Option-implied probability distributions: how reliable? how jagged?

Marco Taboga*

Bank of Italy

December 4, 2014

Abstract

Estimates of option-implied probability distributions are routinely used in central banks, as well as in other institutions, but their reliability is often difficult to assess. To address this issue, we propose a semi-nonparametric model that allows to compute exact confidence intervals around estimated distributions. By analyzing a panel of S&P 500 options, we find that the estimates of the distributions are quite precise. We also provide evidence that the multi-modality often found in option-implied distributions could be an artifact due to over-fitting, and that models with uni-modality constraints have high posterior odds.

JEL classification: C14, C58, G13.

Keywords: implied state prices, implied risk-neutral distributions, state price estimation, option-implied distributions.

*Marco Taboga, Economic Outlook and Monetary Policy Department, Bank of Italy. Via Nazionale 91, 00184, Roma, Italy. e-mail: marco.taboga@bancaditalia.it

Introduction

Since the seminal contributions of Ross (1976), Breeden and Litzenberger (1978), and Banz and Miller (1978), the problem of estimating risk-neutral probabilities (or, equivalently, state prices[1]) from cross-sections of option prices has received much attention in the literature. Furthermore, several papers have demonstrated the importance of estimating state prices for various empirical applications, including the estimation of investors' risk preferences, the assessment of market beliefs about events of interest, the pricing of exotic derivatives, the estimation of parametric asset pricing models, and the calibration of risk management models (see Bondarenko 2003 for a review of applications).

We propose a new model to estimate state prices, which belongs to the class of semi-nonparametric models (see Gallant 1987, and Fengler and Hin 2014). In our model, the vector of state prices is assumed to be interpolated by a spline function. We show that this assumption is equivalent to imposing a set of linear equality restrictions on state prices. We propose two estimators of the vector of state prices. The first one is a least absolute deviations (LAD) estimator that can be obtained as the solution of a computationally inexpensive linear programming problem. The second one is a Bayesian estimator. We prove that the Bayesian estimation problem boils down to updating the prior on the coefficients of a linear regression equation.

We use our model to tackle two issues that, to our knowledge, have been seldom addressed in the literature on the nonparametric estimation of state prices.

The first issue concerns the uncertainty that surrounds the estimates of state prices. While some of the existing models allow to derive the asymptotic distributions of state price estimators, asymptotics are a poor guide for predicting the behavior of the estimators because of the small sample sizes that typically characterize the cross sections of option

[1] State prices are the prices of Arrow-Debreu securities, that is, of securities that pay one unit of account in a specific state of nature and nothing in any other state. They coincide with risk-neutral probabilities when the risk-free interest rate is equal to zero. Otherwise, when the risk-free rate is different from zero, risk-neutral probabilities are proportional to state prices.

prices (Aït-Sahalia and Duarte 2003). Our Bayesian estimation procedure instead allows to derive exact confidence intervals around estimated state prices.

We believe that computing confidence intervals precisely is important not only from a statistical viewpoint, but also in view of the economic applications of state price estimation. Estimated state prices and risk-neutral densities are used to infer market beliefs about economic events of interest, for instance, in central banks, where they are routinely used for monetary policy purposes (e.g., Söderlind and Svensson 1997, Malz 2014). To understand how much weight these estimates should have in influencing policy decisions, it is essential to assess their reliability and statistical precision.

Furthermore, the statistical precision of the estimates has significant consequences for the pricing of derivatives. For example, suppose that a dealer bank uses estimated state prices to price bespoke derivatives traded with clients. It can be rational (e.g., Cont 2006, Routledge and Zin 2009) for the bank to adjust the spread between bid and ask quotes so as to seek protection from the uncertainty about the fair value of the bespoke derivative. Because this uncertainty stems from the uncertainty about the true value of the state prices, the bid and ask quotes ultimately depend on the statistical precision of the estimates of the state prices. Furthermore, in the absence of a proper quantification of the precision of the estimates, the fair value of the bespoke derivative is uncertain in the Knightian sense of the term, and the uncertainty can give rise to ambiguity (e.g., Ellsberg 1961), with the well-known behavioral implications in terms of pricing (e.g., Dana 2004, Easley and O'Hara 2010).

By analyzing a panel of options on the S&P 500 stock index with our model, we find that the posterior distributions of the state prices and of the risk-neutral probabilities have low dispersion (the standard deviation of the state prices is on average 1.7% per cent of their value). The posterior dispersion is even lower for the quantiles of the risk-neutral probability distributions. In particular, while it has been conjectured (e.g., Lee 2014) that tail quantiles might be difficult to estimate precisely with nonparametric methods, we find that the confidence bands around the estimated tail quantiles are quite tight.

The second issue we address is the multi-modality of the risk-neutral probability distributions derived from estimated state prices[2]. Estimated risk-neutral distributions are often reported to be multi-modal (e.g., Melick and Thomas 1997, Taylor, Tzeng and Widdicks 2012, Fengler and Hin 2014). Some studies have advanced possible explanations for multi-modality: heterogeneity in investors' beliefs (Ziegler 2002), jumps and correlation between volatility and returns (Frachot, Laurent and Pichot 1999), regime-switches (see Chabi-Yo, Garcia and Renault 2008 for a review). Despite the attempts to rationalize it, multi-modality is still deemed economically implausible by many economists, on the grounds of a *natura non facit saltum* kind of argument (the more extreme an outcome is, the less likely it should be). Furthermore, as suggested by Malz (2014) and Fengler and Hin (2014), multi-modality could be authentic, but it could also be an artifact of the techniques used to estimate risk-neutral distributions. Given the diversity of views on this issue, it is then desirable to have models that allow to impose the prior of uni-modality and to see what impact it has on estimated risk-neutral distributions and option pricing functions. For example, the argument in favor of uni-modality could be stronger if the prior had a minor impact on pricing errors, that is, on the ability to accurately fit option prices. To our knowledge, how to impose uni-modality restrictions within a nonparametric option pricing framework has been seldom discussed in the literature (although Fengler and Hin 2014, in suggesting directions for future research, hint to a possible extension of their model that could accommodate uni-modality). For this reason, we thoroughly discuss how to incorporate the prior of uni-modality in the Bayesian estimation of our model and what consequences the prior has.

We find that the risk-neutral probability distributions estimated from our data are often multi-modal, and sometimes display several modes both in the center and in the tails of the distribution. Imposing the prior of uni-modality can significantly change the shape of the estimated risk-neutral densities, but it has limited effects on the pricing errors and on the estimated quantiles of the risk-neutral distribution of returns. Furthermore, by computing

[2] While in this paper we propose estimators of state prices, our estimators can be easily rescaled to produce estimators of risk-neutral probability distributions.

posterior odds ratios between uni-modal and multi-modal models, we find that uni-modality is, a posteriori, much more likely than multi-modality. We interpret our findings as empirical support for the hypothesis that the multi-modality often found in option-implied distributions could be an artifact due to over-fitting.

As reported by Lee (2014), the smoothed implied volatility smile (SML) method, first proposed by Shimko (1993), is one of the most widely used techniques for estimating risk-neutral densities and state prices. The SML method, in its several variants (e.g., Malz 1997, Campa, Chang and Reider 1998, Bliss and Panigirtzoglou 2002, Bu and Hadry 2007), is a flexible and computationally efficient non-parametric method that allows to smooth option prices so that they can be numerically differentiated in a reliable manner to obtain state prices[3]. The model we propose has a computational tractability that is comparable to that of the SML method. However, while the SML method is not able to ensure that fitted option prices are arbitrage-free and that estimated state prices are positive, our model produces arbitrage-free estimates of all quantities of interest, because it allows to estimate state prices directly and to impose positivity restrictions on them.

Other non-parametric models that allow to impose no-arbitrage constraints on fitted option prices and estimated state prices have been proposed in the literature, for example, by Aït-Sahalia and Duarte (2003), Bondarenko (2003), Birke and Pilz (2009), and Fan and Mancini (2009). These models usually allow to derive only the asymptotic distributions of the state price estimators. However, as we already mentioned above, with the relatively small sample sizes usually available in real data sets of option prices, the estimators tend to behave much differently than predicted by asymptotics. As a matter of fact, the majority of the aforementioned studies analyze the properties of the estimators by means of Monte Carlo studies in which the true distribution of state prices is assumed to belong to a parametric family. While the results from the Monte Carlo studies are generally indicative of the behavior to be expected in practical cases, they do not allow to compute exact confidence

[3] As proved by Breeden and Litzenberger (1978), the state price density is equal to the second partial derivative of the option price function with respect to the strike price.

intervals around a given point estimate when the true distribution is unknown. Our model, instead, allows to do so.

A number of studies (e.g., Jarrow and Rudd 1982, Madan and Milne 1994, Bahra 1997, Melick and Thomas 1997, Rosenberg 1998) have proposed parametric methods to model risk-neutral distributions (see Jackwerth 2004 for a review). As noted by Birke and Pilz (2009), there are no generally accepted parametric forms for asset price dynamics, for volatility surfaces or for call and put price functions, and parametric approaches incur the risk of mis-specification and of introducing systematic errors in the estimates of state prices. While there can be situations when a parametric approach is preferable (e.g., when option price data is very sparse or noisy), the current consensus is that nonparametric models generally provide more reliable estimates.

Finally, let us mention that the literature on the estimation of risk-neutral distributions is vast and there are several important contributions that do not fall into any of the strands that we have mentioned above (such as, for example, the regularization approach of Jackwerth and Rubinstein 1996). We do not review these contributions here, but we refer the reader to the excellent surveys in Bondarenko (2003), Jackwerth (2004), and Fengler and Hin (2014).

The paper is organized as follows: Section 1 presents the probabilistic framework used to price options; Section 2 introduces the spline methodology used to interpolate state prices; Section 3 derives the LAD estimator; Section 4 discusses Bayesian estimation; Section 5 contains the empirical analysis; Section 6 concludes.

1 The pricing setting

In this section we set out the main elements of the probabilistic setting that we use to price options and estimate state prices.

Denote by S the stochastic value of an asset at a future date. We are going to assume that S has a discrete probability distribution with finite support

$$R_S = \{s_1, \ldots, s_n\} \tag{1}$$

While many asset pricing models assume a continuous probability distribution, the discreteness assumption is not restrictive in practice because any continuous distribution can be approximated arbitrarily well by a discrete one. Furthermore, continuous models used to estimate risk-neutral densities and state price densities are often discretized at some stage of the estimation process, so that the choice between a continuous and a discrete framework becomes a matter of convenience and analytical tractability, and has no relevant consequences in practice. Moreover, asset prices are in fact discrete because of the discreteness of monetary units and of the minimum tick rules often imposed by exchanges.

In the absence of arbitrage opportunities (or if you prefer, in an economic equilibrium), there exist n positive state prices $\pi_1, \ldots, \pi_n$ such that the price $\Pi(f)$ of any derivative security with payoff $f(S)$ can be written as

$$\Pi(f) = \sum_{j=1}^{n} \pi_j f(s_j) \tag{2}$$

where $f : \mathbb{R} \to \mathbb{R}$.

In particular, the price of a European call option with strike K can be written as

$$c_K = \sum_{j=1}^{n} \pi_j (s_j - K)^+ \tag{3}$$

and that of a put option as

$$p_K = \sum_{j=1}^{n} \pi_j (K - s_j)^+ \tag{4}$$

Suppose there are N_C call options with strikes $K_1^C, \ldots, K_{N_C}^C$ and N_P put options with strikes $K_1^P, \ldots, K_{N_P}^P$. If we denote by π the $n \times 1$ vector of state prices and by C and P the

$N_C \times 1$ vector of call prices and $N_P \times 1$ vector of put prices respectively, then

$$C = F_C \pi \tag{5}$$

$$P = F_P \pi \tag{6}$$

where F_C and F_P are $N_C \times n$ and $N_P \times n$ matrices of payoffs such that their generic (i,j)-th elements are

$$F_{C,ij} = \left(s_j - K_i^C\right)^+ \tag{7}$$

$$F_{P,ij} = \left(K_i^P - s_j\right)^+ \tag{8}$$

We use a more compact notation for equations (5-6):

$$Y = X\pi \tag{9}$$

where

$$Y = \begin{bmatrix} C \\ P \end{bmatrix} \tag{10}$$

is a $(N_C + N_P) \times 1$ vector of prices, and

$$X = \begin{bmatrix} F_C \\ F_P \end{bmatrix} \tag{11}$$

is a $(N_C + N_P) \times n$ matrix of payoffs. In practice, put and call prices are observed with error, so that the empirical counterpart to (9) is

$$Y^O = X\pi + \varepsilon \tag{12}$$

where Y^O is a $(N_C + N_P) \times 1$ vector of observed prices and ε is a $(N_C + N_P) \times 1$ vector of pricing errors that can arise for numerous reasons, among which the bounce between bid and ask quotes, price discreteness and stale prices due to illiquidity.

The problem that needs to be solved is to estimate π given that we observe Y^O and we know X. In mathematical terms, it is a regularization problem because n is usually larger than $N_C + N_P$. It is also complicated by the fact that π needs to be positive in order to rule out arbitrage opportunities. Note that the positivity of π is sufficient to rule out arbitrage opportunities.

As a final note, the risk-neutral probability distribution of S can be easily computed as

$$\rho = \frac{\pi}{\sum_{j=1}^{n} \pi_j} \tag{13}$$

where ρ is a vector whose entries are equal to the risk-neutral probabilities of the states.

2 The spline method

In this section we briefly outline the spline methodology that constitutes the principal building block of our model. The main result demonstrated in this section is that parametrizing state prices with a spline is equivalent to imposing a set of linear equality restrictions on state prices (eq. 22).

Without loss of generality, we assume that

$$s_1 < s_2 < \ldots < s_n \tag{14}$$

Morover, we assume that the points are equally spaced with distance δ, that is,

$$s_j = s_1 + (j-1)\,\delta \qquad j = 2, \ldots, n \tag{15}$$

and that there exists a spline function $\overline{\pi} : [s_1, s_n] \to \mathbb{R}_+$ that interpolates the state prices:

$$\pi_j = \overline{\pi}(s_j) \qquad j = 1, \ldots, n \tag{16}$$

In line with the majority of the literature on splines, we require that $\overline{\pi}(s)$ be piece-wise cubic and twice continuously differentiable on $[s_1, s_n]$ in order to obtain minimal curvature.

By using De Boor's (1978) B-spline construction, it can be proved that the first derivative of $\overline{\pi}(s)$ is piece-wise quadratic, the second derivative is piece-wise linear, the third is a step-wise constant function, and the fourth is a function that is zero everywhere except at knot points (this is illustrated in Figure 1 for a spline that interpolates a Gaussian curve). We will use the latter result on fourth derivatives to build our model.

In particular, assume that the set of knots of the spline is a subset of R_S, formed by $N_T < n-4$ points, and that the smallest four points of R_S cannot be knots (this is necessary because we are going to work with fourth differences). If we denote by σ the $(n-4) \times 1$ vector that contains all points of R_S (in increasing order) except the smallest four, then the $N_T \times 1$ vector of knot points T satisfies

$$T = M\sigma \tag{17}$$

where M is a $N_T \times (n-4)$ selection matrix whose rows are vectors of the Euclidean basis of $\mathbb{R}^{n-4}$. For example, if σ_j (the j-th element of σ) is also the i-th element of T, then the i-th row of M is a row vector containing a 1 in position j and 0 in all other positions.

In a similar manner, if we define the $(n-4-N_T) \times 1$ vector U of points of σ that are not knot points, we have

$$U = N\sigma \tag{18}$$

where N is a $(n-4-N_T) \times (n-4)$ selection matrix whose rows are also vectors of the Euclidean basis of $\mathbb{R}^{n-4}$.

Now, for any $k \in \mathbb{N}$ such that $k > 1$, denote by D_k the $(k-1) \times k$ first-difference matrix defined as

$$D_k = \begin{bmatrix} -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ & & \ddots & & & \\ & & & \ddots & & \\ 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix} \tag{19}$$

Then, the fourth-differences of state prices can be obtained as

$$D\pi \tag{20}$$

where

$$D = D_{n-3}D_{n-2}D_{n-1}D_n \tag{21}$$

The fact that the fourth derivative of $\overline{\pi}(s)$ is zero everywhere except at knot points[4] implies a linear restriction on the vector of state prices:

$$ND\pi = 0 \tag{22}$$

This restriction provides the regularization that we use to estimate the state prices π from equation (12). In other words, modelling state prices as a spline is tantamount to imposing a set of linear restrictions on π.

We conducted several experiments to test whether the proposed spline function is able to reproduce shapes that might resemble those of state price functions, and we found that the spline is almost always indistinguishable from the true function. We do not provide a detailed account of the results here, but we collect some plots of common density functions and their spline interpolations in Figure 2.

[4] At knot points, the fourth derivative is a Dirac delta (i.e., it is infinite), but fourth differences are finite.

3 LAD estimation

In this section we present the first estimator of state prices, based on the pricing equation (12) and on the set of linear restrictions (22). The main result presented in this section is that the LAD estimator of the state prices can be obtained as the solution of a simple linear program (eq. 28).

We propose a LAD estimator, obtained from the minimization of absolute pricing errors, instead of a more traditional least squares (LS) estimator, obtained from the minimization of squared pricing errors, because of the computational convenience and the robustness to outliers (e.g., Koenker and Hallock 2001) of the former. In particular, we believe that statistical robustness is an important advantage of the method, in light of the abundant evidence that data on option prices is often contaminated by outliers (e.g., Corrado and Su 1997).

Furthermore, the LAD estimator, besides being interesting in its own right, is instrumental to the Bayesian estimator that we propose in the next section. As we explain in greater detail below, the LAD estimates can be conveniently employed to initialize the MCMC algorithm used in the Bayesian estimation. Furthermore, they are used to analyze the heteroskedasticity of pricing errors, and the results of the analysis are also employed in the Bayesian estimation.

Our estimator $\widehat{\pi}_{LAD}$ of the state prices is the LAD estimator obtained by minimizing weighted absolute pricing errors subject to constraints:

$$\widehat{\pi}_{LAD} = \arg\min_{\pi} \sum_{i=1}^{N_C+N_P} w_i \left| Y_i^O - X_i \pi \right| \tag{23}$$

$$\text{s.t.} \quad ND\pi = 0 \; , \; \pi \geq 0 \tag{24}$$

where Y_i^O and X_i are the rows of Y^O and X respectively, and w_i are the weights assigned to pricing errors.

To our knowledge, there is no standard way to assign weights (but see the discussion by

Bliss and Panigirtzoglou 2002). A possibility is to simply set

$$w_i = 1 \tag{25}$$

for all i, so as to minimize absolute pricing errors in dollar terms.

Another possibility is to choose

$$w_i = \frac{1}{Y_i^O} \tag{26}$$

so as to minimize percentage pricing errors.

However, one might want to take into account the fact that the more an option is out-of-the money, the more it tends to trade infrequently, to have a large bid-ask spread, and, as a consequence, to be characterized by large pricing errors (in percentage terms). Therefore, a dampening factor could be applied to (26) in order to take heteroskedasticity into account. This can be accomplished, for example, by setting

$$w_i = \frac{1}{\sqrt{Y_i^O}} \tag{27}$$

Yet another possibility is to employ a feasible generalized scheme: i) run a first-stage estimation (by assigning weights with any of the methods proposed above); ii) compute estimated pricing errors; iii) estimate a weighting function from the estimated pricing errors; iv) run a second-stage estimation with the weights thus estimated.

All of the above choices will be discussed in the empirical part of the paper.

The minimization problem can be written as a linear programming (LP) problem:

$$\min_z d^\intercal z \tag{28}$$

$$\text{s.t.} \quad Az = b \ , \ z \geq 0 \tag{29}$$

where

$$d = \begin{bmatrix} w \\ w \\ 0 \end{bmatrix} \qquad z = \begin{bmatrix} \varepsilon^+ \\ \varepsilon^- \\ \pi \end{bmatrix} \tag{30}$$

are $(2N_C + 2N_P + n) \times 1$ vectors, and

$$A = \begin{bmatrix} I & -I & X \\ 0 & 0 & ND \end{bmatrix} \qquad b = \begin{bmatrix} Y^O \\ 0 \end{bmatrix} \tag{31}$$

In the definitions (30) and (31) we have denoted by w the $(N_C + N_P) \times 1$ vector of weights, by I a $(N_C + N_P) \times (N_C + N_P)$ identity matrix, and we have defined two $(N_C + N_P) \times 1$ artificial vectors ε^+ and ε^- that are constrained to be equal to the positive and negative parts of the pricing errors.

Of course, as any LP problem, the minimization problem in (28) is guaranteed to have a solution, which can be found by means of standard (and computationally inexpensive) LP algorithms, such as the simplex algorithm.

4 Bayesian estimation

In this section we discuss the Bayesian estimation of the vector of state prices. In particular, we prove that the estimation problem boils down to updating the prior on the coefficients of a linear regression equation.

As demonstrated in Section 2, the vector of state prices satisfies

$$ND\pi = 0 \tag{32}$$

while the fourth differences of π at the knot points are not necessarily zero:

$$MD\pi = d_\pi \tag{33}$$

where d_π is a $N_T \times 1$ vector of free parameters. The equations (32) and (33) completely characterize the fourth differences of π, but four initial conditions are needed in order to recover π as a function of the free parameters. We specify the initial conditions as follows:

$$E_4 \pi = i_\pi \tag{34}$$

where

$$E_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \end{bmatrix} \tag{35}$$

and i_π is a 4×1 vector of parameters. In other words, initial conditions are specified as equality conditions on the first four entries of π.

Denote by β the $(N_T + 4) \times 1$ vector of all the free parameters, obtained by stacking d_π and i_π:

$$\beta = \begin{bmatrix} d_\pi \\ i_\pi \end{bmatrix} \tag{36}$$

Denote by H the $n \times n$ invertible matrix defined by

$$H = \begin{bmatrix} ND \\ MD \\ E_4 \end{bmatrix} \tag{37}$$

Then, by solving equations (32), (33) and (34), the vector of state prices π can be expressed as a function of the vector of parameters β as follows:

$$\pi = H^{-1} J \beta \tag{38}$$

where

$$J = \begin{bmatrix} 0 & I \end{bmatrix}^\top \tag{39}$$

is a $n \times (N_T + 4)$ matrix constructed by adjoining a zero matrix to the $(N_T + 4) \times (N_T + 4)$ identity matrix.

Given this parametrization of state prices, the observed option prices (eq. 12) can be expressed as

$$Y^O = \xi\beta + \varepsilon \tag{40}$$

where

$$\xi = XH^{-1}J \tag{41}$$

Bayesian estimation of the regression equation (40) is carried out by assuming that the pricing errors have a multivariate normal distribution with mean 0 and covariance matrix

$$qW \tag{42}$$

where $q \in \mathbb{R}_{++}$ and W is a diagonal matrix whose diagonal elements depend on the weights assigned to the options (see previous Section):

$$W_{ij} = \begin{cases} 1/w_i^2 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{43}$$

In order to keep the analysis objective, the following uninformative and mutually independent priors are assigned to the parameters:

- q: uniform improper prior on $[0, \infty)$;

- β: uniform improper prior on the set of admissible values

$$A_\beta = \left\{\beta \in \mathbb{R}^{N_T+4} : \pi = H^{-1}J\beta \geq 0\right\} \tag{44}$$

The set of admissible values A_β contains all vectors π that are compatible with the absence of arbitrage opportunities.

There are several methods that one could use to generate draws from the posterior distributions of the parameters q and β, and of the vector of state prices π. We propose to use the random walk Metropolis-Hastings algorithm with block structure[5] (e.g., Bagasheva et al. 2008). Details on the Markov Chain Monte Carlo (MCMC) algorithm can be found in the Appendix. However, notice that the LAD estimation procedure presented in the previous section facilitates Bayesian estimation. As a matter of fact, we can use the LAD estimate $\widehat{\pi}_{LAD}$ as the initial value of the Markov Chain. Since $\widehat{\pi}_{LAD}$ usually lies in a region having high posterior probability, this initialization helps the chain to converge quickly.

The proposed Bayesian estimation method also allows to easily impose uni-modality on the risk-neutral probability distribution. Because the risk-neutral distribution ρ is proportional to π (eq. 13), uni-modality can be imposed directly on π. Define

$$\varphi(\pi) = \arg\max_i \pi_i \tag{45}$$

and a $(n-1) \times 1$ vector $\omega(\pi)$ such that

$$\omega_i(\pi) = \begin{cases} 1 & \text{if } i < \varphi(\pi) \\ -1 & \text{if } i \geq \varphi(\pi) \end{cases} \tag{46}$$

for $i = 1, \ldots, n-1$.

Then, the set of vectors that satisfy uni-modality is

$$\Omega = \left\{ \pi \in \mathbb{R}^n_+ : (D_n \pi) \circ \omega(\pi) \geq 0 \right\} \tag{47}$$

where $\circ$ denotes the Hadamard product and D_n is the first-difference matrix defined in

[5] q and each of the entries of β constitute separate blocks.

equation (19). As a consequence, uni-modality can be imposed by replacing A_β with[6]

$$A'_\beta = \left\{ \beta \in \mathbb{R}^{N_T+4} : \pi = H^{-1} J\beta \geq 0 \text{ and } \pi \in \Omega \right\} \tag{48}$$

5 Empirical results

This section presents the results of some empirical exercises carried out with our model.

5.1 The data

We employ data on the European options written on the S&P 500 stock market index, published by the Options Reporting Authority. Our dataset comprises daily observations, for the semester going from December 31st 2013 to June 30th 2014, of the settlement (closing) prices of all the quoted options with maturity December 20th 2014. We consider only the options whose strikes are multiples[7] of 25, for a total of 76 call and 76 puts. The closing prices of all of these options are recorded on each day in our sample period (in other words, we have no missing values and our panel is balanced). The lowest strike price is 100 and the highest is 3000. For comparison, the closing prices of the underlying S&P 500 index ranged from 1742 to 1963 during our sample period.

5.2 Results from the LAD estimation

In order to carry out the LAD estimation of the state prices, we set the grid increment $\delta = 25$ and the extremes of the grid $s_1 = 25$ and $s_n = 3200$, so that the number of grid points is $n = 128$. Furthermore, we place a knot point every ten grid points (and at the last grid

[6] Also in this case it is possible to obtain a LAD estimate belonging to A'_β, which can be used as the initial state in the MCMC algorithm. This is accomplished as follows: i) run a first LAD estimation to obtain an estimate $\widehat{\pi}_{LAD} \in A_\beta$; ii) set $\varphi(\pi) = \overline{\varphi} = \arg\max_i \widehat{\pi}_{LAD,i}$; iii) run a second LAD estimation where the constraint $\pi \in \Omega$ is imposed as a set of $n-1$ additional linear inequality constraints in the linear program (28).

[7] There are also options whose strikes are multiples of 5 but not of 25. However, these options start trading only if certain conditions on the price of the underlying are satisfied, and they are generally less liquid.

point), so that the knot points are $s_5, s_{15}, \ldots, s_{125}, s_{128}$. Increasing the number of knot points has no visible effects on the results, while decreasing it worsens the ability of the model to fit observed option prices.

For each cross-section of option prices, we estimate the model twice. In the first estimation (LAD), we set weights as in equation (27). In the second estimation (Feasible LAD), we follow the feasible generalized procedure explained in Section 3. The weighting function is estimated by fitting a spline function to the absolute pricing errors computed in the first estimation.

We find that LAD and Feasible LAD provide estimates of the state prices that are, to all practical purposes, indistinguishable. Figure 3 illustrates this finding for only one cross-section of option prices (those recorded on the first day of our sample), but similar results hold also for the other cross-sections. The results obtained with other weighting schemes (see, e.g., eq. 25 and 26) are not reported here, but they are also similar to those we report. However, we note that other weighting schemes can sometimes cause convergence problems (such as slowness or inaccuracy of the solution) in the LP algorithm used to find the LAD estimates.

To further analyze pricing errors, we pool the estimated errors across dates and we plot their mean absolute value by the moneyness of their corresponding option (Figure 4). The moneyness of a call option with strike price K is computed as

$$\ln\left(S_F/K\right) \tag{49}$$

where S_F is the estimated forward price of the underlying:

$$S_F = \frac{1}{\sum_{j=1}^{n} \widehat{\pi}_{LAD,j}} \sum_{j=1}^{n} \widehat{\pi}_{LAD,j} s_j \tag{50}$$

The moneyness of a call option is instead computed as

$$\ln\left(K/S_F\right) \tag{51}$$

As conjectured, deep out-of-the-money options, that is, options with very low moneyness, tend to have pricing errors that are smaller in dollar terms, but larger in percentage terms. As a matter of fact, mean absolute pricing errors on deep-out-of the money options can be as large as 50 per cent of the price of the option, while they are less than 0.01 per cent for deep-in-the-money options. Note that the large percentage errors on deep-out-of-the-money options are not surprising because these options are traded very seldom an can have ask prices that are twice as large as the bid prices. Furthermore, for the majority of days in our dataset, the reported prices of deep-out-of-the-money options are not arbitrage-free[8] (e.g., put prices are often not monotonically increasing in the strike). To sum up, we find homoskedasticity neither in dollar nor in percentage terms, which motivates our choice of the weighting function proposed in equation (27).

Finally, to get a sense of the accuracy of our LAD estimator, we compare it to other two popular nonparametric methods: the positive convolution approximation (PCA) estimator of Bondarenko (2003) and the nonparametric locally linear estimator (NLLE) of Aït-Sahalia and Duarte (2003). We run a Monte Carlo study in which the true state price vector π is assumed to be equal to the state price vector displayed in Figure 3 (i.e., the LAD estimate of the state price vector for the first day in our sample) and observed option prices are generated according to equation (12). The entries of the vector of pricing errors ε are assumed to be independent normal random variables with variances equal to the variances estimated in the first step of the Feasible LAD estimation. Both PCA and NLLE require to set a bandwidth parameter that determines the smoothness of the nonparametric estimate of the state price function. We report the results obtained with the bandwidths that maximize the accuracy of

[8] Several other authors find that cross-sections of reported option prices are often not arbitrage-free (e.g., Aït-Sahalia and Duarte 2003).

the two methods. Note that PCA and NLLE assume a continuous distribution of states, so that they produce estimates of a state price density. We integrate the density by intervals in order to obtain a discrete vector of state prices comparable to our LAD estimate. We check that this conversion has no significant effects on the results of our study by also trying an alternative scheme: we convert both the true state price vector and its LAD estimate into a density[9] and we compare the latter to PCA and NLLE estimates; we find that results are almost unchanged.

The mean absolute errors committed in the estimation of the state prices (computed over 500 Monte Carlo draws, and expressed as a percentage of the average state price[10]) are 0.5, 1.4 and 2.2 per cent for LAD, PCA and NLLE, respectively. Figure 5 plots the mean absolute errors grouped by underlying price intervals. On most intervals, LAD is more accurate than PCA and NLLE, but we find that PCA is sometimes slightly more accurate than LAD in the tails of the distribution. This finding might be explained as follows: the PCA estimator is obtained by fitting a mixture of log-normals to state prices; mixtures of log-normals can be reproduced almost perfectly also by the spline function used in LAD (we tested this in experiments[11] that are not reported here), but the spline has the added flexibility of being able to assign negative weights to the components of the mixture (without violating positivity constraints on state prices); as a consequence, LAD is more flexible and it has the ability to reproduce a wider range of shapes, which helps to reduce estimation errors in the central part of the distribution, but this can cause over-fitting problems in the tails of the distribution, where information from option prices is noisier. Over-fitting problems that concern only a small fraction of the data points can usually be fixed by making the spacing between the knot points of the spline uneven (larger where there are over-fitting problems). We leave this extension of the model for future research.

[9] Accordingly, option prices are calculated as integrals.

[10] Note that it is not possible to express each error as a percentage of its corresponding state price because there are state prices (those corresponding to extreme realizations) that are numerically close to zero.

[11] The spline is able to fit a single log-normal distribution almost perfectly. The ability to fit mixtures descends from the linearity properties of the spline.

5.3 Results from the Bayesian estimation

For the Bayesian estimation of the model, we keep the same values of δ, s_1, s_n and the same knot points as in the LAD estimation. As a consequence, the vector of parameters to be estimated has 19 entries (14 values of the fourth derivative at the knot points, 4 initial conditions, for a total of 18 entries in the vector β, plus the parameter q that determines the scale of the pricing errors). The weights are again assigned as in equation 27.

We perform 250,000 draws for each block (the first 50,000 are used as a burn-in sample and discarded), for a total of 4,750,000 iterations of the Markov chain. Raftery and Lewis' (1995) run length control diagnostic[12] indicates that the sample size is more than 10 times the minimum required size.

At each draw of the parameter vector from the posterior distribution, we compute its associated state price vector. We use the posterior distribution of π thus obtained to also compute a posterior distribution over risk-neutral probabilities ρ (see equation 13) and over the quantiles of the distribution of returns

$$r^{\chi} = \left\{ \frac{s_j + \nu_j^{\chi}}{S_0} - 1 : j = \max \left\{ k \leq n : \sum_{i=1}^{k} \rho_i \leq \chi \right\} \right\} \tag{52}$$

Here, r^{χ} denotes the χ-quantile of the risk-neutral probability distribution of the returns on the underlying, S_0 is the current value of the underlying, calculated as

$$S_0 = \sum_{i=1}^{n} \pi_i s_i \tag{53}$$

and the linear interpolation term

$$\nu_j^{\chi} = \frac{\chi - \sum_{i=1}^{j} \rho_i}{\rho_{j+1}} \left(s_{j+1} - s_j \right) \tag{54}$$

[12] The parameters of the diagnostic are set in such a way that the minimum required size allows to estimate the 2.5% quantile of the posterior distribution of each entry of the parameter vector with an error <1% with probability 95%.

is added in order to avoid discontinuities in the quantile function.

For each cross-section of option prices, we compute posterior distributions twice, once with the prior of uni-modality (eq. 48; henceforth UNI) and once without it (eq. 44; henceforth MULTI).

The average (across states and time periods) standard deviation of the posterior distribution of the state prices is 1.7% (expressed as a percentage of the average posterior mean) for UNI and 2.3% for MULTI. A qualitative judgment on the level of precision is of course subjective, but bear in mind that these values of the standard deviations imply that the posterior medians of the state prices are often almost indistinguishable from the first and ninth deciles when they are plotted together (see Figure 6, for the first day in our sample). As a consequence, we believe it is fair to say, as far as our dataset is concerned, that there seems to be low posterior uncertainty about the true value of the state prices. In other words, estimates are quite accurate. Moreover, the posterior uncertainty decreases by imposing the prior of uni-modality.

To assess the posterior dispersion of the quantiles of the distribution of returns, we consider, for each period and each draw, all the percentiles, that is, $\chi = 0.01, 0.02, \ldots, 1$. The average (across values of χ and time periods) posterior standard deviation of the quantiles is 0.09% and 0.11% for UNI and MULTI respectively. Also in this case the plots of the median and of the first and ninth deciles of the posterior are almost indistinguishable (see Figure 6).

As anticipated in the introduction, some authors (e.g., Lee 2014) have conjectured that upper and lower quantiles might be difficult to estimate precisely with nonparametric methods. To explore this hypothesis, we analyze the posterior distribution of these quantiles in more detail. The average (across time periods) standard deviation of the posterior distribution of the 0.01, 0.05 and 0.10 quantiles is 0.50%, 0.28% and 0.19% respectively (for UNI) and 0.66%, 0.29% and 0.22% respectively (for MULTI). Thus, the posterior uncertainty about the lower quantiles (i.e., in the left tail of the distribution) is above the average (across all the quantiles). However, note that these standard deviations are about two orders of magnitude smaller than the average posterior means of the quantiles (45.1%, 25.9% and 17.6% for the 0.01, 0.05 and 0.10 quantiles respectively, in the UNI case). We find similar results for the right tail of the distribution. We conclude that, while the posterior uncertainty about the tail quantiles is indeed higher than for the other quantiles, it is still quite low.

We find that the prior of multi-modality often has a noticeable impact on the posterior distribution of the state prices. Without imposing it, the draws from the posterior of the risk-neutral probability distribution ρ frequently display several local modes, both in the tails and in the center of the distribution.

The average absolute difference (across periods and states) between the posterior means of the state prices obtained from the UNI prior and those obtained from the MULTI prior is 5.70% (expressed as a percentage of the average posterior mean under UNI). However, we find that most of the dispersion in the posterior of π and ρ is local and is averaged out when the risk-neutral probabilities of the single states are cumulated to produce a distribution function and to compute risk-neutral quantiles (eq. 52): the average absolute difference (across values of χ and time periods) between the posterior means of the quantiles obtained from the UNI prior and those obtained from the MULTI prior is 0.3%.

By imposing the more restrictive prior of uni-modality, the mean absolute pricing error (computed in correspondence of the maximum a posteriori estimate of π, averaged across periods and options, and expressed as a percentage of the average option price) increases from 0.041% (in the MULTI case) to 0.046% (in the UNI case).

To further compare UNI and MULTI, we compute their posterior odds ratio[13]. For each period, we use one-quarter of the observations[14] on option prices as a training sample, so as to avoid indeterminacy[15]. In other words, we use as priors the posteriors formed using

[13] The prior odds ratio is assumed to be equal to one, so that the posterior odds ratio and the Bayes factor coincide.

[14] There are 76 calls and 76 puts (see Section 5.1), indexed by $j = 1, \ldots, 76$. As a training sample, we use those having indices that are multiples of 4, that is, $4, 8, \ldots, 76$.

[15] We need to use a training sample to avoid indeterminacy (Jeffreys 1961) in the computation of the posterior odds ratios because we use improper priors. For more details on training samples and their use in the construction of non-subjective Bayes factors see, for example, Ghosh, Delampady and Samanta (2006).

one-quarter of the observations. The posterior odds ratio is defined as the ratio between the marginal density under UNI of the option prices not used in the training sample and their marginal density under MULTI. Denote by θ the vector of parameters, by Y^{O_1} the training sample, by Y^{O_2} all the observations not included in the training sample, and use the letter f to denote probability densities. Then, the posterior odds ratio R is

$$R = \frac{\int f\left(Y^{O_2}\left|\theta, UNI\right.\right) f\left(\theta\left|Y^{O_1}, UNI\right.\right) d\theta}{\int f\left(Y^{O_2}\left|\theta, MULTI\right.\right) f\left(\theta\left|Y^{O_1}, MULTI\right.\right) d\theta} \tag{55}$$

In practice, the integrals are approximated by summing over the MCMC samples of θ extracted from $f\left(\theta\left|Y^{O_1}, UNI\right.\right)$ and $f\left(\theta\left|Y^{O_1}, MULTI\right.\right)$.

We find that, for all periods, the posterior odds ratio is greater than $1,000$. According to Jeffreys' (1961) scale, this constitutes strong evidence in favor of the hypothesis of uni-modality.

To sum up, imposing the prior of uni-modality can significantly change the posterior distribution of the state prices, but it has a minor impact on the pricing errors and on the estimated quantiles of the risk-neutral distribution of returns. Furthermore, the posterior odds ratios computed after updating the priors with a training sample indicate that there is strong evidence in favor of uni-modality. We interpret our findings as empirical support for the hypothesis, advanced, for example, by Malz (2014) and Fengler and Hin (2014) that multi-modality could be an artifact of the techniques used to estimate risk-neutral distributions (due to over-fitting) and that imposing uni-modality might be a sensible choice.

6 Conclusions

We have proposed a semi-nonparametric model to estimate state price vectors from option prices. The state prices are assumed to be interpolated by a spline function, which we prove to be equivalent to a set of linear restrictions on the state prices. Thanks to this equivalence, it is straightforward to derive computationally inexpensive estimators. Least absolute deviations estimators of the state prices are the solution of a simple linear programming problem, and Bayesian estimation boils down to updating the prior on the coefficients of a linear regression equation. Our model is able to tackle two issues that have been so far left largely unaddressed: i) how to compute exact confidence bands around the estimated state prices; ii) how to impose uni-modality restrictions on the risk-neutral probability distribution derived from the state prices. We have used our model to analyze a panel of options on the S&P 500 stock index. We have found that the state prices are estimated quite precisely. The confidence bands around the median of the posterior distribution of the state prices (and of the risk-neutral probabilities) are tight. This is true also of the quantiles of the risk-neutral probability distribution of the returns. Contrary to the conjecture (e.g., Lee 2014) that tail quantiles might be difficult to estimate precisely with nonparametric methods, we have found that also the confidence bands around the estimated tail quantiles are quite tight. As far as the issue of multi-modality is concerned, the risk-neutral probability distributions estimated from our data are often multi-modal (as it is frequently found in the literature), and sometimes display several modes both in the center and in the tails of the distribution. Imposing the prior of uni-modality can significantly change the shape of the estimated risk-neutral distributions, but it has limited effects on the pricing errors and on the estimates of the quantiles of the distribution of returns. Furthermore, models with uni-modality constraints have high posterior odds. We conclude that the multi-modality often found in option-implied distributions could be an artifact due to over-fitting, and that imposing uni-modality might be a sensible choice.

Appendix

The MCMC algorithm

This appendix explains how we generate a sample from the posterior distribution of the parameters by using a random-walk Metropolis Hastings algorithm with block structure.

Denote the vector of parameters by θ:

$$\theta = \begin{bmatrix} \beta \\ q \end{bmatrix} \tag{56}$$

and its entries by $\theta_1, \ldots, \theta_{N_T+5}$ (each entry constitutes a block).

The vector of parameters is randomly drawn $J_B + J_K$ times and the first J_B draws are discarded (they are a so-called burn-in sample, also used for tuning the transition density of the chain). The last J_K draws are instead kept and constitute a sample of serially dependent draws (following a Markov chain) from the posterior distribution of θ. The value of θ at the j-th iteration of the Markov Chain is denoted by θ^j and its i-th entry by θ_i^j. Furthermore,

$$\theta^j = \begin{bmatrix} \beta^j \\ q^j \end{bmatrix} \tag{57}$$

The posterior density of a generic draw θ^j, denoted by $f\left(\theta^j \left| Y^O \right.\right)$, is known up to a constant of proportionality that does not depend on θ^j:

$$f\left(\theta^j \left| Y^O \right.\right) \propto f\left(Y^O \left| \theta^j \right.\right) f\left(\theta^j\right) \tag{58}$$

where

$$f\left(Y^O\left|\theta^j\right.\right) = (2\pi)^{-(N_C+N_P)/2}\left|\det\left(q^jW\right)\right|^{-1/2} \cdot \exp\left(-\frac{1}{2q^j}\left(Y^O-\xi\beta^j\right)^\top W^{-1}\left(Y^O-\xi\beta^j\right)\right) \tag{59}$$

and

$$f\left(\theta^j\right)=\begin{cases}1 & \text{if } \beta^j\in A_\beta \text{ and } q^j>0\\ 0 & \text{otherwise}\end{cases} \tag{60}$$

Define a $(N_T+5)\times 1$ vector κ of standard deviations of the random-walk increments that will be adaptively adjusted to target an acceptance rate between 30 and 40 per cent; the starting value for κ is

$$\kappa_i = 0.0001 \quad i=1,\ldots,N_T+5 \tag{61}$$

The chain starts from

$$\theta^0=\begin{bmatrix}\widehat{\beta}_{LAD}\\ \widehat{q}_{LAD}\end{bmatrix} \tag{62}$$

where $\widehat{\beta}_{LAD}$ and $\widehat{q}_{LAD}$ are the estimates of β and q obtained from the least absolute deviations estimation. The j-th iteration is made up of the following steps:

1. set $l=j-(N_T+5)\left[j/(N_T+5)\right]$ where $\left[j/(N_T+5)\right]$ denotes the integer part of $j/(N_T+5)$.
2. draw a random number z_j from a standard normal distribution;
3. build a new $(N_T+5)\times 1$ vector $\overline{\theta}$ such that $\overline{\theta}_i=\theta_i^{j-1}$ for $i\neq l$ and $\overline{\theta}_i=\theta_i^{j-1}+\kappa_i z_j$ for $i=l$;
4. compute the acceptance probability a_j as follows:

$$a_j=\min\left(1,\frac{f\left(Y^O\left|\overline{\theta}\right.\right)f\left(\overline{\theta}\right)}{f\left(Y^O\left|\theta^{j-1}\right.\right)f\left(\theta^{j-1}\right)}\right) \tag{63}$$

5. draw a random number u_j from the uniform distribution on $[0, 1]$;

6. if $u_j \leq a_j$ then set $\theta^j = \overline{\theta}$; otherwise, set $\theta^j = \theta^{j-1}$;

7. if $j \leq J_B$, adjust k_l [16];

8. if $j = J_B + J_K$ end the algorithm, otherwise go back to step 1.

[16] The adjustment is done as follows: at each iteration, if an exponentially weighted moving average (with forgetting factor equal to 0.99) of past acceptance indicators (1 in case of acceptance and 0 otherwise) is below 30 per cent for block θ_i, we decrease κ_i by a factor of 0.99; if the same moving average is above 40 per cent, we increase κ_i by a factor of 1.01. This choice of parameters for adjusting κ, although admittedly arbitrary, maintained the acceptance rate broadly on target over a number of repetitions of the algorithm.

References

[1] Aït-Sahalia, Y., and J. Duarte (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, 116, 9-47.

[2] Bagasheva, B. S., J. S. J. Hsu, F. J. Fabozzi and S. T. Rachev (2008) *Bayesian methods in finance*, John Wiley & Sons, Inc., Hoboken, New Jersey.

[3] Banz, R., and M. Miller (1978). Prices for state-contingent claims: some estimates and applications. *Journal of Business*, 51, 653–672.

[4] Bahra, B. (1997). Implied risk-neutral probability density functions from option prices: theory and application. Working paper, Bank of England.

[5] Birke, M., and K. F. Pilz (2009). Nonparametric option pricing with no-arbitrage constraints. *Journal of Financial Econometrics*, 7, 53-76.

[6] Bliss, R. R., and N. Panigirtzoglou (2002). Testing the stability of implied probability density functions. *Journal of Banking & Finance*, 26, 381-422.

[7] Bondarenko, O. (2003). Estimation of risk-neutral densities using positive convolution approximation. *Journal of Econometrics*, 116, 85-112.

[8] Breeden, D. T., and R. H. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51, 621-651.

[9] Bu, R., and K. Hadry (2007) Estimating option implied risk-neutral densities using spline and hypergeometric functions. *Econometrics Journal*, 10, 216-244.

[10] Campa, J. M., P. H. K., Chang, and R.L. Reider (1998) Implied exchange rate distributions: evidence from OTC option markets. *Journal of International Money and Finance*, 17, 117-160.

[11] Chabi-Yo, F., R. Garcia, and E. Renault, E. (2008). State dependence can explain the risk aversion puzzle. *Review of Financial Studies*, 21, 973-1011.

[12] Cont, R. (2006). Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical finance*, 16, 519-547.

[13] Corrado, C. J., and T. Su (1997). Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *European Journal of Finance*, 3, 73-85.

[14] Dana, R. A. (2004). Ambiguity, uncertainty aversion and equilibrium welfare. *Economic Theory*, 23, 569-587.

[15] De Boor, C. (1978). A practical guide to splines.

[16] Easley, D., and M. O'Hara (2010). Liquidity and valuation in an uncertain world. *Journal of Financial Economics*, 97, 1-11.

[17] Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 75, 643-669.

[18] Fan, J., and L. Mancini (2009). Option pricing with model-guided nonparametric methods. *Journal of the American Statistical Association*, 104:488, 1351-1372.

[19] Fengler, M. R., and L-Y Hin (2014) Semi-nonparametric estimation of the call-option price surface under strike and time-to-expiry no-arbitrage constraints, *Journal of Econometrics*, forthcoming.

[20] Frachot, A., J. P. Laurent, and O. Pichot (1999), Multimodal implied risk-neutral densities, mimeo, CREST.

[21] Gallant, A. R. (1987). Identification and consistency in semi-nonparametric regression. *Advances in Econometrics*, Fifth World Congress, vol. 1, Cambridge University Press, Cambridge.

[22] Ghosh, J. K., M. Delampady and T. Samanta (2006) *An introduction to Bayesian analysis: theory and methods*, Springer Science, New York.

[23] Jackwerth, J. C. (2004). Option-Implied Risk-Neutral Distributions and Risk Aversion. Charlottesville, VA: The Research Foundation of AMIR.

[24] Jackwerth, J. C., and M. Rubinstein (1996). Recovering probability distributions from option prices. *Journal of Finance*, 51, 1611-1631.

[25] Jarrow, R., and A. Rudd (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10, 347-369.

[26] Jeffreys, H. (1961) *Theory of Probability*, Oxford University Press, New York.

[27] Koenker, R., and K. Hallock (2001). Quantile regression: An introduction. *Journal of Economic Perspectives*, 15, 43-56.

[28] Lee, S. H. (2013). Estimation of risk-neutral measures using quartic B-spline cumulative distribution functions with power tails. *Quantitative Finance*, 14, 1857-1879.

[29] Madan, D., and F. Milne (1994). Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance*, 4, 223-245.

[30] Malz, A. M. (1997). Estimating the probability distribution of the future exchange rate from options prices. *Journal of Derivatives*, 5, 18–36.

[31] Malz, A. M. (2014). A Simple and Reliable Way to Compute Option-Based Risk-Neutral Distributions. *FRB of New York Staff Report*, 677.

[32] Melick, W.R., and C. P. Thomas, 1997. Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis. *Journal of Financial and Quantitative Analysis*, 32, 91-115.

[33] Raftery, A. E., and S. M. Lewis (1995). The Number of Iterations, Convergence Diagnostics and Generic Metropolis Algorithms, in W. R. Gilks, D. J. Spiegelhalter and S. Richardson (eds), Practical Markov Chain Monte Carlo. London, UK: Chapman and Hall, 115-30.

[34] Rosenberg, J. (1998). Pricing multivariate contingent claims using estimated risk-neutral density functions. *Journal of International Money and Finance*, 17, 229-247.

[35] Ross, S. A. (1976). Options and efficiency. *The Quarterly Journal of Economics*, 90, 75-89.

[36] Routledge, B. R., and S. E. Zin (2009). Model uncertainty and liquidity. *Review of Economic Dynamics*, 12, 543-566.

[37] Shimko, D. C. (1993). Bounds of probability. *Risk*, 6, 33-37.

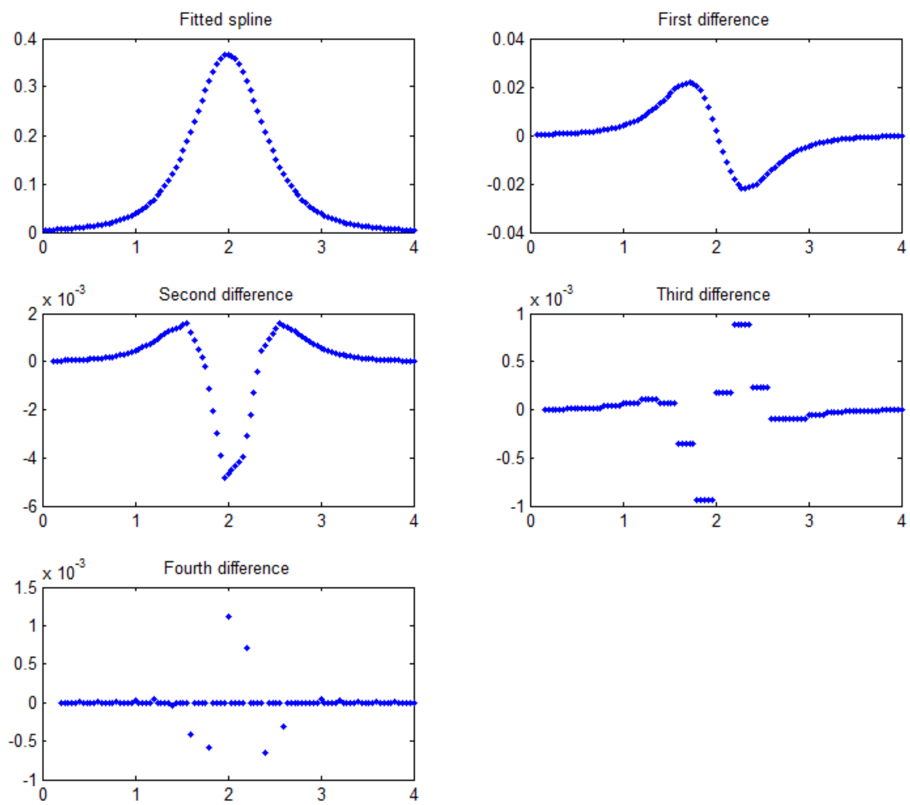
[38] Söderlind, P., and L. Svensson (1997). New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics*, 40, 383-429.

[39] Taylor, S., C. F. Tzeng, and M. Widdicks (2012). Bankruptcy probabilities inferred from option prices. Technical report, Lancaster University Management School, Lancaster, UK.

[40] Ziegler, A. (2002). State-price densities under heterogeneous beliefs, the smile effect, and implied risk aversion. *European Economic Review*, 46, 1539-1557.

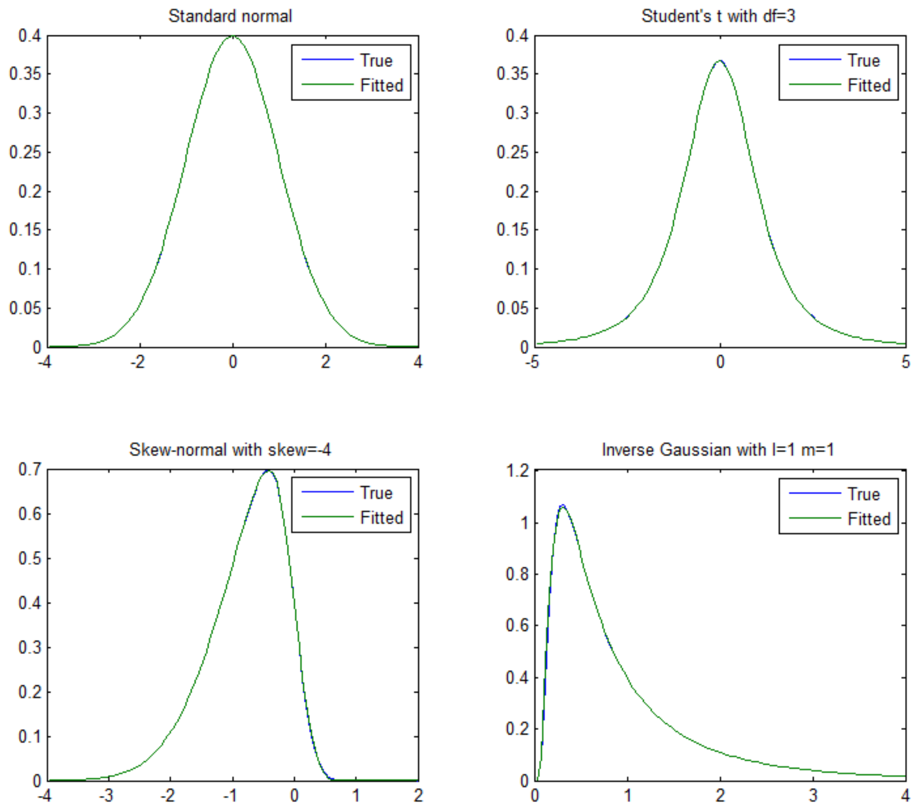
Figures

Figure 1 - The spline and its derivatives[17]



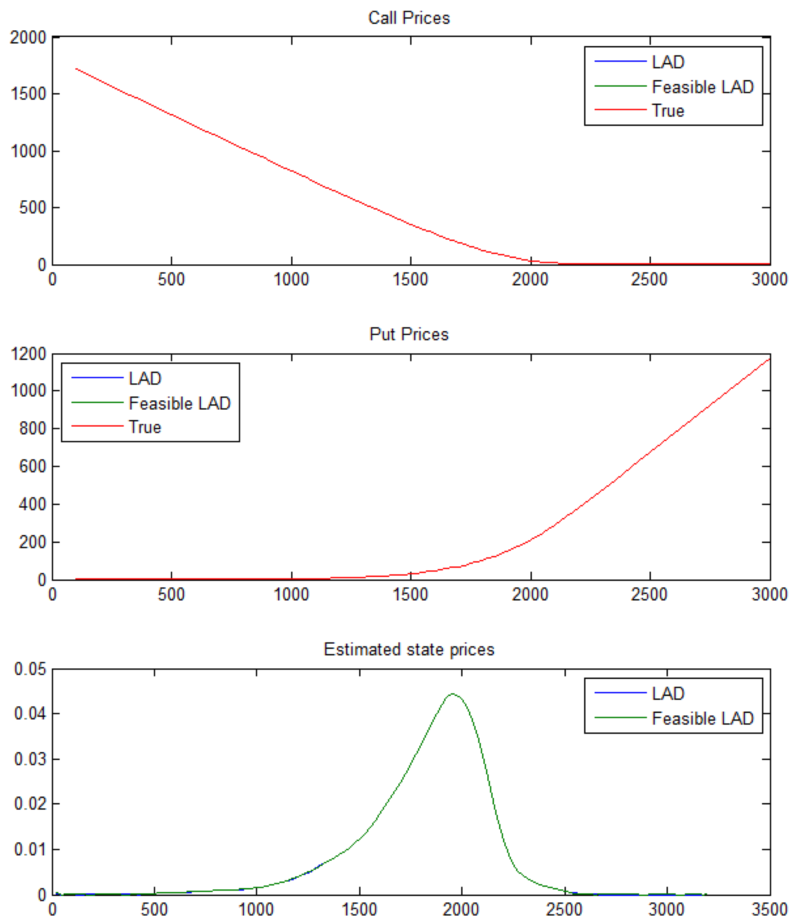
[17] This figure displays a spline fitted to the density of a standard normal distribution, sampled at uniformly spaced intervals, and its finite differences.

Figure 2 - Densities fitted with splines[18]



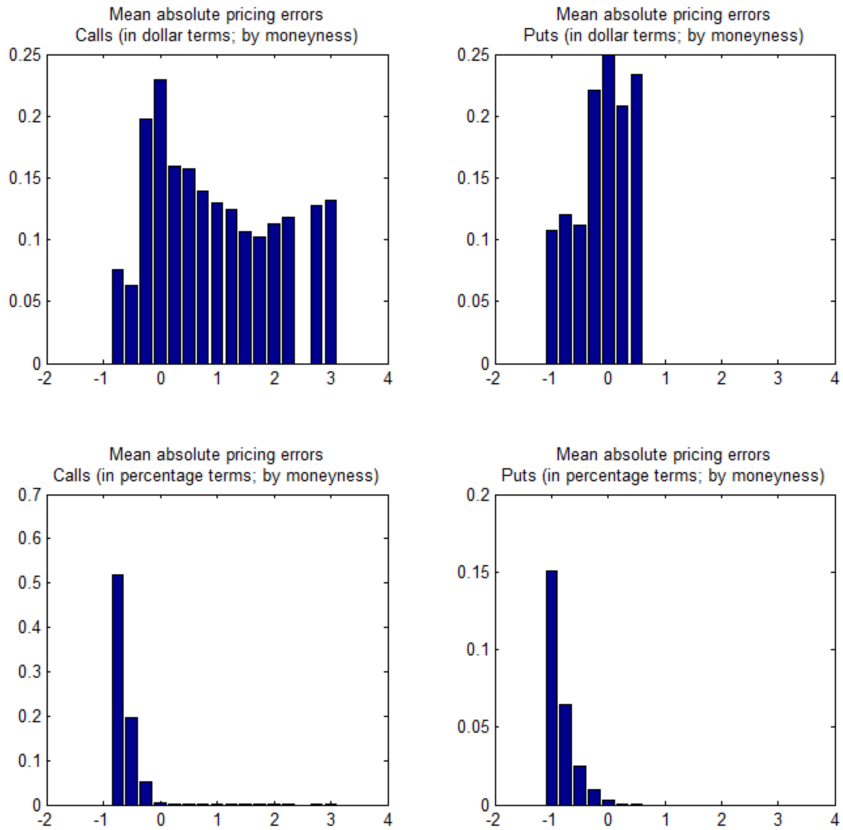
[18] This figure displays four probability density functions, sampled at uniformly spaced intervals, and their spline interpolations.

Figure 3 - Estimated option prices and state prices[19]



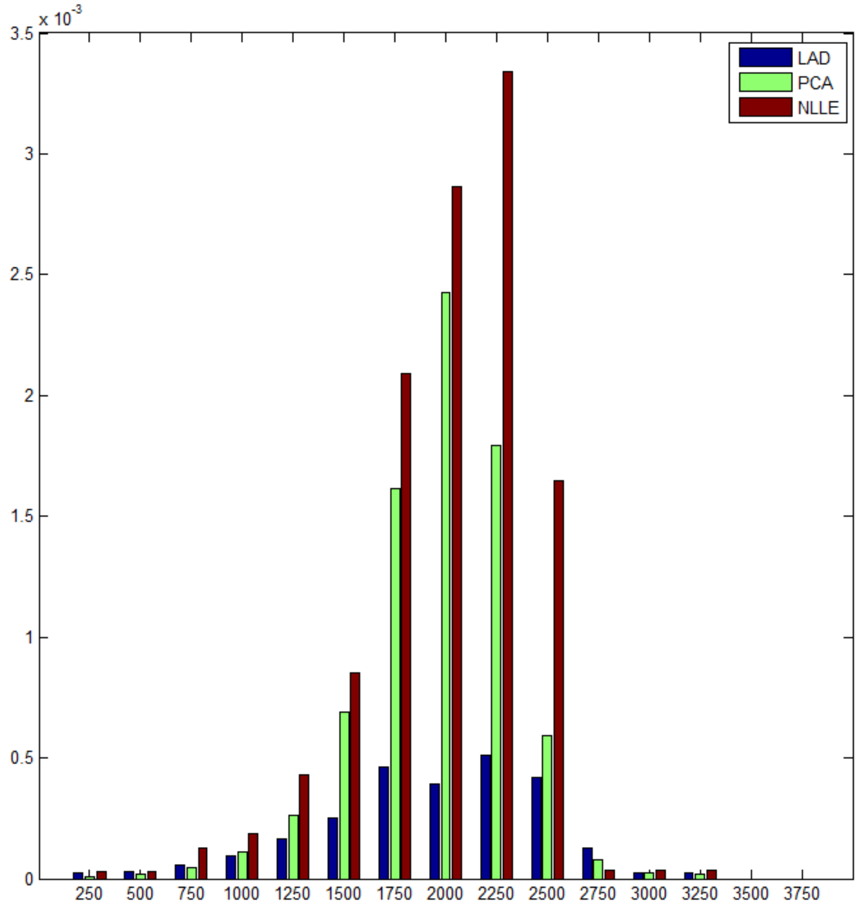
[19] The upper and middle panel of this figure display the closing prices of the put and call options on the S&P 500 stock index as of December 31st 2013, together with the fitted values obtained with the LAD and feasible LAD procedures. The lower panel displays the corresponding estimates of the state prices. In all the three panels, the prices of the underlying stock index are on the x-axis.

Figure 4 - Mean absolute pricing errors (by moneyness)[20]



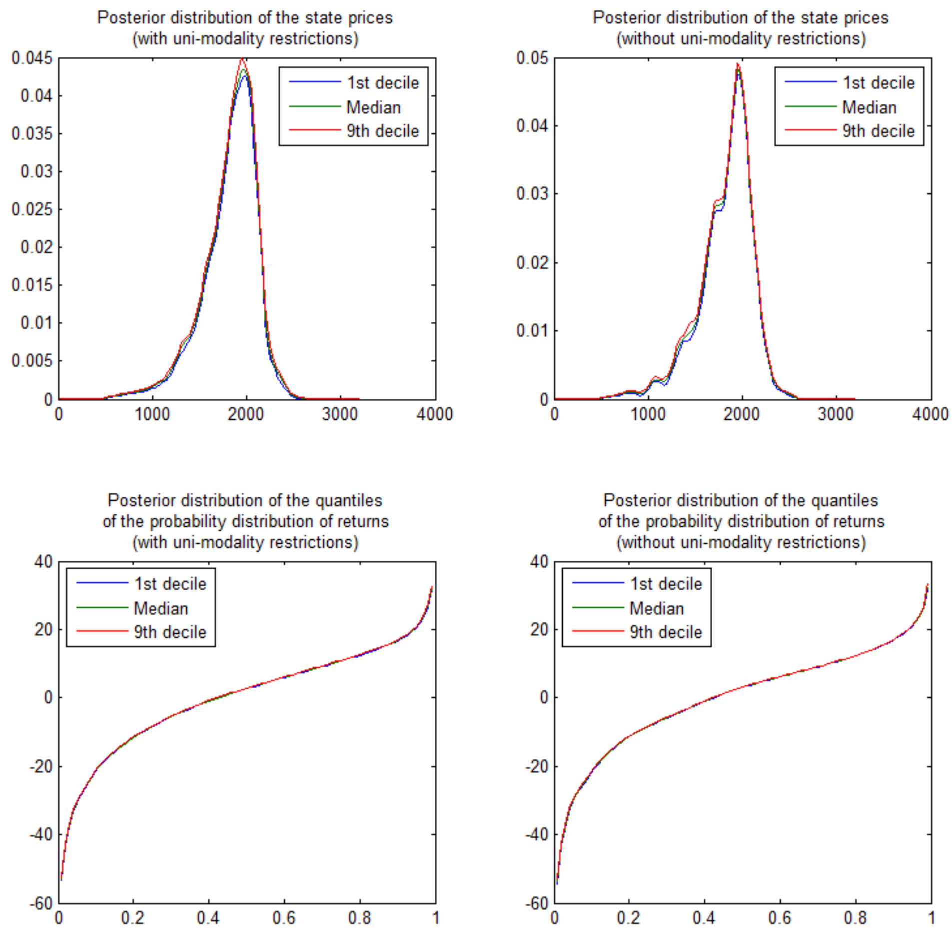
[20] These figures plot the mean absolute pricing errors (on call and put options) estimated with the LAD method. The estimated errors are pooled across dates and their mean absolute value is plotted by the moneyness of their corresponding option (on the x-axis).

Figure 5 - Mean absolute errors of the estimated state prices[21]



[21] This figure plots the mean absolute errors committed in the estimation of the state prices in a Monte Carlo experiment. The errors are grouped by underlying price intervals (on the x-axis). LAD is the Least Absolute Deviations method proposed in this paper, PCA is the Positive Convolution Approximation method, and NLLE is the Nonparametric Locally Linear Estimator.

Figure 6 - Results from the Bayesian estimation[22]



[22]These figures plot the 1st, 5th and 9th deciles of the posterior distributions of the state prices and of the quantiles of the distribution of returns, obtained with the Bayesian model proposed in this paper. These distributions are obtained from the closing prices of the put and call options on the S&P 500 stock index as of December 31st 2013. In the upper panels, the prices of the underlying stock index are on the x-axis. In the lower panels, cumulated probabilities are on the x-axis and quantiles of the distribution of returns are on the y-axis.