Alternative methods to evaluate the arbitrage opportunities

NOUREDDINE KOUAISSAH [1,2], SERGIO ORTOBELLI LOZZA [1,2]
[1]Department SAEQM
[2]Department of Finance
[1]University of Bergamo
[2]VSB Technical University of Ostrava
[1]Via dei Caniana, 2 – 24127 Bergamo, ITALY
[2]Sokolská, 33 - 70121 Ostrava, CZECH REPUBLIC
noureddine.kouaissah@unibg.it , sergio.ortobelli@unibg.it

Abstract: - In this paper, we present alternative methods to evaluate the presence of the arbitrage opportunities in the market. In particular, we investigate empirically the well-known put-call parity no-arbitrage relation and the state price density. First, we measure the violation of the put call parity as the difference in implied volatilities between call and put options that have the same strike price, the same maturity and the same underlying asset. Then, we examine the nonnegativity of the state price density since its negative values immediately correspond to the possibility of free-lunch in the market. We evaluate the effectiveness of the proposed approaches by an empirical analysis on S&P 500 index options data. Moreover, we propose different approaches to estimate the state price density under the classical hypothesis of the Black and Scholes model. In this context, we use two different methodologies to evaluate the conditional expectation and its relationship with the state price density. Firstly, we examine the real mean return function using local polynomial smoothing technique. Then, we evaluate the conditional expectation using the real probability density. According to the hypothesis of the Black and Scholes model, we are able to derive a closed formula for approximating the conditional expectation with the risk neutral probability. Finally, we propose a comparison among two estimators of the state price density.

Key-Words: - arbitrage opportunities, put-call parity, state price density, conditional expectation estimators

1 Introduction

The option-pricing theory has had a central role in modern finance ever since the pioneering work of Black and Scholes [3] (hereinafter BS). The main idea behind BS option pricing model is that the price of an option is defined as the least amount of initial capital that permits the construction of a trading strategy whose terminal value equals the payout of the option. BS model has a great importance for improving research on the option pricing techniques. Unfortunately, widespread empirical analyses point out that a set of assumptions under which BS model built, particularly normally distributed returns and constant volatility, result in poor pricing and hedging performance. However, different generalizations of the BS model have been proposed in literature - see e.g. Merton [16], Heston [17] and Bates [7] for more details. Generally, most models that have been proposed so far mainly relax some assumptions of BS model and then trying to be justified via general fundamental theorem of asset pricing-FTAP, Harrison and Kreps [12]. This theorem provides many challenges in asset pricing theory. In particular, it asserts that the absence of arbitrage in a frictionless financial markets if and only if there exist an equivalent martingale measure under which the price process is a martingale.

One fundamental entity in asset pricing theory is the so called State Price Density (hereinafter SPD). Among no-arbitrage models, the SPD is frequently called risk-neutral density, which is the density of the equivalent martingale measure with respect to the Lebesgue measure. The existence of the equivalent martingale measure follows from the absence of arbitrage opportunities, while its uniqueness demands complete markets. Breeden and Litzenberger [4] proposed an excellent framework to fully recover the SPD in an easy way. In this method, the SPD is simply equal to the second derivative of a European call option with respect to the strike price, see among others Brunner and

Hafner [5] for other estimation technique. Furthermore, it is well known that option prices carry important information about market conditions and about the risk preferences of market participants. In this context, the SPD function derived from observed standard option prices have gained considerable attention in last decades. Indeed, an estimate of the SPD implicit in option prices can be useful in different contexts, see among others Ait-Sahalia and Lo [1]. The most significant application of the SPD is that it allows us computing the no-arbitrage price of complex or illiquid option simply by integration techniques.

The first fundamental contribution of this paper is to evaluate the presence of arbitrage opportunities in the market. To do so, we focus on the violation of the put-call parity no-arbitrage relation and the nonnegativity of the SPD. Firstly, we measure the violation of put-call parity as the difference in implied volatility between call and put options that have the same strike price, the same expiration date and the same underlying asset. Secondly, we compare this result with that obtained from the violation of the nonnegativity of the SPD. This is important, because negative values of the SPD immediately correspond to the possibility of free-lunch in the market.

The second crucial contribution of the paper is to propose different approaches to estimate the SPD. We deviate from previous studies in that we estimate SPD directly from the underlying asset under the hypothesis of the BS model. To this end we follow two distinguished approaches to recover the SPD, the first one based on nonparametric estimation techniques "kernel" which are natural candidates (see among others [1], [2]), then a new method based on conditional expectation estimator proposed by [18]. Firstly, we examine the so called real mean return function using local polynomial smoothing technique. Then, we estimate the conditional expectation under real probability density. According to the hypothesis of BS model, we are able to derive a closed formula for approximating the conditional expectation under risk neutral probability. The main goal of this contribution is to examine and compare the conditional expectation method and the nonparametric technique. These methods allow us extrapolating arbitrage opportunities and relevant information from different markets (futures and options) consistently with the analysis of the underlying.

The rest of this paper is organized as follows. Section 2 presents some methods to evaluate the arbitrage opportunities. Section 3 illustrates the first empirical analysis. Section 4 proposes alternative methods to estimate the SPD. Section 5 includes the second empirical analysis. Concluding remarks are contained in Section 6.

2 Methods to evaluate the arbitrage opportunities

2.1 Black and Scholes methodology

Fisher Black and Myron Scholes [3] achieved a major breakthrough in European option pricing. In this model we assume that the price process follows a standard geometric Brownian motion defined on filtered probability space $(\Omega, \Im, \mathrm{P}, \{\Im_t\}_{t\geq 0})$, where $\{\Im_t\}_{t\geq 0}$ is the natural filtration of the process completed by the null sets. Under these assumptions we know that $\mathrm{E}(\mathrm{S}_T \mid \Im_t) = \mathrm{E}(\mathrm{S}_T \mid \mathrm{S}_t)$ as consequence of Markovian property. The model of stock price behavior used is defined as:

$$dS = \mu S dt + \sigma S dB, \qquad (1)$$

where, μ is the expected rate of return, σ is the volatility of stock return and B denotes a standard Brownian motion. Under this hypothesis we know that the log price is normally distributed:

$$\ln S_T \sim \phi\left\lfloor \ln S_0 + \left(\mu - 0.5\sigma^2\right)T, \sigma^2 T \right\rfloor, \quad (2)$$

where, S_T is the stock price at future time T, S_0 is the stock price at time 0 and ϕ denotes a normal distribution. Please note that μ in equation (1) represents the expected rate of return in real world, while in BS model (risk neutral world) it becomes risk-free rate r.[1]

2.2 Put-call parity

We recall an important relationship between the prices of European put and call options that have the same strike price and the same time to maturity. This relationship is known as put-call parity, see Stoll [20]. In particular, it shows that the value of a European call option with a certain strike price and expiration date can be deduced from the value of a European put option with the same strike price and expiration date, and vice versa. Formally, in perfect

[1] For more details about BS assumptions we refer to Hull [2015]

markets, the following equality must hold for European options on non-dividend-paying stocks:

$$C - P = S_0 - Ke^{-rt}, \tag{3}$$

where, S_0 is the current stock price, C and P are the call and put prices, respectively, that have the same strike price K , the same expiration date and the same underlying asset.

To illustrate the arbitrage opportunities when equation (3) does not hold, we measure the violation of put-call parity as the difference in implied volatility between call and put options that have the same strike price, the same expiration date and the same underlying asset. In this context, it is well known that the BS model satisfies put-call parity for any assumed value of the volatility parameter σ. Hence,

$$C^{BS}(\sigma) - P^{BS}(\sigma) = S_0 - Ke^{-rt} \quad \forall \sigma > 0, \tag{4}$$

where, $C^{BS}(\sigma)$ and $P^{BS}(\sigma)$ denotes BS call and put prices, respectively, as a function of the volatility parameter σ. At this point, from equation (3) and (4) we can deduce that:

$$C^{BS}(\sigma) - P^{BS}(\sigma) = \mathrm{C} - \mathrm{P} \quad \forall \sigma > 0, \tag{5}$$

By definition, the implied volatility (IV) of a call option (IV^{call}) is that value of the volatility of the underlying asset, which matches the BS price with the price actually observed on the market. In formal way:

$$C^{BS}(IV^{call}) = C, \tag{6}$$

Now, it is straightforward form equation (5) that:

$$P^{BS}(IV^{call}) = P, \tag{7}$$

this in turn implies that:

$$IV^{call} = IV^{put}. \tag{8}$$

Put-call parity holds only for European options. Thus, for this type of options, put-call parity is equivalent to the statement that the BS implied volatilities of pairs of call and put options must be equal. Therefore, any violation of put-call parity may contain useful information about the presence of tradable arbitrage opportunities. No attempt will be made to formulate the case of American option, which beyond the scope of this study. However, it possible to derive some results for American options price, where put-call parity takes the form of an inequality.

In this paper, we will carry the analysis on the European options style. Since put-call parity is one of the best known no-arbitrage relations, we use the difference in implied volatility between pairs of call and put options in the spirit of equation (8) in order to detect the presence of arbitrage opportunities in the market. Intuitively, lower call implied volatilities relative to put implied volatilities means that calls are less expensive than puts, and lower put implied volatilities with respect to call implied volatilities suggest the opposite.

We compute the difference in implied volatilities between call and put options that have the same strike price, the same maturity and are written on the same underlying asset. Hence, we refer to such difference as volatility spread (VS) which may represent a valid indicator of the presence of arbitrage opportunities in the market, especially close to at-the-money options. Formally, given call and put options with the same strike price and expiration date, we compute the VS as**:**

$$VS = \max | IV^{call} - IV^{put} | \tag{9}$$

Of course, higher volatility spread is a significant indicator of arbitrage opportunities since put-call parity is a fundamental relation of no-arbitrage. A simple example illustrates intuitively this result.

Example: Consider a put option on S&P 500 index with strike price $K = 2200$ and has 6 months to maturity. The current underlying asset price is $S_0 = 2100$ and the 6-month risk free rate of return is $r = 0.08\%$. Let us assume that the price of this put option is $P = 160$ and the price of the call option on the S&P 500 index with the same strike price and the same maturity is $C = 120$. It is very simple to verify that the put-call parity does not hold and that the volatility of call option is greater than the volatility of the put option. Indeed,

$$P + S_0 = 2240 < C + Ke^{-rT} = 2299.1,$$

$$IV^{call} = 0.2723, \ IV^{put} = 0.1699 \text{ and } VS = 0.1024$$

Arbitrage position: Buy the put option at $P = 160$ and the stock at $S_0 = 2100$, then sell the call option at $C = 120$. To finance this position, borrow:

$$D = P + S_0 - C = 160 + 2100 - 120 = 2140 \text{ at } r = 0.08\%.$$

Payoff to this arbitrage position:

- If $S_T < 2100$, the trader exercises the put option and the payoff is:

$$(\mathrm{K}-S_T)+S_T-De^{rT}=59.14$$

- If $S_T>2100$, the short call option exercised and the payoff is:

$$S_T-(S_T-K)-De^{rT}=59.14$$

In both cases, the trader ends up with a payoff of 59.14 and selling the stock at $K=2200$.

This example illustrates the situation when the call implied volatility is greater than the put implied volatility, such that the option has the same strike price, the same maturity and is written on the same underlying asset. On the opposite, lower call implied volatility relative to put implied volatility means that call option is less expensive than put option. Therefore, one may follow the simplest strategy that involves buying the call option and shorting both the put option and the stock.

The efficacy of this theoretical arbitrage mechanism in maintaining put and call price parity will be examined empirically. However, several papers argue that violations of the put-call parity can be justified via the short sale constraint, data-related issues or even the payment of dividend streams, see among others Ofek et al [9]. To overcome these issues and to have a valid confirmation of this approach we will proceed as follows. We combine the IV smoothing with SPD estimation which requires some properties in order to be consistent with no-arbitrage argument. In particular, we evaluate the nonnegativity property of the SPD since its negative values immediately correspond to the possibility of the arbitrage opportunities in the market. To this end we follow a relatively conservative approach adopted by Benko et al [2].

2.3 State price density

SPDs derived from cross-sections of observed standard option prices have gained considerable attention during last decades. Since given an estimate of SPD, one can immediately price any path independent derivative. Clearly, the well-known arbitrage free pricing formula is of vital practical importance. In this approach, the option price is given as the expected value of its future payoff with respect to the *risk-neutral measure Q* discounted back to the present time *t*. Formally, the price $\Pi_t(\mathrm{H})$ at time *t* of a derivative with expiration date T and payoff –function $\mathrm{H}(\mathrm{S}_T)$ is given by:

$$\Pi_t(\mathrm{H})=e^{-r(T-t)}E^Q\left[\mathrm{H}\mid\Im_t\right]=e^{-r(T-t)}\int_0^{\infty}\mathrm{H}(\mathrm{s})q_{S_T}(\mathrm{s})ds,$$

$$\forall t\in[0,T] \tag{10}$$

where, $q_{S_T}(\mathrm{s})$ denotes the SPD. In this context, one fundamental founding in literature is the relationship between SPD and implied volatility (IV) e.g. see among others Hafner and Brunner [5].

In this paper, in line with Benko et al [2], we apply local polynomial smoothing technique to estimate IVs, and then the SPD. To this end, we first establish the relation between SPD and IVs and then we summaries some properties that SPD demands in order to be consistent with no-arbitrage argument.

It is well known that in the BS model the SPD is assumed to be a lognormal density with mean $(r-0.5\sigma^2)\tau$ and variance $\sigma^2\tau$. In this context, the price of a European call option C_t with expiration date T and strike price K can be obtained: see Black and Scholes [3]

$$C_t(S_t,K,\tau,r,\sigma)=S_t\Phi(d_1)-Ke^{-r\tau}\Phi(d_2), \tag{11}$$

where, $d_1=\dfrac{\ln(S_t/K)+(r+0.5\sigma^2)\tau}{\sigma\sqrt{\tau}}$,

$d_2=d_1-\sigma\sqrt{\tau}$, $\tau=T-t$ is time to maturity, r is a riskless interest rate and $\Phi(\cdot)$ is the standard normal distribution function. The BS formula provides a correspondence between the price of a plain option and the underlying asset volatility. However, it is well known that implied volatilities of quoted European options are not constant and depend on the strike price and the maturity of the option.

Breeden and Litzenberger [4] derived an elegant formula for obtaining an explicit expression for the SPD from option prices. In fact, they observed that the second derivative of the call price function $C_t(K,T)$ with respect to the strike price K is proportional to the SPD. Formally:

$$q_{t,S_T}(x,\tau)=e^{r(T-t)}\left.\frac{\partial^2 C_t(K,T)}{\partial K^2}\right|_{K=x}. \tag{12}$$

The last formula is of great practical importance. Since for any fixed time T, the relation between SPD and IV can be obtained simply by a successive application of (11) and (12). After some algebra, applying chain rule for derivatives one get:

$$q_{t,S_T}(x,\tau)=$$

$$e^{r\tau}S_t\sqrt{\tau}\varphi(\mathrm{d}_1(x,\tau))\left\{\begin{array}{l}\dfrac{1}{x^2\sigma(K,\tau)\tau}+\dfrac{2d_1(x,\tau)}{x\sigma(x,\tau)\sqrt{\tau}}\left.\dfrac{\partial\sigma(K,\tau)}{\partial K}\right|_{K=x}+\\ \dfrac{d_1(x,\tau)d_2(x,\tau)}{\sigma(x,\tau)}\left(\left.\dfrac{\partial\sigma}{\partial K}\right|_{K=x}\right)^2+\left.\dfrac{\partial^2\sigma}{\partial K^2}\right|_{K=x}\end{array}\right\}, \quad (13)$$

where, $d_1(x)=\dfrac{\ln\left(S_t/x\right)+\left(r+0.5\sigma^2(x,\tau)\right)\tau}{\sigma(x,\tau)\sqrt{\tau}}$

$d_2=d_1(x,\tau)-\sigma(x,\tau)\sqrt{\tau}$ and $\varphi(\bullet)$ is the p.d.f. of a standard normal random variable, we refer the reader to Benko et al [2] and Brunner and Hafner [5] for further details.

Following Carr [6] and Brunner and Hafner [5] the SPD has to satisfy a set of properties. Formally:

- Nonnegativity property: the SPD is nonnegative, i.e.:

$$q_{t,S_T}(x)\geq 0,\quad x\in[0,\infty) \qquad (14)$$

- Integrability property: the SPD integrate to one, i.e.:

$$\int_0^\infty q_{t,S_T}(x)dx=1 \qquad (15)$$

- Martingale property: the SPD reprices all calls, i.e.:

$$\int_0^\infty \max\{x-K,0\}q_{t,S_T}(x)dx=e^{r(T-t)}C_t(K,T),$$
$$K\geq 0. \qquad (16)$$

The first two properties ensure that the SPD is indeed a probability density. Furthermore, if q_{t,S_T} satisfies the three properties, it is a well-defined SPD and the market is free of arbitrage opportunities with respect to maturity T.

3 First empirical analysis

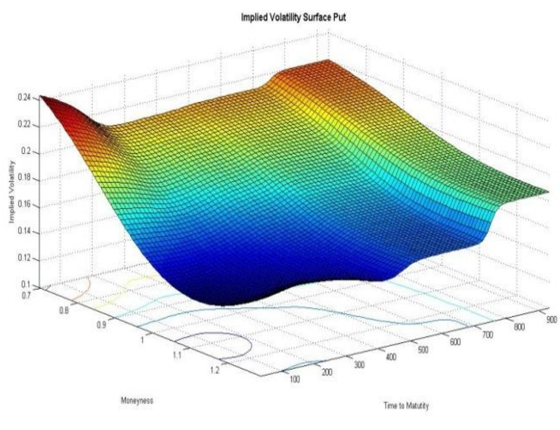
In this section, we report numerical experiments obtained using the methods introduced to detect the presence of arbitrage opportunities in the market. To evaluate the empirical importance of these techniques and the corresponding SPD estimate, we present some applications to the S&P 500 index using daily data obtained from DataStream for the sample period December 26, 2012 to May 13, 2015. Of course, S&P 500 Index options are among the most actively traded financial derivatives in the world.

In the first empirical application to S&P 500 index options we present the analysis concerning the estimation of IVs. For this purpose we use as dataset all options listed on May 13, 2015. The options are European style and the average daily volume during the sample day was 82.65 and 179.01 contracts for call and put respectively. Strike price is at 130 percent and barrier at 70 percent of the underlying spot price at 2098.48, while strike price intervals are 5 points. During sample period, the mean and standard deviation of continuously compounded daily returns of the S&P index are 1.078 percent and 11.268 percent, respectively. Throughout this period short-term interest rates exhibit a very low level. They range from 0.01 percent monthly to 0.89 percent in almost three years. The options in our sample vary significantly in price and terms, for example the time-to-maturity varies from 2 days to 934 days.

The row data present some challenges that must be addressed. Clearly, in-the-money (ITM) options are rarely traded relative to at-the -money (ATM) and out-the-money (OTM) options. For example, the average daily volume for puts that are 25 points OTM is 2553 contracts, in contracts, the volume for puts that are 25 points ITM is 2. This can be justified by the strong demand of portfolio managers for protective puts.

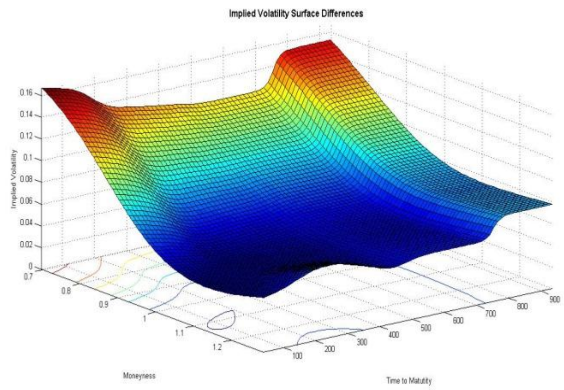
Figure 1 shows the IV surface estimated using put options for the daily data on May 13, 2015. The IV smile is very clear for small maturities and still evident as time to maturity increases.

Figure 1: Implied volatility surface of S&P 500 put options



To evaluate the presence of arbitrage opportunities, we compute the difference in implied volatilities between call and put options that have the same strike price, the same maturity and are written on the same underlying asset. In particular, we consider the differences that are greater than 80 percent of the maximum absolute value of the differences between call and put implied volatilities. In this way, we rule out some differences due to the noisy data or transaction costs. Figure 2 shows the differences in implied volatilities between call and put options.

Figure 2: Implied volatility surface differences



In Figure 2, it is clear that the differences are significant at lower moneyness which corresponds to OTM put options and ITM call options. However, since the market increases and it is well known that OTM put options and ITM call options are not reliable data to evaluate arbitrage opportunities, we focus on at ATM options. From figure 2, we observe even at ATM option there are small differences, which may represent arbitrage opportunities. In particular, the differences increase as the maturities increase.

To evaluate the size of the arbitrage opportunities, we combine the IV smoothing with SPD estimation. This is important, because the SPD requires some properties in order to be consistent with no-arbitrage argument. In particular, we examine the nonnegativity property of the SPD since its negative values immediately correspond to the possibility of the arbitrage opportunities in the market. To this end we follow a relatively conservative approach adopted by Benko et al [2]. The last approach is of great practical importance and mainly confirms the result obtained via the violation of the put call parity relation.

The second contribution of this paper is to propose different methods to estimate SPD under the classical hypothesis of BS model. In particular, we use two different methodologies to evaluate the conditional expectation. Namely, the nonparametric estimator based on kernel estimator and a new alternative technique the so called OLP estimator proposed by Ortobelli et al [18]. Differently from previous studies we estimate SPD directly from the underlying asset under the hypothesis of the BS model. To do so, firstly we examine the real mean return function using local polynomial smoothing technique. Then, we estimate the conditional expectation under real probability density. According to the hypothesis of the BS model, we are able to derive a closed formula for approximating the conditional expectation under risk neutral probability. Now, we describe in details our alternative approach towards estimating the SPD.

4 Alternative methods to estimate the SPD

For the sake of clarity, denote S^{RW} for a real world price and S^{RN} for the risk neutral price. Under the hypothesis of the BS model it is straightforward to write:

$$S_T^{RN} = S_T^{RW} e^{-(\mu - r)T}, \qquad (17)$$

Since $S_t = e^{-r(\mathrm{T}-\mathrm{t})} E(S_T^{RN} \mid \Im_t)$, we can write $S_t = e^{-r(\mathrm{T}-\mathrm{t})} E(S_T^{RW} e^{-(\mu-r)T} \mid \Im_t)$ from which we obtain:

$$E(S_T^{RN} \mid \Im_t) = e^{-\mu T + rt} E(S^{RW} \mid \Im_t), \qquad (18)$$

If we assume μ changes over time in model (1), then equation (18) becomes

$$e^{-\int_0^T (\mu(\tau)-r)d\tau} E(S_T \mid \Im_t) = E^Q(S_T \mid \Im_t), \qquad (19)$$

where, $E^Q(S_T \mid \Im_t)$ denotes expectation under risk neutral world and $E(\mathrm{S}_T \mid \Im_t)$ the conditional expected price under real world. Moreover, (19) is equivalent to:

$$e^{-\int_0^T \mu(\tau)d\tau} \int_0^\infty s q_{RW}(\mathrm{s}) ds = e^{-rT} \int_0^\infty s q_{RN}(\mathrm{s}) \mathrm{ds}, \qquad (20)$$

where, $\mathrm{q}_{RW}(\mathrm{s})$ and $q_{RN}(s)$ denotes SPDs under real and risk neutral world respectively. Please note that under the BS hypothesis S_{T-t} has the same distribution as $S_T e^{-\mu t}$.

4.1 Local polynomial regressions

The first step in this approach is to propose a direct method of estimating the real mean return function. Therefore, we use a local estimator that automatically provides an estimate of the real mean function and its derivatives. The input data are daily prices. Denoting the intrinsic value by $\tilde{\mu}_i$ and the true function by $\mu(t_i)$, $i=1,....,n$, we assume the following regression model:

$$\tilde{\mu}_i = \mu(t_i) + \varepsilon_i, \tag{21}$$

where, ε_i models the noise, n denotes the number of data considered. The local quadratic estimator $\hat{\mu}(t)$ of the regression function $\mu(t)$ in the point t is defined by the solution of the following local least squares criterion:

$$\min_{\alpha_0,\alpha_1,\alpha_2} \sum_{i=1}^{n} \left\{ \tilde{\mu}_i - \alpha_0 - \alpha_1(t_i - t) - \alpha_2(t_i - t)^2 \right\} k_h(t - t_i), \tag{22}$$

where, $k_h(t-t_i) = \frac{1}{h} k\left(\frac{t-t_i}{h}\right)$ is *kernel* function, see Fan and Gijbels [10] for more details. Comparing the last equation with the Taylor expansion of μ yields:

$$\alpha_0 = \hat{\mu}(t_i), \alpha_1 = \hat{\mu}'(t_i), 2\alpha_2 = \hat{\mu}''(t_i), \tag{23}$$

which make the estimation of the regression function and its two derivatives possible. The second step towards estimating state price density is to use two methodologies, namely OLP estimator and kernel estimator, to estimate the quantity $E(S_T \mid S_t)$.

4.2 Nonparametric conditional expectation estimators

Regression analysis is surely one of the most suitable and widely used statistical techniques. In general, it explores the dependency of the so-called dependent variable on one (or more) explanatory or independent variables. Without significant loss of generality, the mathematical notation changes in this section (the distinction of the variables will always be clear from context). Interpret Y as S_T, while X as S_t.

$$Y = E(Y \mid X = x) + \varepsilon = g(x) + \varepsilon. \tag{24}$$

It is well known that, if we know the form of the function $g(x) = E(Y \mid X = x)$, (e.g. polynomial, exponential, etc.), then we can estimate the unknown parameters of $g(x)$ with several methods (e.g. least squares). In particular, if we do not know the general form of $g(x)$, except that it is a continuous and smooth function, then we can approximate it with a non-parametric method, as proposed by [8] and [11]. The aim of the non-parametric technique is to relax assumptions on the form of regression function and to allow data search for an appropriate function that represents well the available data, without assuming any specific form of the function. Thus, $g(x)$ can be estimated by:

$$\hat{g}_n(x) = \frac{\sum_{i=1}^{n} y_i k\left(\frac{x-x_i}{h(n)}\right)}{\sum_{i=1}^{n} k\left(\frac{x-x_i}{h(n)}\right)}, \tag{25}$$

where, $k(\bullet)$ is a density function such that: i) $k(x) < C < \infty$, ii) $\lim_{x\to\pm\infty} |xk(x)| = 0$, iii) $h(n) \to 0$ when $n \to \infty$. h is a bandwidth, also called a smoothing parameter, which controls the size of the local averaging. The function $k(x)$ is called the kernel; observe that kernel functions are generally used for estimating probability densities non-parametrically (see [21]). An overview of nonparametric regression or smoothing techniques may be found, e.g. among others Fan and Gijbels [10].

An alternative non-parametric approach for approximating the conditional expectation denoted by "*OLP*" has been given in [18]. Define by $\mathfrak{I}_X$ the σ-algebra generated by X (that is $\mathfrak{I}_X = \sigma(X) = X^{-1}(\mathcal{B}) = \{X^{-1}(B): B \in \mathcal{B}\}$, where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$). Observe that the regression function is just a "pointwise" realization of the random variable $E(Y|\mathfrak{I}_X)$, which can equivalently be denoted by $E(Y|X)$. The following methodology is aimed at estimating $E(Y|X)$ rather than $g(x)$. For this reason, we propose the following consistent estimator of the random variable $E(Y|X)$.

Let $X: \Omega \to \mathbb{R}$ and $Y: \Omega \to \mathbb{R}$ be integrable random variables in the probability space $(\Omega, \mathfrak{I}, P)$. Notice that: $E(Y|X)$ is equivalent to $E(Y|\mathfrak{I}_X)$. We can approximate $\mathfrak{I}_X$ with a σ-algebra generated by a suitable partition of Ω. In particular, for any $k \in \mathbb{N}$, we consider the partition $\{A_j\}_{J=1}^{b^k} = \{A_1, \dots, A_{b^k}\}$ of Ω in b^k subsets, where b is an integer number greater than 1 and:

- $A_1 = \{\omega: X(\omega) \le F_X^{-1}\left(\frac{1}{b^k}\right)\}$,
- $A_h = \{\omega: F_X^{-1}\left(\frac{h-1}{b^k}\right) < X(\omega) \le F_X^{-1}\left(\frac{h}{b^k}\right)\}$, for h=2,…,b^k–1,

- $A_{b^k} = \Omega - \bigcup_{j=1}^{b^k-1} A_j = \{\omega: X(\omega) > F_X^{-1}\left(\frac{b^k-1}{b^k}\right)\}$.

Thus, starting with the trivial σ-algebra $\mathfrak{I}_0 = \{\emptyset, \Omega\}$, we can generate a sequence of sigma algebras generated by these partitions obtained by varying k (k=1,…,m,…). Thus, $\mathfrak{I}_1 = \sigma\{\emptyset, \Omega, A_1, \dots, A_b\}$ is the sigma algebra generated by $A_1 = \{\omega: X(\omega) \le F_X^{-1}(1/b)\}$, $A_s = \left\{\omega: F_X^{-1}\left(\frac{s-1}{b}\right) < X(\omega) \le F_X^{-1}\left(\frac{s}{b}\right)\right\}$, s=1,...,b-1 and $A_b = \{\omega: X(\omega) > F_X^{-1}((b-1)/b)\}$, moreover:

$$\mathfrak{I}_k = \sigma\left(\{A_j\}_{j=1}^{b^k}\right), k \in \mathbb{N} \tag{26}$$

Under these hypotheses [18] proved that:

$$E(Y|X) = \lim_{k\to\infty} E(Y|\mathfrak{I}_k) \text{ a.s.} \tag{27}$$

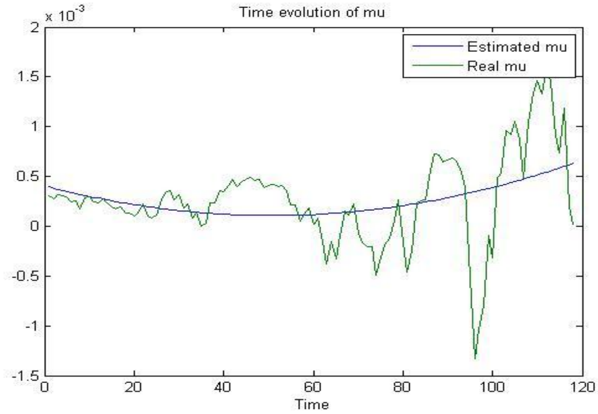
where, $E(Y|\mathfrak{I}_k)(\omega) = \sum_{j=1}^{b^k} E(Y|A_j) 1_{A_j}(\omega)$ a.s. and $1_{A_j}(\omega) = \begin{cases} 1 & \omega \in A_j \\ 0 & \omega \notin A_j \end{cases}$.

When b is large enough, even $E(Y|\mathfrak{I}_1)$ can be a good approximation of the conditional expected value $E(Y|X)$. On the one side, given N i.i.d. observations of Y, we get that $\frac{1}{n_{A_j}} \sum_{y\in A_j} y$ (where n_{A_j} is the number of elements of A_j) is a consistent estimator of $E(Y|A_j)$. On the other side, if we know that the probability p_i is the probability of the i-th outcome y_i of random variable Y, we get $E(Y|A_j) = \sum_{y_i\in A_j} y_i p_i / P(A_j)$, otherwise, we can give a uniform weight to each observation, which yields the following consistent estimator of $E(Y|A_j)$=$\frac{\mathbf{1}}{\boldsymbol{n_{A_j}}} \sum_{\boldsymbol{y_i \in A_j}} \boldsymbol{y_i}$, where n_{A_j} is the number of elements of A_j. Therefore, we are able to estimate $E(Y|\mathfrak{I}_k)$, that is a consistent estimator of the conditional expected value $E(Y|X)$ as a consequence of Proposition 1 in [18].

5 Second empirical analysis

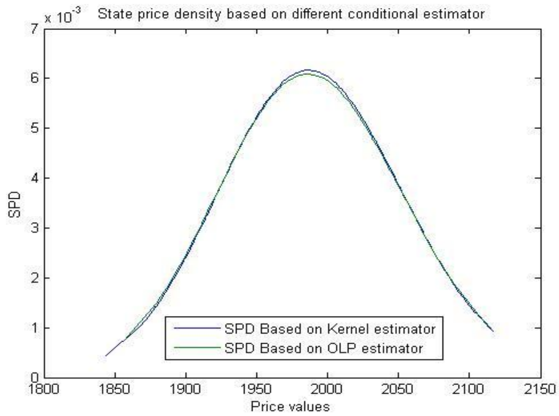
In the second empirical analysis, we present an application to the S&P 500 index using daily data for the sample period April 28, 2014 to April 28, 2015. In this context, we use Treasury Bond 3 months as a riskless interest rate for a period matching our selecting data. Firstly, we examine the real mean return function using local polynomial smoothing technique (22). The results of this analysis are reported in Figure 3.

Figure 3: Real mean return function estimation over time



Secondly, we evaluate the conditional expected price using both estimators, namely kernel estimator and OLP, to estimate $E(S_T \mid S_t)$ as described above. Finally, we use the relationship (19) in order to recover the SPD. The results of this analysis are reported in Figure 4.

Figure 4: State Price Densities obtained with Kernel and OLP estimators



From Figure 4 we note a slight difference in the result obtained from both estimators. This result can be explained by the nature of the two methodologies. In particular, the OLP method proposed by [18] yields a consistent estimator of the random variable $E(X \mid Y)$, while the generalized kernel method proposed in equation (25) yields a consistent estimator of the distribution function of $E(X \mid Y)$.Thus, OLP method that yields consistent estimators of random variables $E(X \mid Y)$ can be used to evaluate the SPD.

6 Conclusion

In this paper, we present alternative methods to evaluate the presence of arbitrage opportunities in the market. In particular, we examine the violation of the well-known put-call parity no-arbitrage relation and the nonnegativity of the SPD. Then, we propose different methods to estimate SPD. Particularly, we use two distinct methodologies for estimating the conditional expectation, namely the kernel method and the OLP method recently proposed by [18]. We deviate from previous studies in that we estimate SPD directly from the underlying asset under the hypothesis of BS model. To this end, firstly we examine the real mean return function using local polynomial smoothing technique. Then, we estimate the conditional expectation under real probability density. Under the hypothesis of BS model, we are able to derive a closed formula for approximating the conditional expectation under risk neutral probability. This analysis allows us extrapolating arbitrage opportunities and relevant information from different markets (futures and options) consistently with the analysis of the underlying.

Acknowledgements

This paper has been supported by the Italian funds ex MURST 60% 2014, 2015 and MIUR PRIN MISURA Project, 2013–2015, and ITALY project (Italian Talented Young researchers). The research was also supported through the Czech Science Foundation (GACR) under project 13-13142S and through SP2013/3, an SGS research project of VSB-TU Ostrava, and furthermore by the European Regional Development Fund in the IT4Innovations Centre of Excellence, including the access to the supercomputing capacity, and the European Social Fund in the framework of CZ.1.07/2.3.00/20.0296 (to S.O.L.).

References:

[1] Ait-Sahalia, Y., and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance* 53(2), 499–547

[2] Benko, M., Fengler, M., Härdle, W., Kopa, M (2007). On Extracting Information Implied in Options. *Computational statistics*, 22, 543–553

[3] Black, F., Scholes, M (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637-659

[4] Breeden, D. T., and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business* 51(4), 621–51

[5] Brunner B, Hafner R (2003) Arbitrage-free estimation of the risk-neutral density from the implied volatility smile. *Journal of Computational Finance*, 7(1), 75–106

[6] Carr, P. (2001). Constraints on implied volatility. Working paper, Bank of America Securities.

[7] D. S. Bates (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 9, 69–107

[8] E. A. Nadaraya, (1964). "On estimating regression," *Theory of Probability and its Applications*, 9(1), 141-142

[9] Eli Ofek, Mathew Richardson, Robert F. Whitelaw. (2004). Limited arbitrage and short sales restrictions: evidence from the options markets. *Journal of Financial Economics*, 74 305–342

[10] Fan J, Gijbels I (1996). Local polynomial modelling and its applications. Chapman and Hall, London

[11] G. S. Watson (1964), "Smooth regression analysis," *Sankhya, Series A*, 26(4), 359-372

[12] Harrison J, Kreps D.M (1979). Martingales and arbitrage in multiperiod securities markets. *Journal Economic Theory* 20,381–408

[13] J. Hull, (2015). Options, Futures, and Other Derivatives, 9th edition. Prentice Hall, New Jersey, USA

[14] Lando T., Ortobelli S. (2015). On the Approximation of a Conditional Expectation. *WSEAS Transactions on Mathematics*, ISSN 1109-2769, Volume 14, pp. 237-247

[15] Petronio F., Ortobelli S., Tamborini, Lando T. (2014). Portfolio selection in the BRICs stocks markets using Markov processes. *International Journal of Mathematical Models and Methods in Applied Sciences* 8, 311-318 (ISSN:1998-0140)

[16] R. C. Merton (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3, 125–144

[17] S. Heston (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6, 327–343

[18] S. Ortobelli, F. Petronio, T. Lando, (2015) "A portfolio return definition coherent with the investors preferences," to appear in IMA-*Journal of Management Mathematics*.

[19] S. Ortobelli, F. Petronio, T. Tichy (2014). Dominance among financial markets. *WSEAS Transactions on Business and Economics* 11, 2014, pp. 707- 717

[20] Stoll, H, R (1969). The relationships between Put and Call options prices. *The journal of finance*, 24(5), 802- 824

[21] V.A. Epanechnikov, (1965). "Non-parametric estimation of a multivariate probability density," *Theory of Probability and its Applications*, 14(1), 153-158