ACRN Journal of Finance and Risk Perspectives
Vol. 1, Issue 2, p. 43-68, Dec. 2012
ISSN 2305-7394

IS THE PRICE KERNEL MONOTONE?

Giovanni Barone-Adesi[1], Hakim DallÒ[2], Volodymyr Vovchak[3]

[1] Swiss Finance Institute at University of Lugano
[2] University of Lugano
[3] Swiss Finance Institute at University of Lugano

Abstract. *The pricing kernel based on SPX option prices and GARCH model is derived and tested for monotonicity. Derivation of the risk neutral distribution is conducted based on the result in Breeden and Litzenberger (1978) and the historical density is estimated by means of our asymmetric GARCH model. Applying two statistical tests we are not able to reject null hypothesis of monotonically decreasing pricing kernel, showing that using a large dataset and introducing non-Gaussian innovations solves the pricing kernel puzzle posed in Jackwerth (2000), both in a single day and over an average of different days with the same options' maturity. We also evaluate the price kernel before and during the recent crisis and we look at the change in the shape in order to evaluate the difference.*

Keywords*: Pricing kernel, State price density per unit probability, Risk neutral, Historical distribution*

Introduction

According to economic theory, the shape of the state price density (SPD) per unit probability (also known as the asset pricing kernel, (Rosenberg & Engle, 2002) or stochastic discount factor (SDF), (Campbell, Lo, & MacKinlay, 1997)) is a decreasing function in wealth.

(Jackwerth J. C., 2000) finds a kernel price before the crash of 1987 in agreement with economic theory, but a discordant result for the post-crash period. After his work, a number of papers have been written on this topic trying to explain the reason for this puzzle. (Rosenberg & Engle, 2002), (Detlefsen, Härdle, & Moro, 2007) and (Jackwerth J. C., 2004) are among the most interesting papers on this subject. Unfortunately, none of them found an answer to this puzzle. In all of these papers the authors found problems in the methodology employed by previous papers and tried to improve them, but the result was the same: the puzzle remained.

An answer to this puzzle has been given in (Chabi-Yo, Garcia, & Renault , 2005), where they argue that the main problem is the regime shifts in fundamentals: when a volatility change, the kernel price is no longer monotonically decreasing. In each regime they prove that the kernel price is consistent with economic theory, but when there is a shift in regime the kernel price changes in its shape and it is no longer consistent with economic theory.

In a recent paper, (Barone-Adesi, Engle, & Mancini, 2008) compute again the kernel price and find kernel prices consistent with economic theory. In particular they find kernel price consistency for fixed maturities. They do not pool different maturities as (Aït-Sahalia & Lo, 1998) and therefore they avoid the problem that arises when maturities are different, but they do not consider the change in fundamentals as a relevant aspect of their computation. Their

result can be explained by the fact that the sample they use is very short (3 years) and that throughout this period (2002 - 2004) the volatility does not change much.

In this paper we compute the kernel price both in a single day and as an average of kernel prices over a period of time, holding maturity constant. We want to understand the implication of the changing regime using two measures of moneyness: in the first case we consider the kernel price as a function of two parameters, the underlying and the interest rate (we do not take into consideration the changing regime) and then we add third parameter – the volatility of the underlying. As argued in (Brown & Gibbons, 1985), under some general assumptions one may substitute (in estimation) consumption with the market index while working with asset pricing models1. That's why in order to evaluate the kernel price we need to take a broad index which attempts to cover the entire economy. As it is common in this kind of literature to use S&P500 index, we also use data on S&P500 index prices and options on the S&P500 index over a period of 12 years (from the 2nd of January 1996 to the 31st December 2007).

Evaluating the kernel price in a period of time, without taking into consideration the change in volatility, should lead to a kernel not consistent with economic theory. Surprising, when we compute the kernel price considering only two parameters (the underlying and the interest rate), the average kernel price is consistent with economic theory, with the exceptions of a few dates.

To check our result we also do kernel smoothing which is similar to averaging, but it has the advantage of producing smooth price kernel without the spikes one might get from simple averaging. Another robustness check for our results is the testing of monotonicity of the obtained kernel price. We take our estimated average price kernel, consider its monotone version and then compare the monotone version with the estimated version by means of Kolmogorov-Smirnov test.

In order to estimate the risk neutral distribution, we use the well-known result in (Breeden & Litzenberger, 1978). The difference with previous works is in the options we use. Instead of creating option prices through nonparametric or parametric models (all the previous research use artificial price of options and this could introduce a bias in the methodology), we use only the options available on the market. We then construct the historical density using the GJR GARCH model with Filtered Historical Simulation already presented in (Barone-Adesi, Engle, & Mancini, 2008).

As discussed in (Rosenberg & Engle, 2002), among the several GARCH models, the GJR GARCH with FHS has the flexibility to capture the leverage effect and the ability to fit daily S&P500 index returns. Then, the set of innovations estimated from historical returns and scaled by their volatility gives an empirical density function that incorporates excess skewness, kurtosis, and other extreme return behavior that is not captured in a normal density function. These features avoid several problems in the estimation of the kernel price. For example, using a simple GARCH model where the innovations are standard normal (0; 1) leads to a misspecification of the return distribution of the underlying index.

Once we have the two probabilities, under the pricing and the objective measures, we take the ratio between the two densities, discounted by the risk-free rate, in a particular day, to compute the kernel price for a fixed maturity. We repeat the same procedure for all the days in the time series which have options with the same maturity and then we take the average of the kernel price through the sample. At the same time we apply kernel smoothing on the estimated values of the price kernel to confirm our result.

1The aggregated consumption is inconvenient in two ways: (1) it is hard to measure, and (2) no options on aggregate consumption are traded.

We also evaluate how the shape of the price kernel changes before and during a crisis (the 2008 crisis). We notice that the three periods before the crisis (2005, 2006 and 2007) exhibit fairly monotonically decreasing paths, while during the crisis, the kernel price remains monotonically decreasing, but has higher values. this is consistent with the idea that during a crisis investors increase the risk aversion.

In order to evaluate the impact of the shifting regime, we repeat the computation of the different kernel prices considering the volatility as a parameter of the kernel function. As expected, results improve, but they are still quite similar, supporting our first intuition that the changing regime is relevant, but our methodological choices have a strong impact on the final result.

The remainder of this paper is organized as follows. In section 2, we present a review of the literature and we define the "pricing kernel puzzle". In section 3, we define our method to estimate the kernel price. We explain our application of the result of (Breeden & Litzenberger, 1978) and we derive the risk neutral distribution. We then estimate the historical density using a GJR GARCH method with FHS and we take the kernel price from a particular day as well as the kernel price over the time series of our sample. In the last part of the section we conduct two statistical tests. Namely, we use two Kolmogorov type tests of the monotonicity of the estimated pricing kernels. In section 4, we provide further evidence of our results. First we plot kernel price with different maturities to prove the robustness of our methodology, then we take the average of these different kernel prices and we show that the average of SPD per unit probabilities with close maturities have a monotonically decreasing path. In section 5, we present the change in the kernel price shape before and during the recent crisis. In section 6, we extend our model, using a kernel price with three parameters (underlying, volatility and risk-free), and in section 7 we offer conclusions.

Review of the Literature

In this section we derive the price kernel as in macroeconomic theory and also as in probability theory. We then present some methods, parametric and non-parametric, to derive the kernel price.

Price kernel and investor preference

The ratio between the risk neutral density and the historical density is known as the price kernel or state price density per unit probability. In order to explain the relationship between the risk-neutral distribution and the historical distribution we need to introduce some basic concepts from macroeconomic theory. In particular, we use a representative agent with a utility function $U(\cdot)$. According to economic theory (the classical von Neumann and Morgenstern economic theory), we have three types of investors: risk averse, risk neutral and risk lover. The utility function $U(\cdot)$ of these investors is a twice differentiable function of consumption c: $U(c)$. The common property for the three investors is the non-satiation property: the utility increase with consumption, e.g. more consumption is preferred to less consumption, and the investor is never satisfied - he never has so much wealth that getting more would not be at least a little bit desirable. This condition means that the first derivative of the utility function is always positive. On the other hand, the second derivative changes according to the attitude the investor has toward risk.

If the investor is risk averse, his utility function is an increasing concave utility function. The risk neutral investor has a second derivative equal to zero, while the risk seeker - a convex utility function.

Defining $u(\cdot)$ as the single period utility function and β as the subjective discount factor, we can write the intertemporal two-period utility function as

$$U(c_t; c_{t+1}) = u(c_t) + \beta u(c_{t+1})$$

We introduce ξ as the amount of an asset the agent chooses to buy at time t, e as the original endowment of the agent, P_t as the price of the asset at time t and x_{t+1} as the future payoff of the asset. The optimization problem is:

$$\max_{\xi}\{u(c_t) + \beta E_t[u(c_{t+1})]\}$$

$$subject\ to$$

$$c_t = e_t - P_t\xi$$

$$c_{t+1} = e_{t+1} + x_{t+1}\xi$$

The first constraint is the budget constraint at time 1, while the second constraint is the Walrasian property, e.g. the agent will consume his entire endowment and asset's payoff at the last period. Substituting the constraints into the objective and setting the derivative with respect to ξ equal to zero we get:

$$P_t = E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}x_{t+1}\right].$$

We define

$$\beta E_t\left[\frac{u'(c_{t+1})}{u'(c_t)}\right] = m_{t,t+1} = MRS, \tag{1}$$

as the *marginal rate of substitution* at time t. The MRS is also known as the Stochastic Discount Factor (SDF) or the price kernel. Therefore the price of any asset can be expressed as

$$P_t = E_t[m_{t,t+1}x_{t+1}].$$

In a continuous case, the price of any asset can be written as

$$P_t^p = \int_{\mathbb{R}} m_{t,T}(S_T)x_T(S_T)p_{t,T}(S_T)\,dS_T \tag{2}$$

where $p_{t,T}(S_T)$ is the physical probability of state S_T (for the rest of the paper we refer to this probability as the historical probability) and $x_T(S_T)$ is the payoff of an asset.

To define the price of an asset at time t, under the risk neutral measure, we can write equation (2) as:

$$P_t^q = e^{-rt}\int_{\mathbb{R}} x_T(S_T)q_{t,T}(S_T)\,dS_T \tag{3}$$

where $q_{t,T}(S_T)$ is the state price density (for the rest of the paper we refer to this probability as the risk neutral probability). At this point, combining equation (2) and (3) we can derive the SDF as:

$$m_{t,T}(S_T) = e^{-rt}\frac{q_t(S_T)}{p_t(S_T)} \tag{4}$$

In this case we consider a two period model where the price kernel is a function only of the underlying, S_T, and the risk free rate, r. In the following part we will see how to have a kernel price with more parameters.

In their papers (Arrow, 1964) and (Pratt, 1964) find a connection between the kernel price and the measure of risk aversion of a representative agent. Arrow-Pratt measure of absolute risk-aversion (ARA) is defined as:

$$A_t(S_T) = \frac{u''(S_T)}{u'(S_t)}$$

The absolute risk aversion is an indicator of willingness to expose some amount of wealth to risk as a function of wealth. An agent's utility function demonstrating decreasing (constant or increasing) absolute risk aversion implies that her willingness to take risk increases (does not change or decreases) as the agent becomes wealthier.

Classic economic theory assumes risk averse economy agents, i.e. the utility function of the economy is concave (mathematically $u''(S_T) \leq 0$). The following argument should unveil an impact of this basic property of pricing kernel behavior.
From (1), the pricing kernel can be written as function of the marginal utility as:

$$m_{t,T}(S_T) = \beta\frac{u'(S_T)}{u'(S_t)},$$

and its first derivative is:

$$m'_{t,T}\,(S_T) = \beta\frac{u''(S_T)}{u'(S_T)} = -\beta A_t(S_T),$$

which (remember $u''(S_T) \leq 0$ and $u'(S_T) > 0\ \forall\ t$) implies $m'_{t,T}\,(S_T) \leq 0$, or in words, the pricing kernel is decreasing as a function of the wealth. We are aiming to check if the pricing kernel is decreasing and, as a consequence, if agents in the economy are risk averse.

Nonparametric and parametric estimation

There are several methods to derive the kernel price. There are both parametric models and nonparametric models. In this section we give a review of the most well-known methods used in literature. We focus particularly on the nonparametric models because they do not assume any particular form for the risk neutral and historical density and also for the kernel price.

One of the first papers to recover the price kernel in a nonparametric way is (Aït-Sahalia & Lo, 1998). In their work they derive the option price function by nonparametric kernel regression and then, applying the result in (Breeden & Litzenberger, 1978), they compute the risk neutral distribution. Their findings are not consistent with economic theory. Because they

look at the time continuity of $m_{t,T}$ across time, one may understand their results as estimates of the average kernel price over the sample period, rather than as conditional estimates.

Other problems in their article are discussed in (Rosenberg & Engle, 2002). In particular they suggest that the non-specification of the investors beliefs about future return probabilities could be a problem in the evaluation of the kernel price. Also they use of a very short period of time, 4 years, to estimate the state probabilities. Moreover, they depart from the literature on stochastic volatility, which suggests that future state probabilities depend more on recent events than past events. In fact, past events remain useful for prediction of future state probabilities. In order to take this into account we use a dataset of 12 years of option prices.

A work close in spirit to (Aït-Sahalia & Lo, 1998) is (Jackwerth J. C., 2000). His article is one of the most interesting pertaining to this literature. Beyond the estimation technique used, his paper is noteworthy because it also opened up the well-known "pricing-kernel puzzle". In his nonparametric estimation of the kernel price, Jackwerth finds that the shape of this function is in accordance with economic theory before the crash of 1987, but not after the crash. He concludes that the reason is the mispricing of options after the crash.

Both articles could incur some problems that cause the kernel price and the relative risk aversion function (RRA) to be not consistent with economic theory. In (Aït-Sahalia & Lo, 1998), we see that, if the bandwidth changes, the RRA changes as well and this means that the bandwidth chosen influences the shape of the RRA; on the other hand, in (Jackwerth J. C., 2000), the use of option prices after the crisis period could influence the shape of the kernel price if volatility is misspecified.

Another nonparametric estimation model for the kernel price is given by (Barone-Adesi, Engle, & Mancini, 2008), where they use a procedure similar to the one used by (Rosenberg & Engle, 2002), but with a nonparametric estimation of the ratio $q_{t,t+\tau}/p_{t,t+\tau}$. While in the papers by (Aït-Sahalia & Lo, 2000) and (Jackwerth J. C., 2000) results are in contrast with the economic theory, (Barone-Adesi, Engle, & Mancini, 2008) find a kernel price which exhibits a fairly monotonically decreasing shape.

Parametric methods to estimate the kernel price are often used in literature. (Jackwerth J. C., 2004) provides a general review on this topic, but for the purpose of our work we do not go into much detail on parametric estimation. As pointed out by (Birke & Pilz, 2009) there are no generally accepted parametric forms for asset price dynamics, for volatility surfaces or for call and put functions and therefore the use of parametric method may introduce systematic errors.

Our goal is to test whether a different nonparametric method, starting from option pricing observed in the market, respects the conditions of no-arbitrage present in (Birke & Pilz, 2009). In particular, we test if the first derivative of the call price function is decreasing in the strike and the second derivative is positive. These conditions should guarantee a kernel price monotonically decreasing in wealth.

It is important to stress that our kernel price is a function of three variables: the underlying price, the risk-free rate and volatility. In the first part, we use only two factors: the underlying and the risk-free rate. In last sections we introduce also volatility.

Empirical kernel price

In this section we compute the kernel price as the ratio of the risk-neutral and the historical density, discounted by the risk-free interest rate. First we describe how we compute the risk-neutral density. Then, we explain our computation of the historical density. In each part we describe the dataset we use and our filter for cleaning it.

Theoretical backgrounds of risk-neutral density

(Breeden & Litzenberger, 1978) shows how to derive the risk-neutral density from a set of call options with fixed maturity. The formula for risk-neutral density is (see Appendix A for derivation):

$$f(K) = e^{rT} \frac{\partial^2 C(S_t, K, T)}{\partial K^2} |_{S_T = K} \tag{5}$$

We can approximate this result for the discrete case as:

$$f(K) \approx e^{rT} \frac{C_{i+1}(S_t, K, T) - 2C_i(S_t, K, T) + C_{i-1}(S_t, K, T)}{(K_{i+1} - K_i)(K_i - K_{i-1})} |_{S_T = K} \tag{6}$$

and for puts

$$f(K) \approx e^{rT} \frac{P_{i+1}(S_t, K, T) - 2P_i(S_t, K, T) + P_{i-1}(S_t, K, T)}{(K_{i+1} - K_i)(K_i - K_{i-1})} |_{S_T = K} \tag{7}$$

Note that in equations (6) and (7) we wrote numerical derivatives for values of K_{i-1} and K_{i+1} which are not symmetric around K_i. (Breeden & Litzenberger, 1978) used symmetric strikes while deriving (5). But having non symmetric strikes does not hurt our estimation in any sense; to the contrary it gives us more observations and as a result may improve our estimation.

Now we discuss some other methods of deriving the risk neutral density and compare it to the one we use.

A recent paper by (Figlewski, 2008) is very close in spirit to our work. In his paper he derives the risk neutral distribution using the same result in (Breeden & Litzenberger, 1978). We differ from him in some aspects. First, we use the bid and ask prices that are given on the market to construct butterfly spreads. e.g. for the long position the ask price is used and for the short position the bid price. Our choice removes negative values in the risk-neutral distribution and we therefore find that the no-arbitrage condition described in (Birke & Pilz, 2009) holds. Second, we do not need to convert the bid, ask, or mid-prices into implied volatility to smooth the transition from call to put because we take the average of butterfly prices from several days with equal maturities and this improves the precision of our result. Other similar works are discussed in (Bahra, 1997), (Pirknert, Weigend, & Zimmermann, 1999) and (Jackwerth J. C., 2004).

In (Bahra, 1997), the author proposes several techniques to estimate the risk neutral density. For every method he explains the pros and the cons. He then assumes that the options prices can be derived either using a parametric method, by solving a least squares problem, or nonparametric only, using kernel regression. In our work, using a time series of options over a sample of 12 years and taking averages we avoid the parametric or nonparametric pricing step and therefore we rely only on pricing available on the market.

In (Pirknert, Weigend, & Zimmermann, 1999), they use a combination approach to derive the risk neutral distribution. They combine the implied binomial tree and the mixture distributions to get the approach called “Mixture Binomial Tree”. Our work differs from their work due to our use of European options. In their work, they use American options and therefore they could have the problem of the early exercise. In our sample, we consider only European options to be sure to have the risk neutral density for that expiration time only.

(Jackwerth J. C., 2004) may be considered as a general review of different methods and problems. He concentrates in particular on nonparametric estimation, but he gives a general

overview also on parametric works, sorting parametric works into classes and explaining the positive and negative aspects of each one.

Historical density

To obtain historical probability density we need to account for features of the empirical returns of S&P500 index. There are lots of evidence suggesting that return innovations are (i) not normal, (ii) volatility is stochastic and, (iii) that positive and negative shocks to return have diverse effect on returns' volatility (see for example (Ghysels , Harvey , & Renault , 1996). That's why we are going to use a GARCH model together with the filtered historical simulation (FHS) approach used by (Barone-Adesi, Bourgoin, & Giannopoulos, 1998). FHS approach allows to model volatility of returns without specifying any assumption on return innovations.

Among variety of GARCH models we are going to use (Glosten, Jagannathan, & Runkle, 1993) (GJR GARCH) model. The choice of the GJR model relies on two properties: 1) its ability to capture an asymmetry of positive and negative returns effect on return volatility and, 2) its fitting ability (Rosenberg & Engle, 2002) document that GJR GARCH model fits S&P500 returns data better than other GARCH models).

Under the historical measure, the asymmetric GJR GARCH model is

$$log\frac{S_t}{S_{t-1}} = \mu + \epsilon_{t,}$$

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\epsilon_{t-1}^2 + \gamma I_{t-1}\epsilon_{t-1}^2,$$

where S_t- the underlying price, $\epsilon_t = \sigma_t z_t$, and $z_t \sim f(0,1)$, and $I_{t-1} = 1$ when $\epsilon_{t-1} < 0$, and $I_{t-1} = 0$ otherwise. The scaled return innovation, z_t, is drawn from the empirical density function $f(\cdot)$, which is obtained by dividing each estimated return innovation, $\hat{\epsilon}_t$, by its estimated conditional volatility $\hat{\sigma}_t$. This set of estimated scaled innovations gives an empirical density function that incorporates excess skewness, kurtosis, and other extreme return behavior that is not captured in a normal density function.

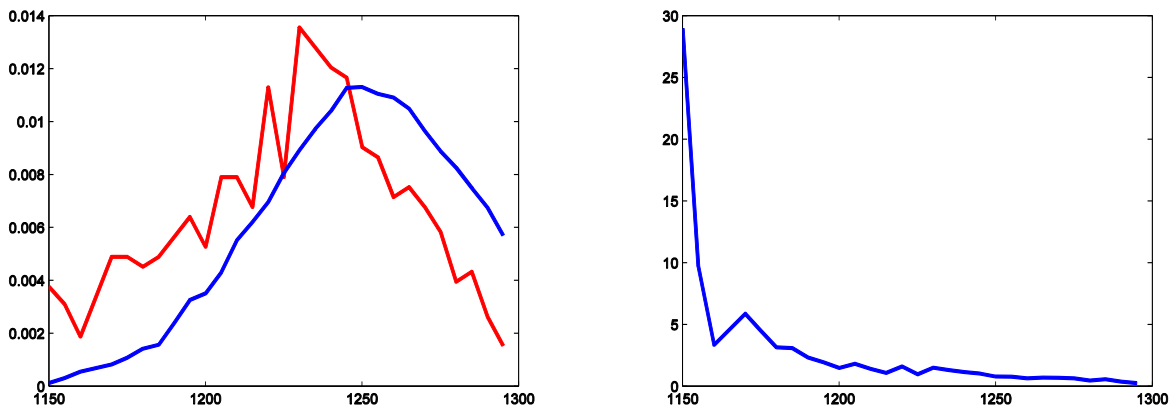


Figure 1. Left: Risk-neutral distribution (red line) and the historical distribution (blue line). We take one day at random from our sample (11 August 2005) with maturity equal to 37 days. Right: Price kernel for this particular day (11 August 2005).

The methodology we use to estimate the historical density is as follows. We have a set of risk neutral densities, $f(K_i)$, for each day over 12 years. They are calculated from S&P500 index options with constant maturities. Our $f(K_i)$ are prices of hypothetical butterfly strategies constructed from two (call or put) options with strike K_i and two long options of the same type with strikes K_{i-1} and K_{i+1}. Our triplets K_{i-1}, K_i and K_{i+1} are not necessary symmetric. For each day, we estimate the parameters of the GJR GARCH using a time series of 3500 returns from the S&P500. Once we have the estimated parameters for each day, we simulate 35000 paths of S&P500 index using as a distribution of z_t the empirical distribution of the normalized innovation (FHS). We estimate the probability that at maturity we exercise the butterfly, e.g. we count the fractions of paths that at maturity are in the range $[K_{i-1}, K_{i+1}]$2:

$$p(K_i) = \frac{\frac{\text{\#of paths in the interval } [K_{i-1}, K_{i+1}]}{\text{total number of paths}}}{K_{i+1} - K_{i-1}}. \tag{8}$$

Once we have computed the probability for each day, we can apply the same methodology we use for the risk-neutral distribution. We round the butterfly moneyness to the second digit after the decimal point and we take the average over the sample period.

We use the mid-strike for the butterfly and we round the moneyness to two decimal places. We take the average throughout the time series and we plot the resulting distribution as a function of moneyness. The historical density is drawn for one day at figure 1, and averaged at figure 2.

In the next subsection we explain the estimation method we use for the risk neutral distribution.

Risk-neutral estimation

We use European options on the S&P 500 index (symbol: SPX) to implement our model. We consider the closing prices of the out-of-the-money (OTM) put and call SPX options from 2nd January 1996 to 29th December 2007. It is known that OTM options are more actively traded than in-the-money options and by using only OTM options one can avoid the potential issues associated with liquidity problems.

Option data and all the other necessary data are downloaded from OptionMetrics. We compute the risk-neutral density at two different maturities: 37, 46, 57 and 72 days3. The choice of maturities is random and the same procedure can be applied for all other maturities. We download all the options from our dataset with the same maturities (we provide analysis and graphs for four maturities: 37, 46, 57 and 72 days; for other maturities results are similar) and we discard the options with an implied volatility larger than 75%, an average price lower than 0.05 or a volume equal to 0. In table 1 we summarized the number of options available for each maturity.

2 In our sample we use intervals with different lengths: most of them are intervals with a length of 10 index points, but we also have some intervals with 25 or 50 points, and these intervals are in some cases overlapping.

3 We work with this four maturities through the paper, although we provide sometimes graphs and p-values for more maturities.

Table 1. The options number for all maturities

Panel A. Main sample (1996-2007)

Maturity	36	37	38	39	43	44	45	46
Calls	2753	2789	2667	2537	2424	2321	2318	2171
Puts	3406	3421	3308	3217	3112	3057	3026	2870

Maturity	54	57	58	59	71	72	73	74
Calls	1879	1873	1853	1934	1460	1466	1414	1281
Puts	2469	2488	2497	2569	1976	1985	1966	1837

Panel B. Around crisis samples

Maturity	37	46	57	72
Aug 12, 2004 to Sep 15, 2005				
Calls	340	250	207	147
Puts	385	297	217	179
Nov 10, 2005 to Oct 10, 2006				
Calls	364	280	256	172
Puts	424	329	303	199
Jun 14, 2006 to Jun 14, 2007				
Calls	486	355	325	187
Puts	608	441	386	250

	Oct 11, 2007 to Aug 14, 2008			
Calls	607	505	403	271
Puts	727	609	478	301

We construct then butterfly spreads using the bid-ask prices of the options. The butterfly spread is formed by two short call options with strike K_i and long two call with strikes K_{i+1} and K_{i-1}, the same for puts. We divide the dataset and we construct a butterfly spread for every day. We try to use the smallest distance possible in the strikes to construct the butterfly spread. Following the quotation for the SPX we use a difference of 5 basis points. However, for the deep-out-of-the-money options we need to take into consideration a larger distance because there are fewer options traded. In that case, we arrive to have spreads of 10 to 50 points. We download option prices, order by strike, from smallest to largest. We take

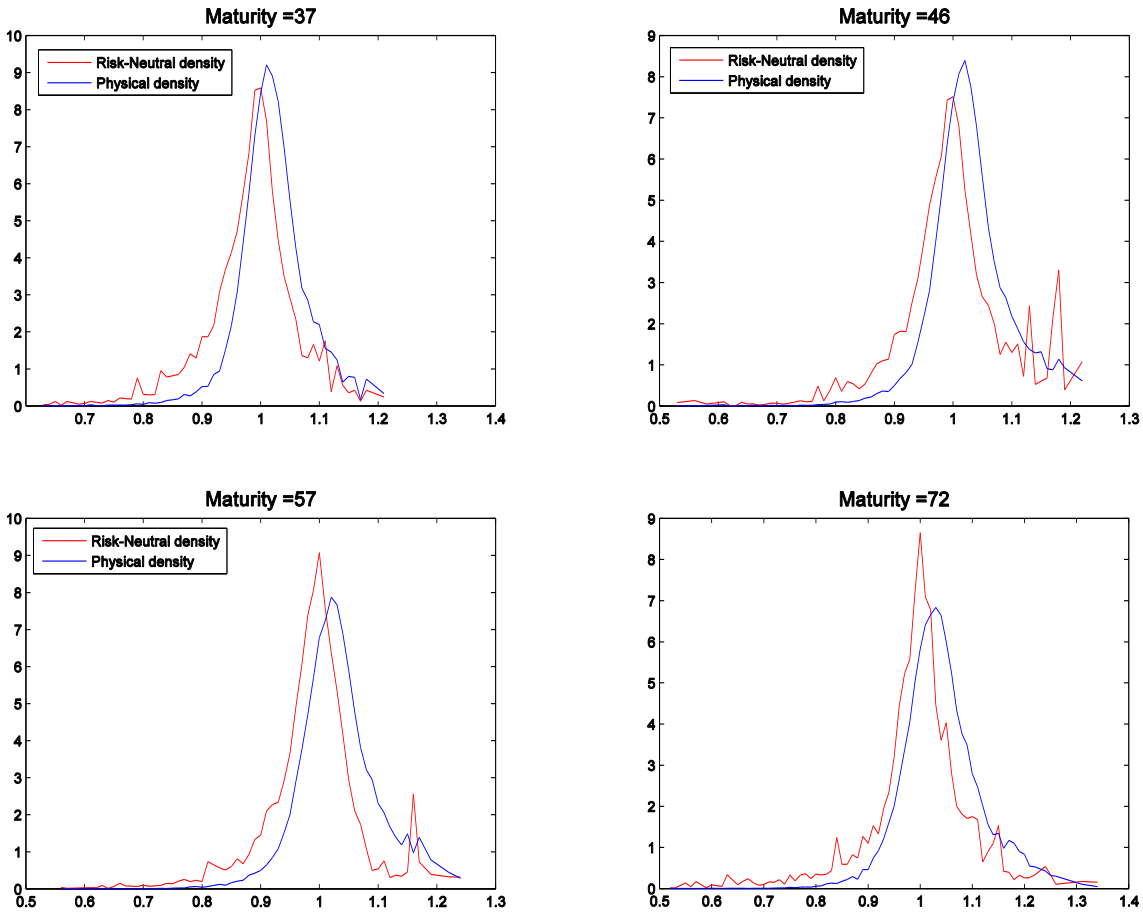


Figure 2. Risk neutral and historical distribution as the 12 year average of risk neutral and historical distributions for a fixed number of days to maturity.

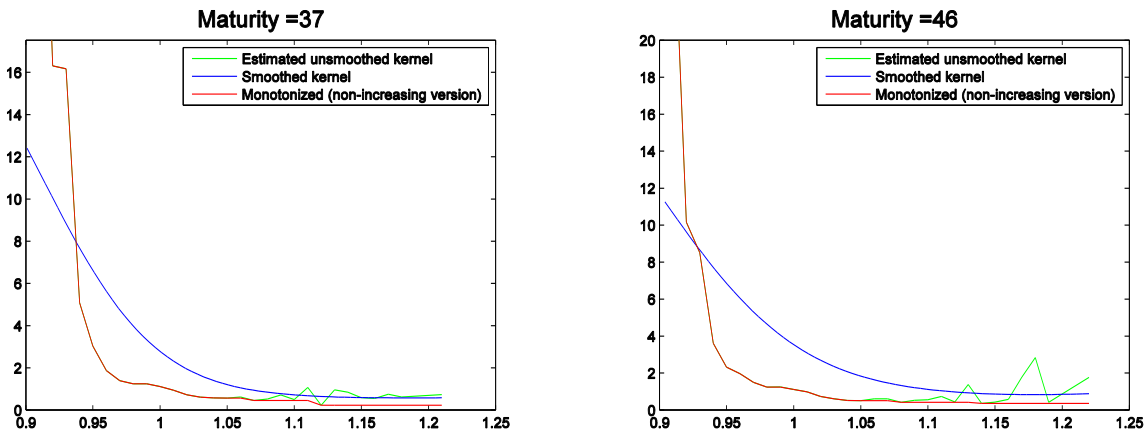
the second difference of option prices using formulas (6) and (7), for calls and puts separately, and then combine them. We do it for each day available in our dataset.

In figure 1 we take a day at random from our sample and we show the risk-neutral distribution, the historical distribution and their ratio as the price kernel. As an example, we take 11 August 2005, and we look at options with a maturity equal to 37 days. We see that for this choice, the kernel price shows a monotonically decreasing path in S_T, with some jumps because we do not smooth the curve.

At this point, we take into consideration the moneyness of each butterfly. As reference moneyness of the butterfly spread, we use the moneyness of the middle strike. We round all the butterfly moneyness to the second decimal digit and we take the average of all the butterfly prices with equal moneyness4. We can now plot the risk-neutral distribution as an average of the butterfly prices for a fixed maturity over a twelve year period. Figure 2 draws the risk-neutral and historical (physical) distributions of the underlying index. As expected, the risk neutral distribution is shifted to the left with respect to the historical distribution.

Kernel price

We apply the definition given in equation (4) in order to get the kernel price. From previous calculations we obtain the average risk-neutral distribution for the fixed maturity and also the average historical density. In order to get the average kernel price we take the kernel price of each day and then we compute the average from all the days in our time series. Averaging across time allows us to increase the otherwise small number of data points.



4 In order to find an equal moneyness it is necessary to round the moneyness values to the second decimal digit. Otherwise we cannot average and we are left with a lower number of points.

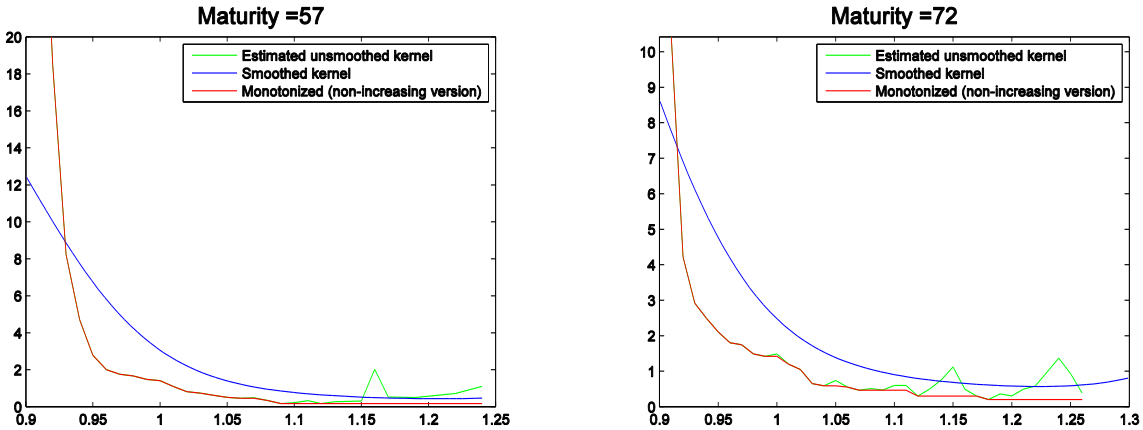


Figure 3: SPD per unit probability as the average of the SPD per unit probability throughout the time series of 12 years and with equal maturity. It is important to keep in mind that this SPD per unit probability is not derived from the two distributions given in Figure 2.

It is important to recall that our kernel price is the average of the kernel prices estimated each day. In other words, we estimate the risk neutral and the historical densities for each day, calculate the price kernel and then average these daily kernels, rather than calculate average densities over the entire sample.

Generally, for all different maturities we get a monotonically decreasing path for the kernel price and all of these are in accordance with economic theory.

We also obtain a kernel-smoothed version of the price kernel by applying to our unaveraged pricing kernel the kernel smoothing. Our smoothed pricing kernels are also monotonic decreasing (see figure 3).

Monotonicity testing

In this section we introduce some monotonicity testing. We do two kinds of monotonicity tests. Both of them, as many nonparametric tests, involve the notion of Kolmogorov distance. The first one is very simple, may be not completely justified by theory but it is very intuitive. The second one is more thorough and is more sound statistically proves and results.

Simple test. Our first test (call it *simple*), considers a monotonized version of the price kernel obtained earlier. We test that the estimated and monotonized versions are equal.

Table 2: The pricing kernel monotonicity testing

Maturity, days	Intuitive test		Durot test	
	$H_0{}^a$	*P- value*	$H_0{}^b$	*P- value*
36	Not rej.	0.3213*	Not rej.	0.2398
37	Not rej.	0.0221**	Not rej.	0.2138
38	Not rej.	0.0259**	Not rej.	0.4809
39	Not rej.	0.3402*	Not rej.	0.3338
43	Not rej.	0.1088*	Not rej.	0.3535
44	Not rej.	0.6976*	Not rej.	0.1824
45	Not rej.	0.1088*	Not rej.	0.4186
46	Not rej.	0.0259**	Not rej.	0.1046
54	Not rej.	0.1844*	Not rej.	0.2855
57	Not rej.	0.0259**	Not rej.	0.2079
58	Not rej.	0.1088*	Not rej.	0.4088
59	Not rej.	0.0343**	Not rej.	0.1097
71	Not rej.	0.1315*	Not rej.	0.1021
72	Not rej.	0.0244**	Not rej.	0.1034
73	Rej.	0.0082	Rej.	0.0493
74	Rej.	0.00013	Not rej.	.05114

^{a} In the intuitive test we use 1% confidence interval. Starred numbers denote that in this case H_0 is not rejected for the 5% confidence level, double starred − for the 1%.

^{b} While using the Durot testing procedure we use 5% confidence level.

We create our monotonized version, $\widehat{m}(x)$, as follows. From estimation of the pricing kernel, $x_i \to m(x_i)$, we inspect each point $m(x_i)$ for monotonicity. If it is between its adjacent points, the monotonized version is defined to be equal to estimated one, otherwise − the monotonized version is defined to be constant and equal to previous value. More precisely, the monotone version is given as:

$$\widehat{m}(x_1) = m(x_1),$$

$$\widehat{m}(x_{i+1}) = \begin{cases} m(x_{i+1}), & if\ m(x_{i+1}) \le m(x_i) \\ \widehat{m}(x_i), & if\ m(x_{i+1}) \ge m(x_i) \end{cases}$$

where $m(x)$ − the estimated price kernel.

After getting the monotone version we compare it with the estimated price kernel, $m(x)$, by means of Kolmogorov-Smirnov test5.

Results of testing are given in the table 2, test result in "Not rej." if $H_0: \widehat{m} \equiv m$ is not rejected and in "Rej." − if it was rejected at the 1% significance level. Also one can observe p-values of H_0 and that comes from the table is that we are not able to reject null hypothesis of monotonicity of the price kernel at the confidence level 1% (for some maturities it is 10%). Only for maturities of 73 and 74 days we reject monotonicity, possibly because of discretization errors. In any case this test is rather weak. Thus we are going to introduce a more powerful test.

Sophisticated test. To test monotonicity of the pricing kernel more thoroughly we use a Kolmogorov-type test for monotonicity of a regression function described in (Durot, 2003). Hypothesis testing is performed within the following regression model

$$y_i = f(x_i) + \varepsilon_i,$$

where, in our case x_i is the moneyness of the option, y_i is the price kernel, and ε_i − random errors with mean 0. Our second test is based on the fact that f is non-increasing (decreasing) if and only if $\hat{F} \equiv F$, here $F(t) = \int_0^t f(s)ds, t \in [0,1]$, and $\hat{F}$ is the least concave majorant (lcm) of F. One should reject H_0 about monotonic decrease of pricing kernel in case the difference between F and $\hat{F}$ corresponding to our price kernel is too large.

Test construction. From previous subsections we can obtain the pricing kernel, so we have function f given as

$$moneyness\ (x_i) \xrightarrow{\quad f \quad} pricing\ \ kernel$$

As mentioned above f is non-increasing on [0,1] if and only if F is concave on[0,1]. Denote i_t integer part of nt and define

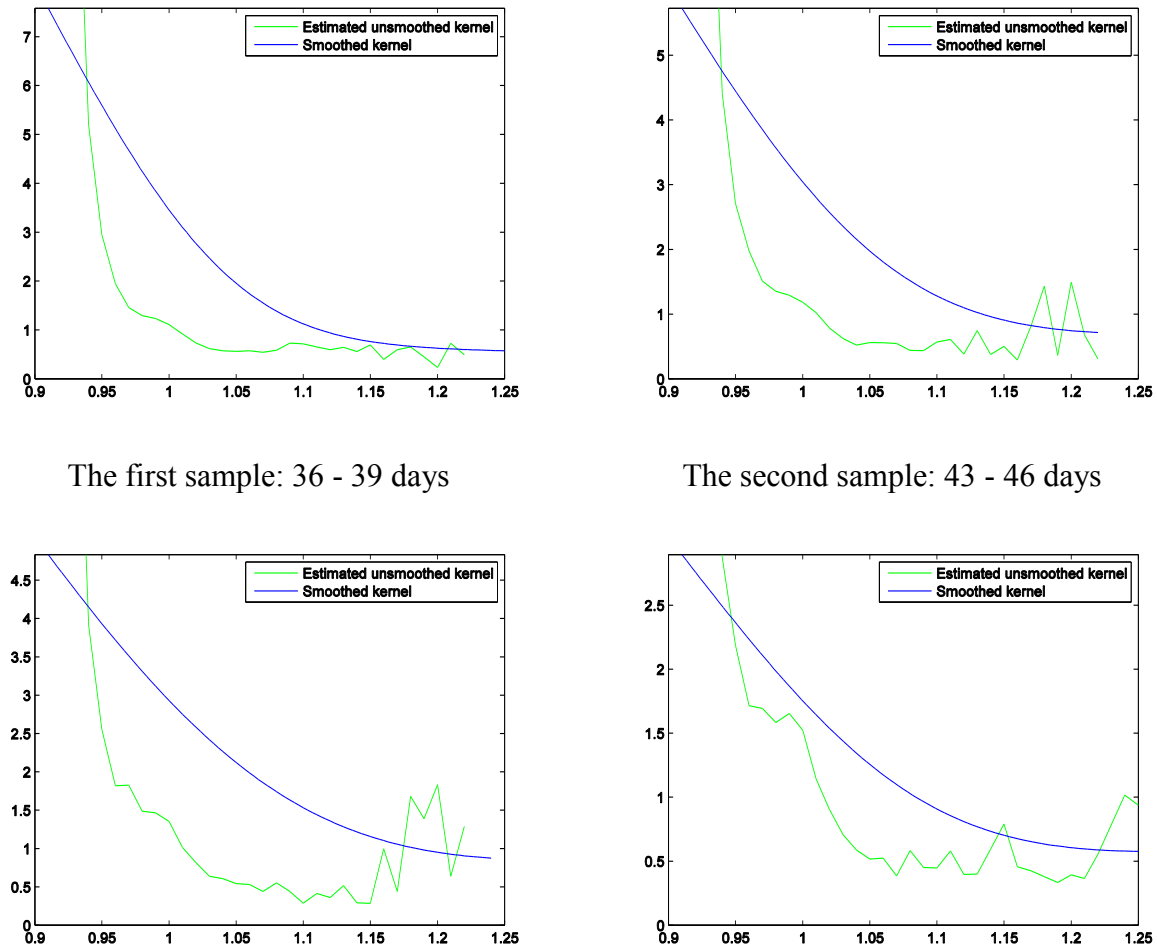
$$F_n(t) = \frac{1}{n}\sum_{j \le i_t} y_i + \left(t - x_{i_t}\right) y_{i_t}, \qquad t \in [0,1]$$

F_n is approximation of F, thus we consider Kolmogorov-type test statistic

$$S_n = \frac{\sqrt{n}}{\widehat{\sigma_n}} \sup_{t \in [0,1]} \left|\widehat{F_n}(t) - F_n(t)\right|, \tag{9}$$

5 One can do it using MatLab standard function kstest2.

where $\widehat{F_n}$ is lcm of F_n and $\widehat{\sigma_n}$consistent estimator of σ_n6. (Durot, 2003) proves that under H_0 S_n converges in distribution to $Z = \|\widehat{W} - W\|$, where W is standard Brownian motion, $\widehat{W}$ its lcm and $\|\cdot\|-$ supremum distance.



The first sample: 36 - 39 days

The second sample: 43 - 46 days

The third sample: 54, 57 - 59 days

The fourth sample: 71 - 74 days

Figure 4: SPD per unit probability over time. It is averaged over close maturities.

6 We use the one provided in Durot paper: $\hat{\sigma}_n^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (y_i - y_{i+1})^2$.

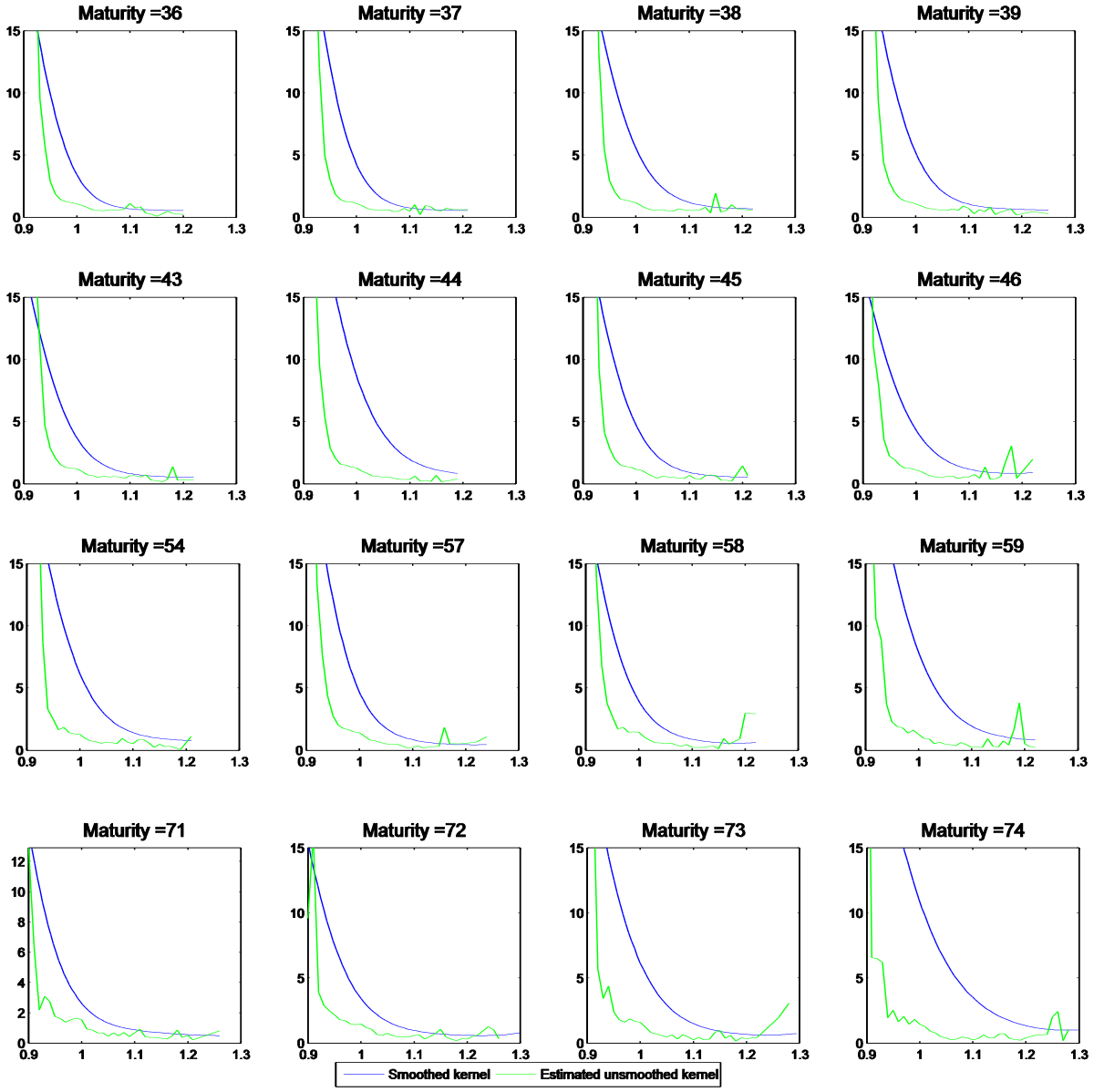
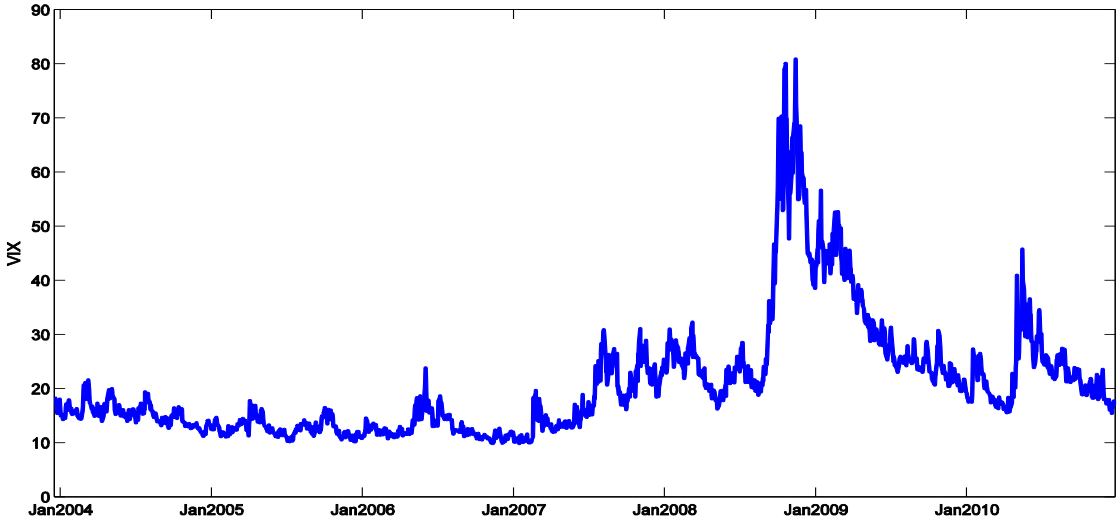


Figure 5: SPD per unit probability for different maturities. Maturities are written above each figure.

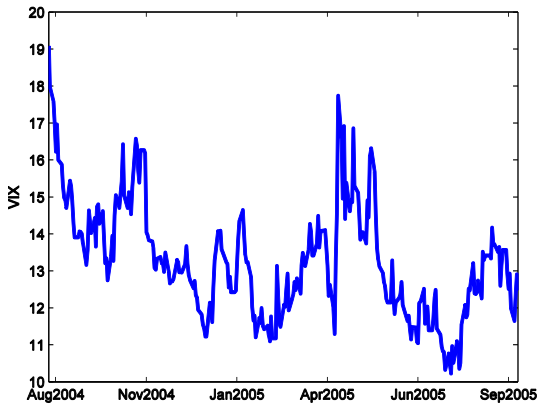
In Table 2 results of described testing procedure are presented on the right. We can see that all p-values support our hypothesis, namely we cannot reject H_0 for 5% confidence level (except for the price kernel obtained from 73 days maturity options where p-value is 0.0493. Even for the 10% confidence level most of the samples would not contradict monotonicity of pricing kernels.

Averaging price kernel over time

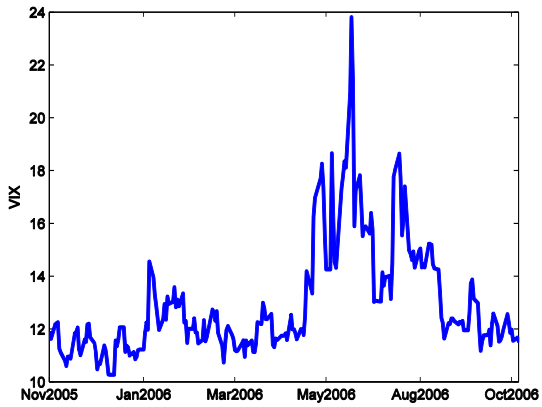
In this section we check the robustness of our methodology and we try to find smoothness criteria for smoothing our price kernels. First of all, we show different price kernels with maturities close to those we showed before. According to economic theory, price kernels with close maturities should have similar shape. Different price kernels with close maturity should be similar to each other. In order to check this, we create four samples: the



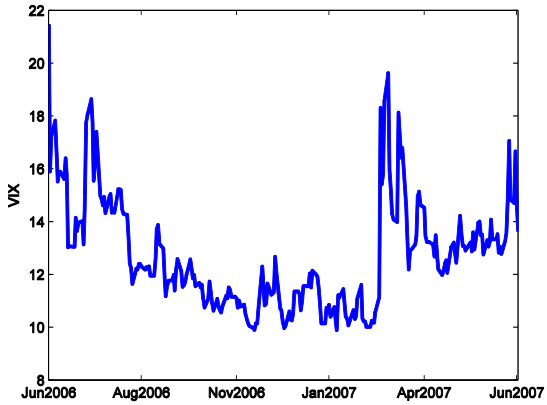
The VIX index between 2nd of January 2004 and 29st of December 2010



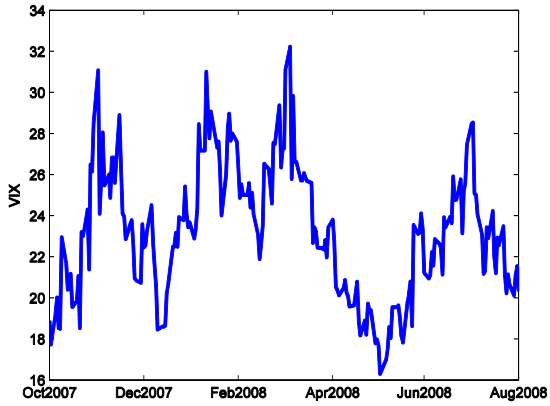
The sample: Aug 12, 2004 to Sep 15, 2005



The sample: Nov 10, 2005 to Oct 10, 2006



The sample: Jun 14, 2006 to Jun 14, 2007



The sample: Oct 11, 2007 to Aug 14, 2008

Figure 6: The four samples we use to compute the different SPD per unit probabilities over different years.

first one has maturities ranging 36 - 39 days, the second − 43 - 46 days, the third − 54, 57-59 days7, and the last one consists of 71 - 74 days maturities.

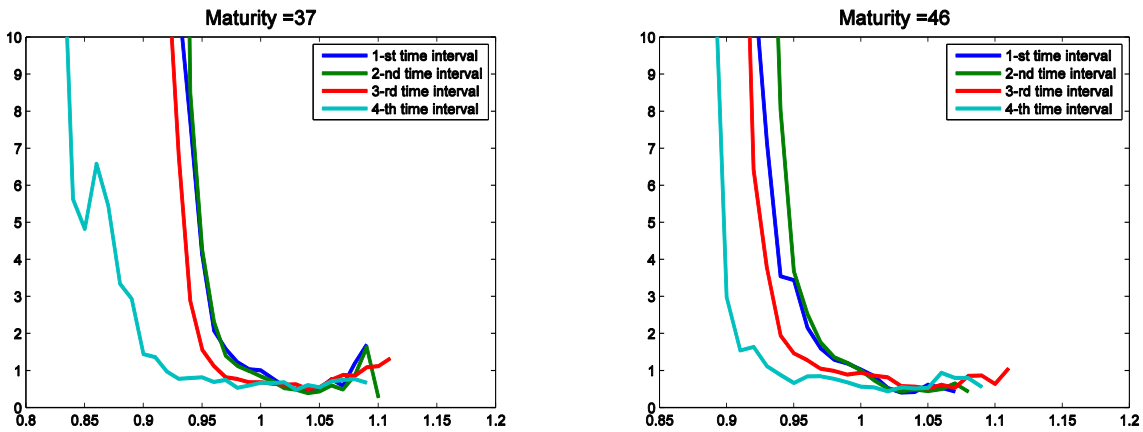
We use the approach explained in previous section. By this method we derive the price kernels for the maturities in all four samples and in Figure 5 we plot the results of our estimations. The unsmoothed kernel prices show a clearly monotonically decreasing path, except in some points that may be due to the discretization of the data. Our smoothed pricing kernels are all smoothly decreasing. In order to verify that the price kernels are monotonous over time, we plot the kernel price as the average of different maturities (see figure 4). In particular, referring to our four samples (the first one is for maturities 36 – 39 days, the second − 43 - 46 days, the third − 54, 57 - 59, and the last − 71 - 74 days), we take the average over the 4 different maturities. We expect to find a kernel price that is monotonically decreasing in wealth, because of the fact that we average over close maturities in our sample.

As we see in figure 5, the kernel prices close in maturity, have similar path, supporting the robustness of our methodology. Test statistics for monotonicity of these price kernels are presented in table 2.

In figure 4, we plot the average for each sample. We find decreasing kernel prices for smoothed estimation, and mainly decreasing kernel prices for unsmoothed estimation, although 50th and 70th samples have some jumps. In this way we were able to have some sort of smoothing criteria without using a method which biases our findings.

Price Kernel around a crisis

In this section we evaluate the change of kernel price during the crisis. In particular, we look



at kernel prices before and during the recent financial crisis. We divide our sample in 4 periods. Every period is from 9 to 12 months and we take periods which show a similar range in volatility according to the VIX index (see Figure 6).

7 There are no options for 55 and 56 days to maturity in OptionMetrix.

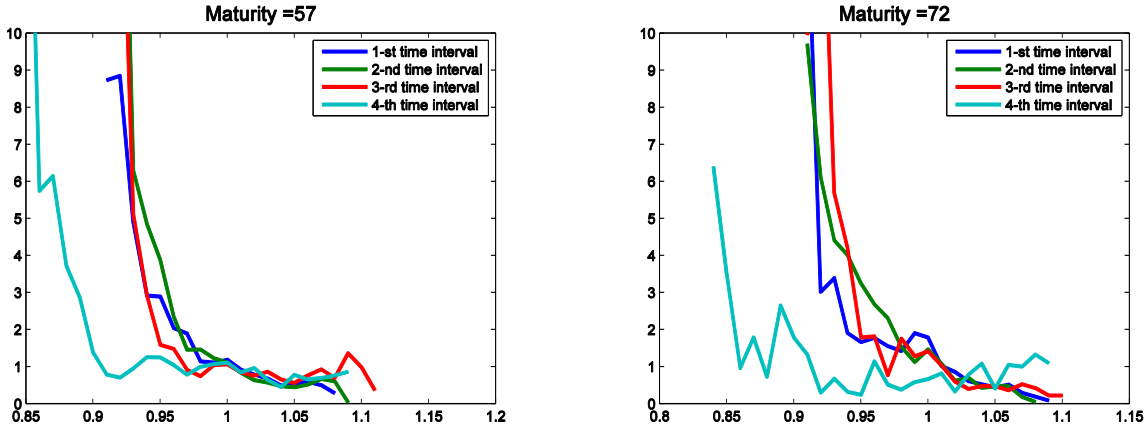


Figure 7: The kernel prices for four samples we create looking at different levels of volatility index.

Estimates of pricing kernels in different periods

We identify four different periods between August 2004 and August 2008. The first period goes from of August 2004 to the 15th of September 2005. In this period volatility is between 10 and 20 points. The second period is between 10 November 2005 to 10 October 2006. In this second period the volatility is again in a fixed range between 10 to 20 points. The third period, which is before the crisis period, is between 14 June 2006 and 14 June 2007. Even here the volatility is in a range of 10 to 20 points. The last period, the period of the beginning of the crisis is between 11 October 2007 and 14 August 2008. In this period the volatility is much higher and it is in a range between 10 and 30 points.

For each period, we compute the price kernel by the methodology presented above. We fix a maturity (in this case we look at maturities of 37, 46, 57 and 72 days) and we plot the kernel price of each period.

As expected, for the three periods before the crisis we get price kernels monotonically decreasing and very similar in shape one each other. For the kernel price of the crisis period, we have a different shape. It is higher for moneyness smaller than 1 and constant for moneyness larger than 1. For the value smaller then 1, this is exactly what we expected to obtain. The probability of negative outcome is higher therefore we give more weight of negative outcomes. On the other hand, we do not expect to have a constant kernel price for moneyness values larger than 1.

In the next section, we focus only on the kernel price of the crisis period.

Kernel Price in Crisis Time

In the previous subsection we show kernel prices for different periods (see Fig 7). We see in figure 7 that the kernel prices in period where the volatility is stable (it remains in a determinate range of 10 to 20 points) the SPD per unit probabilities exhibit a monotonically decreasing shape.

When we enter in a period of the crisis, volatility changes dramatically. In this case, we observe a kernel price that is no more monotonically decreasing, but decreasing on the left, with constant value after the moneyness equal to 1.

However, it is interesting to notice that the method we use to derive kernel prices is sufficiently robust to guarantee that even in a crisis we get kernel prices in agreement (in part of the graph) with economic theory.

Kernel price as a function of volatility

In this section we would like to extend our model and consider the kernel price as a function of more variables. In fact, as explained in (Chabi-Yo, Garcia, & Renault , 2005), one possible explanation for the non-monotonicity of the price kernel is volatility. In a previous section we compute the price kernel as a function of one variable: the underlying, $m_{t,T}(S_T)$. We know from (Pliska, 1986), (Karatzas, Lehoczky, & Shreve, 1987), and (Cox & Huang, 1989) that the kernel price is characterized by at least two factors: the risk-free rate and the market price of risk. In our analysis we would like to consider the kernel price as a function of three different factors: the risk-free rate, the underlying price and the volatility. We have already introduced the underlying price and the risk free-rate. Now we want to introduce also the volatility, so $m_{t,T}(S_T, r_f, \sigma)$.

We saw that the risk-free rate is a parameter that does not enter in our analysis for as a decisive factor. Both probabilities are forward looking. In fact, the historical probability is seen as the probability to exercise a butterfly spread at maturity, while the risk-neutral is seen as the probability of a particular state.

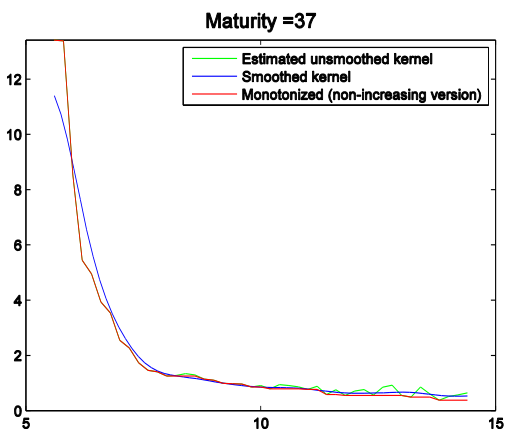
The moneyness is nothing else than the K/S_t. In order to introduce the volatility we take as a reference the idea by (Carr & Wu, 2003). They use a moneyness defined as:

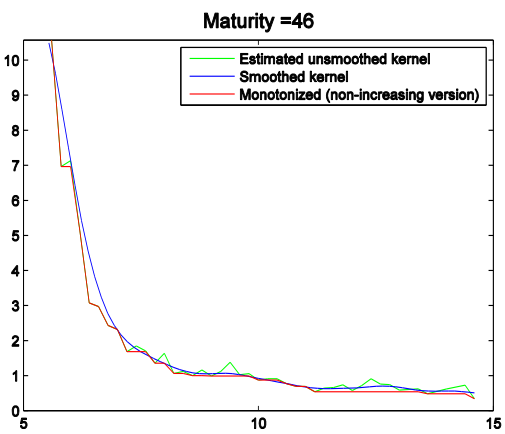
$$\text{moneyness} = \frac{\log(K/F)}{\sigma\sqrt{T}},$$

where F is the futures contract price, T is the maturity time and σ is the average volatility of the index.

For our propose, we can change this formula to look:

$$\text{moneyness} = \frac{K}{S_t \cdot \sigma},$$





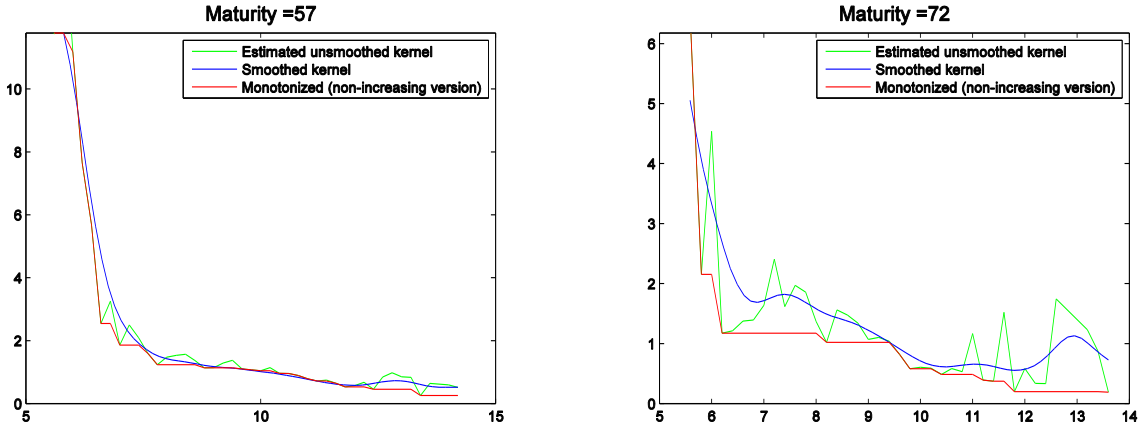


Figure 8: The Kernel prices when we use moneyness factor that take into consideration underlying price and volatility.

In fact the futures price is already considered for the above explanation of the forward looking probabilities, while the time to maturity is constant over the sample we consider. In our case the volatility is not anymore the average volatility, but the implied volatility of each option.

The procedure to derive the kernel price is again the same we have seen in the previous sections8 and therefore our result for maturities equal to 37, 46, 57 and 72 are as at the figure 8.

At figure 8 we plotted pricing kernels with volatility accounted in the moneyness parameter.

8 There is only a small difference when we round the new moneyness in order to average different periods. We do not take the second digit after the point, but we arrive only at the first one.

Table 3: The pricing kernel monotonicity testing in case of volatility being additional parameter for moneyness. Table provides p-values for monotonicity tests.

Maturity	Intuitive	Durot	Maturity	Intuitive	Durot
36	0.7651	1.0000	54	0.1489	1.0000
37	0.2951	1.0000	57	0.4075	0.5313
38	0.2823	0.9401	58	0.5480	0.9037
39	0.0569	1.0000	59	0.2436	0.0058
43	0.1078	0.9715	71	0.0042	0.0000
44	0.0372	0.5398	72	0.0042	0.0000
45	0.2823	0.9363	73	0.6403	0.1268
46	0.0569	0.8069	74	0.0000	0.0446

In this case the results are consistent with the economic theory. In table 3 we give p-values for all maturities we have seen above. One can see that these p-values confirm our graph, namely for smooth pricing kernels p-values are high in both test. Only for maturities equal to 71, 72 p-values suggest that pricing kernels are not monotone which one can also notice at the graph.

Conclusion

We propose a method to evaluate the kernel price in a specific day for a fixed maturity as well as the average of different kernel prices in a time series of 12 years for a fixed maturity. Using option prices on the S&P 500, we derive the risk-neutral distribution through the well-known result in (Breeden & Litzenberger, 1978).

We compute the risk neutral distribution in each day where we have options with a fixed maturity. Then, we compute the historical density, for the same maturity, in each day, using a GARCH method, based on the filter historical simulation technique. We then compute the ratio between the two probabilities in order to derive the kernel price for that given day. We show that in a fixed day (chosen at random in our sample) the risk-neutral distribution implied in the option prices satisfies the no-arbitrage condition.

We provide a smoothed version of the pricing kernel, to test its monotonicity. Our tests support the pricing kernel monotonicity. Therefore, we show that the ratio between the two probabilities, is monotonically decreasing in agreement with economic theory (see figure 1). We then show how the average of the different kernel prices across 12 years display the same monotonically decreasing path (see figure 3 and p-values in table 2).

We also prove that average price kernels over time, if we take close maturities, exhibit a monotonically decreasing path in agreement with economic theory.

Furthermore, we try to explain the reason why with different methods it could be possible to incur into the "pricing kernel puzzle" and have a different shapes for the kernel price. We can conclude that in most cases the model used to estimate the kernel price or the sample taken into consideration could introduce some errors in the estimation of the kernel price.

In the last part, we show the changing in shape of different price kernels before and during the recent crisis. We see that before the crisis the price kernels are monotonically decreasing while during the crisis it becomes decreasing in a part and then constant for moneyness value higher than 1. We understand this result in a very simple way: the risk neutral probability changes faster with respect to the historical one and therefore the ratio between the two remain constant.

References

Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial assets prices. *Journal of Finance, 53*(2), 499-548.

Aït-Sahalia, Y., & Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics, 94*, 9-51.

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *The Review of Economic Studies*, 91-96.

Bahra, B. (1997). Implied risk-neutral probability density functions from options prices: Theory and application. *Working Paper*.

Barone-Adesi, G., Bourgoin, F., & Giannopoulos, K. (1998). Don't look back. *Risk, 11*, 100-103.

Barone-Adesi, G., Engle, R. F., & Mancini, L. (2008). A garch option pricing model with Filtered Historical Simulation. *Review of Financial Studies, 21*(3), 1223-1258.

Birke, M., & Pilz, K. F. (2009). Nonparametric Option Pricing with No-Arbitrage Constraints. *Journal of Financial Econometrics, 7*(2), 53-76.

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business, 51*(4), 621-651.

Brown, D. P., & Gibbons, M. R. (1985). A Simple Econometric Approach for Utility-Based Asset Pricing Models. *The Journal of Finance, 40*(2), 359-381.

Campbell, J. Y., Lo, A. W., & MacKinlay, C. A. (1997). *The Econometrics of Financial Markets.* Princeton University Press.

Carr, P., & Wu, L. (2003). The FiniteMoment Log Stable Process and Option Pricing. *Journal of Finance, 58*(2), 753-777.

Chabi-Yo, F., Garcia, R., & Renault , E. M. (2005). State Dependence in Fundamentals and Preferences Explains Risk-Aversion Puzzle. *Working Paper*, 33.

Cox, J. C., & Huang, C.-f. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory, 49*(1), 33-83.

Detlefsen, K., Härdle, W., & Moro, R. (2007). Empirical Pricing Kernels and Investor Preferences. *Working Paper*, 33.

Durot, C. (2003). A kolmogorov-type test for monotonicity of regression. *Statistics & Probability Letters, 63*, 425-433.

Figlewski, S. (2008). The risk neutral probability distribution for the US stock market. *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*.

Ghysels , E., Harvey , A. C., & Renault , E. M. (1996). Stochastic volatility. In G. S. Maddala, & C. R. Rao, *Handbook of Statistics 14: Statistical Methods in Finance* (pp. 119-192). Butterworth Heinemann.

Glosten, L. E., Jagannathan, R., & Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance, 48*(5), 1779-1801.

Jackwerth, J. C. (2000). Recovering risk aversion from option prices and realized returns. *Review of Financial Studies, 13*(2), 433-451.

Jackwerth, J. C. (2004). *Option-implied risk-neutral distributions and risk aversion.* Charlotteville: Research Foundation of AIMR.

Karatzas, I., Lehoczky, J. P., & Shreve, S. E. (1987). Optimal Portfolio and Consumption Decisions for a "Small Investor" on a Finite Horizon. *Journal of Control and Optimization, 25*(6), 1557-1586.

Pirknert, C. D., Weigend, A. S., & Zimmermann, H. (1999). Extracting risk-neutral density from option prices using mixture binomial trees. *Proc. IEEE/IAFE 1999 Conf. Comput Intell. Financial Eng.*, 135-158.
Pliska, S. R. (1986). A Stochastic Calculus Model of Continuous Trading: Optimal Portfolios. *Mathematics of Operations Research, 11*, 371-382.
Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica, 32*(1/2), 122-136.
Rosenberg, J. V., & Engle, R. E. (2002). Empirical pricing kernels. *Journal of Financial Economics, 64*(3), 341–372.

Appendix A. Derivation of risk-neutral density

We start from a portfolio with two short call options with strike K and long two call with strikes $K-\epsilon$ and $K+\epsilon$ and consider $1/2\epsilon$ shares of this portfolio. The result is a butterfly spread which pays nothing outside the interval $[K-\epsilon, K+\epsilon]$. Letting ϵ tend to zero, the payoff function of the butterfly tends to a Dirac delta function with mass at K9, i.e. this is nothing else than an Arrow-Debreu security paying \$1 if $S_T = K$ and nothing otherwise (see (Arrow, 1964)). In this case, define K as the strike price, S_t the value of the underlying today, r as the interest rate, and T as the maturity time, the butterfly price is given by

$$P_{butterfly} = \frac{1}{2\epsilon}[2C(S_t, K, T, r) - C(S_t, K-\epsilon, T, r) - C(S_t, K+\epsilon, T, r)]$$

taking limit of this expression as $\epsilon \to 0$ we get

$$\lim_{\epsilon \to 0} P_{butterfly}(S_T) = \frac{\partial^2 C(S_T, K, T)}{\partial K^2}. \tag{1}$$

Now substitute the butterfly payoff, $x_{\text{butterfly}}(S_T) = \mathbb{I}_{[K-\epsilon, K+\epsilon]}(S_T)$, into equation (3) we get that the price of the butterfly is:

$$P_{butterfly} = e^{-rT} \int_{K-\epsilon}^{K+\epsilon} q_t(S_T) dS_T.$$

If we take limit as $\epsilon \to 0$ and calculate this integral using properties of Dirac delta function, we get that

$$\lim_{\epsilon \to 0} P_{butterfly} = e^{-rT} q_t(S_T)|_{S_T = K}. \tag{2}$$

Rearranging equations (10) and (11) we can have that

$$\lim_{\epsilon \to 0} P_{butterfly} = e^{-rT} q_t(S_T) = \frac{\partial^2 C(S_T, K, T)}{\partial K^2}|_{S_T = K}. \tag{3}$$

This result suggests that the second derivative of a call price (we will see that it is also true for a put price) with respect to the strike price gives the risk neutral distribution10

$$q_t(S_T) = e^{rT} \frac{\partial^2 C(S_t, K, T)}{\partial K^2}|_{S_T = K}.$$

9 More formally, payoff of the butterfly is $x_{\text{butterfly}}(S_T) = \mathbb{I}_{[K-\epsilon, K+\epsilon]}(S_T)$, or when $\epsilon \to 0$ it is $\lim_{\epsilon \to 0} x_{\text{butterfly}}(S_T) = \delta_K(S_T)$.

10 This can also be obtained by differentiating $C(S_t, K, T) = \int_K^\infty e^{-rT}(S_T - K) f(S_T) dS_T$ w.r.t. S_T as in (Birke & Pilz, 2009)

In the next part of this section we see how to apply this result in the discrete case.
In the following, we consider three call options with strikes K_i, K_{i-1}, K_{i+1}, where $K_{i+1} > K_i > K_{i-1}$. We have seen that the price of a call option can be written as:

$$C(S_t, K, T) = \int_K^\infty e^{-rT}(S_T - K)f(S_T)dS_T$$

We define $F(x)$ as the cumulative distribution function, $\mathrm{f}(x)$ as the probability density, $C(S_t, K, T)$ as the price of a European call option, $P(S_t, K, T)$ as the price of a European put option, and K as the strike price of the reference option. According to the result in (Breeden & Litzenberger, 1978) taking the first derivative with respect to the strike price, we get:

$$\begin{aligned}\frac{\partial\, C(S_t, K, T)}{\partial K} &= \frac{\partial}{\partial K}\left[\int_K^\infty e^{-rT}(S_T - K)f(S_T)dS_T\right] \\ &= e^{-rT}\left[-(K-K)f(K) + \int_K^\infty -f(S_T)dS_T\right] = e^{-rT}\int_K^\infty -f(S_T)dS_T \\ &= -e^{-rT}(1 - F(K))\end{aligned}$$

Solving for $F(x)$ one gets:

$$F(K) = e^{rT}\frac{\partial C(S_t, K, T)}{\partial K} + 1$$

Now, taking the second derivative, we have equation (5).