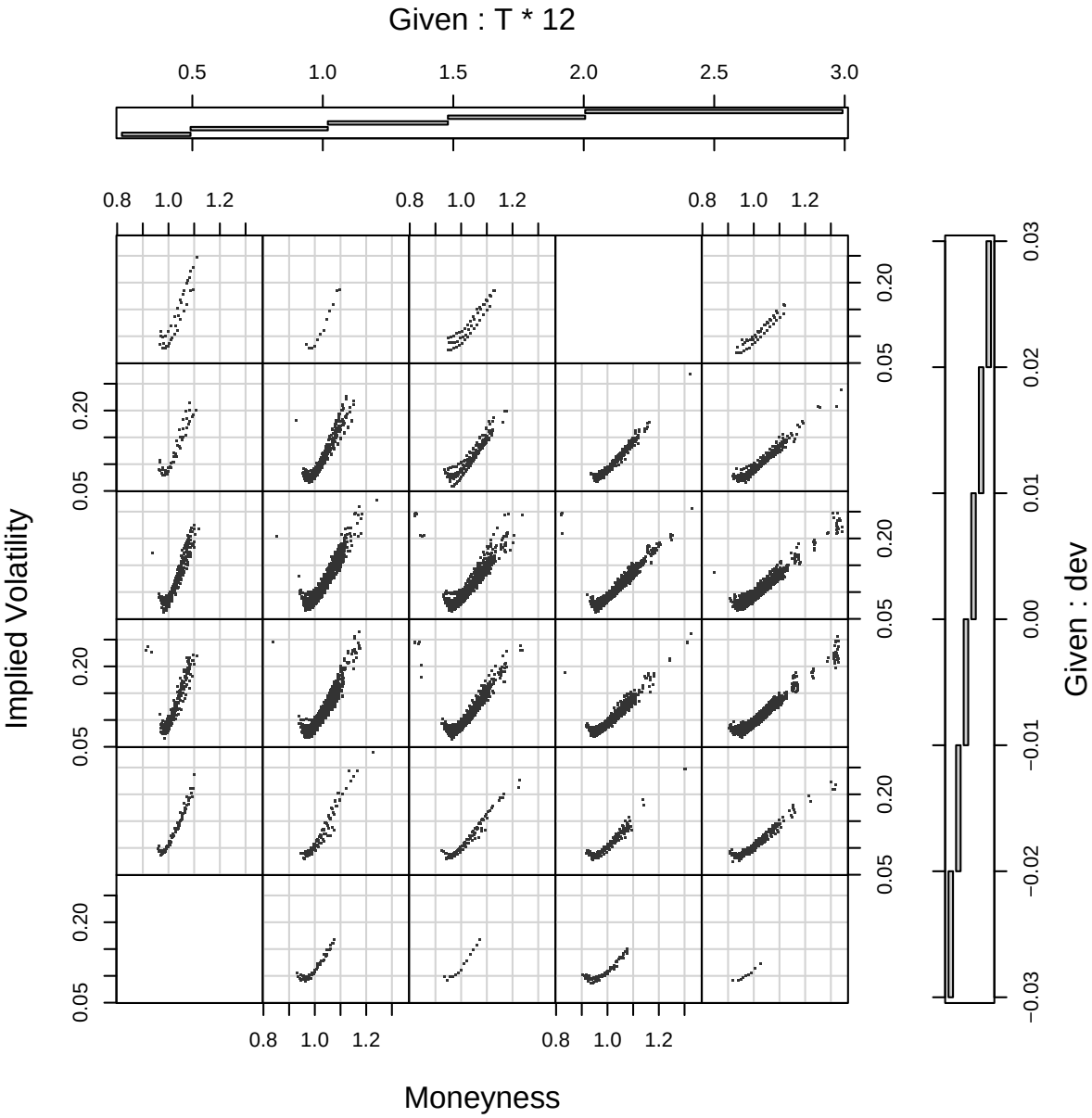

Figure 2: Implied volatilities for the Options on the S&P 500 index.

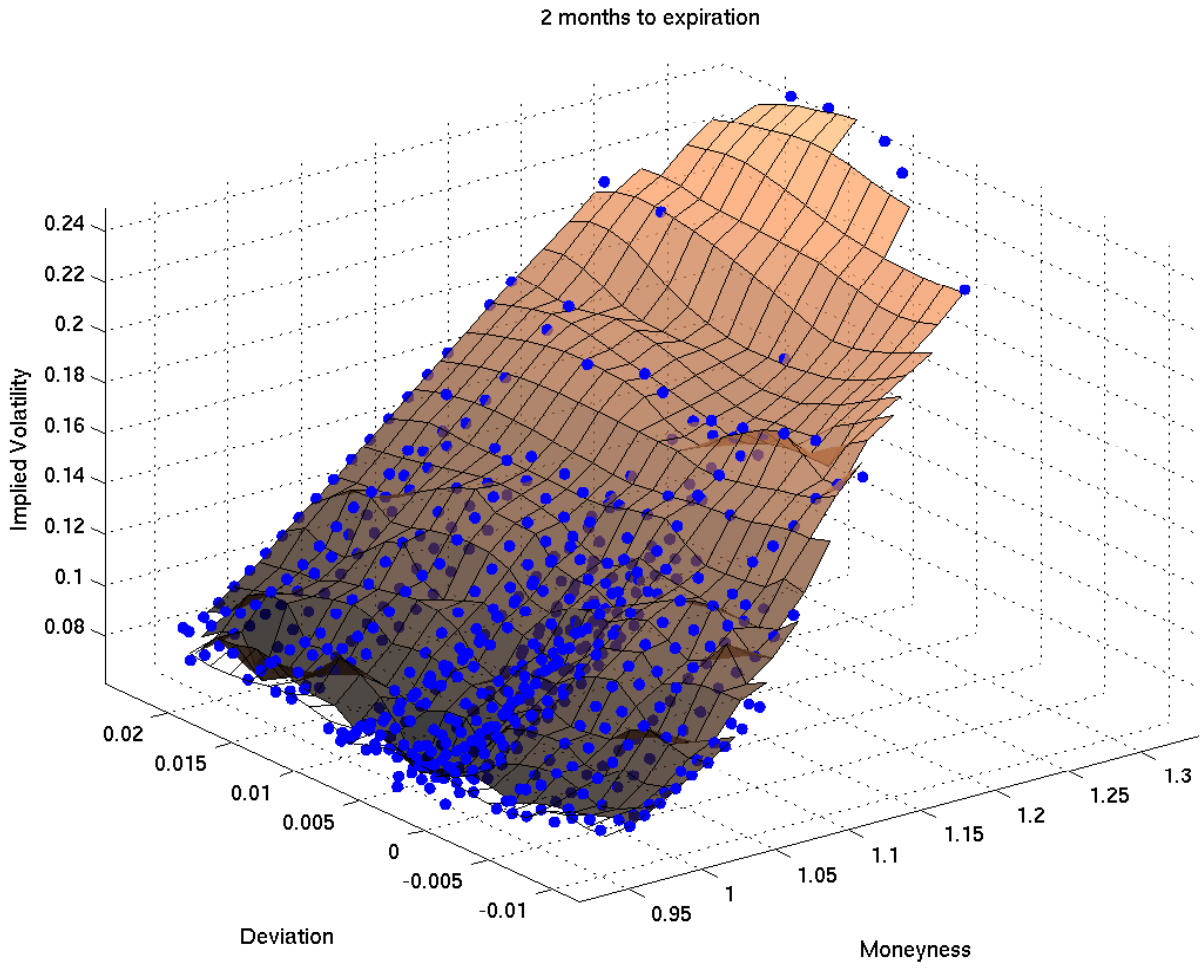

Figure 3: Implied volatility surface. The volatility, with time to maturity equal to two months, as a function of moneyness and deviation.

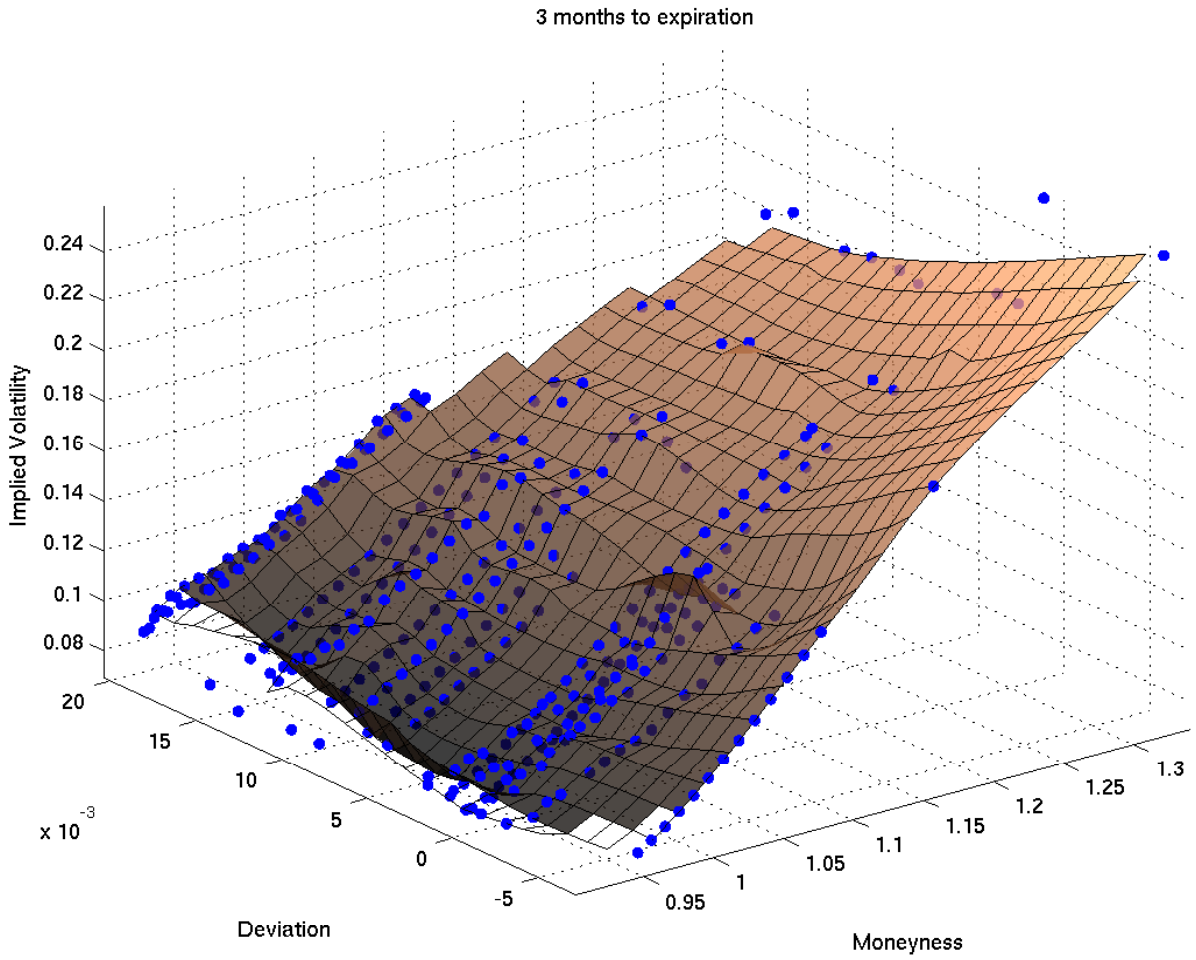

Figure 4: Implied volatility surface. The volatility, with time to maturity equal to three months, as a function of moneyness and deviation.

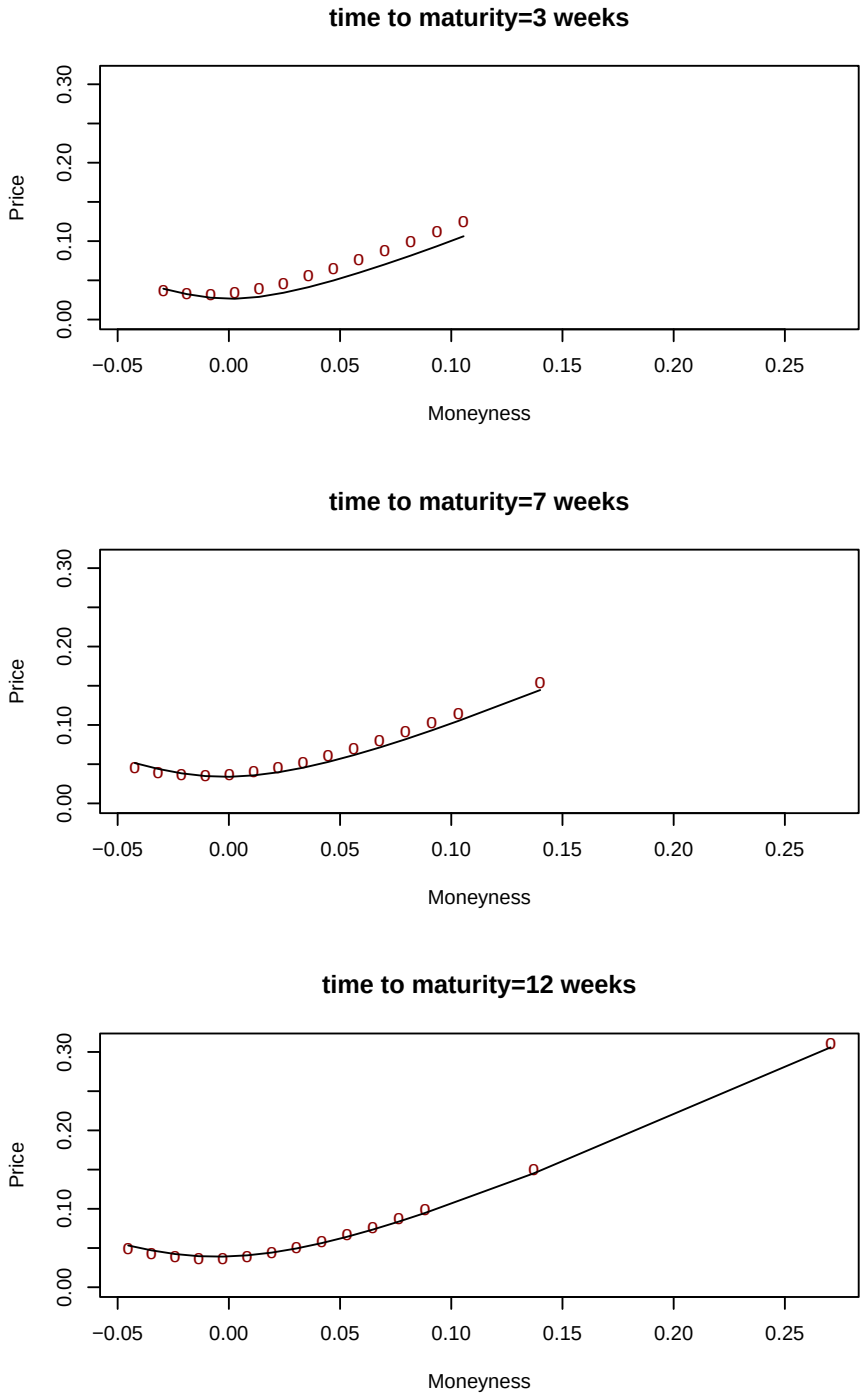

Figure 5: In-sample test: market and calibrated prices of straddles on August, 30. Calibration performed over three months: August, September and October.

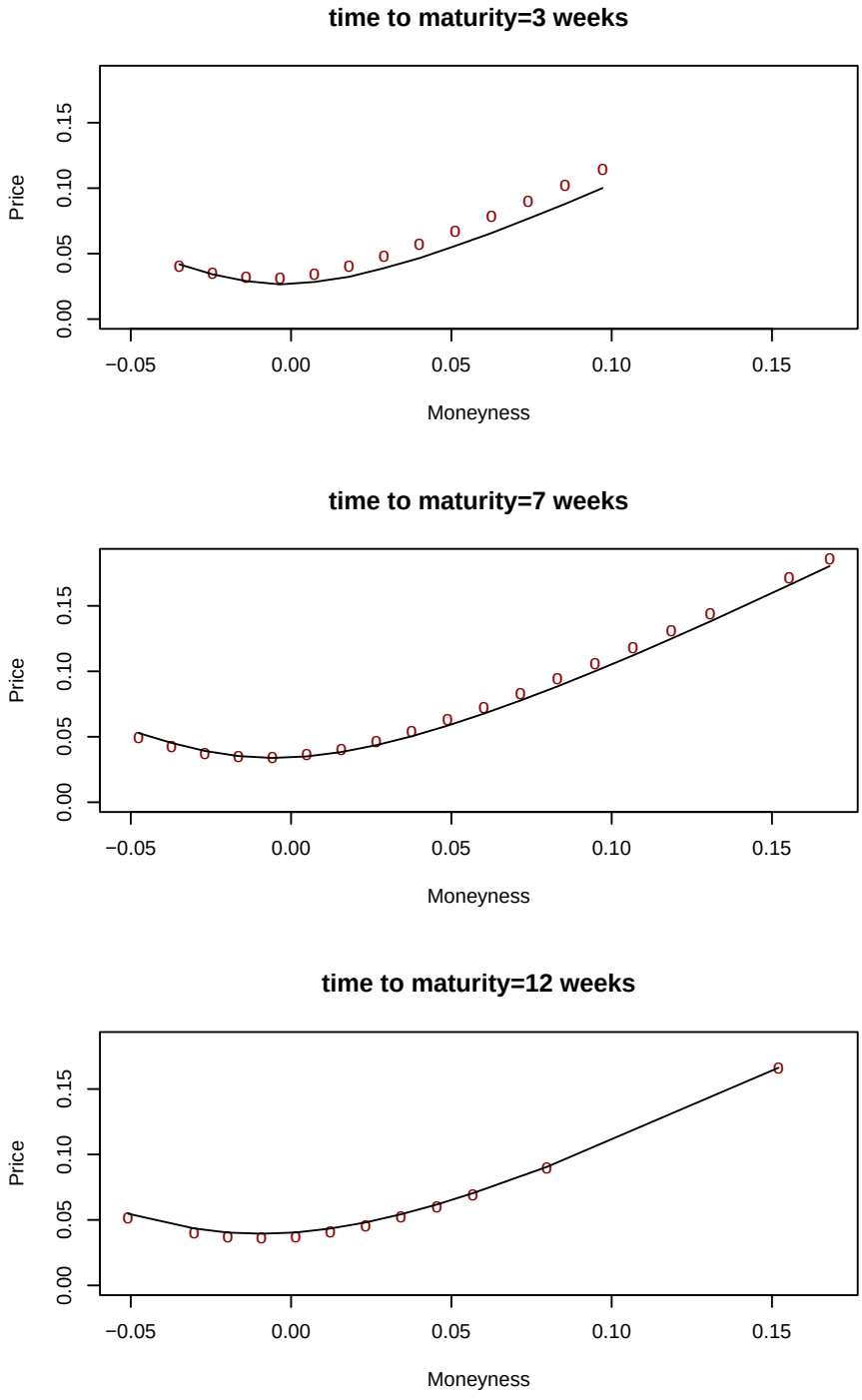

Figure 6: Out-of-sample test: market and calibrated prices of straddles on November, 1. Calibration performed over three months: August, September and October.

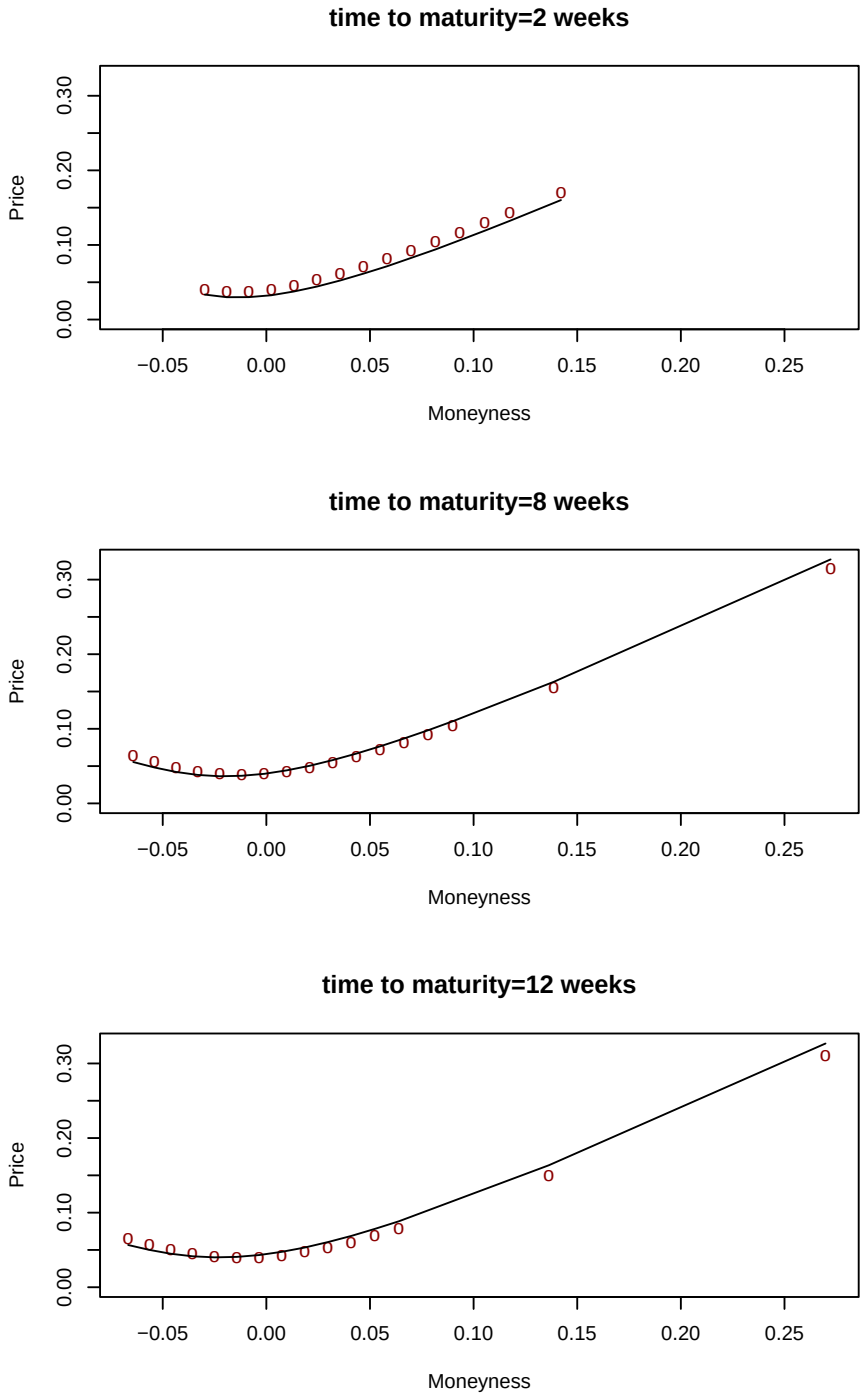

Figure 7: Out-of-sample test: market and calibrated prices of straddles on November, 30. Calibration performed over three months: August, September and October.

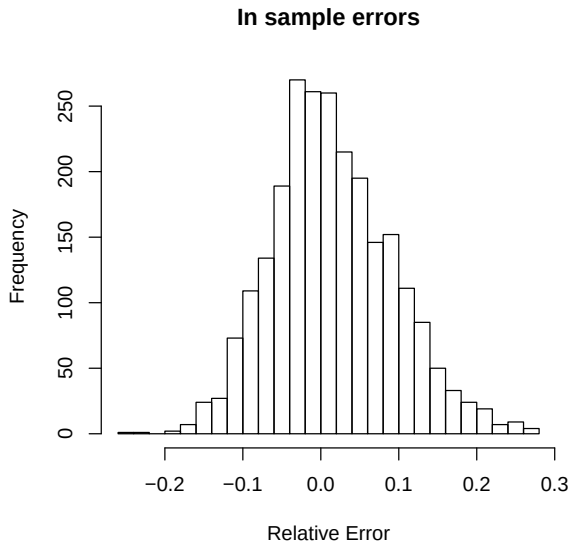

Figure 8: Distribution of relative errors of prices for the period August-October obtained with the calibrated Hobson&Rogers model in the same period.

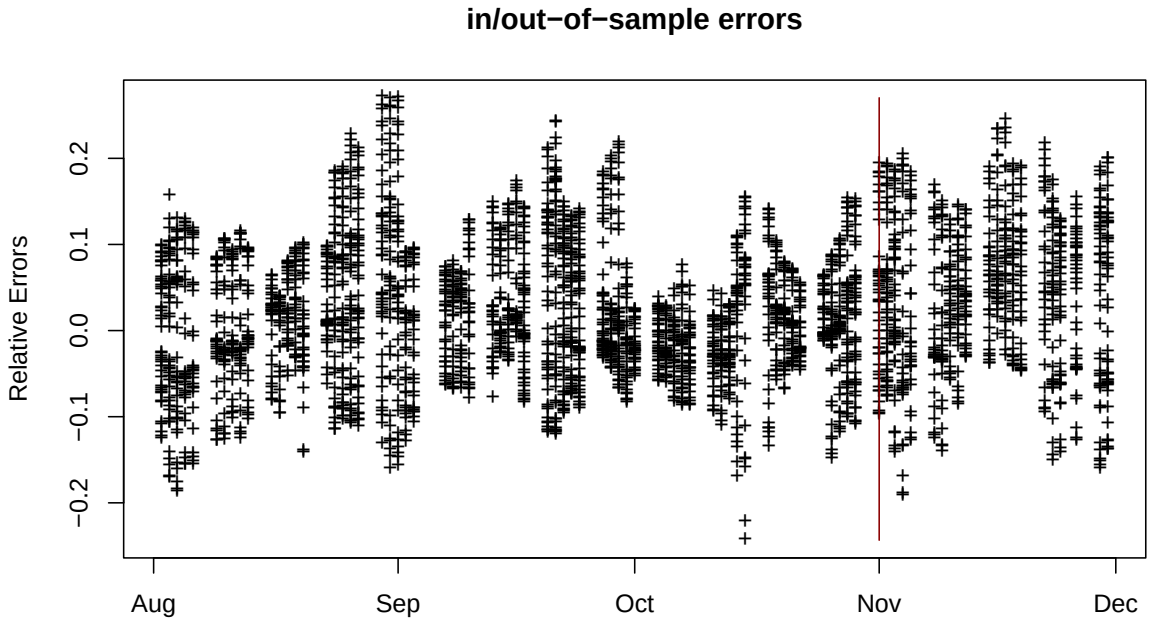

Figure 9: Relative errors for the whole period considered: the red line distinguishes in-sample from out-of sample errors.

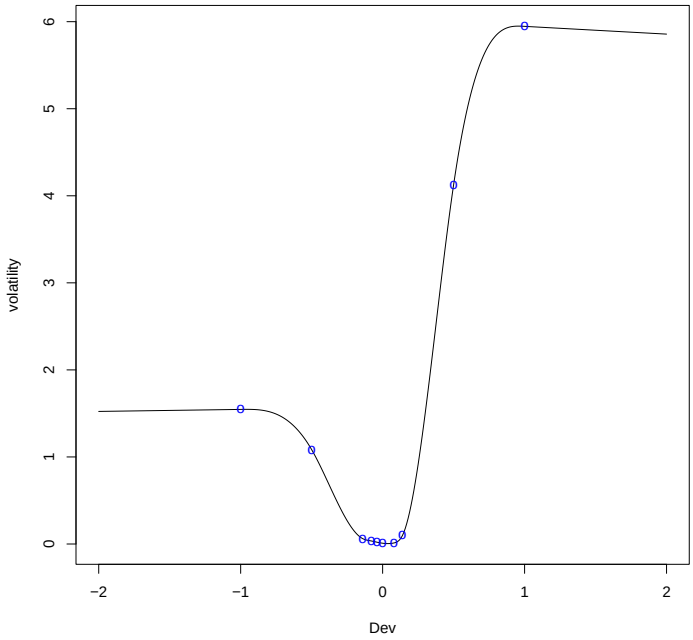

Figure 10: Calibrated volatility function $\sigma(D)$.

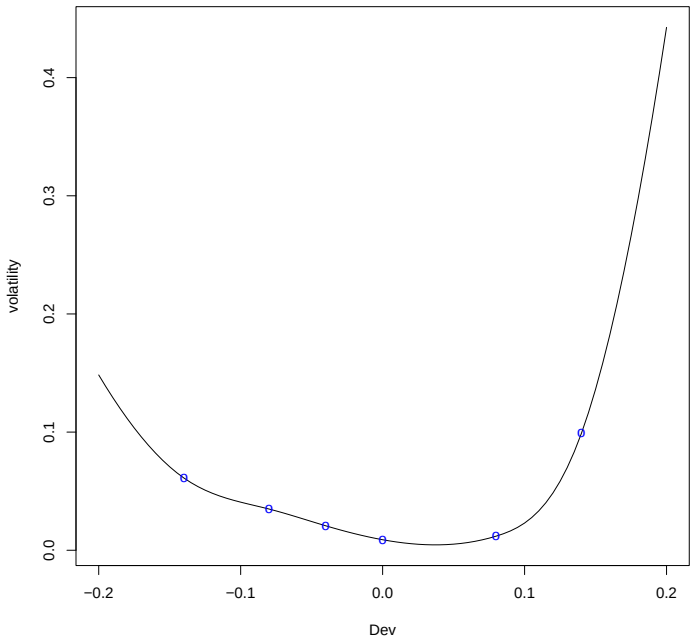

Figure 11: Calibrated volatility function $\sigma(D)$ zoomed near the origin.

References

[1] Y. Aït-Sahalia and A. W. Lo, *Nonparametric estimation of stete-price densities implicit in financial asset prices*, The Journal of Finance, LIII (1998), pp. 499–547.

[2] V. Blaka Hallulli and T. Vargiolu, *Financial models with dependence on the past: a survey*, preprint, (2004).

[3] C. Chiarella and K. Kwon, *A complete markovian stochastic volatility model in the HJM framework*, Asia-Pacific Financial Markets, 7 (2000), pp. 293–304.

[4] M. Di Francesco, P. Foschi, and A. Pascucci, *Analysis of an uncertain volatility model*, Journal of Economic Dynamics and Control, (2004). Submitted.

[5] M. Di Francesco and A. Pascucci, *On the complete model with stochastic volatility by Hobson and Rogers*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 460 (2004), pp. 3327–3338.

[6] M. Di Francesco and A. Pascucci, *On a class of degenerate parabolic equations of Kolmogorov type*, preprint, (2005).

[7] B. Dupire, *Pricing and hedging with smiles*, in Mathematics of derivative securities (Cambridge, 1995), vol. 15 of Publ. Newton Inst., Cambridge Univ. Press, Cambridge, 1997, pp. 103–111.

[8] G. Figà-Talamanca and M. L. Guerra, *Complete models with stochastic volatility: further implications*, Working Paper, Università della Tuscia, Facoltà di Economia, 5 (2000).

[9] R. Fletcher, *Practical Methods of Optimization*, John Wiley and Sons, Inc., 1987.

[10] D. G. Hobson and L. C. G. Rogers, *Complete models with stochastic volatility*, Math. Finance, 8 (1998), pp. 27–48.

[11] F. Hubalek, J. Teichmann, and R. Tompkins, *Flexible complete models with stochastic volatility generalising Hobson-Rogers*, preprint, (2004).

[12] E. Lanconelli, A. Pascucci, and S. Polidoro, *Linear and nonlinear ultraparabolic equations of Kolmogorov type arising in diffusion theory and in finance*, in Nonlinear problems in mathematical physics and related topics, II, vol. 2 of Int. Math. Ser. (N. Y.), Kluwer/Plenum, New York, 2002, pp. 243–265.