Is the Price Kernel Monotone?

Giovanni Barone-Adesi          Hakim Dallo

First version: April 2010
Current version: April 2010

This research has been carried out within the NCCR FINRISK project on
“Interest Rate and Volatility Risk”
Is the Price Kernel Monotone?

April, 2010

Abstract

We provide a new method to derive the state price density per unit probability based on option prices and GARCH model. We derive the risk neutral distribution using the result in Breeden and Litzenberger (1978) and the historical density adapting the GARCH model of Barone-Adesi, Engle, and Mancini (2008). We find that using a large dataset and introducing non-Gaussian innovations, the pricing kernel puzzle that arises in Jackwerth (2000) disappears both in a single day and over an average of different days with options expiring at the same maturity.

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

JEL Classification: G12, G13, G14.
According to economic theory, the shape of the kernel price or the state price
density (SPD) per unit probability (also known as the asset pricing kernel (Rosen-
berg and Engle (2002)) or stochastic discount factor (SDF) (Campbell, Lo, and
MacKinlay (1997))) is a decreasing function in wealth.

In his paper, Jackwerth (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 ex-
and Detlefsen et al. (2007) are among the most interesting papers on this topic.
Unfortunately, none of them found an answer to this puzzle. In all of their papers
they found problems in the methodology previous papers presented 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 et al. (2008), where they
argue that the main problem in this puzzle is the regime shifts in fundamentals:
when volatility changes, 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, and Mancini (2008) compute again
the kernel price and they find kernel prices consistent with economic theory. In
particular they find kernel price consistency for fixed maturities. They do not
average kernel estimates over different maturities as Aït-Sahalia and Lo (1998)
and therefore they avoid the problem that arises when the maturity is different,
but they do not consider the change in volatility as a relevant aspect of their
classification. Nevertheless they find monotone kernels, possibly because their
sample is very short in time (3 years) and in that period (2002 - 2004) the volatility
does not exhibit big changes.

In this paper we use market option prices to describe the shape of the kernel price. We compute the kernel price both in a single day and as an average of kernel prices over a period of time (considering a fixed maturity). 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 only two parameters (we do not take into consideration the changing regime) and then as a function of underlying asset price, interest rate and volatility. In order to evaluate the kernel price we consider a broad index which attempts to cover the entire economy. We use index options on the S&P500 with a time series of 12 years (from the 2nd of January 1996 to the 31st December 2007).

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

In order to estimate the risk neutral distribution, we use the well known result in Breeden and Litzenberger (1978). An important difference with previous works is in the option prices we use. Instead of creating option prices through nonparametric or parametric models (all the previous research use artificial price of options that may 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 Filter Historical Simulation as in Barone-Adesi, Engle, and Mancini (2008).

As discussed in Barone-Adesi et al. (2008), among the several GARCH models, the GJR GARCH with FHS has the flexibility to capture the leverage effect and
has the ability to fit daily S&P500 index returns best. Then, the 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. These features avoid several problems in the estimation of kernel prices. For example, using a simple GARCH model where the innovations are standard normal \((0, 1)\) leads to a mis specification of the return of the underlying that affects kernel estimates.

Once we have the two probabilities, we take the ratio between the two densities, discounted by the risk-free rate, in a particular day, to estimate the kernel price for a fixed maturity. We repeat the same procedure for all the days in the time series which have options with a given maturity and then we take the average of the kernel price through the sample.

We also evaluate how the shape of the price kernel changes before and during a crisis (the 2008 crisis). We notice that the three years before the crisis (2005, 2006 and 2007) exhibit fairly monotonically decreasing paths. During the crisis, the kernel price remains monotonically decreasing, but shows higher values (consistently with the idea that during a crisis investors risk aversion increases).

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. Our results are quite similar to the ones we obtained previously, with some improvements, that supports our first intuition that the changing volatility regime is important.

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 and Litzenberger (1978) for a discrete case 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 section 4, we provide further evidence of our results. First we plot kernel prices 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 SPD per unit probabilities for adjacent 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 3 parameters (underlying, volatility and risk-free rate), and in section 7 we offer some conclusions.

I. Review of the Literature

In this section we review the price kernel as discussed in microeconomic theory and also as in probability theory literature.

A. Price kernel and investor preferences

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 recall some basic concepts from macroeconomic theory. In particular, we use a representative agent with a utility function $U(c)$. 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 $U(c)$ 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 wealth e.g. more wealth is preferred to less wealth, 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 towards the risk.

If the investor is risk averse, his utility function is an increasing concave utility function which displays a strictly negative second derivative. The risk neutral investor has a second derivative equal to zero while the risk seeker has a second derivative strictly positive, which means a convex utility function.

Defining \( u(\cdot) \) as the single period utility function and \( \beta \) 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 \( \xi \) 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:

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

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 all of his endowment and
the asset’s payoff is the last period. Substituting the constrains into the objective and setting the derivative with respect to $\xi$ 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

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

as the Marginal rate of Substitution at time $t$. The MRS is also known as the Stochastic Discount Factor (SDF) or Price Kernel. Therefore the price of any asset can be expressed as

$$P_t = E_t [m_{t,t+1} x_{t+1}]. \quad (1)$$

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

$$P^p_t = \int_0^\infty m_{t,T}(S_T)x_T(S_T)p_{t,T}(S_T)dS_T, \quad (2)$$

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

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

$$P^n_t = e^{-rt} \int_0^\infty x_T(S_T)q_{t,T}(S_T)dS_T, \quad (3)$$

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:

\[ m_{t,T}(S_T) = e^{-r_t q_t(S_T)} \frac{p_t(S_T)}{p_t(S_T)} \]  \hspace{1cm} (4)

In this case we consider a two period model where the price kernel is a function only of the underlying and of the risk-free rate.

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. We can define the agent’s coefficient of relative risk aversion (RRA) as:

\[ \rho_t(S_T) = - \frac{S_T u''(S_T)}{u'(S_T)} \]. \hspace{1cm} (5)

A decreasing relative risk aversion indicates that the percentage of wealth exposed to risk increases with wealth. Constant relative risk aversion implies that the percentage of wealth exposed to risk remains unchanged as wealth increases or decreases. Increasing relative risk aversion means that the percentage of wealth one is willing to expose to risk falls as wealth increases.

The absolute risk aversion is the absolute amount of wealth an individual is willing to expose to risk as a function of changes in wealth. The absolute risk aversion can be decreasing, constant or increasing in wealth. Decreasing absolute risk aversion means that the amount of wealth someone is willing to expose to risk increases as wealth increases. Constant absolute risk aversion means that the amount of wealth exposed to risk remains unchanged as wealth increases or decreases. Increasing absolute risk aversion means that one’s tolerance for absolute risk exposure falls as wealth increases.

The pricing kernel can be written as function of the marginal utility as in
Aït-Sahalia and Lo (2000):

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

(6)

and the first derivative is:

\[ m'_{t,T}(S_T) = \beta \frac{u''(S_T)}{u'(S_t)}, \]  

(7)

Using the first and the second derivatives of the utility function from equation 6 and equation 7 we can write the RRA as:

\[ \rho_t(S_T) = - \frac{S_T \beta m'_{t,T}(S_T) u'(S_t)}{\beta m_{t,T}(S_T) u'(S_t)} = - \frac{S_T m'_{t,T}(S_T)}{m_{t,T}(S_T)}. \]  

(8)

Using the definition of MRS, we can write the RRA as:

\[ \rho_t(S_T) = - S_T \left[ \beta \frac{q_t(S_T)}{p_t(S_T)} \right]' \left[ \beta \frac{q_t(S_T)}{p_t(S_T)} \right], \]  

(9)

\[ = - S_T \frac{[q_t(S_T)p_t(S_T) - q_t(S_T)p'_t(S_T)]}{[q_t(S_T)/p_t(S_T)]}, \]  

(10)

\[ = S_T \left[ \frac{p'_t(S_T)}{p_t(S_T)} - \frac{q'_t(S_T)}{q_t(S_T)} \right]. \]  

(11)

B. Nonparametric and parametric estimation

There are several parametric and nonparametric methods to derive kernel prices. In this section we give a review of the most well-known method used in literature. We focus particularly on the nonparametric model because it does not assume any particular form for the risk neutral and historical density or kernel prices.

One of the first papers to recover the price kernel in a nonparametric way is
Aït-Sahalia and Lo (1998). In their work they derive the option price function by nonparametric kernel regression and then, applying the result in Breeden and Litzenberger (1978), they compute the risk neutral distribution. Their result is not consistent with classical economic theory, but, because they look at $m_{t,T}$ across time, we may understand their results as estimates of the average kernel price over the sample period, rather than as conditional estimates.

Additional problems in their article are discussed in Rosenberg and 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 price kernel. Also their use of a very short period of time, 4 years, to estimate the state probabilities may be problematic. Moreover, they depart from the literature on stochastic volatility, which suggests that future state probabilities depend more on recent events than past events.

A work close in spirit to Aït-Sahalia and Lo (1998) is Jackwerth (2000). His article is one of the most interesting pertaining to this literature. Beyond the estimation technique he used, his paper 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 for the period after the crash. He concludes that the reason is the mispricing of options after the crash.

Both articles incur in some problems that cause the kernel price and the relative risk aversion function to be not consistent with economic theory. In Aït-Sahalia and Lo (1998), we see that, if the bandwidth changes, the RRA changes as well and this means that the methodology used influences the shape of the RRA; on the other hand, in Jackwerth (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, and Mancini (2008), where they relax the normality assumption in Rosenberg and Engle (2002) and provide a nonparametric estimation of the ratio \( q_{t,T}/p_{t,T} \). While in the papers by Aït-Sahalia and Lo (2000) and Jackwerth (2000) results are in contrast with the economic theory, Barone-Adesi, Engle, and 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 (2004) provides a general review on this topic, but for the purpose of our work we do not go into detail on parametric estimation. As pointed out by Birke and 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 methods may introduce systematic errors.

Our goal is to test whether a different nonparametric method, starting with option pricing observed in the market, respects the conditions of no-arbitrage present in Birke and 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 a positive function. These conditions should guarantee a kernel price monotonically decreasing in wealth.

It is important to underlying that our kernel price is a function of three variables: the underlying price, the risk-free rate and the volatility. In the first part, we use only the first two of them: the underlying asset and the risk-free rate. At the end of the article we extend our methodology introducing volatility.
II. Empirical kernel price

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

As the risk-free we use the Unsmoothed Fama-Bliss zero-coupon rate. The methodology followed for the estimation of these rates has been described in Bliss (1997).

A. Risk-neutral density

Breeden and Litzenberger (1978) show how to derive the risk-neutral density from a set of call options with fixed maturity. They start from a portfolio with two short call options with strike $K$ and two long call with strikes $K - \epsilon$ and $K + \epsilon$ and they consider $\frac{1}{2\epsilon}$ shares of this portfolio. The result is a butterfly spread which pays nothing outside the interval $[K - \epsilon; K + \epsilon]$. Letting $\epsilon$ tend to zero, the payoff function of the butterfly tends to a Dirac delta function with mass at $K$, i.e. this is simply 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 $\tau$ as the maturity time, the butterfly price is given by

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

(12)

for $\epsilon \to 0$ and substituting equation 8 in equation 3 we find that the price of the
butterfly is:

\[
\lim_{\varepsilon \to 0} P_{\text{butterfly}} = \lim_{\varepsilon \to 0} e^{-rT} \int_{k+\varepsilon}^{k} x_{\text{butterfly}}(S_T)q_t(S_T)dS_T.
\] (13)

In the limit the right side becomes:

\[
\lim_{\varepsilon \to 0} P_{\text{butterfly}} = e^{-rT}q_t(S_T).
\] (14)

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:

\[
C(S_t, K, T) = \int_{K}^{\infty} e^{-rT}(S_T - K)f(S_T)dS_T.
\] (15)

We define \(F(x)\) as the cumulative distribution function, \(f(x)\) 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 and Litzenberger (1978) the first derivative with respect to the strike price is:

\[
\frac{\partial C(S_t, K, T)}{\partial K} = \frac{\partial}{\partial K} \left[ \int_{K}^{\infty} e^{-rT}(S_T - K)f(S_T)dS_T \right] =
\]

\[
e^{-rT} \int_{K}^{\infty} - f(S_T)dS_T = -e^{-rT}[1 - F(K)]
\]

or, rearranging

\[
F(K) = e^{rT} \frac{\partial C(S_t, K, T)}{\partial K} + 1
\] (16)

Where \(F(K)\) is the cumulative distribution function of the underlying asset price.
In order to find $F(x)$ in the discrete case, we can use the following approximation:

$$F(K_i) \approx e^{rT} \left[ \frac{C_{i+1}(S_t, K, T) - C_{i-1}(S_t, K, T)}{K_{i+1} - K_{i-1}} \right] + 1$$  \hspace{1cm} (17)$$

Now, taking the second derivative, we have:

$$f(K) = e^{rT} \frac{\partial^2 C(S_t, K, T)}{\partial K^2} \bigg|_{S_T=K}.$$  \hspace{1cm} (18)

We can approximate this result in the discrete case as:

$$f(K_i) \approx e^{rT} \frac{C_{i+1}(S_t, K, T) - 2C_i(S_t, K, T) + C_{i-1}(S_t, K, T)}{(K_{i+1} - K_i)^2} \bigg|_{S_T=K}.$$  \hspace{1cm} (19)

The same is also true for put options. In that case the two approximate distributions (cumulative and density) are given as:

$$F(K_i) \approx e^{rT} \left[ \frac{P_{i+1}(S_t, K, T) - P_{i-1}(S_t, K, T)}{K_{i+1} - K_{i-1}} \right],$$  \hspace{1cm} (20)$$

and

$$f(K_i) \approx e^{rT} \frac{P_{i+1}(S_t, K, T) - 2P_i(S_t, K, T) + P_{i-1}(S_t, K, T)}{(K_{i+1} - K_i)^2} \bigg|_{S_T=K}.$$  \hspace{1cm} (21)$$

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 and Litzenberger (1978). We differ from him in some aspects. First, we use the market bid and ask prices to construct butterfly spreads: e.g. we use for the long position the ask price for the short position the bid price. This way we avoid negative values in the risk-neutral distribution and we respect the no-arbitrage condition described in Birke and Pilz (2009). 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 with equal maturities
from several days. This improves the precision of our result (see Figure 2). Other
similar works are discuss in Bahra (1997), Pirkner, Weigend, and Heinz (1999)
and Jackwerth (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 risk neutral density can be derived either parametrically, by
solving a least squares problem, or with a nonparametric method, using kernel
regression. In our work, using a time series of options over a sample of 12 years
and taking the average of them, we avoid the parametric or nonparametric pricing
step and therefore we rely only on market prices.

In Pirkner et al. (1999) a combination approach is used to derive the risk
neutral distribution. They combine the implied binomial tree and the mixedtured
distributions to get the approach called “Mixture Binomial Tree”. The main differ-
ence with our work is that in their work, they use American options and therefore
they could have the problem of early exercise. In our sample, we consider only
European options to be sure to have the risk neutral density for the appropriate
expiration time.

He concentrates in particular on nonparametric estimation, but he gives a general
overview also on parametric works, dividing every parametric work in a particular
class and explaining the positive and negative aspects.

In the next section we explain the estimation method we use to obtain our risk
neutral distribution.
B. 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 mitigate the potential issues associated with liquidity problems.

Option data and all the other necessary data are downloaded from Option-Metrics. We compute the risk-neutral density at two different maturities: 37 days and 72 days. Our choice of maturities is random and the same procedure can be applied for other maturities. We download all the options from our dataset with the same maturities (e.g. 37 days and 72 days) and we discard the options with implied volatilities larger than 70%, average price lower than 0.05 or volume equal to 0. In Table I we summarize the number of options that we have for each maturity.

[Table I about here]

Next, we construct butterfly spreads using the bid-ask prices of our 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 $K_{i-1}$. We divide the dataset and we construct a butterfly spread for every day (clearly the butterfly spread is symmetric around $K_i$ and maturity time must be the same for all options). By dividing the dataset within each day we keep the maturity times constant. Furthermore, we have to ensure that there is symmetry around the middle strike. 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 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. Having a spread of 5 to 50 points does not impact our probability computation because we adjust it by the spread.

Our methodology to construct butterflies is very simple. We download the option prices, order by strike, from smallest to largest, we check if the first three options are symmetric around the second one. If so, we construct the butterfly, then we take into consideration the second to fourth values and we check again if they are symmetric around the third, repeating this process for every value. In this way, we form butterfly spreads which overlap for some options. We repeat the construction of the butterfly spread with the same maturity for each day available in our dataset.

At the end of this procedure we get a number of butterfly prices summarized in table I. In figure 1, we take a day at random from our sample and we apply equations (20) and (21) to compute the risk-neutral distribution, the historical distribution and their ratio as the SPD per unit probability. We take as an example the 11 August 2005, and we look at options with a maturity equal to 37 days. We see that for one day, the kernel price shows a fairly monotonically decreasing path in $S_T$ with some jumps. These jumps are present because we do not smooth our results.

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 digit after the decimal point and we take the average of all the butterfly prices with equal moneyness.\(^1\)

\[^1\text{In order to find an equal moneyness it is necessary to round the moneyness values, to the}^1\]
In table I we summarize the results after this procedure. At this point we are ready to plot the risk-neutral distribution as an average of the butterfly prices for a fixed maturity over a twelve year period.

C. Historical density

In order to construct the historical density we use a GARCH approach. Specifically, we use the asymmetric GJR GARCH model. As discussed in Barone-Adesi, Engle, and Mancini (2008), among the several GARCH models, the GJR GARCH has the flexibility to capture the leverage effect and has the ability to fit daily S&P500 index returns. 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 \( \epsilon_t = \sigma_t z_t, \) 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.

Our estimation of the historical density is as follows. We have butterfly prices for each day in the time series of twelve years, with equal maturity. For each day, we estimate the parameters of the GJR GARCH using a time series of 3500 returns second decimal digit. At this point we are left with a lower number of price calls that we average.
from the S&P500. Once we have the parameters for each day, we simulate 35000 paths and we look at the probability that at maturity we exercise the butterfly. e.g. we count the number of paths that at maturity are in the range $[K_{i-1}, K_{i+1}] = 1$ (the length of the interval)$^2$. To compute the probability density under the historical measure, we apply the following relation:

$$p = \frac{\# \text{ of paths in the interval } l}{\text{total number of paths}} \cdot$$

(24)

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.

[Figure 2 about here]

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 distribution in respect to the moneyness.

\section*{D. Kernel price}

We apply the definition given in equation 4 to compute the kernel price. From previous calculations we obtain the average risk-neutral distribution for each fixed maturity and also the average historical density. We then discount this ratio by the risk-free rate. To estimate 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.

We compute the average across time because of the low number of data at each

$^2$In our sample we use intervals with different lengths: most of them are intervals with a length of 10 points, but we also have some intervals with 25 or 50 points especially in the tails. These intervals are in some cases overlapping.
time.

We must keep in mind, throughout the process, that the risk-neutral distribution and the historical distribution we plotted in Figure 2, are not the same ones we used to derive the kernel price. The kernel price is the average of the kernel price of each day. The distributions are the averages of each day distribution.

[Figure 3 about here]

For all different maturities we get a monotonically decreasing path for the kernel price in accordance with classical economic theory\(^3\).

### III. Averaging price kernel over time

In this section we check the robustness of our methodology and we try to find a smoothness criteria for smoothing our price kernels. First of all, we show different price kernels with maturity close to the one we showed before. According to economic theory, adjacent maturity price kernels should have similar shape and therefore, averaging close maturities, could be a good criteria for smoothing.

In order to verify the robustness of our methodology, we create two samples: the first one has maturities equal to 36, 37, 38 and 39 days. The second one, 71, 72, 73 and 74 days.

We use the approach explained in the previous section to derive the price kernel for a fixed maturity. By this method we derive the price kernels for the maturities in the two samples and in the following figures we plot the results of the two samples.

[Figure 4 about here]

\(^3\)The presence of some jumps in our empirical analysis is expected, because we do not introduce any smoothing criteria.
The kernel prices show a clearly monotonically decreasing path, except in some points, likely due to the discretization of the data. In order to verify that the price kernels are monotonous over time, we plot the kernel price as the average of different days. In particular, referring to the two samples (the first one is for maturities equal to 36, 37, 38 and 39 days and the second sample for maturities equal to 71, 72, 73 and 74 days), we take the average over 4 different maturities. We expect to find a kernel price that is monotonically decreasing in wealth, because of the close maturities in our sample.

As we see in Figure 4, kernel prices close in maturity show similar paths. This result supports the robustness of our methodology.

In Figure 5, we plot the averages for the two samples and we find decreasing kernel prices. In this way we were able to have a sort of smoothing criteria without using methods which bias our dataset.

IV. Price Kernel around the financial crisis

In this section we evaluate the change of kernel price at the onset of the recent crisis. We divide our sample in 4 periods. Every period is between 9 to 12 months long.

A. Estimate pricing kernels in different periods

From the VIX index, we identify four different periods between August 2004 and August 2008. The first period runs from the 12th of August 2004 to the 15th of September 2005. The second period runs between the 10th of November 2005 to 10th
of October 2006. The third period, which is before the crisis period, runs between 14\textsuperscript{th} of June 2006 to 14\textsuperscript{th} of June 2007. In these periods 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\textsuperscript{th} of October 2007 and 14\textsuperscript{th} of August 2008. In this period the volatility is in a range between 10 to 30 points.

[Figure 6 about here]

[Figure 7 about here]

For each period, we compute the price kernel using the methodology presented in sections 3 and 4. We fix a maturity (in this case we look at a maturity of 37 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 to each other.

[Figure 8 about here]

An interesting result we obtain in the last period is the following. While the pricing kernel is still monotone, it reaches higher values at the money, possibly because of learning effects about volatility in the option market, that lead to a higher $q_{t,T}/p_{t,T}$ ratio in that range.

V. 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, and Renault (2008), one way to overcome the problem of non-monotonicity of the price kernel is to consider another factor: volatility.
In a previous section we compute the price kernel as a function of two variables: the underlying and the risk-free rate. We know from Pliska (1986), Karatzas et al. (1987), and Cox and 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, \( m_{t,T}(S_T, r_f, \sigma) \). Because we look at a price kernel in a two-period model, we are not interested in looking at dividends, but an extension with multi-period kernel price would be possible.

We have already introduced the underlying price in our kernel price when we use, as strike price, the moneyness. The moneyness is nothing else then the \( k/S_t \). In order to introduce the volatility we follow Carr and Wu (2003). They use a moneyness defined as:

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

where \( F \) is the futures price, \( T \) is the maturity time and \( \sigma \) is the average volatility of the index.

For our propose, we can change this formula in:

\[
moneyness = \frac{K}{S_t * \sigma},
\]

Time to maturity is constant over the sample we consider because we fix the maturity at the beginning. Therefore, we do not include the square of the time in the analysis. Furthermore, the volatility is not anymore the average volatility, but the implied volatility of each option.

Our procedure to derive the kernel price is again the same we have seen in the previous sections \(^4\) and therefore our result for maturities equal 36, 37, 38 and 39

\(^4\)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 only the first one.
as well as 71, 72, 73 and 74 are:

[Figure 9 about here]

Also in this case the result are consistent with economic theory. In particular we can see how the kernel prices for close maturities are very similar, in agreement with our previous results.

VI. Conclusion

We proposed 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 and 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, each day, using a GARCH method, based on the filter historical simulation technique. We then compute the ratio between the two probabilities discounted by the the risk-free rate, in order to derive the kernel price for that given day. We show that, for a date chosen at random in our sample, the risk-neutral distribution taken from the option prices respects the no-arbitrage condition proven in Birke and Pilz (2009).

Therefore, we find that the ratio between the two probabilities, in that particular day, is monotonically decreasing, in agreement with economic theory (see figure 1). We also show how the average of the different kernel prices across 12 years display the same monotonically decreasing path (see figure 3).

We also find that average price kernels over time, if we take closing maturities, exhibit a monotonically decreasing path in agreement with classic economic theory.
Moreover, we parametrize kernel prices with volatility. Averages across several years result more smooth and still coherent with classic economic theory.
References


Figure 1: **Kernel Price in one day.** Left: Risk-neutral distribution (thick line) and the historical distribution (thin line). We take one day at random from our sample with option maturity equal to 37 days. Right: SPD per unit probability for this particular day (11 August 2005).
Figure 2: **Non-standardize risk neutral and historical distribution.** Non-standardize risk neutral and historical distribution as the average of 12 years risk neutral and historical distribution for a fixed maturity. Left: 37 days to maturity, right 72 days to maturity.

Figure 3: **Kernel prices for two different maturities.** SPD per unit probability as the average of the SPD per unit probability throughout the time series of 12 years and with equal maturity. Left: 37 days, right 72 days. It is important to bear in mind that this SPD per unit probability is not derived from the two distributions given in figure 2.
Figure 4: **Kernel prices for different maturities.** SPD per unit probability for different maturities. Left: from top to bottom, SPD per unit probability for maturity equal to 36, 37, 38 and 39. Right: from top to bottom, SPD per unit probability for maturity equal to 71, 72, 73 and 74.
Figure 5: **SPD per unit probability over time.** Left: SPD per unit probability for the first sample (from 36 to 39 days). Right: SPD per unit probability for the second sample (from 71 to 74 days).

Figure 6: **VIX index.** The VIX index between 2nd of January 2004 and 21st of April 2009
Figure 7: **Different regime samples.** The VIX samples we use to compute the different SPD per unit probabilities over different years.
Figure 8: **Kernel prices over time.** The kernel prices for the four samples we create looking at different levels of volatility index. The solid line, dot line and the dot-dash line are the pre-crash kernel prices, while the dashed line is the kernel price in the crisis period.
Figure 9: **Kernel prices standardized by volatility.** SPD per unit probability for different maturities. Left: from top to bottom, SPD per unit probability for maturity equal to 36, 37, 38 and 39. Right: from top to bottom, SPD per unit probability for maturity equal to 71, 72, 73 and 74.
<table>
<thead>
<tr>
<th>Maturity and type of option</th>
<th>Sample</th>
<th>Number of butterfly spreads</th>
<th>Number of butterfly spreads after rounding</th>
</tr>
</thead>
<tbody>
<tr>
<td>37 days call</td>
<td>2903</td>
<td>2500</td>
<td>54</td>
</tr>
<tr>
<td>37 days put</td>
<td>3544</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72 days call</td>
<td>1497</td>
<td>1196</td>
<td>58</td>
</tr>
<tr>
<td>72 days put</td>
<td>2025</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table I: **Summary of the number of option data used.** The second column is the number of options from our dataset once the options with implied volatility larger than 70%, average price lower than 0.05 and a volume equal to 0 have been discarded. The third column is the number of butterfly spreads we can construct. The fourth column is the number of butterfly spreads after we round the moneyness to two decimal places. The first two rows refer to the call and put options with maturity equal to 37 days and the third and fourth rows with maturity equal to 72 days.