

Estimating implied volatility surfaces using Bayesian splines under shape restrictions

Piantino, Guilherme[†]

2023

Abstract This work develops a statistical model for estimating implied volatility surfaces, using information about the expectations of market agents contained in the market prices of options. The implied volatility curves are estimated by shape-constrained splines, using a Bayesian method (MCMC) that imposes no-arbitrage conditions on the price curve using shape restrictions.

Keywords: Implicit Volatility, Non-parametric Estimation, Bayesian Methods, Market Expectations.

JEL Codes: E3, C41, C43.

Resumo Este trabalho desenvolve um modelo estatístico para estimação de superfícies de volatilidade implícita, utilizando informações sobre as expectativas dos agentes de mercado contidas nos preços de mercado de opções. As curvas de volatilidade implícita serão estimadas por splines com restrição de forma, usando um método bayesiano (MCMC) que impõe condições de não arbitragem na curva de preço usando restrições de formato.

Palavras-chave: Volatilidade Implícita, Estimativa Não Paramétrica, Métodos Bayesianos, Expectativas de Mercado.

Códigos JEL: E3, C41, C43.

1. Introduction

The implied volatility can be considered as a measure of the expected future risk of a given underlying asset, as it is calculated using market prices of derivatives and thus incorporates the expectations of market participants. Therefore, accurately modeling and predicting implied volatility is extremely important, both theoretically and practically, as this would make it possible to predict future movements in assets or even more accurately price derivative instruments.

Based on this assumption, several studies were carried out with the aim of verifying whether the information contained in the implied volatility would make it possible to predict movements in the stock market, as well as its

[†]Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo

volatility. In (DAY and LEWIS, 1992), the predictive power of implied volatility was studied in comparison with that of historical data, using a GARCH model with the inclusion of implied volatility as an explanatory variable. They concluded that both have an impact on future volatility. In (XING et al., 2010), the impact that the shape of the implied volatility curve (skew and kurtosis) has on future stock market returns was verified. The results were consistent with the hypothesis that there is additional information embedded in the volatility curve. Considering an article that deals with the Brazilian market, we have the study by (VICENTE and GUEDES, 2010) in which a relationship between implied volatilities and realized volatilities was identified.

Therefore, the idea of this work is to build a statistical model to estimate the implied volatility, contemplating the information contained in the options market about the expectations of market participants in relation to future movements of stock prices and thus, consequently, be able to use this model to predict future volatility and more accurately price option contracts.

The main original contribution is in the method used to estimate the implied volatility, with the use of Bayesian inference, through a new non-parametric formulation for the estimation of the implied volatility, through *splines* with shape restrictions, whose application to the Brazilian options market is still unprecedented. As the method is non-parametric, it is not necessary to assume a functional form or to have a previously established distribution, making the model more flexible, but imposing the non-arbitrage conditions necessary for the construction of the price curve and implicit volatility. Added to this, the shape constraints make the model more robust to the choice of interpolation nodes and parsimonious in the fit to the data, making it balanced between a good fit while avoiding *over-fitting*. Furthermore, when using Bayesian inference, it is possible to go beyond point estimates, with the construction of posterior distributions and confidence intervals for the regressions, which facilitates the quantification of uncertainty for the estimated values.

This work is divided into 5 parts. In the second section we review the existing literature on the construction of implied volatility curves and surfaces, and the use of Bayesian methods for this purpose. In the third section, we approach the methodology used to calculate the implied volatilities from market prices, and subsequently the implied volatility curves using *splines* with shape restriction via Bayesian estimation. In the fourth section, we present the results obtained and finally, in the last one, the conclusions.

1.1 Objective and justification of relevance

The Black–Scholes–Merton (BSM) ((BLACK and SCHOLES, 1973)) model serves as an important conceptual reference for the pricing of options, including being a kind of “language” among *traders*, academics and practitioners who work in this market. However, although it is the most known and widely used, this model has some simplifications in its assumptions, deviating from what is observed in the price generating mechanism in the financial markets. As a result, the model ends up generating incorrect pricing mainly for out-of-the-money and in-the-money options, as observed in (RUBINSTEIN, 1985).

Among the main assumptions assumed by the model, we can mention that volatility is constant, not being a function of time or the underlying asset; and that log-returns are normally distributed, that is, prices follow a lognormal distribution. When using the model considering these assumptions in the presence of time-varying volatility and non-log-Gaussian distributions, incorrect pricing and distortions in calculated prices are obtained, mainly for out-of-the-money and in-the-money options.

One of the most notorious evidences that volatility is not constant throughout the exercise prices is the shape of the implied volatility curve, which is characteristic, with higher volatility values at the extremes, forming what is conventionally called the volatility’ smile. That is, options with strike prices farther from the asset’s current value tend to have higher volatilities than those with strike prices closer to the price of the underlying asset.

A first alternative to the Black and Scholes model was developed in (COX et al., 1979). This is the Cox-Ross-Rubinstein (CRR) model, which became popularly known as the binomial model. Such a model assumes that asset prices move according to a discrete binomial process. A price tree is formed, representing the different paths through which the price of the underlying asset can follow throughout the life cycle of the option. At the end of each node, the option price is calculated, and from there, the value of the option at the preceding nodes is calculated sequentially. This model is recommended when the asset is paying dividends and also for the pricing of American options, but can be modified to the presence of a non-constant volatility.

(HESTON, 1993) developed a model with stochastic volatility, assuming that volatility follows a stochastic process defined by an additional stochastic differential equation in the model. Through this model, the price of an option can be derived as an integral of the future density of asset prices, which in turn can be calculated through an inverse Fourier transform. A rule of thumb used to correct prices in the BSM model is, instead of using historical volatility for all strike prices of a given maturity, prices are quoted in terms of another parameter, which is whether of implied volatility.

In (SCHMALENSSEE and TRIPPI, 1978), a numerical method was used to calculate the volatility of the underlying asset based on the prices reflected in the options market, a volatility that became known as implied volatility. Conceptually, implied volatility is defined as the volatility that, when included in the BSM model formula, arrives at the option prices observed in the market. The usual way to calculate it is through inversion using numerical methods, given that the model formula is nonlinear. The implied volatility surface can be defined as a set of implied volatilities obtained via BSM based on market prices for each option strike price and maturity, thus forming a three-dimensional structure.

Thus, by estimating implied volatility more precisely, it is possible to arrive at prices with smaller errors in relation to market prices. Based on this idea, this work will propose a model to estimate the implied volatilities, with the objective of, when using the information contained in the prices of the options, arriving at pricing with smaller errors than the classic models mentioned above, since the expectations will be considered of market agents.

Therefore, the justification for this work is to contribute to the evolution of methods for estimating volatility curves and, ultimately, for pricing options using a non-parametric Bayesian approach with the imposition of no-arbitrage restrictions. This model has the advantage of not assuming a functional form, and as such it is more robust to misspecification problems and allowing a more accurate fit to the implied price and volatility curves, while being consistent with the no-arbitrage constraints needed to asset pricing. Furthermore, after estimating the volatility for each strike price, expiration date and period, the model's forecast level will be tested, calculating the implied volatility for one period ahead of the estimated surface.

In the next section, the main methods already used for the construction of curves and surfaces of implied volatility are revised, considering their positive and negative points, and we the methodology that be used in our work is presented.

2. Literature review

2.1 Implied volatility curves

Over the years, different types of techniques have been used in the literature to model implied volatilities. In (SHIMKO, 1993) it was proposed to use a simple quadratic polynomial to adjust the volatility curve in relation to the strike prices. In (MALZ, 1997) the curve was adjusted in relation to the delta of the options using the same polynomial method. In (CAMPA et al., 1998), instead of using a polynomial method, a natural *spline* was used to model

the implied volatility in relation to the strike price, in order to better capture the *smoothing* of the curve. In (BLISS and PANIGIRTZOGLU, 2002) they used a *smoothing cubic spline* to adjust the implied volatility in relation to the option's delta. In (PANIGIRTZOGLU and SKIADOPOULOS, 2004) the same *smoothing cubic spline* was used, but in relation to the strike prices. Finally, in (LAURINI, 2011) a *smoothing B-splines* method was used under monotonicity and convexity restrictions, adjusting the implied volatility in relation to the *moneyness* of the options.

2.2 Construction of implied volatility surfaces

In (ALEXANDER, 2001), *principal components analysis* (PCA) was applied to the daily variation of the implied volatility differential between a given strike price and *at-the-money* (ATM) volatility in each term. In addition, a second approach was made with the application of the PCA to the daily variation of volatility as a function of the delta in each term. (KAMAL and DERMAN, 1997) also used PCA, but with the application of PCA to the daily variation of the volatility surface, that is, we worked with volatility as a function of delta and maturity. In (CONT and FONSECA, 2002) they applied principal component analysis to the volatility surface as a function of term and *moneyness* for the SP500 and FTSE100 indices. The volatility surface for these indices was built from option prices on the market, being subjected to a smoothing process before principal component analysis.

(HOMESCU, 2011) revise the main methodologies for building an implied volatility surface. Different topics related to the construction of such surfaces in practice are addressed, such as arbitrage-free conditions for both the strike price and the time, the problem of carrying out extrapolations beyond the central region, choosing functions for calibration, as well as the selection of algorithms for numerical optimization.

2.3 Construction of surfaces using Bayesian methods

Finally, we revise the use of Bayesian methods for the estimation of function surfaces related to financial problems. In (AVELLANEDA et al., 2000), the volatility surface was constructed using a non-parametric method known as *feed-forward neural network*. The formulation of the problem was done in a Bayesian way, imposing a prior structure on the components of the neural network. After constructing the volatility surface, Dupire's equation ((Dupire, 1994) was used to estimate the local volatility surface.

In (CALDEIRA et al., 2010) a dynamic model was proposed for the term structure of the interest rate, considering two different types of specification for the (NELSON and SIEGEL, 1987) model. In the first specification is

introduced conditional heteroscedasticity, using a stochastic volatility model with common factors. In the second, a model was made with individually latent factors following an auto-regressive process with stochastic volatility. Bayesian inference was conducted by the method of *Markov chain Monte Carlo* (MCMC).

In (LAURINI, 2013), to predict the evolution over time of the interest rate term structure, a non-parametric model based on *penalized splines* with Bayesian estimation made by a combination of Gibbs and Metropolis Hastings was used, again using a *Markov chain Monte Carlo* (MCMC) method.

In (ALMEIDA and QIN, 2020) a non-parametric Bayesian method was used to predict and estimate the volatility surface from market data. Gaussian processes with different *kernels* characterizing covariance functions were used. Posterior distributions and confidence intervals were obtained for the implied volatilities. In the study it was concluded that the Bayesian method is a powerful alternative to the existing pricing models.

Based on the analysis of the available literature on the subject, as well as considering the strengths of each method, a semi-parametric Bayesian model will be proposed in this work to predict and estimate the implied volatility surface. In this method, regressions via *splines* with shape constraints will be used, more specifically *C-splines*, that is, *splines* with convexity constraint. The functions will be estimated using *Markov chain Monte Carlo* (MCMC) algorithms. By using a Bayesian *framework* it will be possible to produce uncertainty estimates much more easily than in comparison with a frequentist paradigm, in addition to enabling comparison and selection between models.

3. Methodology

3.1 Black-Scholes model

To calculate the implied volatilities from market prices, the Black-Scholes model will be used, following the main convention of the implied volatility literature. In (BLACK and SCHOLE, 1973), it is assumed that S_t , that is, the price of the underlying asset, follows a geometric Brownian motion (GBM), in the form of the following equation:

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

where μ is the drift, σ is the volatility and B_t is a stochastic process that follows a Brownian motion. Considering the payoff of a call as $\max(S_t - K, 0)$ and that of a put as $\max(K - S_t, 0)$, when resolving the EDPs respecting the boundary conditions, one arrives the option prices solution for a non-dividend asset:

$$\begin{aligned}
c_t &= S_t N(d_1) - Ke^{-rT} N(d_2) \\
p_t &= Ke^{-rT} N(-d_2) - S_t N(-d_1) \\
d_1 &= \ln\left(\frac{S_t}{K}\right) + (r + 1/2\sigma^2)T \\
d_2 &= d_1 - \sigma\sqrt{T}
\end{aligned}$$

Where:

c_t - European call premium in t.

p_t - European put premium in t.

S_t - Price of the asset defined in t.

t – current period.

K – Call option strike price.

T – Time to maturity of the call option.

σ – Implicit volatility.

r - Risk-free interest rate.

$N(\cdot)$ – The standard normal cumulative function

The option pricing formula developed for this model is used to calculate the value of European options to call or put stocks. The formula assumes that the volatility of the underlying asset is constant. With the definition of the other parameters of the Black-Scholes equation, such as asset price, interest rate, strike price, maturity and volatility, it is possible to arrive at the value of the premium for the option. This premium at time t of a European-type call option, which is exercised only on the expiration date.

The model has the following assumptions:

(1) the risk-free rate is constant until option expiration (2) returns have a log normal distribution and volatility is constant, (3) there are no dividends (Merton Model includes), (4) options are exercised only at maturity, (5) there are no transaction costs, (6) there are no risk-free arbitrage opportunities, and (7) there are no short sales.

Of these assumptions we will particularly delve into the non-existence of arbitrage opportunities, as this has a direct impact on the shape of the options price surface and, therefore, on the shape of the implied volatility surface. As discussed by (LAURINI, 2011):

“In the classic option pricing model of Black and Scholes, the non-arbitrage conditions are equivalent to the existence of a risk-neutral density (density of price states), which gives the base price of the asset in each possible state of nature in a risk-neutral measure. Fundamental conditions for the imposition of non-arbitrage are given by the relationship between the price of the options and the strike.”

Therefore, consider $C(K,T)$ as the price of a call option on a stock with

strike price K and expiry date T , with a given expiration date T_f . In (CARR and WU, 2010), the non-arbitrage conditions for a call option with a fixed strike price and term to maturity are:

$$\begin{aligned} C(K,0) &= \max[0, S_{T_f} - K] \\ \max[0, S_t - Ke^{-rT}] &\leq C(K,T) \leq S_t \\ \text{Put - Call - Parity : } C(K,T) + Ke^{-rT} &= P(K,T) + S_t \end{aligned}$$

For options with different expiration dates and strike prices:

$$\begin{aligned} \text{If } K_1 > K_2, \text{ then } 0 \leq C(K_2,T) - C(K_1,T) &\leq (K_1 - K_2)e^{-rT} \\ C(K,T) \text{ is a convex function for } K & \\ C(K,T) \text{ is a monotonically increasing function for } T & \end{aligned}$$

Thus, when building the curves, any points violating these constraints generate arbitrage opportunities. Therefore, such restrictions must be taken into account in the development of the model.

3.2 Implied Volatility

Implicit volatility is the volatility obtained by using the market prices of options in the Black-Scholes model formula, that is, given the market price of an option, by inverting the price function one can arrive at the corresponding implicit volatility.

Assume an observed price of a call option C^* , the implied volatility value σ^* is the one that satisfies the following relation:

$$C^* = C(S, K, r, \sigma^*, T)$$

However, as the volatility parameter of the Black and Scholes (1973) formula above cannot be isolated to obtain the implied volatility, given the non-linearity of the formula, it is necessary to use numerical methods for extracting the roots of the price functions.

According to (CONT and FONSECA, 2002), the implied volatility $\sigma_t^*(K, T)$ exists and is unique because the price of a call (or put) option as a function of volatility is a monotonic mapping of $[0, +\infty)$ to $[0, S_t - Ke^{rT}]$ or $[0, Ke^{rT} - S_t]$. If we stick to (K, T) , $\sigma_t^*(K, T)$ is a stochastic process. If we fix it at t , it is dependent on maturity T and strike K .

According to (ORLANDO and TAGLIALATELA, 2017), the most accurate calculation of implied volatility is by the Newton-Raphson numerical method, which is also the one that converges most quickly. The process is fast and simple, as the function c_t (or p_t) is monotonic increasing with respect to volatility.

One can describe the Newton-Raphson method as below. Suppose a continuous and differentiable function $f(x)$ and it is known that the root of the function is close to the point $x = x_0$. By Newton's method the best approximation to the root is as follows:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

This process is repeated as many times as necessary until the desired accuracy is achieved. In general, for any x_n value, the next value is given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

If x_0 is in the neighborhood of 0 and $f'(x) \neq 0$ then the method will converge.

Implicit volatility surfaces or curves cannot have an arbitrary shape and therefore, to build admissible surfaces or curves, one must follow the restrictions imposed by the non-arbitrage conditions. With that in mind, we will impose the restrictions that all calculated implicit volatilities must be greater than or equal to zero, and the volatility curves, for a given maturity period, must have a convex shape in relation to the strike price.

3.3 Shape constrained splines

Splines are piecewise polynomials that are continuously differentiable to a certain degree and are connected by a sequence of points, known as nodes. Each segment is polynomials that connect smoothly to each other at the nodes. A *spline* function can be represented as a linear combination of its basis functions.

Thus, before talking about base *spline* functions, we need to define a sequence of nodes. Suppose K is a set of nodes indexed by m , a positive integer, we have the following sequence (k_1, k_2, \dots, k_m) so that the interior nodes are between the nodes of boundary L and U . Defining d as the degrees of freedom of the basis function and $q = d + 1$ as the order. We define the sequence of nodes with additional nodes at the ends, $L = k_1 < k_2 < \dots < k_{m+2} = U$. In this way, it is possible to construct a *B-spline*, that is, a base *spline*, by means of a recursive formula. Thus, the i th basis function *B-spline* of degree d and order q can be defined recursively as follows:

$$\begin{aligned} B_{1,q}(x) &= 1 \text{ for } k_i \leq x < k_{i+1} \\ B_{1,q}(x) &= 0 \text{ for all other cases} \\ B_{i,q}(x) &= \frac{x - k_i}{k_{i+q-1} - k_i} B_{i,q-1}(x) + \frac{k_{i+q} - x}{k_{i+q} - k_{i+1}} B_{i+1,q-1}(x) \text{ for } q > 1 \end{aligned}$$

A *M-spline* can be considered as a *B-spline* normalized according to (CURRY and SCHOENBERG, 1966). Given the sequence of nodes, the i -th basis function *M-spline* of degree d and order q , denoted by $M_{i,q}(x)$, can be considered as a basis function *B-spline* normalized, if it satisfies the following relation:

$$M_{i,q}(x) = \frac{q}{(k_{i+q}-k_i)} B_{i,q}(x)$$

Furthermore, a *M-spline* can be defined analogously to a *B-spline*, through a recursive formula, as follows:

$$M_{i,q}(x) = \frac{q}{(q-1)(k_{i+q}-k_i)} [(x-k_i)M_{i,q-1}(x) + (k_{i+q}-x)M_{i+1,q-1}(x)]$$

A *I-spline* of order $q+1$ is found by integrating a *M-spline* of order q . In this way, when we integrate a *M-spline* from L to x , we arrive at a *I-spline*, as shown below:

$$I_{i,q+1}(x) = \int_L^x M_{i,q}(u) du$$

For the interval $L \leq x \leq U$ *M-splines* are strictly positive. Thus, by definition, *I-splines* are monotonically non-decreasing from L to U . Therefore, (RAMSEY, 1988) proposed the use of *I-splines* for monotonic regressions. A monotonically increasing function can be obtained by a linear combination of *I-splines* and an additional intercept term, so that monotonicity is ensured by constraining the coefficients of the *I-splines* to be positives. Analogously, a regression can be performed for a monotonically decreasing function, simply restricting the coefficients of the *I-splines* to be negative.

In (MEYER, 2008) *C-splines* of degree $q+1$ and order $q+2$ were formed as the integration of *I-splines* of order $q+1$, so to be estimated with curvature constraints. The relationship between *I-splines* and *M-splines* can be seen as below:

$$C_{i,q+2}(x) = \int_L^x I_{i,q+1}(u) du$$

Functions with concave or convex shape constraints can be approximated using a linear combination of *C-splines*, an identity function and an intercept term. To estimate a convex regression function, just restrict the *C-splines* coefficients to non-negative values.

Similarly, to estimate a concave regression function, a linear combination of *C-splines* with coefficients constrained to non-positive values is performed, plus a linear combination of the identity vector with unrestricted coefficients.

Estimates for regression functions under a combination of constraints can also be done analogously. However, as in this work only curvature constraints will be used, this topic will not be addressed.

3.4 Regression via shape-constrained splines

Non-parametric regression methods allow estimating a function with the minimum of assumptions and, therefore, are the most indicated when a parametric form is not available. In the case of regressions by *splines* it is necessary that the number of nodes and their positions are specified beforehand. If the function estimate is not robust to these choices, the inference of the regression function is impaired.

Making estimates exclusively using shape constraints such as those of monotonicity or curvature do not require previously defining parameters, however, in general they are not smooth, nor parsimonious, causing the degrees of freedom of the model to end up having values, in a certain way, elevated. *splines* regressions are smooth, flexible, and parsimonious estimators of nonparametric functions. However, the number and position of nodes are sensitive. Therefore, one way to make them robust in choosing nodes is by imposing shape constraints. Thus, by joining the regression by *splines* with the shape constraint, the method becomes both flexible and robust.

Spline regressions basically work as follows: a set of basis functions is formed and used as regressors in an OLS model. The base functions are smooth and have segments with polynomials of a certain degree, positioned between the nodes specified by the user. The set of linear combinations of the basis functions is large enough to provide high flexibility to the model, which also raises a certain concern about *over-fitting*.

One way to avoid *over-fitting* is to include a penalty term to reduce the flexibility of the regression estimator. In the case where shape constraints are imposed, however, this technique is not necessary, as the constraints themselves do not allow the formation of peaks, valleys or sinuosity in the fit, remedying the *over-fitting*.

In this work, given the convex shape of the implied volatility curves, *C-splines* will be specifically used to estimate the regression functions. Below there is a description of the method proposed in Meyer (2008) to estimate regression functions with convexity restriction.

Consider a regression model, where X and Y are the predictor variables and the response variables, respectively. The observed data are independent pairs (x_i, y_i) , such that $i \in 1, \dots, n$. It is assumed that the true function $f(\cdot)$ is continuous and has a certain known shape restriction, in this specific case, convexity.

A convex curve can be estimated through a linear combination of *C-spline* basis functions, an identity function, and an intercept term, whose *C-spline* coefficients are restricted to non-negative values. In convex regression, the coefficients are estimated by restricted least squares in order to minimize the

squared error loss, as follows:

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n [y_i - (\alpha_0 + \alpha_1 x_i + \beta^\top C(x_i))]^2, \text{ s.a. } \beta \geq 0$$

where y_i is each observed value of the independent variable, α_0 is the intercept, α_1 is the coefficient of the linear term, and $C(x_i)$ is a vector of *C-splines* evaluated at each x_i .

3.5 Bayesian inference for shape-constrained splines

Non-parametric and semi-parametric regression estimation can also be modeled using the Bayesian paradigm. The advantages of the Bayesian method are that it makes it possible to generate joint posterior distributions for the coefficients and, therefore, allows inference through various sampling methods, including for *frameworks* with shape constraints, as is our case.

In (HACKSTADT, 2011) an adaptation of the method developed in (MEYER, 2008) was made, in which a model is proposed, using a Bayesian approach, for estimation regression splines with shape restrictions. The proposed methodology uses the Gibbs sampler, an MCMC algorithm, which makes it possible to infer the posterior distribution from random draws, making this process easier, since it is not trivial to obtain the variance of the vector of coefficients of a regression with constraints in a frequentist framework. In addition, through this method it is also possible to use model selection tools that are more difficult or impossible to work with under a frequentist paradigm.

In this model, gamma distributions will be adopted as priors for coefficients that are real positive, with hyperparameters chosen such that the variance is large and the mean relatively small. When using C-splines, shape restrictions will be imposed simply by requiring that the coefficients of the base functions of splines be non-negative.

Consider X and Y as a series of predictor variables and response variables, respectively. The observed data are independent pairs (x_i, y_i) , such that $i \in 1, \dots, n$. Assume the model $y = \eta + \varepsilon$, where

$$\varepsilon \sim N(0, \tau^{-1})$$

$$\eta = \sum_{j=1}^m \beta_j \delta_j + \sum_{j=1}^p \alpha_j v_j$$

where δ_j are the basis vectors corresponding to the shape constraint of the function $f(\cdot)$; v_j are the vectors of the observed values of the covariates to be modeled parametrically; m is the number of chosen nodes. When using a Bayesian method for inference prior densities are defined for the coefficients

α and β . To enforce the shape restrictions, the values of β will be restricted to the set of positive real numbers.

The priors for each parameter will be given as follows:

$$\begin{aligned}\tau &\sim \text{Gamma}(d_1, d_2) \\ \beta_j &\sim \text{Gamma}(c_1, c_2) \\ \alpha &\sim N(0, MI)\end{aligned}$$

Where the parameters d_1 and d_2 were chosen such that the mean of the prior d_1/d_2 is the inverse of the model variance. The mean c_1/c_2 was chosen to be a prior of R (range) divided by m, where R is the range of the function f, that is $R = f(\max(x) - \min(x))$ and m is the number of interior nodes of the function. And finally, M is a parameter chosen discretionarily. In (HACKSTADT, 2011) $M = 1000$ was used, a value that will also be adopted in this work.

As the posterior distribution is analytically intractable, the way to obtain it is through a Markov Chain Monte Carlo (MCMC) method, more specifically the Gibbs sampling algorithm.

At each iteration t, random drawings of the complete conditional distributions are generated for the parameters in the Gibbs sampler, as follows:

$$\eta^{(t)} = \sum_{j=1}^m \beta_j^{(t)} \delta_j + \sum_{j=1}^p \alpha_j^{(t)} v_j^{(t)}$$

The estimator (posterior mean) of the function \hat{f} in x_i is calculated by averaging the $\eta_i^{(t)}$ values obtained in N iterations, after removing the burn-in, as follows:

$$\hat{f}(x_i) = \frac{1}{N-B} \sum_{t=B+1}^N \eta_i^{(t)}$$

Where N is the total number of iterations of the MCMC algorithm and B is the Burn-in.

3.6 Selecting models and hyperparameters

The comparison between the different possible models is an important component of any modelling process. Below are listed some metrics to be used in the work to define and select which model is most appropriate for the dataset under study.

The *Akaike Information Criteria* (AIC) measures the relative quality of statistical models for a given set of data, that is, it estimates the quality of

each model in comparison with other models for the same data, penalizing for the complexity of the model. The measurement of quality is done through the amount of information lost by the model, so that the more information the model loses, the lower its quality. However, this measure is valid only asymptotically, requiring some kind of correction if the data set is small. The methodology used in the AIC aims to indicate the best balance for the model, between quality of fit to the data and simplicity. In other words, you are dealing with the risk of *overfitting* and *underfitting*.

Assume a given statistical model, for a given set of data. Let k be the number of parameters estimated in the model and L be the maximized value of the *likelihood* of the model, we have that the value for the AIC is the following:

$$AIC = -\log(L) + 2(k)$$

with k a measure of model complexity, usually the number of parameters in the model.

Within the set of models analyzed for the same data, the preferred model is the one with the lowest value for the AIC. Thus, the AIC measures the quality of fit to the data through the likelihood function, at the same time that it includes a penalty related to the number of parameters estimated in the function. The penalty aims to control *overfitting*, by penalizing the increase in the number of model parameters, balancing with the quality of the fit to the data.

It is worth mentioning that the AIC is not informative regarding the absolute quality of the model, only being a tool for comparing one model with the other. Therefore, it is important to use metrics other than the AIC when performing model selection, in order to validate the absolute quality of the selected model. One way to carry out this validation is through metrics related to the residuals generated by the models under analysis.

Another indicator that we will use in our work is the *Bayesian information criterion* (BIC), also known as *Schwarz information criterion* (SIC). Like the AIC, it is a relative model selection indicator, in which models with lower values for the indicator are preferred. The measurement of the quality of the models is based on the likelihood function, similarly to the AIC and there is also a penalty term, however the BIC also takes into account the sample size for the penalty.

Assume a given statistical model, for a given set of data. Let k be the number of estimated parameters in the model, L be the maximized value of the likelihood of the model and N be the number of observations in the sample, we have that the value for the BIC is the following:

$$BIC = -\log(L) + \log(N)(k)$$

When comparing various models, in general, the best is the one with the lowest value for the BIC. However, it is important to consider that the BIC can be used to compare estimated models only when the number of dependent variables are identical in all models being compared. Furthermore, it is worth noting that, like the AIC, the BIC should not be used to measure the absolute quality of the model.

In view of this, to measure the absolute quality of fit to the model data, metrics related to the residuals of each model will also be used. The *mean error* (ME), in general, refers to the average of all errors of a given estimation. An error in this context is the difference between the value measured by the model and the correct value, that is, the observed value. This measure can be calculated according to the formula below:

$$ME = \frac{1}{N} \sum_{i=1}^N residuals$$

The *mean error* (ME) usually results in a value that is not very useful analytically, because when adding negative and positive errors, one cancels the other, resulting in an ME close to zero. A more widely used metric that corrects this problem is *mean absolute error* (MAE).

The *mean absolute error* (MAE) is calculated by the sum of the absolute difference between the observed values and the estimated values of a given model, divided by the sample size, that is, it is the arithmetic mean of the absolute errors. This measure can be calculated according to the formula below:

$$MAE = \frac{1}{N} \sum_{i=1}^N |residuals|$$

Among the advantages of using the *Mean Absolute Error* (MAE) are the easy interpretability and the fact that each error influences the MAE in the same way.

4. Results

4.1 Database

For this work, it was decided to specifically use Petrobras preferred shares (PETR4) and their respective option contracts, as it is the stock with the highest number of option contracts, as well as having the most liquid contracts in the Brazilian variable income market. In addition, it has an extensive history, having been, throughout this period, an extremely representative action of the

market movements of the Brazilian stock exchange in the most varied scenarios faced by the country, given the characteristics of the Brazilian economy intrinsically linked to *commodities*, as well as the fact that it is a state-owned company and its market performance is also associated with country risk.

As a database, daily quotations available on the B3 website were used, for shares and for all option contracts with trades carried out in the period. The period considered for this work comprises daily quotations of these securities ranging from 2010-01-04 to 2022-06-30. It is worth mentioning that given the nature of option contracts, which expire monthly and have variable strike prices at each expiration depending on the current price of the underlying asset, it was not possible to obtain relevant historical data that included intraday prices for the options. In this way, we only obtain the daily opening, high, low and closing prices of stocks and options.

First of all, a filter was made on the options contracts, considering only those with a minimum daily financial volume of R\$ 200,000.00. Such a filter was made with the intention of discarding those contracts that did not have many trades carried out on the day, thus not being representative for the formation of market prices and, therefore, would probably generate outliers for the calculated implicit volatilities. After the filter, 75.49% of the number of original contracts in the database remained for call options and 75.54% for put options.

In order to calculate the implied volatility, it is necessary that the prices of the options occurred at the same time as the observed price of the share. As only contracts with high liquidity were selected, it is reasonable to assume that these had a continuous flow of trades throughout the day and, therefore, it was considered that the maximum and minimum prices of the options were observed at the same moment as the maximum prices and minimum of the action. In addition, analogously, we consider that the closing prices of the liquid options occurred simultaneously with those of the stock.

As for the opening prices, there is evidence of a difference in trading hours between the stock and the options, given the large distortions and disparities seen in the calculated implied volatilities, which is why we discard the opening prices.

Therefore, for the purpose of constructing the curves and surfaces of implicit volatilities, it was considered that the maximum, minimum and closing quotations for shares and options, for a given day, occurred in the same period t . That is, each daily curve will be built using the implicit volatilities calculated for the maximum, minimum and closing quotes within the same day.

For the model's risk free interest rate, the daily Selic rate provided by the

Central Bank via API was used.

4.2 Implied volatility

First, the daily implied volatilities were calculated for each option contract traded on that day. A restriction was imposed for implied volatilities not to assume negative values. In cases where this happened, the value of implied volatility was replaced by zero.

Examples of implied volatilities calculated for call option contracts can be seen in table A1 of the appendix and in table A2 for put option contracts. For each option contract, the term to maturity was calculated on an annual basis, and the moneyness, that is, the value of the exercise price divided by the price of the underlying asset.

As there are large distortions in options with very short or very long terms, only contracts with a minimum maturity of 3 days and a maximum of 63 days were considered in this work. This fact also occurs in relation to the moneyness of the options, so only contracts with a minimum moneyness of 0.50 and a maximum of 1.50 will be considered.

4.3 Selection of parameters for the model

Shape constraints by themselves, when imposed on the model already result in estimators that are typically more robust compared to a pure smoothing, however, it is still important to test and evaluate which parameter for the number of nodes makes the more efficient model. In addition, another parameter that we will test will be the size of the training dataset that we will use to estimate the curves with greater precision and better quality in the data fit.

Table 1
Average errors per period - Calls

Period	RMSE	MAE	AIC	BIC
5 days	0.052205	0.032982	-119.3492	-110.745277
10 days	0.063512	0.03825	-231.751148	-218.178936
21 days	0.052966	0.033768	-124.516339	-114.546634
63 days	0.067962	0.047895	-235.763484	-221.234706
126 days	0.122034	0.094138	-146.435863	-130.784203
252 days	0.128492	0.098518	-218.124663	-199.714877

Note: Table shows the average RMSE, MAE, AIC and BIC calculated for 6 different models using calls. Each model assume a different value for n, that is the quantity of previous days utilized to estimate the curve for the following day. The values assumed by n are 5, 10, 21, 63, 126 and 252 days. The sample encompasses the period from 2010-01-04 to 2015-12-30.

We divided the period considered in the work into an in-sample period, which encompasses a period from 2010-01-04 to 2015-12-30, and an out-of-sample period which runs from 2016-01-04 to 2022-06-30. The in-sample period will be used in order to determine the values of the model parameters that minimize errors in the estimation and generate better quality of fit to the model, while the out-of-sample sample will be used to estimate the curves that will be used for the analysis of the results.

We will estimate a $\hat{f}^{(t)}(x_i)$ for each maturity day, that is, for each t that consists of a specific day in our sample, curves of the function will be estimated regression at each maturity date. The tested models will have different data sets for training, and there will be models with the implicit volatilities calculated for the n days prior to t . For the in-sample sample, values for n , i.e., the length of the period, will range from 5, 10, 21, 63, 126 to 252 days. And so, it will be defined which value for the hyper parameter n generates the model with the best results. For this analysis, we will set the number of nodes to 5 for both calls and puts.

Table 2
Average errors per period - Puts

Period	RMSE	MAE	AIC	BIC
5 days	0.122819	0.098015	-56.039674	-45.800943
10 days	0.123624	0.09874	-55.766396	-45.529878
21 days	0.123546	0.098447	-58.13886	-47.712096
63 days	0.118931	0.092418	-98.562609	-85.42091
126 days	0.122034	0.094139	-146.436189	-130.784529
252 days	0.128492	0.098518	-218.125029	-199.715244

Note: Table shows the average RMSE, MAE, AIC and BIC calculated for 6 different models using puts. Each model assumes a different value for n , that is the quantity of previous days utilized to estimate the curve for the following day. The values assumed by n are 5, 10, 21, 63, 126 and 252 days. The sample encompasses the period from 2010-01-04 to 2015-06-30.

Similarly, the number of interior nodes k are analyzed to define the optimal value of this parameter for the model, considering errors, as well as the level of complexity of the model in the in-sample sample. For this analysis, the number of periods n will be fixed at 63 days, both for call and put options. The values of k will range from 1, 2, 3, 4, 5, 6, 7, 8, 9 to 10, that is, these values will be tested for the number of nodes in the model.

As criteria for comparing the quality of the model's fit for the different proposed specifications, the ME, the RMSE, and the MAE will be calculated.

Additionally, the AIC and BIC will be calculated, as they are metrics that also take into account the degree of complexity of the model, not just its errors.

We compared the values of these metrics for the set of models considered, and the best model is the one with the lowest values for the AIC and BIC, that is, the one that has the least loss of information relative to the true model. It is worth mentioning that in this work the comparison between models will be performed using identical samples.

For the absolute values, the RMSE and MAE error metrics will be considered, in order to verify whether the models actually present small errors, that is, meaning that they have a good quality of fit to the data.

Table 3
Average errors by knots - Calls

Knots	RMSE	MAE	AIC	BIC
k=1	0.069285	0.04876	-240.00057	-235.157644
k=2	0.119724	0.093049	-103.580244	-97.009394
k=3	0.049992	0.031512	-102.186912	-94.769097
k=4	0.11905	0.092487	-100.396481	-89.445065
k=5	0.067962	0.047895	-235.763484	-221.234706
k=6	0.049272	0.03123	-96.728731	-83.747554
k=7	0.049115	0.031178	-94.819492	-79.983861
k=8	0.049007	0.031107	-93.048068	-76.335781
k=9	0.049075	0.031143	-90.974948	-72.387085
k=10	0.11904	0.092531	-89.210195	-65.042001

Note:Table shows the average RMSE, MAE, AIC and BIC calculated for 10 different models using calls. Each model assumes a different value for k, that is the number of knots utilized to estimate the curve. The values assumed by k are 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 knots. The sample encompasses the period from 2010-01-04 to 2015-06-30.

Table 1 shows the average error values, as well as the AIC and BIC calculated for the call options in each period. As we can observe from the results obtained, the number of periods that generates better results, considering both the error values and the AIC and BIC values, would be $n = 10$ days, since it presents lowest errors, at the same time that it is between the models with the lowest AIC and BIC values.

In Table 1, displaying the results obtained for the put options, it can be seen that the period that presents the most promising results taking into account all metrics would be $n = 63$ days, as it has the lowest values for errors, as well as significantly low values for the AIC and BIC.

Observing the data showed in Table 3, which considers the metrics for

number of nodes for the call options, it can be seen that the model with the best results would be $k = 5$, considering jointly the values for the error metrics, for the AIC and for the BIC. In the case of put options, in Table 4, the values are very close, but when considering both types of metrics together, it was concluded that $k = 5$ is the choice with the best results.

Thus, we can say that $k = 5$ for the number of nodes proved to be a robust choice, as both the call options and the put options were the selected models.

Table 4
Average errors by knots - Puts

Period	RMSE	MAE	AIC	BIC
k=1	0.120172	0.0934	-105.096917	-100.71635
k=2	0.119724	0.093049	-103.580495	-97.009645
k=3	0.119262	0.09264	-102.128461	-93.367328
k=4	0.119049	0.092487	-100.396931	-89.445515
k=5	0.118931	0.092418	-98.562609	-85.42091
k=6	0.118843	0.09236	-96.680068	-81.348086
k=7	0.118777	0.092323	-94.765592	-77.243327
k=8	0.118843	0.092376	-92.922906	-73.191079
k=9	0.118985	0.092489	-91.120693	-69.162501
k=10	0.119009	0.092506	-89.271186	-65.102201

Note:Table shows the average RMSE, MAE, AIC and BIC calculated for 10 different models using puts. Each model assumes a different value for k , that is the number of knots utilized to estimate the curve. The values assumed by k are 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 knots. The sample encompasses the period from 2010-01-04 to 2015-06-30.

Finally, after testing different models in the in-sample period, the following parameters were defined for call options: $n = 10$ days as the number of days used to estimate the curves and $k = 5$ as the number of nodes on each curve. As for the put options, the defined parameters are the following: $n = 63$ days as the number of days used to estimate the curves and also $k = 5$ for the number of nodes in each curve. The other exogenous parameters of the model will be selected according to (HACKSTADT, 2011), both for call and put options: $d_1 = 1/10$, $d_2 = 1/10$, $M = 1000$ and $N = 5000$.

4.4 Construction of implied volatility surfaces

In this section, the implied volatility curves estimated using the proposed model are presented, with the optimal parameters defined according to the values found in the in-sample period. For each day t , implied volatility curves are estimated for all the different maturities, forming an implied volatility

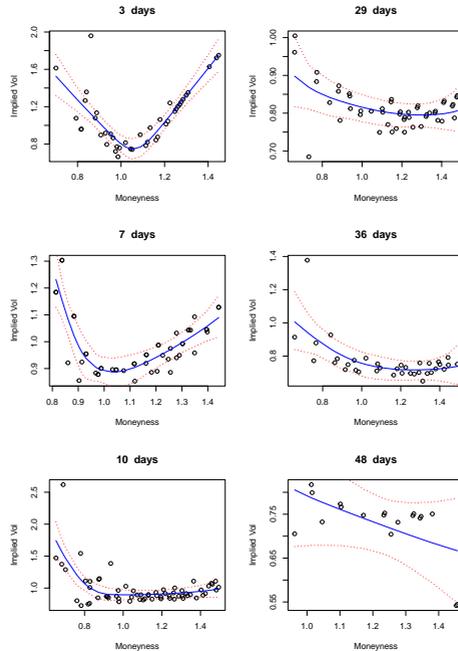


Figure 1

Examples of curves by maturity on 2016-02-12 - Call options

Note: Figure provides examples randomly selected for the curves estimated on 2016-02-12. We have call options curves for 3, 7, 10, 29, 38, 43 days to maturity. The curves present the implied volatility as a function of the moneyness. The blue line corresponds to the estimated curve, while the dotted red lines correspond to the 95% credible intervals. The parameters used are $n=10$ and $k=5$.

surface on each given day t . Finally, the error metrics for these estimates will be calculated, analyzing the quality of fit to the data by term to maturity.

The Bayesian method used to estimate the curves was done as follows: for each set of parameters β and α produced in an iteration of the Gibbs sampler, after Burn-in, the function estimator was computed to each value of x and thus, we arrived at $\eta_i^{(t)}$. We used a total of 5000 iterations for each run of the MCMC algorithm, with a burn-in of 500. Using the values of $\eta_i^{(t)}$, at each value of x the prediction interval of 95% for the posterior density is estimated.

By the posterior joint distribution generated in the Bayesian model, it is also possible to calculate credible intervals of 95% for the error vector, which are represented as the red dotted lines in the graphs.

In Figure 1, there are examples of estimated call option curves for some

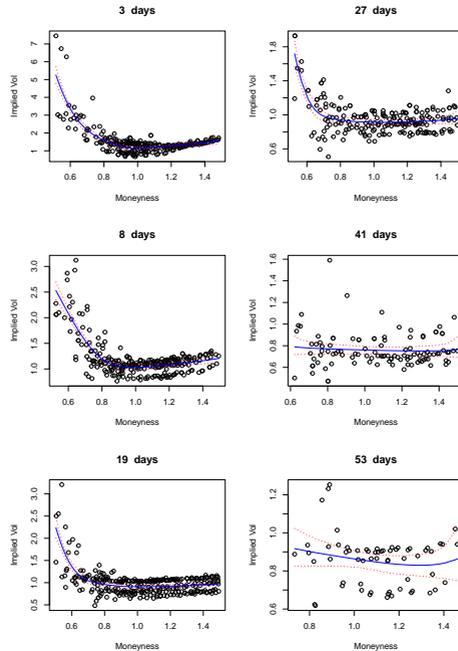


Figure 2

Examples of curves by maturity on 2016-04-29 - Put options

Note: Figure provides examples randomly selected for the curves estimated on 2016-04-29. We have put options curves for 3, 8, 19, 27, 41, 53 days to maturity. The curves present the implied volatility as a function of the moneyness. The blue line corresponds to the estimated curve, while the dotted red lines correspond to the 95% credible intervals. The parameters used are $n=10$ and $k=5$.

specific maturities on the same day (2016-02-12). In Appendix B, all curves estimated on that day were included. In blue is the estimated regression function and in red the calculated 95% credible intervals. Based on these results, it can be seen that the narrowest intervals were generated in the intermediate maturities, showing that for these maturities there is a greater degree of confidence in the estimated function.

In Figure 1, we show examples of estimated curves for put options for some specific maturities on the same day (2016-04-29). Again, in appendix B all estimated curves for that day can be found. In blue is the estimated regression function and in red the calculated 95% credible intervals. Based on these results, similarly to what happened with the call options, it is possible to notice that the narrowest intervals were generated in the intermediate maturities, showing that for these maturities there is a greater degree of confidence

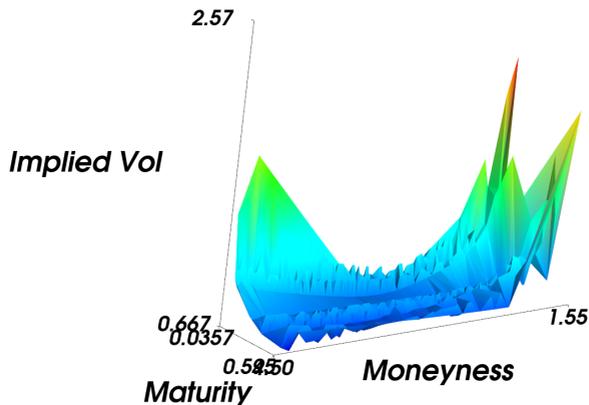


Figure 3

Example of estimated volatility surface on day 2016-02-12 - Call options

Note: Figure shows an example of implied volatility surface estimated on 2016-02-12 for call options. The method to construct the surface was the Delaunay Triangulation algorithm applied to the curves estimated for each maturity. The x and y axis correspond to moneyness and maturity, while the z axis is the implied volatility.

in the estimated function in relation to the shorter ones and the longer ones. longer.

To construct the implied volatility surface, a set of points S was considered where each point p of S is represented by a triple (x,y,z) , where x , y encode moneyness and maturity, respectively, and z is the implied volatility. The algorithm used was Delaunay triangulation, which constructs the surface by finding the nearest neighbor of each point. Basically, the algorithm works as follows: for a set of points P in the plane, a triangulation $DT(P)$ is formed where no point in P is inside the circle formed by any triangle in $DT(P)$. The Delaunay Triangulation maximizes the smallest angle of all triangles in the triangulation, that is, it tends to avoid triangles with very small interior angles.

In Figure 3, the surface for call options on 2016-02-12 was built. And in Figure 4, there is the volatility surface for put options on 2016-04-29. It can be seen that, in both curves, the implied volatility assumes higher values at the extreme points of the curve, that is, at the points where the options contracts are out-of-the-money (OTM) or in-the-money (ITM). Figure 4 displays the estimated volatility surface for put options on 2016-04-29.

In general, on the surfaces analyzed, put options even show higher values

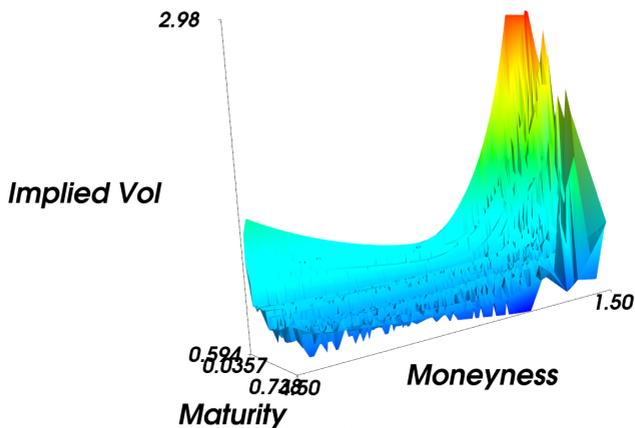


Figure 4

Example of estimated volatility surface on day 2016-04-29 - Put options

Note: Figure shows an example of implied volatility surface estimated on 2016-04-29 for put options. The method to construct the surface was the Delaunay Triangulation algorithm applied to the curves estimated for each maturity. The x and y axis correspond to moneyness and maturity, while the z axis is the implied volatility.

for implicit volatilities, which can be explained by the aversion to downside risk being greater among market participants. Another factor that may possibly influence is the fact that downward shocks are usually more intense than upward movements and, therefore, put options are imbued with a higher risk premium than call options, reflecting in volatilities higher implied charges for these contracts.

Another important point to be noted is that the curves become less and less convex as the maturities become longer, that is, the shorter maturities have more pronounced curvatures, while the long ones become flatter.

This is a common and expected pattern to find on implied volatility surfaces, as the volatility smile becomes more prominent as option contracts approach expiration, while it is smaller for longer expiry dates. This effect can be visualized both in the curves for call options and for put options.

In this work, curves were estimated for each day in the sample, but it would be unfeasible to analyze surface by surface the entire out-of-sample period, which has 1607 days. Thus, to analyze the period as a whole, the aggregated data of the estimated surfaces will be used.

Table 5 shows the average error metrics calculated considering all sur-

faces estimated for the call options. The errors were discriminated according to the maturity of each curve. In the table, it is possible to notice, in general, that the biggest errors are concentrated in the shorter and longer maturities, that is, they occur in the initial and final portions of the surface. The region of intermediate maturities is where the best estimates for implied volatilities are obtained. It is noteworthy that these results are consistent with what was observed in the regressions, through the estimated credible intervals.

Table 5
Fit Measures by Maturity - Call options

	ME	RMSE	MAE	Maturity
1	-0.00029578	0.04959469	0.03408946	0.01190476
2	-0.00019231	0.05942243	0.03583769	0.01587302
3	-0.00032002	0.04851704	0.03726949	0.01984127
4	-0.00029510	0.05790880	0.03804669	0.02380952
5	-0.00017984	0.06905411	0.04320385	0.02777778
6	-0.00022597	0.06898533	0.04256784	0.03174603
7	-0.00008615	0.02998500	0.02335438	0.06746032
8	-0.00006069	0.04209598	0.03359188	0.07142857
9	-0.00013128	0.05255464	0.03811689	0.07539683
10	-0.00013922	0.05360920	0.04351660	0.07936508
11	-0.00000438	0.05884527	0.03978898	0.08333333
12	-0.00006945	0.05588821	0.03872878	0.08730159
13	0.00027295	0.03378409	0.02271625	0.09126984
14	0.00016483	0.05930876	0.02970889	0.09523810
15	0.00000458	0.04993858	0.02978554	0.09920635
16	0.00016410	0.04191529	0.02369235	0.10317460
17	-0.00006764	0.03393187	0.02389310	0.16666667
18	0.00008035	0.09284850	0.06058907	0.17063492
19	-0.00013294	0.03963977	0.03380105	0.17460317
20	0.00017265	0.09067822	0.04790037	0.17857143
21	-0.00060011	0.04749486	0.03441645	0.18253968
22	0.00024507	0.04864295	0.04233694	0.18650794
23	0.00003121	0.02694296	0.02366308	0.19047619
24	-0.00012803	0.01978534	0.01612023	0.19444444
25	-0.00017289	0.02848108	0.01889085	0.19841270
26	-0.00010706	0.01771308	0.01374104	0.20238095
27	0.00000971	0.07667624	0.06080340	0.25396825

Note: Table presents the average ME, RMSE and MAE calculated by maturity for call options, using n=10 and k=5. The sample encompasses a period from 2016-01-04 to 2022-06-30.

In Table 6 are the fit measures calculated for all estimated surfaces of the

put options. In the case of put options, the highest errors are mostly concentrated in the shortest expiry dates, however, in general, the pattern found in call options is repeated, with intermediate maturities showing smaller errors. Such results are in line with what was seen in the regression credible intervals.

Table 6
Fit Measures by Maturity - Put options

	ME	RMSE	MAE	Maturity
1	-0.000008	0.276572	0.195396	0.011905
2	-0.000204	0.251432	0.182027	0.015873
3	0.000333	0.290231	0.223977	0.019841
4	-0.000059	0.200795	0.149301	0.023810
5	-0.000088	0.161452	0.125182	0.027778
6	0.000024	0.194903	0.148540	0.031746
7	0.000097	0.250040	0.183533	0.035714
8	0.000073	0.122411	0.095864	0.039683
9	0.000004	0.129246	0.090720	0.043651
10	0.000056	0.157789	0.111927	0.047619
11	-0.000017	0.099822	0.074453	0.051587
12	0.000135	0.165126	0.107631	0.055556
13	-0.000007	0.091990	0.067100	0.059524
14	0.000051	0.096774	0.067579	0.063492
15	0.000068	0.109930	0.076456	0.067460
16	-0.000011	0.118498	0.090706	0.071429
17	-0.000005	0.103198	0.085422	0.075397
18	-0.000061	0.138264	0.100970	0.079365
19	0.000103	0.099218	0.078693	0.083333
20	0.000061	0.108247	0.086854	0.087302
21	0.000094	0.130126	0.106181	0.091270
22	-0.000014	0.131674	0.100771	0.095238
23	0.000079	0.104110	0.079705	0.099206
24	-0.000039	0.101306	0.083401	0.103175
25	-0.000090	0.103731	0.081645	0.107143
26	-0.000001	0.137527	0.100907	0.111111
27	0.000031	0.073661	0.055950	0.115079

Note: Table presents the average ME, RMSE and MAE calculated by maturity for put options, using n=63 and k=5. The sample encompasses a period from 2016-01-04 to 2022-06-30.

A possible hypothesis for this is that intermediate options contracts are the most traded and, therefore, have the highest number of points available to form the curve, making the estimates more accurate. In addition, because they have greater liquidity, they generate market prices with less distortions.

A second possible hypothesis is that shorter and longer maturities concentrate more risk than intermediate ones. The short ones because they suffer more abrupt and intense oscillations, while the long ones have a higher level of uncertainty regarding the possible price trajectories of the underlying asset. This causes greater ranges of implied volatility levels, with more heterogeneous prices, as well as greater oscillation, culminating in more dispersed values in relation to the estimated curves, which in turn generates a model with greater errors.

Another relevant point to highlight is that, in general, errors, when considering all metrics, were higher in put options than in call options. This pattern repeats itself for the aggregated data, as seen earlier when analyzing the surfaces separately. Again, the explanation may be related to the liquidity of these contracts, since call options are more liquid than put options in general and, therefore, have a greater number of points available for building the curve, forming more accurate estimates, as well as presenting prices with lower distortions given the greater volume of business.

Another possible hypothesis, which has already been raised for the case of individual surfaces, is the fact that downward shocks in the stock market are more intense and, consequently, cause greater volatility and greater variation in volatility levels, thus causing put options have greater dispersion in the implied volatilities of the different contracts.

Finally, when analyzing the results as a whole, it can be seen that the model was efficient and estimated both the curves and the surfaces as expected. Absolute errors had low values, although for some maturities and strike prices they showed non-negligible values. Such problems result from the non-verification, in practice, of an important assumption made by the BSM model, which is the no-existence of arbitrage.

Analyzing the market data, it is possible to note that at those points where the estimation presented more significant errors, based on the values found, there were, even if on a small scale, arbitrage opportunities. In order to resolve these incurrences, prices with low liquidity were discarded, as well as the most extreme values, both for exercise prices and for maturity periods.

Although these opportunities are short-lived, they are enough to create momentary distortions in prices and, consequently, in the implied volatility of some contracts. However, it is worth mentioning that these situations are increasingly rare, and have been decreasing as the sample has advanced to more recent periods. Which certainly indicates that the markets have become increasingly efficient, mainly due to the introduction of electronic trading, increased liquidity and trading volume and the more regular and continuous performance of market makers.

5. Conclusion

The advantages of using a Bayesian method for estimation is that it makes it possible to perform a statistical inference based on the observed sample, without the need of asymptotic properties and permitting to obtain a full characterization of the posterior distribution of the estimated quantities. This property is useful for this specific study, since the variance of the vector of regression coefficients with shape restrictions would not be easily obtained through a frequentist paradigm, since the imposition of restrictions alters the property of estimators, as discussed by example by (Gourieroux and Monfort, 2010). As the shape restrictions are imposing the no-arbitrage restrictions, we obtain full posterior inference under no-arbitrage conditions, and thus constructing flexible non-parametric estimators respecting the necessary financial assumptions needed for derivative pricing.

Another positive point was that it was possible to generate, from the same drawings of the posterior distribution, probability intervals for the regression function. In Bayesian statistics, a posterior probability interval is called a credible interval, This made it easier to look at the uncertainty level of the estimated curves in the model.

Finally, it is worth mentioning that the model presented good fit results, in a way that it would be possible to use it for future studies, as well as to expand the scope of the theme addressed in this work. A possible application of the model would be to use the estimated implied volatility surfaces as an input in the BSM model formula, replacing the historical volatility that is commonly used and thus be able to calculate options prices one day ahead of the date on which the implied volatilities were estimated.

From there, it would be possible to compare the results obtained with the classic pricing models mentioned in this work (BSM, binomial and Heston), verifying if there is a reduction in measurement errors and in which situations this occurs (For longer or longer maturities short, for OTM, ATM or ITM options). The idea, in general, is to verify whether the proposed model produces smaller pricing errors compared to the other models, confirming that the options market, through implied volatility, would be a good proxy for measuring the volatility to be used to price options.

References

ALEXANDER, C. (2001). Principal Component Analysis of Volatility Smiles and Skews, *Technical report*, University of Reading Working Paper in Finance, UK.

- ALMEIDA, C. and QIN, Z. (2020). A Bayesian nonparametric approach to option pricing, *Brazilian Review of Finance (Online)* **18**(4): 115–137.
- AVELLANEDA, M., CARELLI, A. and STELLA, F. (2000). A Bayesian approach for constructing implied volatility surfaces through neural networks, *Journal of Computational Finance* **4**(1): 83–107.
- BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy* **81**(3): 637–654.
- BLISS, R. and PANIGIRTZOGLU, N. (2002). Option-implied risk aversion estimates, *Journal of Finance* **59**(407-446).
- CALDEIRA, J. F., LAURINI, M. P. and PORTUGAL, M. S. (2010). Bayesian Inference Applied to Dynamic Nelson-Siegel Model with Stochastic Volatility, *Brazilian Review of Econometrics* **30**(1): 123–161.
- CAMPA, J. M., CHANG, P. K. and REIDER, R. L. (1998). Implied exchange rate distributions: evidence from otc option markets, *Journal of International Money and Finance* **17**(1): 117–160.
- CARR, P. and WU, L. (2010). A new simple approach for constructing implied volatility surfaces, *Technical report*, SSRN.
- CONT, R. and FONSECA, J. d. (2002). Dynamics of implied volatility surfaces, *Quantitative Finance* **2**: 45–60.
- COX, J. C., ROSS, S. A. and RUBINSTEIN, M. (1979). Option pricing: A simplified approach, *Journal of Financial Economics* **7**(3): 229–263.
- CURRY, H. and SCHOENBERG, I. (1966). On Pólya frequency functions iv: The fundamental spline functions and their limits, *Journal d'Analyse Mathématique* **17**: 71–107.
- DAY, T. and LEWIS, C. (1992). Stock market volatility and the information content of stock index options, *Journal of Econometrics* **52**(1-2): 267–287.
- Dupire, B. (1994). Pricing with a Smile, *Risk* **7**(1): 18–20.
- Gourieroux, C. and Monfort, A. (2010). *Statistics and Econometric Models*, Cambridge University Press, London.
- HACKSTADT, A. J. (2011). *Bayesian Shape restricted splines*, PhD thesis, Colorado State University.

- HESTON, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options , *Review of Financial Studies* **6**(2): 327–334.
- HOMESCU, C. (2011). Implied volatility surface: construction methodologies and characteristics , *Technical report*, SSRN.
- KAMAL, M. and DERMAN, E. (1997). The Patterns of Change in Implied Index Volatilities, *Technical report*, Quantitative Strategies Research Notes.
- LAURINI, M. (2011). Imposing no-arbitrage conditions in implied volatilities using constrained smoothing splines , *Applied Stochastic Models in Business and Industry* **27**(6): 649–659.
- LAURINI, M. (2013). Dynamic functional data analysis with non-parametric state space models , *Journal of Applied Statistics* **41**(1): 142–163.
- MALZ, A. (1997). Estimating the probability distribution of future exchange rates from option prices , *Journal of Derivatives* **5**(2): 18–36.
- MEYER, M. C. (2008). Inference using shape-restricted regression splines , *Annals of Applied Statistics* **2**(3): 1013–1033.
- NELSON, C. R. and SIEGEL, A. F. (1987). Parsimonious modeling of yield curves , *The Journal of Business* **60**(4): 473–89.
- ORLANDO, G. and TAGLIALATELA, G. (2017). A Review on Implied Volatility Calculation , *Journal of Computational and Applied Mathematics* **320**(15): 202–220.
- PANIGIRTZOGLU, N. and SKIADOPOULOS, G. (2004). A new approach to modeling the dynamics of implied distributions: Theory and evidence from the SP 500 options , *Journal of Banking & Finance* **28**(7): 1499–1520.
- RAMSEY, J. O. (1988). Monotone regression splines in action , *Statistical Science* **3**(4): 425–441.
- RUBINSTEIN, M. (1985). Nonparametric Tests of Alternative Option Pricing Models , *The Journal of Finance* **40**(2): 45–480.
- SCHMALENSEE, R. and TRIPPI, R. (1978). Common Stock Volatility Expectations Implied by Option Premia , *Journal of Finance* **33**(1): 129–147.

SHIMKO, D. (1993). Bounds of probability , *Risk* **6**(4): 33–37.

VICENTE and GUEDES (2010). A Volatilidade implícita contém informações sobre a volatilidade futura? Evidências do mercado de opções de ações , *Brazilian Business Review* **7**(1): 48–65.

XING, ZHANG and ZHAO (2010). What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns? , *Journal of Financial and Quantitative Analysis* **45**(3): 641–662.

A. Additional tables and figures

This appendix contains some additional tables and figures. Examples of implied volatilities calculated for call and put option contracts, respectively, can be seen in tables A1 and A2. Figures A1 and A2 include all implicit volatility curves for call options estimated for 2016-02-12. Figures A3 to A7 include all the implied volatility curves for put options estimated for 2016-04-29.

Table A1
Implied volatilities - Call options

	Date	Contract	Maturity	Time to Maturity	Moneyness	Implied Vol
1	2010-01-04	PETRA28	2010-01-18	0.04	0.76	0.93
2	2010-01-04	PETRA32	2010-01-18	0.04	0.84	0.55
3	2010-01-04	PETRC50	2010-03-15	0.19	1.34	0.29
4	2010-01-04	PETRC50	2010-03-15	0.19	1.35	0.30
5	2010-01-04	PETRC50	2010-03-15	0.19	1.34	0.29
6	2010-01-04	PETRC48	2010-03-15	0.19	1.29	0.29
7	2010-01-04	PETRC48	2010-03-15	0.19	1.30	0.30
8	2010-01-04	PETRC48	2010-03-15	0.19	1.29	0.29
9	2010-01-04	PETRC46	2010-03-15	0.19	1.23	0.29
10	2010-01-04	PETRC46	2010-03-15	0.19	1.24	0.28
11	2010-01-04	PETRC46	2010-03-15	0.19	1.23	0.27
12	2010-01-04	PETRC44	2010-03-15	0.19	1.17	0.29
13	2010-01-04	PETRC44	2010-03-15	0.19	1.19	0.29
14	2010-01-04	PETRC44	2010-03-15	0.19	1.17	0.28
15	2010-01-04	PETRA34	2010-01-18	0.04	0.90	0.31
16	2010-01-04	PETRA34	2010-01-18	0.04	0.91	0.33
17	2010-01-04	PETRA36	2010-01-18	0.04	0.96	0.27
18	2010-01-04	PETRA36	2010-01-18	0.04	0.97	0.29
19	2010-01-04	PETRA36	2010-01-18	0.04	0.96	0.25
20	2010-01-04	PETRA38	2010-01-18	0.04	1.00	0.26
21	2010-01-04	PETRA38	2010-01-18	0.04	1.02	0.29
22	2010-01-04	PETRA38	2010-01-18	0.04	1.00	0.25
23	2010-01-04	PETRA39	2010-01-18	0.04	1.04	0.30
24	2010-01-04	PETRA39	2010-01-18	0.04	1.05	0.27
25	2010-01-04	PETRA39	2010-01-18	0.04	1.04	0.30
26	2010-01-04	PETRA40	2010-01-18	0.04	1.06	0.27
27	2010-01-04	PETRA40	2010-01-18	0.04	1.07	0.28
28	2010-01-04	PETRA40	2010-01-18	0.04	1.06	0.26
29	2010-01-04	PETRA42	2010-01-18	0.04	1.11	0.31
30	2010-01-04	PETRA42	2010-01-18	0.04	1.12	0.31
31	2010-01-04	PETRA42	2010-01-18	0.04	1.11	0.29
32	2010-01-04	PETRA44	2010-01-18	0.04	1.17	0.38
33	2010-01-04	PETRA44	2010-01-18	0.04	1.19	0.37
34	2010-01-04	PETRA44	2010-01-18	0.04	1.17	0.34
35	2010-01-04	PETRA46	2010-01-18	0.04	1.23	0.47

Note: Table presents some examples of implied volatilities calculated for call options on 2010-01-04.

Table A2
Implied volatilities - Put options

	Date	Contract	Maturity	Time to Maturity	Moneyneyness	Implied Vol
1	2010-01-04	PETRM34	2010-01-18	0.04	0.90	0.40
2	2010-01-04	PETRM34	2010-01-18	0.04	0.91	0.37
3	2010-01-04	PETRM34	2010-01-18	0.04	0.90	0.40
4	2010-01-04	PETRM36	2010-01-18	0.04	0.96	0.34
5	2010-01-04	PETRM36	2010-01-18	0.04	0.97	0.27
6	2010-01-04	PETRM36	2010-01-18	0.04	0.96	0.32
7	2010-01-04	PETRM38	2010-01-18	0.04	1.01	0.37
8	2010-01-04	PETRM38	2010-01-18	0.04	1.03	0.17
9	2010-01-04	PETRM38	2010-01-18	0.04	1.01	0.29
10	2010-01-04	PETRM40	2010-01-18	0.04	1.07	0.50
11	2010-01-04	PETRM40	2010-01-18	0.04	1.07	0.32
12	2010-01-04	PETRN36	2010-02-08	0.10	0.96	0.32
13	2010-01-04	PETRN36	2010-02-08	0.10	0.97	0.29
14	2010-01-04	PETRN36	2010-02-08	0.10	0.96	0.32
15	2010-01-04	PETRN37	2010-02-08	0.10	0.98	0.32
16	2010-01-04	PETRN37	2010-02-08	0.10	1.00	0.27
17	2010-01-04	PETRN37	2010-02-08	0.10	0.98	0.32
18	2010-01-05	PETRN38	2010-02-08	0.10	1.01	0.37
19	2010-01-05	PETRN38	2010-02-08	0.10	1.03	0.22
20	2010-01-05	PETRN38	2010-02-08	0.10	1.02	0.29
21	2010-01-05	PETRN36	2010-02-08	0.10	0.96	0.33
22	2010-01-05	PETRN36	2010-02-08	0.10	0.97	0.26
23	2010-01-05	PETRN36	2010-02-08	0.10	0.97	0.30
24	2010-01-05	PETRM40	2010-01-18	0.04	1.06	0.47
25	2010-01-05	PETRM38	2010-01-18	0.04	1.01	0.38
26	2010-01-05	PETRM38	2010-01-18	0.04	1.03	0.16
27	2010-01-05	PETRM38	2010-01-18	0.04	1.02	0.29
28	2010-01-05	PETRM36	2010-01-18	0.04	0.95	0.37
29	2010-01-05	PETRM36	2010-01-18	0.04	0.97	0.23
30	2010-01-05	PETRM36	2010-01-18	0.04	0.96	0.33
31	2010-01-06	PETRP43	2010-04-19	0.28	1.14	0.35
32	2010-01-06	PETRP43	2010-04-19	0.28	1.16	0.27
33	2010-01-06	PETRP43	2010-04-19	0.28	1.14	0.35
34	2010-01-06	PETRP37	2010-04-19	0.28	0.97	0.33
35	2010-01-06	PETRP37	2010-04-19	0.28	0.99	0.30

Note: Table presents some examples of implied volatilities calculated for put options on

2010-01-04.

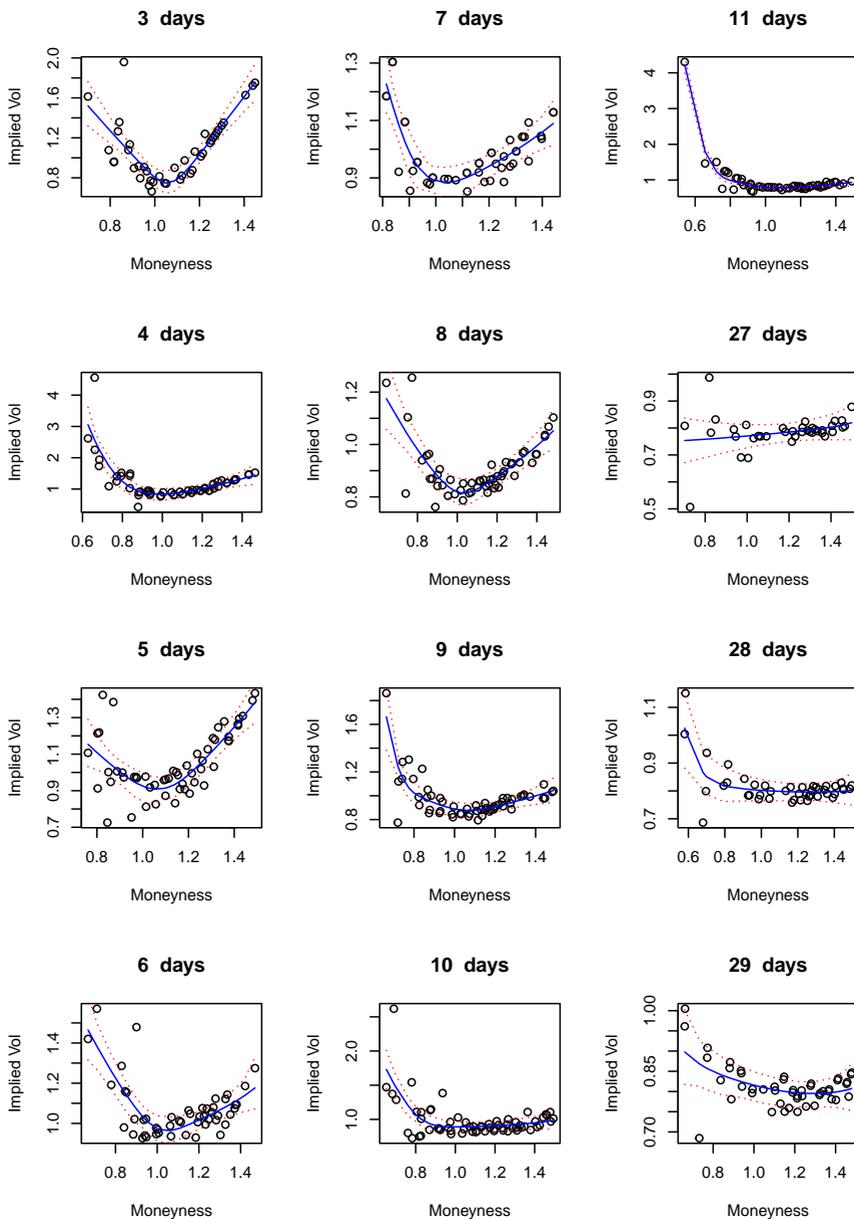


Figure A1

Call options by maturity - 2016-02-12

Note: Figure provides graphs for the curves estimated on 02/12/2016 by maturities from 3 to 29 days for call options.

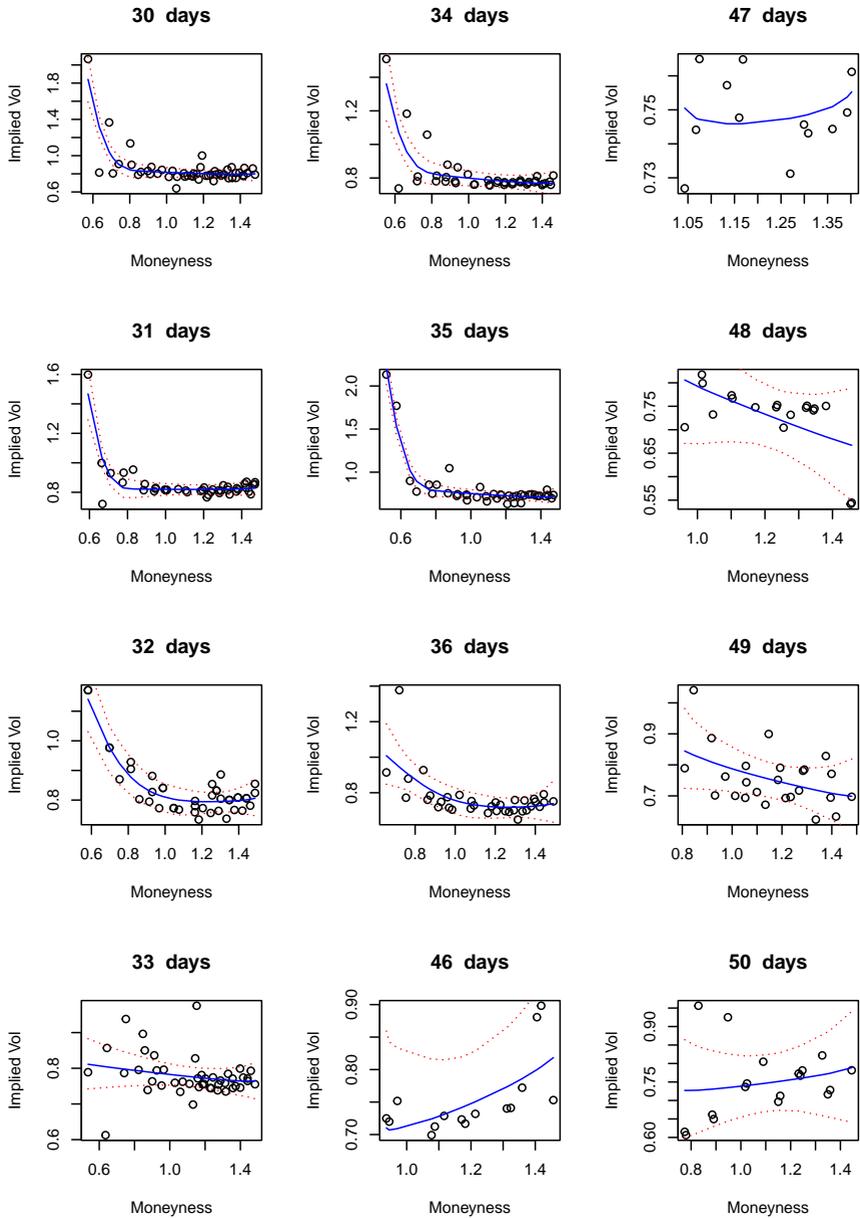


Figure A2

Call options by maturity - 2016-02-12

Note: Figure provides graphs for the curves estimated on 02/12/2016 by maturities from 30 to 50 days for call options.

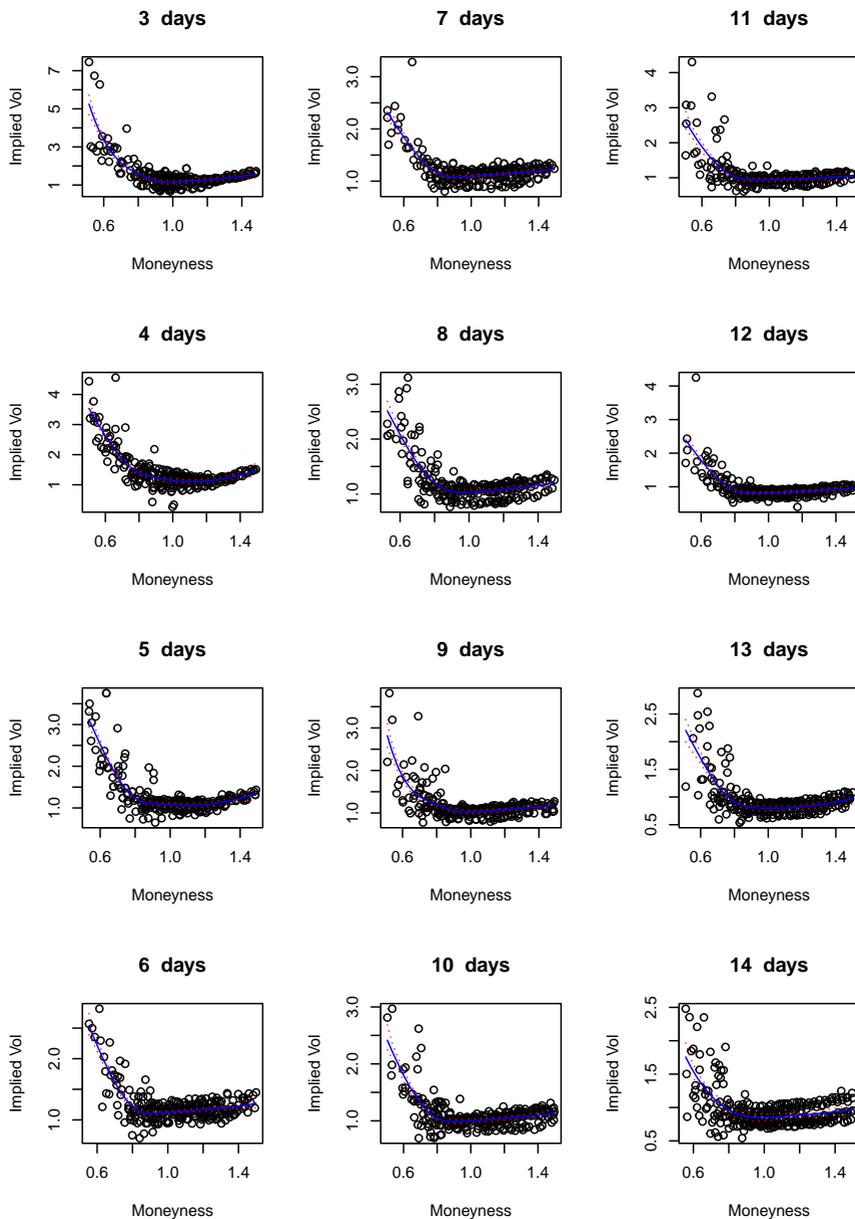


Figure A3

Put options by maturity - 2016-04-29

Note: Figure provides graphs for the curves estimated on 04/29/2016 by maturities from 3 to 14 days for put options.

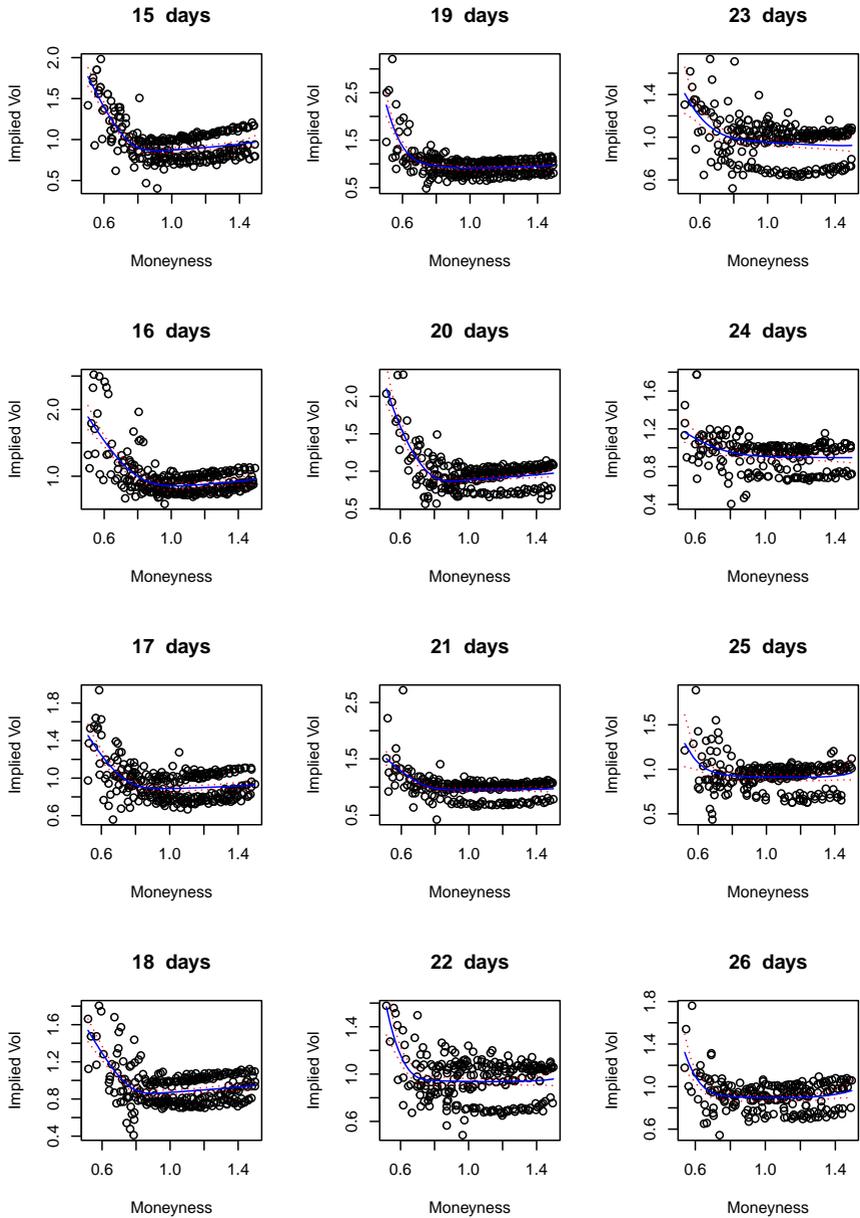


Figure A4

Put options by maturity - 2016-04-29

Note: Figure provides graphs for the curves estimated on 04/29/2016 by maturities from 15 to 26 days for put options.

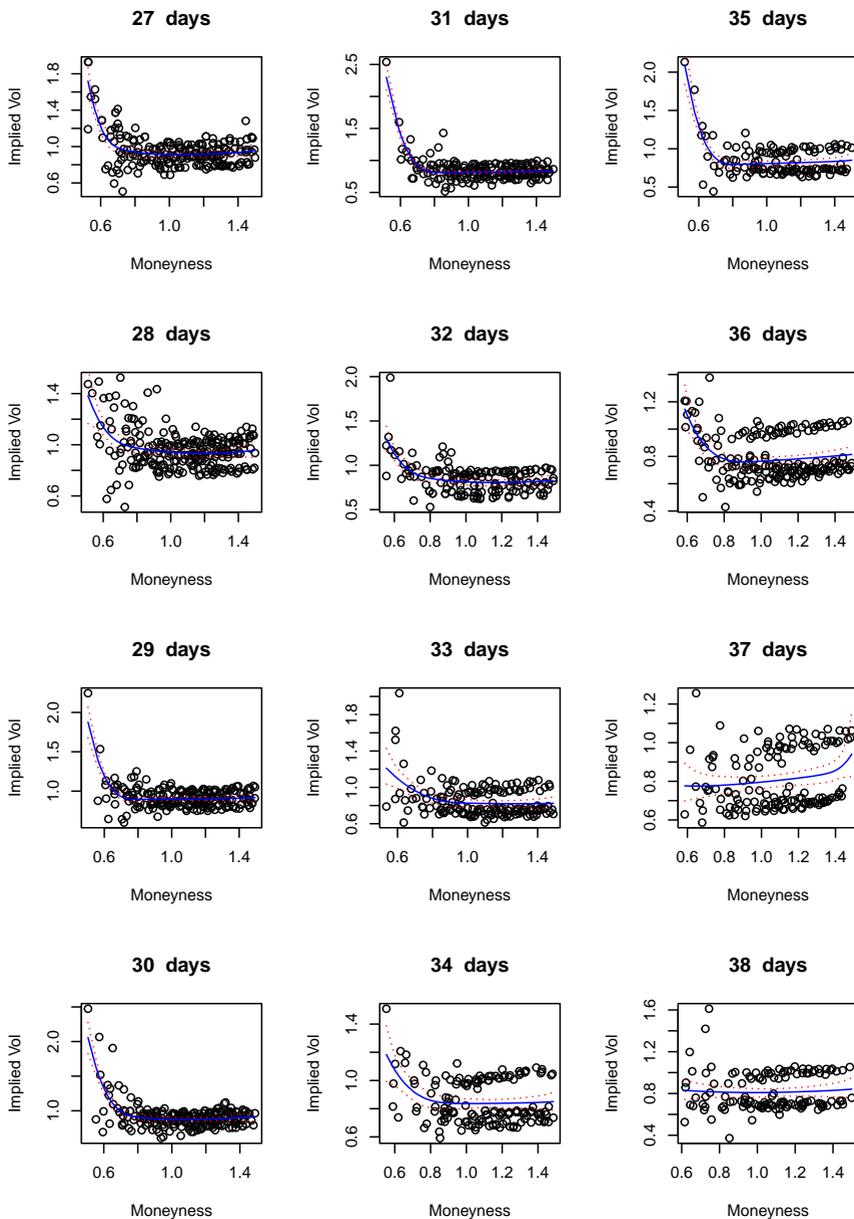


Figure A5

Put options by maturity - 2016-04-29

Note: Figure provides graphs for the curves estimated on 04/29/2016 by maturities from 27 to 38 days for put options.

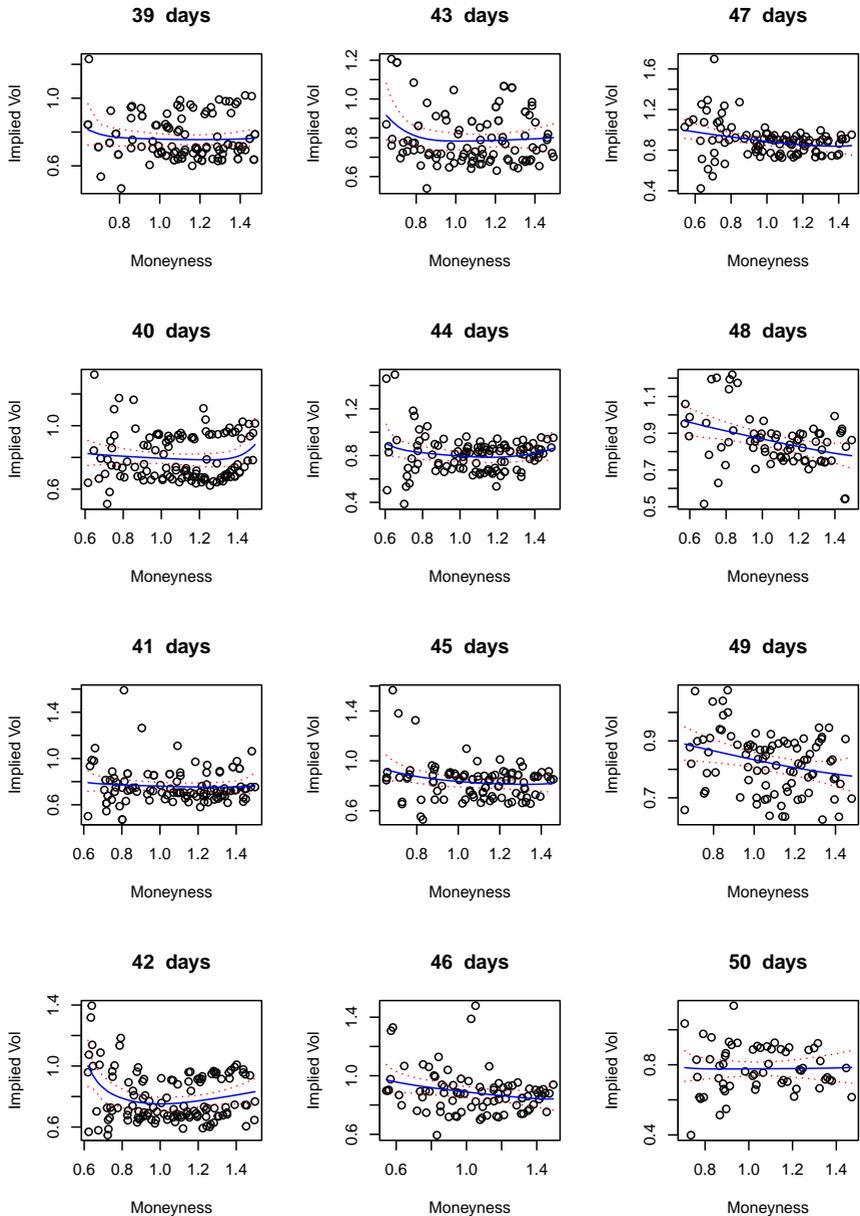


Figure A6

Put options by maturity - 2016-04-29

Note: Figure provides graphs for the curves estimated on 04/29/2016 by maturities from 39 to 50 days for put options.

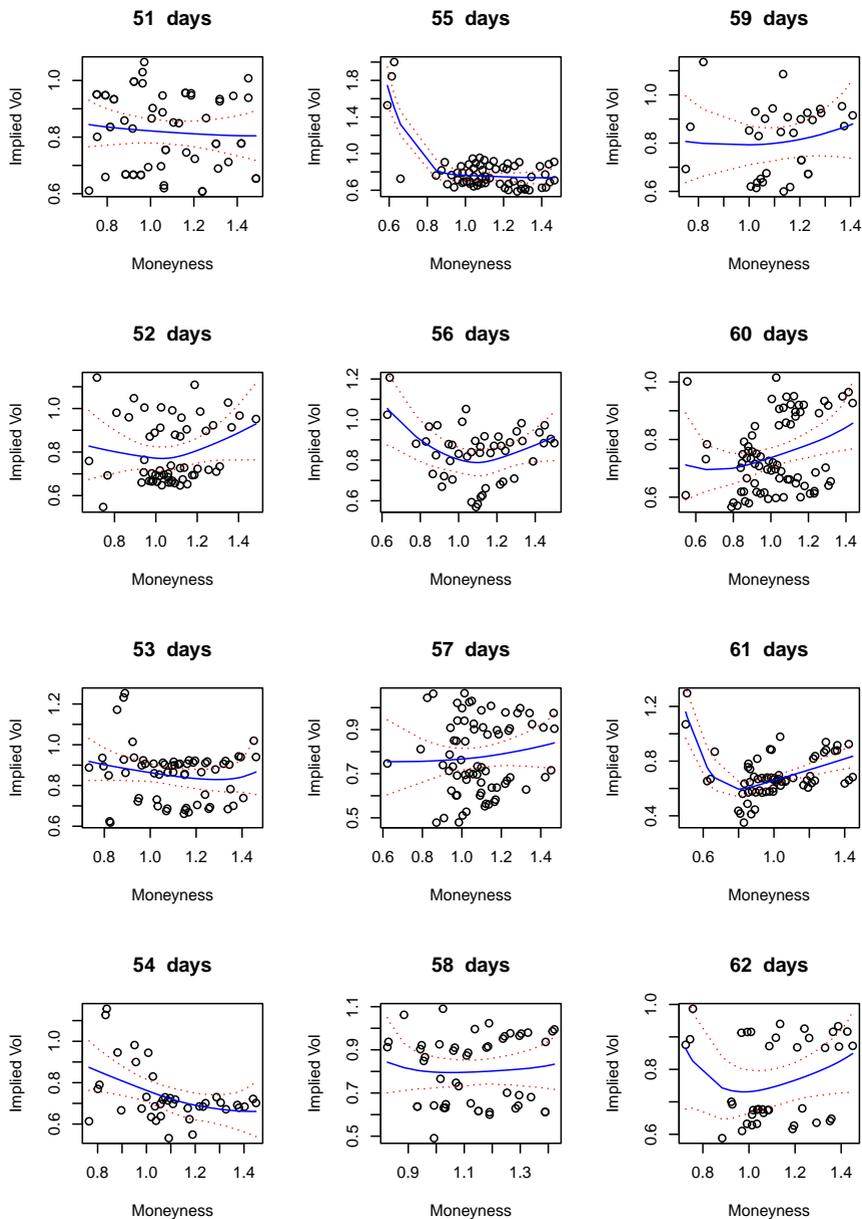


Figure A7

Put options by maturity - 2016-04-29

Note: Figure provides graphs for the curves estimated on 04/29/2016 by maturities from 51 to 62 days for put options.