DYNAMICS OF IMPLIED DISTRIBUTIONS: EVIDENCE FROM THE CAC 40 OPTIONS MARKET

Lamya Kermiche

Presses universitaires de Grenoble | « Finance »

2009/2 Vol. 30 | pages 63 à 103

ISSN 0752-6180
ISBN 9782706115899

Revue de l'association française de finance, vol. 30, n° 2/2009

Dynamics of Implied Distributions: Evidence from the CAC 40 Options Market

Lamya KERMICHE*

1. Introduction

Researchers and market participants have shown a growing interest in implied density probabilities calculated from options prices, primarily due to their usefulness in both risk management and options valuation. Risk managers are responsible for measuring and controlling risk; this requires knowledge of the probabilities of possible future conditions in the financial markets, and more generally the probability densities associated with future returns. Due to their prospective nature, options appear to be the best instruments for measuring this probability density.

Risk-neutral densities can be used for many purposes. One direct application is the valuation of exotic options with the same maturity as a given density. For instance, one could use the prices of highly-liquid options to calculate the implied distribution of an underlying asset, and then determine the price of illiquid options on this underlying asset. Central banks also use risk-neutral densities to gauge investors' expec-

* Grenoble École de Management, 12 rue Pierre Sémard – BP 127 38003 Grenoble cedex 01 – France lamya.kermiche@grenoble-em.com

This article is part of my PhD Dissertation, and I am particularly grateful to Pascal Louvet (my supervisor) for his helpful comments. I am also grateful to the editor and referees of the review *Finance* for their contribution to the quality of this paper. I am responsible for all remaining errors and omissions. This publication has benefited from financial support from the French Finance Association (AFFI).

tations of future interest rates, share prices, and exchange rates [see for example Bahra (1997) in a paper for the Bank of England and Coutant *et al.* 2001 in a paper for Banque de France]. Finally, risk-neutral densities obtained from options prices can be used in association with subjective probability densities obtained from historical data to estimate a risk aversion function that is representative of investor preferences [see for example Aït-Sahalia and Lo (1998); Coutant 1999; Pérignon and Villa 2002; and more recently, Bliss and Panigirtzoglou 2004].

Because of these many applications, abundant literature is available on the various methods for estimating risk-neutral density from options prices [see for example Bahra (1997); Cont (1997); or Jackwerth (1999) for a summary of this literature]. This paper focuses on the method introduced by Shimko (1993) that involves interpolating or smoothing the implied volatility smile. This method is based on the observation that Black-Scholes implied volatilities are smoother than options prices. From a mathematical point of view, this observation facilitates the application of Breeden and Litzenberger's (1978) finding that the risk-neutral density is equal to the second derivative of an option's price with respect to its exercise price (after adjustment for the discount factor). Bliss and Panigirtzoglou (2002) compared several different methods for calculating risk-neutral densities, and concluded that smoothing the volatility smile is the most robust method.

Although there is extensive literature on methods for estimating implied risk-neutral densities at a given moment, few authors have looked at the evolution of this density over time. Some approaches, highly data-consuming, even require market prices over several days to ensure convergence. For example, Aït-Sahalia and Lo's method (1998) uses a full year of market prices to calculate a single risk-neutral density, therefore assuming that the risk-neutral density is constant. Data of such frequency can be used to study the general characteristics of risk-neutral densities, but are not practical for option valuations or risk management. Indeed, such data show day-to-day changes in risk-neutral densities – changes which are used to analyze major events and assess market expectations in response to such events. These events could be political [Coutant *et al.* (2001), for the dissolution of the French parliament and the 1997 legislative elections, and Gemmill and Saflekos (2000), for general elections in the UK], economic [Beber and Brandt (2006), for macroeconomic announcements], or geopolitical [Melick

and Thomas (1997), for the effect of the Gulf War on risk-neutral densities derived from crude oil options prices]. These various studies show that the options market tends to anticipate information through changes in risk-neutral density moments.

When considering these changes, it is imperative to look at the dynamics of risk-neutral densities. Lynch and Panigirtzoglou (2003) studied time series of descriptive statistics of risk-neutral densities, but it is important to consider the entire distribution over time. Skiadopoulos and Hodges (2001) proposed an alternative to existing models (Black-Scholes, stochastic volatility, local volatility, etc.) by simulating changes in implied risk-neutral densities over time, so as to prevent arbitrage opportunities. However, their approach does not account for the empirical properties of risk-neutral densities. Panigirtzoglou and Skiadopoulos (2004) were the first to study implied distribution dynamics empirically by first carrying out a Principal Component Analysis (PCA hereafter) on the cumulative distribution function, and then using a diffusion process to develop a model based on the Monte Carlo method.

In our study we used an approach similar to Panigirtzoglou and Skiadopoulos (2004). This approach is motivated by the analogy between implied distribution and implied volatility, and in light of recent developments in research on implied volatility dynamics. Cont and Da Fonseca (2002) studied changes in implied volatility surfaces and found three factors that can shift such surfaces. Fengler, Härdle, and Mammen (2007) developed a semi-parametric factorial model for changes in implied volatility; Giacomini, Härdle, and Kratschmer (2008) also used this model to estimate risk-neutral densities. Kermiche (2008) employed a method similar to that of Cont and Da Fonseca, and found a jump component in the first factor that can shift implied volatility surfaces. Alternatively, Jones (2006) linked average option returns to three fundamental risk factors (market elements, volatilities, and jumps). Although implied distributions and implied volatilities reveal the same information at any given moment (Panigirtzoglou and Skiadopoulos, 2004), these variables are not used by the same market participants. Therefore, it would be useful to be able to reflect changes in implied volatilities in risk-neutral densities, and to pinpoint the factors influencing risk-neutral densities and the way they interact with density moments.

Our research differs from previous studies in several ways. First, we carried out a PCA directly on risk-neutral densities (which to our knowledge has never been done before), using a metric that ensures that the densities are fully represented on a daily basis. Our PCA revealed two main shock factors, which we then modelled using a mean-reversion process incorporating a jump component. Second, we calculated risk-neutral densities from intraday CAC 40 option prices, a market that has been little-studied until now. Third, we applied our findings directly to options portfolio hedging.

The rest of this paper is organised as follows: Section 2 presents the theoretical framework and the various methods for calculating risk-neutral densities. Section 3 outlines the data we used in our study and the risk-neutral density calculation method, including the procedure for smoothing the implied volatility surface. Section 4 uses a PCA to analyze risk-neutral density dynamics, and discusses the important factors to be considered. Section 5 gives a model of the factors obtained from the Ornstein-Uhlenbeck process, with and without jumps, along with a comparison of these approaches. Section 6 describes how our findings can be applied to risk management, by reviewing the performance of options portfolios that have been protected against the two risk-neutral density shock factors identified in the PCA. Section 7 summarises our findings and their significance, and suggests areas for further research.

2. Risk-neutral density

2.1. The relationship between options prices and risk-neutral densities

Cox and Ross (1976) have shown that the price of a European option on an asset S is the current value of the cash flow expected from the option at expiration. For a call option with exercise price K and maturity T, this gives:

$$\begin{aligned} C_t(S_t,t;K,T) &= e^{-r(T-t)}E_t\left[max(S_T-K,0)\right] \\ &= e^{-r(T-t)}\int_K^{\infty}(S_T-K)q(S_T/S_t)dS_T \end{aligned} \tag{1}$$

Similarly, for a put option with exercise price K and maturity T, this gives:

$$\begin{aligned} P_t(S_t,t;\,K,T) &= e^{-r(T-t)}E_t\,[max(K-S_T,0)] \\ &= e^{-r(T-t)}\int_0^K (K-S_T)q(S_T/S_t)dS_T \end{aligned} \tag{2}$$

Where $q(S_T/S_t)$ is the risk-neutral probability density function of the underlying asset at time T, conditional on S_t. This density will hereafter be referred to as $q(S_T)$ for simplicity.

If we take the first derivative of Equation 1 with respect to the exercise price, we obtain:[1]

$$\frac{\partial C(K,T)}{\partial K} = -e^{-r(T-t)}\int_K^\infty q(S_T)dS_T \tag{3}$$

Or if we use $Q(.)$ to represent the cumulative distribution function of S_T, we get:

$$\frac{\partial C(K,T)}{\partial K} = -e^{-r(T-t)}\,[1-Q(K)]$$

This indicates that the first derivative with respect to the exercise price of a European call option is equal to the negative probability (discounted at the risk-free rate) that the option is in-the-money when it expires ($S_T > K$). In other words, if the option is far in-the-money ($S_t \gg K$), it will almost surely be so at expiration ($Q(K) = 0$), in which case it will most likely be exercised. If the exercise price increases by one unit, the option holder must pay one more unit to acquire the underlying asset, which reduces the option price by the current value of this unit. Inversely, if the option is far out-of-the-money ($S_t \ll K$), it will most likely not be exercised ($Q(K) = 1$); an increase in the exercise price will not affect the option price (since it would not be exercised). If the option is at-the-money, the effect of an increase in the exercise price will depend on the probability of the option eventually becoming in-the-money.

1. This example uses a call option, but the same method would apply to a put option.

If we take the second derivative with respect to the exercise price, we obtain Breeden and Litzenberger's (1978) result:

$$\frac{\partial^2 C(K,T)}{\partial K^2} = e^{-r(T-t)} q(K) \tag{4}$$

Therefore, the second derivative with respect to the exercise price of a European call option is equal to the current value of the risk-neutral probability density of S_T. This is the result we used to estimate risk-neutral probability densities from options prices.

2.2. Black-Scholes risk-neutral density

This section reviews the Black-Scholes model (1973) assumptions and the resulting risk-neutral density, and then shows how this model and the resulting density are not reflective of actual market conditions. It also discusses the importance of developing new, more accurate methods for calculating risk-neutral probability densities.

The Black-Scholes model assumes that the underlying asset follows a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{5}$$

Where μ and σ are assumed constant, and W_t is the Wiener process. Applying the Ito lemma to Equation 5 gives the following:[2]

$$\ln S_T \sim \phi \left[\ln S_t + \left(\mu - \frac{\sigma^2}{2} \right) \tau, \sigma \sqrt{\tau} \right]$$

Or:

$$\ln\left(\frac{S_T}{S_t}\right) \sim \phi \left[\left(\mu - \frac{\sigma^2}{2} \right) \tau, \sigma \sqrt{\tau} \right]$$

Where $\phi\,(\alpha,\beta)$ is a normal distribution with parameters α and β. This means that the risk-neutral density function $q(S_T)$ of S_T is a log-normal distribution with parameters α and β. The log-normal density function is:

2. *See* Bahra (1997) and Hull (2006, Chapter 13).

$$q(S_T) = \frac{1}{S_T \beta \sqrt{2\pi}} e^{-\frac{(\ln S_T - \alpha)^2}{2\beta^2}}$$

Where:

$$\alpha = \ln S_t + \left(\mu - \frac{\sigma^2}{2}\right)\tau$$

$$\beta = \sigma\sqrt{\tau}$$

The Black-Scholes model assumes that the implied volatility of an underlying asset is the same for all options of a given maturity, regardless of the exercise price. However, in reality the implied volatility does vary with the exercise price; the resulting curve is known as the volatility smile or skew depending on the shape. The volatility curve illustrates that investors do associate higher volatility with options that are far in-the-money or out-of-the-money, and that they give these conditions different probabilities than those indicated by the Black-Scholes model. Therefore, the final distribution (at maturity) is not a log-normal distribution.

In fact, the sharper the volatility smile curve (i.e., the more convex), the farther the distribution is from log-normal. High volatilities indicate that traders give high probabilities to extreme S_T values, resulting in fat distribution tails. In addition, the direction of the smile's slope is reflected in the risk-neutral density asymmetry; a positive slope indicates a shift to the right (relative to a log-normal distribution), while a negative slope indicates a shift to the left. Changes in the smile curve's shape lead to changes in the corresponding risk-neutral density (through changes in the price function shape and convexity).

2.3. Estimating risk-neutral densities

There is a great deal of literature on the different methods for estimating risk-neutral densities from options prices; this literature has been summarised by Bahra (1997), Cont (1997), and Jackwerth (1999), among others. The methods can be divided into four types: (i) those where risk-neutral densities are calculated from assumptions on

the evolution of the underlying asset (continuous using stochastic models or discontinuous using binomial trees); (ii) those where risk-neutral densities are estimated directly in line with observed prices; (iii) those where a price function is interpolated, and its second derivative is taken according to Breeden and Litzenberger's (1978) procedure; and (iv) those where an implied volatility smile curve is interpolated.

The first approach assumes that the underlying asset follows a specific stochastic process whose parameters can be determined from options prices. The process is generally chosen so as to give an analytical formula for calculating risk-neutral densities[3], which are fully defined once the distribution parameters are known. The Black-Scholes model is an example of this method, since it assumes that the underlying asset follows a geometric Brownian motion with a constant trend and volatility, which implies that the risk-neutral density is log-normal.

In the second approach, the risk-neutral density function is estimated using either a parametric or non-parametric method. The parametric approach requires that the shape of the risk-neutral density curve be selected beforehand, and the function parameters subsequently estimated so that the options prices calculated from the function are as close as possible to those on the market. The maximum entropy method is an example of a non-parametric approach which aims to find a probability density that is as close as possible (in terms of informative content) to a chosen distribution, and which correctly values options on the market.

The third approach uses the findings of Breeden and Litzenberger's (1978) work, and requires a continuum of options prices. But because such continuous price data cannot be obtained in practice, a continuous function must be interpolated from observed prices. This interpolated function must respect uniformity and convexity conditions, and be twice derivable.

The fourth approach is an alternative to an interpolated price function. Shimko (1993) noted that Black-Scholes implied volatilities are smoother than options prices, and concluded that it may be better to

3. Monte Carlo simulation methods could also be used to obtain risk-neutral densities if an analytical formula is not available.

interpolate implied volatilities rather than prices. This method uses the Black-Scholes model as a way to pass from prices to implied volatilities, without making any assumptions as to the model's validity. The risk-neutral density can then be obtained by taking the second derivative of the price function, as in the previous approach.

Bliss and Panigirtzoglou (2002) tested the stability of the implied risk-neutral density function by comparing two methods (double log-normal and implied volatility) and found that densities calculated from a smoothed smile curve are the most stable. Therefore, this is the method that we used in our research. We started by using a non-parametric approach to construct smoothed volatility surfaces, as described in the next section. Even though we needed to take the derivative only with respect to the exercise price, we decided to determine the entire surface so that we could obtain the smile curve for any maturity. This was important for our study of dynamics because it let us compare daily risk-neutral densities for a given maturity[4].

3. Constructing risk-neutral density curves

3.1. Data description

Our initial database covered all the transactions on CAC 40 index options (call and put) registered on the Paris traded options market (MONEP) between 1 April 1999 and 28 December 2001. The data included the date and time of each transaction, option features (call or put, exercise price, and maturity), option price, and the value of the underlying asset (CAC 40) at the same time. We used intraday data to avoid the imperfect synchronization problem raised by Harvey and Whaley (1991), and to have enough data points for each day to allow for convergence of the implied volatility surface smoothing method [Aït-Sahalia and Lo (1998)].

4. Lynch and Panigirtzoglou (2003) encountered this problem in their work on descriptive statistics for risk-neutral densities, since their data sample contained options with fixed maturities. To resolve this problem, they created an artificial option at a constant maturity by interpolating between the maturities in their sample. Panigirtzoglou and Skiadopoulos (2004) also faced this problem when they tried to construct a distribution function at a constant maturity, and similarly interpolated between different maturities according to the method developed by Clews *et al.* (2000).

The CAC 40 options traded on MONEP are European options (no early exercise) and include eight maturities: the first three months, the next three quarters after the first three months on a sliding basis (quarters ending March, June, September, and December), and two six-month periods (ending in March and September). The options can be exercised on the last working day of each month (market hours are from 9:00am to 5:30pm), with 50 index points separating two consecutive exercise prices.

We screened the maturity, exercise price, option price, and liquidity data to purge outliers and unreliable entries. We eliminated options with very short maturities (and therefore sensitive to minor calculation errors) and very long maturities (for which few exercise prices are available), keeping only those with maturities of between one week and one year. We eliminated options whose forward moneyness (exercise price/forward price) was outside of the range 0.5 to 1.5. Options falling outside this range are illiquid and have a great deal of uncertainty in the implied volatility calculation. We eliminated options whose price is below 0.10 index points or 0.10 euros, and we replaced in-the-money (and therefore highly illiquid) call and put options by their matching put or call, respectively, based on the put-call parity[5].

We used daily EURIBOR data from Datastream for the term structure of interest rates, constructing a curve of data points at one week, one month, two months, and so on up to one year. We then interpolated the curve (using Matlab's cubic spline interpolation) to obtain interest rates at particular points. We also used futures contract closing prices from Datastream (for all maturities) along with CAC 40 closing values to calculate the implied dividend yield q for each maturity with the formula:

$$F_t = S_t e^{(r-q)(T-t)}$$

With these dividend yields we were easily able to obtain future values at any time of day, using the corresponding CAC 40 value at that time.

5. More specifically, we got rid of in-the-money call and put options ($m < 1$ for calls and $m > 1$ for puts), and replaced out-of-the-money puts ($m < 1$) with their matching call according to the put-call parity. We kept out-of-the-money calls ($m > 1$). This gave us a sample comprised only of call options, but which retains all of the necessary implied volatility information related to out-of-the-money put options.

After screening the data, the number of observations eligible for our study fell from 478,290 to 357,100, meaning that we kept 75% of the initial data points and got rid of most of the noise from outliers and unreliable entries. These observations occurred over a period of 687 days, giving an average of 519 observations per day.

We used the Black-Scholes model on the screened data sample to calculate implied volatilities. We then filtered the data one last time based on these volatilities, since non-arbitrage constraints set upper and lower bounds for the slope of the implied volatility surface relative to the exercise price and maturity. Some smile curve and term structure shapes are not possible. However, these upper and lower bounds are difficult to determine precisely, so we removed the data points where the implied volatility is more than four standard deviations away from the mean[6]. This enabled us to eliminate outlying volatilities while minimising the chances of losing a valid data point. This removed a further 1,154 observations from our data sample.

Therefore, our final data set contained 355,946 observations taking place over 687 working days. The next section describes how we constructed a daily implied volatility surface from these data using a smoothing procedure.

3.2. Constructing a smoothed surface

We used the implied volatilities obtained by inverting the options prices in our final data set to construct a surface for each day from a grid of data points (moneyness m, time to maturity τ), incomplete and variable with time. We then smoothed this surface to be able to obtain a volatility surface for any data point (m, τ).

We decided to use the non-parametric smoothing method developed by Aït-Sahalia and Lo (1998), even though surfaces are usually inter-

6. Although we did not make any assumptions as to the shape of the implied volatility distribution curve, we decided to use four standard deviations to be consistent with a normal distribution curve in which 99.7 % of the data points are within three standard deviations of the mean (95 % within two standard deviations and 68% within one standard deviation). Our maximum volatility is 97% despite this constraint. For instance, Fengler, Härdle, and Villa (2003) eliminated options with over 80% volatility in their study of options on the DAX index.

polated or smoothed via a parametric method. In fact, there is no theoretical basis for using a parametric method; it is generally used for practical reasons only. Moreover, a parametric approach limits the possible curve shapes to the selected parametric family, which, although flexible, cannot be used to construct all curve shapes. Aït-Sahalia and Lo's method does not impose parametric restrictions on the shape of the implied volatility surface, thereby enabling us to capture all the characteristics of each day's data.

The basic concept behind Aït-Sahalia and Lo's method is that if a surface is sufficiently smooth, then the value at any given point can be calculated by taking the weighted average of all neighbouring points. The weight given to each neighbouring point decreases as the point is located further away from the target point. We used a kernel density function to model this series of weights; this function includes a parameter called the smoothing parameter that specifies the neighbouring points to be included in the weighted average.

We used the Nadaraya-Watson estimator to construct an implied volatility function:

$$\hat{\sigma}(m,\tau) = \frac{\sum_{i=1}^{n} \sigma_i g_m\left(\frac{m-m_i}{h_m}\right) g_\tau\left(\frac{\tau-\tau_i}{h_\tau}\right)}{\sum_{i=1}^{n} g_m\left(\frac{m-m_i}{h_m}\right) g_\tau\left(\frac{\tau-\tau_i}{h_\tau}\right)}$$

Where $\sigma_i = \sigma(m_i,\tau_i)$ for $i \in [1..n]$ are the points of the discrete surface and n is the number of data points for that day.

g_m and g_τ are the kernel functions representing the moneyness m (m = strike price/future price) and time to maturity τ ($\tau = T - t$). We chose the Gaussian kernel for g_m and g_τ defined by:

$$g(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

The smoothing parameter h has a great importance because it determines the degree of smoothing. If h is too big, the surface will be too smooth and important surface characteristics will be lost; if h is too small, the surface will be too irregular and incorporate too much noise.

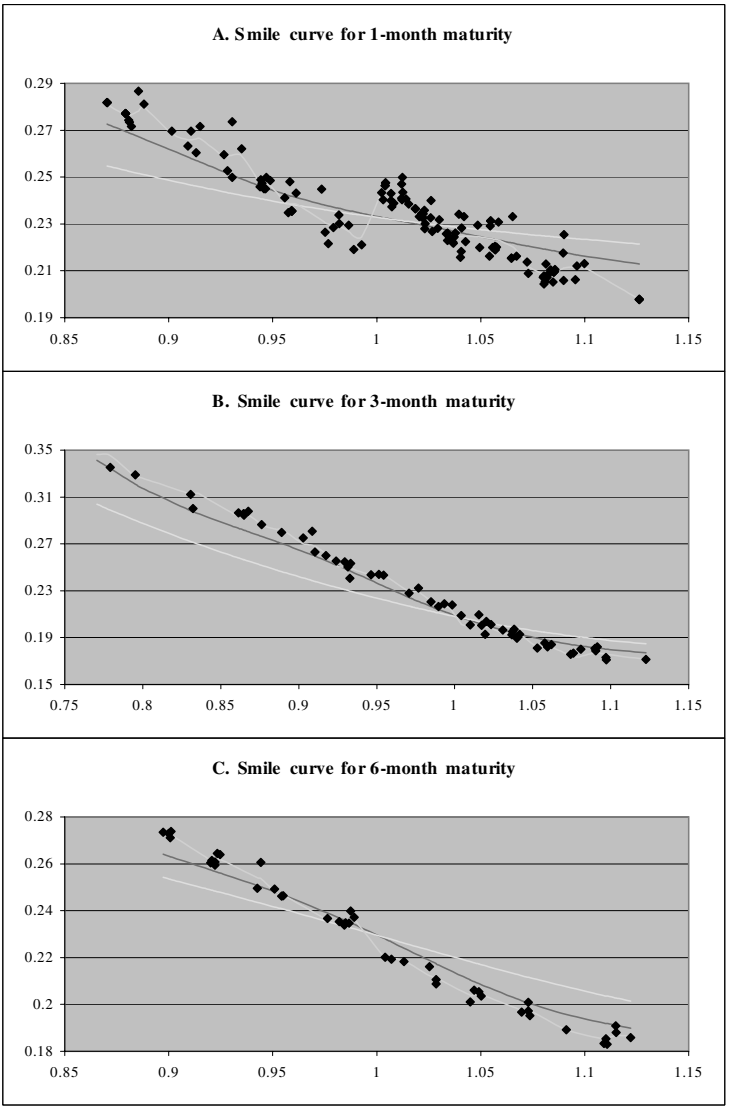

Figure 1. – Volatility smile curves and smoothing parameters.

Graphs A, B, and C show the smile curves for one-month, three-month, and six-month maturities, respectively. We calculated these curves using the optimal smoothing parameter and smoothing parameters equal to one-tenth and twice the optimal value.

We used Aït-Sahalia and Lo's method (1998, 2000) to calculate the optimal parameter value. This method is based on obtaining the best convergence while reducing variance (see their paper for more details about the calculation).

Figure 1 shows the importance of the smoothing parameter. It gives the volatility smile on a given day for maturities of one month, three months, and six months. The smile curve at each point was calculated using the optimal h value (from Aït-Sahalia and Lo's method) and h values of one-tenth the optimal value (for a high-variance curve) and twice the optimal value (for a curve with little variance but a large bias). The bias in the volatility smile will be reflected in the risk-neutral density curve. If the smile is too smooth, the risk-neutral density will be close to a normal distribution; if the smile is too irregular, the price curve will not be smooth enough to be able to take a second derivative. Only the optimal smile curve can be used to correctly represent the data and estimate risk-neutral densities.

3.3. Estimating risk-neutral densities

This section describes how we used the smoothed implied volatility surfaces calculated above to construct risk-neutral density functions for each day, at one-month, three-month, and six-month maturities[7].

We calculated the daily smile curves at each maturity with 10,000 exercise points (using the non-parametric method described above), covering a range of values wide enough to ensure that the integral is close to one. We then used the Black-Scholes model to transform these curves into price curves[8], and took the second derivative of the price curves relative to the exercise price according to the following formula, in order to obtain the risk-neutral densities:

$$q(K) = e^{r(T-t)} \frac{C(K+\varepsilon,T) - 2C(K,T) + C(K-\varepsilon,T)}{\varepsilon^2}$$

7. We used relatively short-term maturities because the most exercise price data is available for these maturities, making the smile curve calculation more reliable.

8. We used the Black-Scholes model solely to convert implied volatilities into prices, without making any assumptions as to the formula's validity [see Shimko (1993); Bahra (1997); and Bliss and Panigirtzoglou 2002)].

Where $C(K-\varepsilon,T)$ and $C(K+\varepsilon,T)$ are the points on either side of $C(K,T)$ on the price curves comprised of 10,000 points.

We used the 10,000 risk-neutral density points thus obtained to numerically calculate the integrals needed for descriptive statistics.

In order to validate our calculation method and the way it was applied to the data, we compared some of the numerically-calculated integrals with expected values. First we calculated the integrals of the risk-neutral densities to ensure that they are sufficiently close to one. This is done to confirm the reliability of our data and calculation method. This also let us ensure that the integration interval we selected $m_{min} = 0.5$ and $m_{max} = 1.5$ does cover the entire curve. We decided to use a moneyness interval (rather than an interval based on the underlying asset value) because it automatically adjusts to changes in the value of the underlying asset, giving us a centred interval for each day. We should have:

$$\int_{m_{min}}^{m_{max}} q(S_T)dS_T = 1$$

The second statistic we calculated was the average, with the following integral:

$$\mu = E[S_T] = \int_{m_{min}}^{m_{max}} S_T q(S_T)dS_T$$

This statistic can also be considered as a check of our method in that the average risk-neutral density should equal today's forward price. Bliss and Panigirtzoglou (2002) noted that the price of a call option with a zero exercise price is the current value of the expected price of the underlying asset at maturity, or the current value of the forward price. Replacing the exercise price in Equation 1 with zero gives:

$$C_t(0,T) = e^{-r(T-t)} \int_0^{\infty} S_T q(S_T)dS_T$$

Here the price of a call option with a zero exercise price is equal to the current value of the average risk-neutral density. This shows that the forward price should indeed be equal to the average risk-neutral density.

Table 1 presents these key statistics: the maximum, minimum, and average density function integral for the three maturities considered (one-month, three-month, and six-month). The average is very close to one for all three maturities, proving that the interval covers almost the entire curve. Moreover, all the integrals are between 0.95 and 1, which is consistent with the fact that a probability cannot exceed one (which means that the event is certain to occur).

Table 1 also gives the maximum, minimum, and average relative difference between the average density and the forward price for the three maturities. This relative difference is very low on average, with a minimum of zero and a maximum ranging from 1% to 6%, values that are quite acceptable.

These statistics confirm that our calculation method is valid and that the interval we selected is appropriate.

Table 1. – Statistics checking the risk-neutral density calculation

The average (avg), minimum (min), and maximum (max) integrals of the risk-neutral density curve ($q(S_T)$), which should be close to one, and of the relative deviation of the expected value of S_T ($E[S_T]$), calculated from the risk-neutral density and the forward price (fwd), which should be close to zero. These statistics are given for one-month, three-month, and six-month maturities.

Maturity		1-month	3-month	6-month
$\int q(S_T)dS_T$	avg	0.999	0.999	0.990
	min	0.959	0.985	0.953
	max	1.000	1.000	0.999
$\frac{\lvert E[S_T] - fwd \rvert}{fwd}$	avg	0.0001	0.0005	0.0066
	min	0.0000	0.0000	0.0003
	max	0.064	0.012	0.041

3.4. Risk-neutral densities of the underlying asset returns

The risk-neutral densities discussed above represent the distribution curves of the underlying assets. This distribution curve is log-normal

under the Black-Scholes model, which means that the log of the underlying asset or its return has a normal density distribution. In order to compare distribution curves calculated from options prices with a normal distribution curve (most notably for the third and fourth moments), we estimated the risk-neutral densities of the underlying asset returns, equivalent to the densities of the underlying assets.

We want to estimate the risk-neutral density of the asset return, which is continuously compounded over the period $\tau = T - t$, where T is the maturity and t is the day of estimation. If u_τ is the return, then:

$$u_\tau = \ln \frac{S_T}{S_t}$$

And the risk-neutral density of u_τ can be obtained by noting that[9]:

$$Proba(\ln(\frac{S_T}{S_t}) \leqslant u) = Proba(S_T < S_t e^u) = \int_0^{S_t e^u} q(S_T) dS_T$$

The risk-neutral density of the return, equivalent to the risk-neutral density over the price, is given by:

$$\frac{\partial}{\partial u} Proba(\ln(\frac{S_T}{S_t}) \leqslant u) = S_t e^u q(S_t e^u)$$

This is the density that we compared with the Black-Scholes normal density distribution curve, $\phi\left[\left(\mu - \frac{\sigma^2}{2}\right)\tau, \sigma\sqrt{\tau}\right]$. Figure 2 shows the difference between these two densities for a given day and maturity (the Black-Scholes density was calculated using an at-the-money volatility for the given day and maturity).

The following section discusses how we calculated statistics for the risk-neutral densities of the underlying asset returns, and compared these statistics with those of a normal distribution.

9. See Aït-Sahalia and Lo (1998, 525-526).

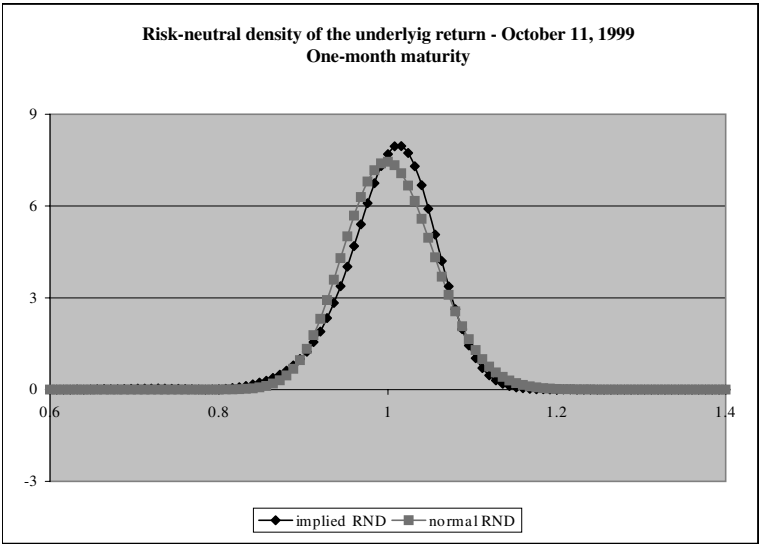

Figure 2. – Estimated and Black-Scholes (normal) risk-neutral densities of the underlying asset return.

The estimated density was calculated from the given day's options prices using a non-parametric method, and the Black-Scholes density was calculated using an at-the-money volatility for the given day and maturity.

3.5. Risk-neutral density statistics

We obtained statistics for the risk-neutral densities by numerically calculating the corresponding integrals. Table 2 gives the results for one-month, three-month, and six-month maturities. We were particularly interested in the skewness (asymmetry) and kurtosis (flatness) variables, which measure the difference between the actual distribution curve of the underlying asset return (estimated from options prices) and a normal distribution curve (which has a skewness of zero and a kurtosis of three).

Table 2 shows that the average skewness is always negative, indicating that the distribution curves are skewed to the left. In addition, the skewness moves further away from zero at higher maturities.

Table 3 also shows that the average kurtosis is always greater than three (or a positive excess kurtosis), which means that the distribution curves have fat tails (or are leptokurtic). The difference between the kurtosis and the "normal" value of 3 also increases at higher maturities. Therefore, the longer the maturity, the more the distribution curve of the underlying asset return shifts away from a normal curve. This is consistent with Aït-Sahalia and Lo's (1998) finding that skewness and kurtosis increase with maturity, which is based on a calculation of a single risk-neutral density for each maturity in their data sample.

Table 2. – Risk-neutral density statistics

The expected value($E[u]$), standard deviation (StdDev), skewness (skew), and kurtosis (kurt) of the underlying asset return $u = \ln(\frac{S_T}{S_t})$. These values were calculated from the daily risk-neutral densities obtained from options prices (687 data points) for one-month, three-month, and six-month maturities.

Variable	Mean	Median	Std Dev	Minimum	Maximum
1-month maturity					
$E[u]$	0.003	0.003	0.0006	0.002	0.004
StdDev	0.069	0.067	0.014	0.045	0.172
skew	– 0.362	– 0.293	0.298	– 3.840	0.093
kurt	3.882	3.533	1.037	2.869	14.082
3-month maturity					
$E[u]$	0.009	0.009	0.002	0.006	0.012
StdDev	0.119	0.118	0.018	0.085	0.208
skew	– 0.693	– 0.527	0.508	– 2.909	0.197
kurt	5.483	4.374	2.786	2.281	18.794
6-month maturity					
$E[u]$	0.019	0.020	0.003	0.012	0.025
StdDev	0.171	0.169	0.020	0.130	0.265
skew	– 1.163	– 1.160	0.550	– 2.801	0.163
kurt	6.981	6.853	2.064	3.561	14.991

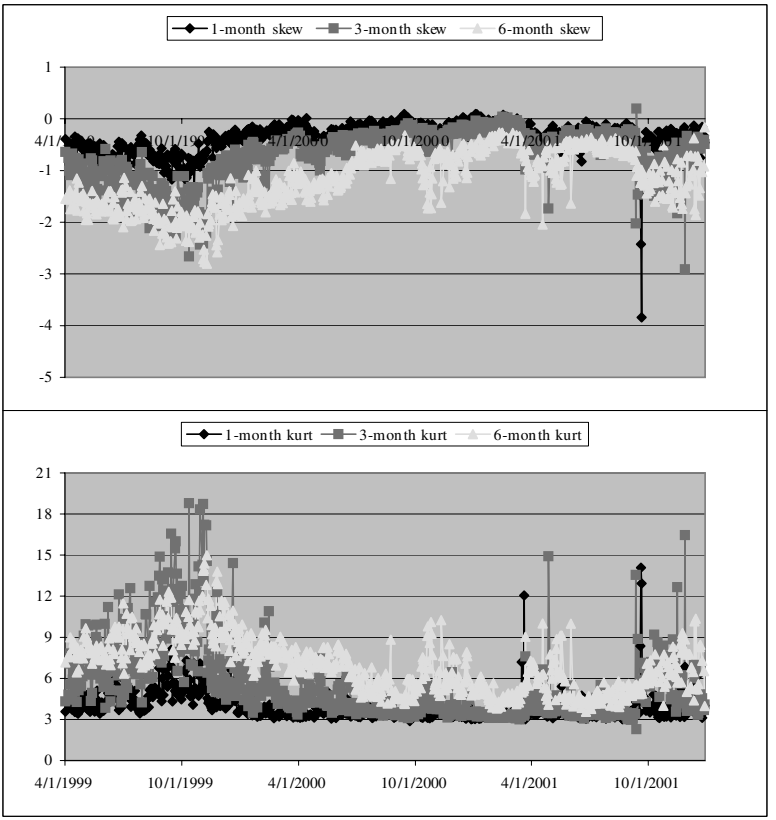

Figure 3. – Time evolution of the skewness and kurtosis of the distribution of the underlying asset return.

These values were calculated from daily risk-neutral densities estimated from options prices using a non-parametric method.

Figure 3 gives the changes in skewness and kurtosis over time for the three maturities. The values fluctuate considerably and differ from the normal distribution values of zero and three. Therefore, the risk-neutral densities of the underlying asset returns are clearly not normal distributions.

4. Risk-neutral density dynamics

To be able to study changes in risk-neutral densities over time and the factors that influence these changes, we needed a set of daily risk-neutral density curves, which we were able to obtain from the data in our sample. However it was difficult to represent the entire curve. Panigirtzoglou and Skiadopoulos (2004) alternatively used a delta metric to calculate cumulative density functions[10] instead of density functions.

We used here a moneyness metric to calculate the risk-neutral density curves of the underlying asset returns. We decided to work with return densities rather than underlying asset densities because it is returns that are of most interest to investors. Indeed, the risk/return ratio is a key concept in finance, and the standard deviation of a return determines a widely-used risk measure, historic volatility. In addition, by using a moneyness metric we ensured that our risk-neutral densities would always be centred. We took a moneyness range of 0.5 to 1.5, with data points at 0.05 intervals, in order to cover the entire curve. This gave us a curve with 21 points, or 21 variables on which we performed a PCA. The results are discussed below.

Table 3 gives the Eigen values, explained variances, and cumulative explained variances of the first three PCA factors for the three maturities considered. Figure 4 gives the variable correlations with these three factors.

Our analysis shows that the first three PCA factors explain around 90% of the total variance. But the correlations indicate that interpreting the results is not straightforward, especially for the second and third factors. The shapes of the correlation curves for the first factor, where the explained variance ranges from 59% to 67%, are the same for all three maturities. This factor has a highly positive effect on the curve ends, with the same magnitude for all three maturities. Its effect subsequently diminishes and becomes highly negative at the curve centre (over an interval centred at-the-money). Therefore, a positive shock on this factor would lower the central portion of the density curve over a

10. Panigirtzoglou and Skiadopoulos (2004) assign 21 delta values and the associated exercise prices, and then calculate the probability of S_T being below these values by taking the integral of the risk-neutral density function.

Table 3. – Principal Component Analysis (PCA) of risk-neutral density curves calculated using a moneyness metric: Eigen values and explained variances

PCA of risk-neutral density curves for moneyness values of 0.5 to 1.5. The three factors were extracted from the first three Eigen values, with an explanation of the system's maximum variance.

Component	1	2	3
1-month maturity			
Eigen value	12.501	4.506	1.663
Explained variance (%)	59.529	21.455	7.919
Cumulative explained variance (%)	59.529	80.984	88.904
3-month maturity			
Eigen value	13.625	3.506	1.918
Explained variance (%)	64.882	16.696	9.131
Cumulative explained variance (%)	64.882	81.578	90.709
6-month maturity			
Eigen value	14.163	3.905	1.076
Explained variance (%)	67.445	18.596	5.123
Cumulative explained variance (%)	67.445	86.041	91.164

given interval, and raise the curve ends. Note that the size of the interval centred at-the-money (where the negative effect is the strongest) changes with different maturities. It is narrower at short maturities and wider at long maturities. This factor could be related to distribution variance or other moments.

As for the second factor, the shapes of the correlation curves are not the same for all three maturities. The one-month maturity curve has a shape that is the inverse of that for the first factor at the centre of the density curve, with highly positive correlation, but also positive correlations at the ends. The correlations are negative only in the intermediate sections. Therefore, a positive shock on this factor would raise the distribution curve at the curve centre and ends, leading to higher extreme probabilities (fat tails) and a narrower centre (sharp bell). This

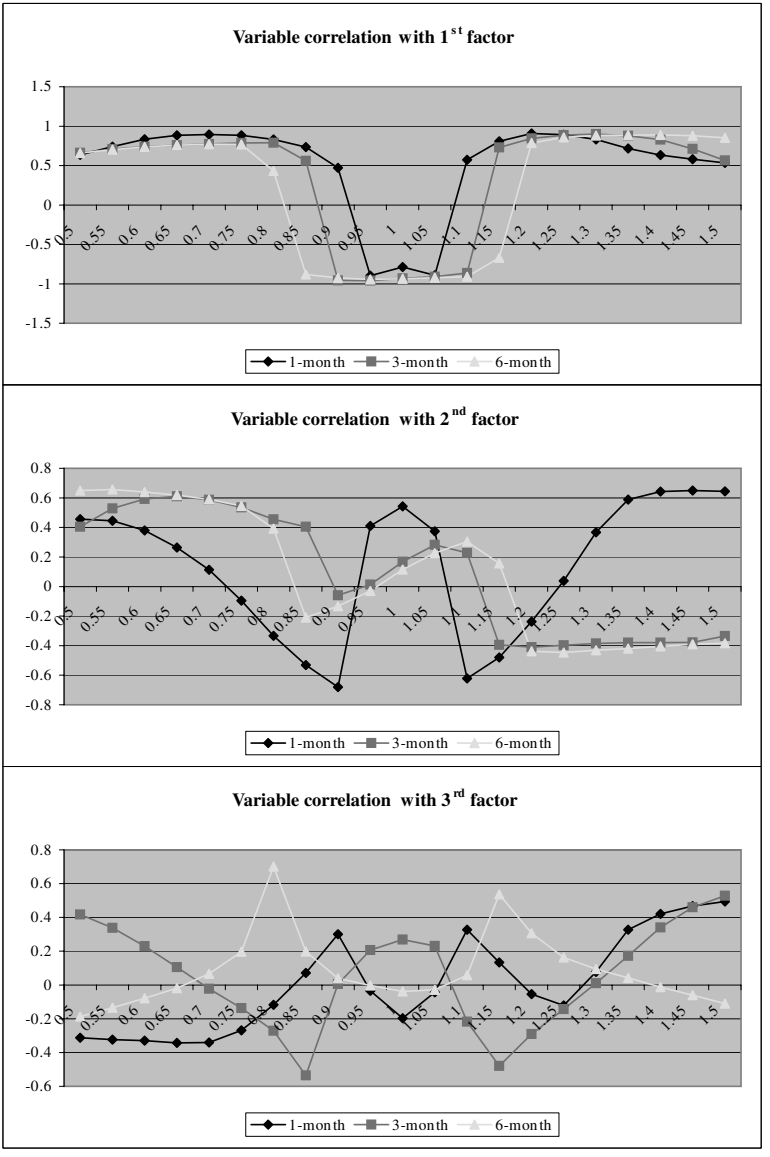

Figure 4. – Correlation of variables with PCA factors.

The correlation of variables describing risk-neutral density curves with factors obtained from a PCA of risk-neutral density curves for moneyness values of 0.5 to 1.5.

comes the closest to kurtosis, but since it is not perfectly symmetrical it probably includes some asymmetry effects. The three-month and six-month maturity curves for the second factor have a similar shape, with a positive correlation on the left portion of the density curve and a negative correlation on the right. A positive shock on this factor would raise one end of the risk-neutral density curve and lower the other, consequently altering the curve's symmetry (or asymmetry). The explained variance on this factor ranges from 16% to 21% of the total variance, depending on the maturity.

The correlation curves for the third factor, which explains 5%-9% of the total variance, are more chaotic. There does not seem to be any connection between the curves of different maturities, nor do they seem to be tied to a density moment. This latter factor could simply be due to noise. As a result, we decided to use only the first two factors in our study since they explain 81%-86% of the total variance, and seem to be fairly well-linked to risk-neutral density moments (even though no formal relationship has been established).

Our results are similar to those obtained by Panigirtzoglou and Skiadopoulos (2004) on the cumulative function, who also decided to use two factors (explaining around 60% and 30% of the variance, respectively) because they felt that the third factor had random behaviour. And like Panigirtzoglou and Skiadopoulos (2004), we were not able to come up with an economic explanation of these shock factors. We could not establish a single, independent link between each factor and the risk-neutral density moments (mean, variance, etc.), even though some trends do appear – nor could we tie the factors to shocks affecting the implied volatility surface, even though the number of shocks is the same.

5. Risk-neutral density modelling

Panigirtzoglou and Skiadopoulos (2004) developed a risk-neutral density modelling algorithm based on the Monte Carlo method. Their approach gives confidence intervals for risk-neutral density values, but these values fall outside the confidence interval in the event of a shock (in their study, the 11 September terrorist attacks). In order to account for such discontinuities, we decided to develop a risk-neutral density model incorporating a jump component.

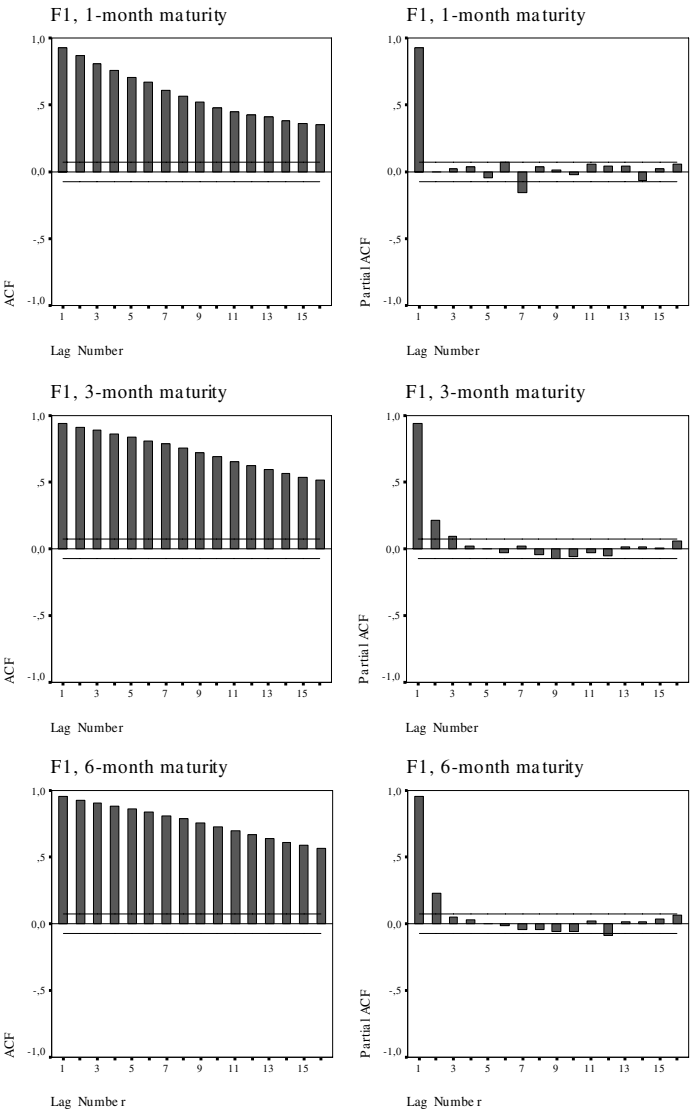

Figure 5. – Autocorrelation and partial autocorrelation functions (ACFs) for the second PCA factor.

The autocorrelation and partial autocorrelation functions for the second factor obtained from a PCA on risk-neutral density curves for moneyness values of 0.5 to 1.5.

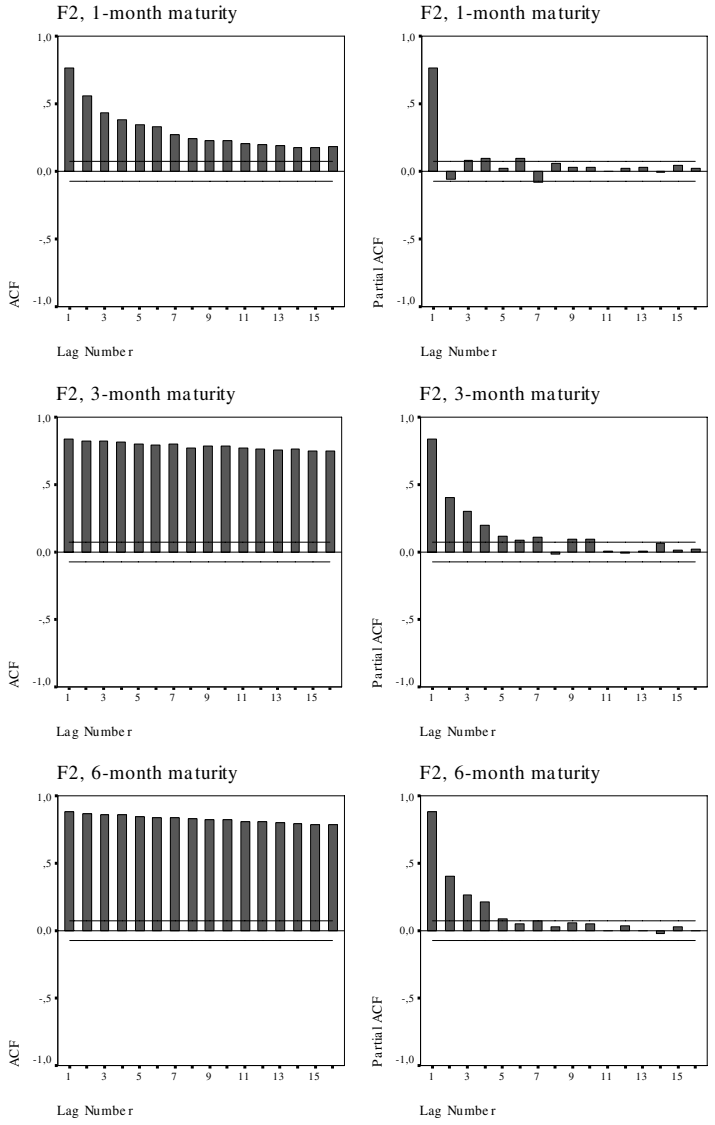

Figure 6. – Autocorrelation and partial autocorrelation functions (ACFs) for the first PCA factor.

The autocorrelation and partial autocorrelation functions for the first factor obtained from a PCA on risk-neutral density curves for moneyness values of 0.5 to 1.5.

As discussed above, we used the first two factors obtained from a PCA on risk-neutral density curves calculated using a moneyness metric.

Figures 5 and 6 give the correlograms and partial correlograms of the first and second factors, respectively, and show that the autocorrelation functions can be represented by a first-order autoregressive process. The autocorrelation functions gradually decrease for all three maturities, while the first-order partial autocorrelation coefficient is the only significant one or considerably larger than the others. The latter remark should be tempered for the second factor, for which the second- and fourth-order autocorrelation coefficients for the three – and six-month maturities are not negligible relative to the first-order. But in order to be consistent, we chose to model all factors with an Ornstein-Uhlenbeck first-order autoregressive process, while keeping in mind the particularities of the three – and six-month maturities for the second factor. We then added a jump component to this process to account for potential discontinuities.

Before beginning the modelling process, we looked more deeply into the factors' correlations with the first four risk-neutral density moments. These factors most likely incorporate the combined effect of several moments (since we were not able to establish a single, formal link between the factors and a particular moment), so we felt it would be useful to identify these effects in order to better understand their interactions. Table 4 gives these correlations for each maturity (one-month, three-month, and six-month). All of these correlations are significant, proving that the factors are not linked to a particular moment. Some correlations are especially strong (above 0.7). The first factor is highly correlated to the first two risk-neutral density moments: the expected value and the standard deviation of the return. The second factor is the most-highly correlated with the third and fourth risk-neutral density moments (skewness and kurtosis) for the three-month and six-month maturities. It does not have any correlations of such a magnitude for the one-month maturity. The other moments are also significantly correlated to the two factors for all maturities, although the correlations are not as high as those mentioned.

All of this indicates that the PCA factors do not represent the same risk-neutral density movements for each maturity. Although some similar behaviours can be seen in the first factor (which accounts for most of the explained variance), the same is not true for the second

factor. The short-term (one-month) maturity seems to differ from the others in its shape and correlation with risk-neutral density moments. We kept the particular characteristics of the second factor in mind during our modelling process.

Table 4. – Correlation of PCA factors with the first four risk-neutral density moments

Correlation of the first two PCA factors with the first four risk-neutral density moments: expected value (μ), standard deviation (σ), skewness (skew), and kurtosis (kurt).

Factor	μ	σ	skew	kurt
1-month maturity				
1	-0.87^{***}	0.91^{***}	-0.41^{***}	0.35^{***}
2	0.20^{***}	-0.34^{***}	-0.37^{***}	0.45^{***}
3-month maturity				
1	-0.88^{***}	0.95^{***}	-0.29^{***}	0.23^{***}
2	-0.28^{***}	-0.18^{***}	-0.82^{***}	0.74^{***}
6-month maturity				
1	-0.84^{***}	0.93^{***}	-0.45^{***}	0.27^{***}
2	-0.30^{***}	-0.12^{***}	-0.81^{***}	0.79^{***}

*** Indicates a critical probability of less than 1%.

5.1. Mean-reversion process with and without jumps

For a stochastic variable $x(t)$ (which in our case represents the PCA factors on the risk-neutral density curve), the mean-reversion process is given by the following equation:

$$dx(t) = \eta(\mu - x(t))dt + \sigma dB(t) \tag{6}$$

Where η is the speed of the return to the long-term mean μ (also called the equilibrium value). σ is the short-term standard deviation, and B(t) is the standard Brownian motion.

Adding a jump component to this equation gives:

$$dx(t) = \eta(\mu - x(t))dt + \sigma dB(t) + dq(t) \tag{7}$$

Where $dq(t)$ represents the jumps such that $dq(t) = \phi(t)$ if a jump occurs (probability of λdt) and $dq(t) = 0$ if no jump occurs (probability of $1 - \lambda dt$).

We assume that the occurrence of jumps follows a Poisson process with frequency λ. The size of the jump $\phi(t)$ will have a normal distribution $N(\theta,\delta^2)$. We approximated this process with a Bernoulli process whose parameters are easier to calibrate.

We used the maximum likelihood method to estimate the parameters η, μ, and σ for the mean-reversion process and λ, θ, and δ for the jump component. We used the numeric procedures in Matlab's optimization library to find the parameter values maximising these functions.

We tested the relevance of the jumps through a likelihood ratio test, by comparing the hypothesis of absence of jumps ($H_0 : \lambda = 0$) against the hypothesis of presence of a jump ($H_1 : \lambda \neq 0$). We used the following statistic:

$$LR = -2ln\Lambda \tag{8}$$

$$\Lambda = \frac{L_0}{L_\lambda} \tag{9}$$

Where L_0 and L_λ are the likelihood function maximums under the null hypothesis and the alternative hypothesis, respectively.

Under the null hypothesis, the LR statistic follows a χ^2 distribution with three degrees of freedom (since there are three more parameters in the general model than in the restricted model).

5.2. Estimating the parameters and the test statistic

Table 5 gives the estimated parameters for the mean-reversion process, with and without jumps, for each maturity considered for the two PCA factors. The table also gives the LR statistic for the test of occurrence of jumps hypothesis, for each maturity and factor.

We started with a mean-reversion process without jumps and calculated the half-life time H, which is the time needed for the process

Table 5. – Estimated parameters and test statistic

Estimated mean-reversion model parameters (with and without jumps) for the two PCA factors. LR is the test statistic of the H_0 hypothesis ($\lambda = 0$) against the H_1 hypothesis ($\lambda \neq 0$).

	F_1		F_2	
Parameter	Without jumps	With jumps	Without jumps	With jumps
1-month maturity				
η	0.0728	0.0915	0.1725	0.0575
μ	– 0.0032	– 0.2605	0.0012	– 0.0339
σ	0.3816	0.1669	0.7382	0.2503
λ	–	0.0735	–	0.0171
θ	–	0.3075	–	0.1233
δ	–	1.1887	–	1.8000
LR	730.21***		1180.69***	
3-month maturity				
η	0.0592	0.0890	0.1797	0.1032
μ	– 0.0201	– 0.4976	0.0010	– 0.0937
σ	0.3428	0.2106	0.5995	0.3960
λ	–	0.1926	–	0.0896
θ	–	0.2143	–	0.1047
δ	–	0.5755	–	1.3510
LR	182.68***		199.75***	
6-month maturity				
η	0.0477	0.0599	0.1203	0.1130
μ	– 0.0412	– 0.4020	0.0168	– 0.2740
σ	0.3008	0.2519	0.4939	0.4227
λ	–	0.0946	–	0.1883
θ	–	0.2273	–	0.1457
δ	–	0.4840	–	0.5339
LR	56.02***		14.35***	

*** Indicates a critical probability of less than 1%.

to reach half of its equilibrium value regardless of the starting point. H is given by the following formula:

$$H = \frac{ln2}{\eta}$$

We found that the half-life time increases with maturity for all factors; it ranges from 9 to 15 days for the first factor and 4 to 6 days for the second. It is always less than the maturity, meaning that an isolated shock will probably be fully absorbed before the option expires. The mean-reversion time, defined as the inverse of the parameter η (the speed of the return to the long-term mean), supports this conclusion. The mean-reversion time is between 14 and 21 days for the first factor and between 5 and 9 days for the second. Although the mean-reversion time increases with maturity, it is always considerably less than the maturity. This seems to indicate that there is no need to include a jump component in risk-neutral density factor models.

However, the LR statistic of the likelihood ratio test indicates that the no-jump hypothesis ($\lambda = 0$) is rejected in all cases (for both factors and all three maturities), and this at the 1% statistical threshold. In addition, σ, which represents the variance explained by the mean-regression portion, is lower in the model with jumps in all cases. This means that some of the variance is absorbed by the jump component.

We then decided to look at jump frequency. If a new shock occurs before the previous one has been fully absorbed, it will automatically increase the time needed to return to equilibrium. Therefore, jumps should be considered in the modelling of density function movements. Jump frequency is given by the inverse of the parameter λ. For the one-month maturity, it indicates that there will be a jump every 14 days on average for the first factor and only every 58 days on average for the second. Thus investors seem to have good short-term expectations, with only a few types of events being significant enough for them to change these expectations. Nevertheless, the high values of the parameter representing jump variance indicate that jumps could occasionally have a considerable effect. So when investors do change their short-term expectations, they do so drastically. Indeed, given how close these changes are made to the time of option maturity, they appear to reflect an urgency nature. In contrast, the three-month and six-month

maturities have a high jump frequency, ranging from 5 days to 11 days on average according to their respective λ parameters. This implies that traders adjust their long-term expectations more often in light of new information. Moreover, jumps have less of an effect for these maturities than for the shorter (one-month) maturity.

Given the above elements and the PCA factors' correlation with the first four risk-neutral density moments, we concluded that the most appropriate process for modelling risk-neutral density dynamics may depend on the maturity being considered. The following section substantiates this conclusion by comparing the forecasting power of the two models.

5.3. Model forecasting power

We used an out-of-sample test to estimate the forecasting power of the two models, in order to compare them and assess the relevance of the jump component. We divided our data sample into two sets: T_1, comprised of all the data up to the last 15 days; and T_2, comprised of the last 15 days. We used the T_1 data set to recalculate the model parameters; the new parameters are given in Table 6. Note the stability of these parameters, as compared to those calculated from the entire data sample.

We used the new parameters to calculate the values for both factors on T_2, and compared these calculated values with the actual values. The differences between the actual and calculated values are the model forecasting residuals. We then calculated the mean quadratic error (MQE) according to the formula:

$$EQM = \frac{1}{T_2}\sum_{i=1}^{T_2}(x(T_1+i) - \widetilde{x}(T_1+i))^2 \qquad (10)$$

Where $x(T_1+i)$ and $\widetilde{x}(T_1+i)$ are the actual and calculated values, respectively, on T_1+i. The calculated value corresponds to:

$$\widetilde{x}(T_1+i) = E_{T_1}\left[x(T_1+i)/x(1),\ldots,x(T_1)\right], i = 1,\ldots,T_2 \quad (11)$$

We obtained the conditional expected value from equation 11 by simulation, using the same procedure as the one described in Kermiche (2008).

Table 6. – Estimated paramaters for the T_1 data set

Estimated mean-reversion model parameters (with and without jumps) for the two PCA factors for the period covered by the T_1 data set (up to the last 15 days).

	F_1		F_2	
Parameter	Without jumps	With jumps	Without jumps	With jumps
1-month maturity				
η	0.0715	0.0903	0.1733	0.0572
μ	– 0.0110	– 0.2696	0.0112	– 0.0179
σ	0.3819	0.1646	0.7460	0.2517
λ	–	0.0680	–	0.0174
θ	–	0.3305	–	0.1219
δ	–	1.2392	–	1.8106
3-month maturity				
η	0.0585	0.0895	0.1724	0.1009
μ	– 0.0329	– 0.5484	0.0091	– 0.0802
σ	0.3443	0.2036	0.5914	0.3978
λ	–	0.2120	–	0.0804
θ	–	0.2125	–	0.1089
δ	–	0.5630	–	1.3912
6-month maturity				
η	0.0452	0.0617	0.1169	0.1146
μ	– 0.0772	– 0.4407	0.0130	– 0.6851
σ	0.2944	0.2592	0.4599	0.4171
λ	–	0.0492	–	0.1172
θ	–	0.4662	–	0.6463
δ	–	0.4441	–	0.5111

We calculated the MQE for both models (with and without jumps) for periods of 5, 10, and 15 days. We expected the MQE to increase with time for both models since the forecasting power of any model should

diminish over increasingly longer periods. Moreover, if diffusion processes with jumps are more appropriate for modelling risk-neutral densities, then the MQE should be lower in the model with jumps than the one without. We compared the MQEs of the two models for each factor and time period (estimation time period and risk-neutral density maturity) separately; the results are given in Table 7.

Table 7 shows that the MQE for both models is quite high for the three-month and six-month maturities for the second factor. Neither model takes into account changes in this factor for these maturities, as indicated by the correlograms. A different model would be needed, probably a third-order or fourth-order autoregressive process with or without jumps. However the two models do seem to represent the one-month maturity for the second factor fairly well. This could imply that

Table 7. – Mean Quadratic Error (MQE)

Estimation of the Mean Quadratic Error (MQE) for the models with and without jumps, for the two PCA factors, for estimation periods of 5, 10, and 15 days.

	F_1		F_2	
MQE	Without jumps	With jumps	Without jumps	With jumps
1-month maturity				
5 days	0.0223	0.0117	0.0317	0.0095
10 days	0.0819	0.0710	0.0670	0.0508
15 days	0.0985	0.0874	0.0902	0.0621
3-month maturity				
5 days	0.0093	0.0087	0.5366	0.4838
10 days	0.0350	0.0233	0.5228	0.4606
15 days	0.0644	0.0597	0.6126	0.3643
6-month maturity				
5 days	0.0111	0.0095	0.2474	0.3947
10 days	0.0471	0.0325	0.4336	0.4571
15 days	0.0913	0.0888	0.8820	0.7536

the most appropriate risk-neutral density model depends on the maturity being considered. For all cases apart from the three-month and six-month maturities for the second factor, the MQE increases at higher maturities (even if some values are very small). Therefore, the forecast reliability diminishes as the maturity extends; a fact that should be considered whenever the two models are being used.

The addition of a jump component to the mean-reversion process significantly improves the model's forecasting power for all maturities for the first factor. This factor is highly correlated to the second risk-neutral density moment, and then to historic volatility. Since historic volatility can be well-estimated from implied volatility [see for example Blair, Poon, and Taylor (2001)], the first factor is probably also correlated to implied volatility. This result is consistent with Kermiche's findings (2008) on implied volatility surfaces.

The addition of a jump component also improves the model's forecasting power for the one-month maturity for the second factor. However it is difficult to reach a conclusion for this factor due to the lack of results for the other maturities.

Therefore, a risk-neutral density model with a jump component can account for discontinuities in the underlying asset, especially for the first factor. Given the fact that this factor represents most of the risk-neutral density variance (using a moneyness metric) with over 60% of the explained variance, it is important to take it into account.

While it is important to incorporate discontinuities when modelling risk-neutral densities, the absolute MQE values we obtained indicate that a mean-regression process may not be the best method for doing so, especially for the second factor. Consequently other processes should also be tested.

6. Application to options portfolio hedging

This section illustrates the usefulness of shock factors affecting risk-neutral densities (as obtained from a PCA) in options portfolio hedging. In this example we hedged a long position on an at-the-money option of a given maturity with in-the-money and out-of-the-money options of the same maturity. The portfolio was built so as to be immune to the two primary risk factors determined by the PCA. We

then compared this portfolio with a delta-gamma neutral portfolio comprised of the same options. More specifically, the hedging portfolio was designed to meet the following criteria:

$$\begin{aligned} & w_{atm} = 1 \\ & w_{in} b_{1,in} + w_{atm} b_{1,atm} + w_{out} b_{1,out} = 0 \\ & w_{in} b_{2,in} + w_{atm} b_{2,atm} + w_{out} b_{2,out} = 0 \end{aligned} \tag{12}$$

Where w_x is the proportion of the portfolio invested in option x; $b_{1,x}$ is the weight of the first PCA factor in the risk-neutral density for option x; and $b_{2,x}$ is the weight of the second PCA factor in the risk-neutral density for option x. The parameter x depends on whether the option is at-the-money ($x = atm$), out-of-the-money ($x = out$), or in-the-money ($x = in$), using the moneyness metric m discussed earlier.

The first equation represents the fact that we want to hedge a long position on an at-the-money option, while the next two indicate that the portfolio is immune to the two primary risk factors.

Similarly, we built a classic delta-gamma neutral portfolio as follows:

$$\begin{aligned} & w_{atm} = 1 \\ & w_{in} \Delta_{in} + w_{atm} \Delta_{atm} + w_{out} \Delta_{out} = 0 \\ & w_{in} \Gamma_{in} + w_{atm} \Gamma_{atm} + w_{out} \Gamma_{out} = 0 \end{aligned} \tag{13}$$

Under normal circumstances a delta-gamma neutral hedge is adjusted regularly, but we used a static approach and kept the portfolio the same until maturity so as to be able to compare it with the portfolio immune to the two primary risk factors affecting risk-neutral densities.

We calculated the volatilities of the portfolios built from our entire data sample. Table 8 gives the average volatility for each maturity (one-month, three-month, and six-month) for both the delta-gamma neutral portfolio and the immune portfolio. The immune portfolio has a lower average volatility than the delta-gamma portfolio[11]. This improvement is observable for all maturities, even if the performance (in terms of risk) of the hedged portfolios is considerably worse for the three-month and six-month maturities. Nevertheless, it is difficult to

11. Note that the portfolio we used is static, which is not the main objective of a delta-gamma neutral portfolio.

reach a conclusion due the relatively small number of test portfolios for these maturities. To prevent the problem of overlapping portfolios, we waited until previous portfolios had matured before building new portfolios. This gave us a total of 33 portfolios for the one-month maturity, 11 portfolios for the three-month maturity, and only 5 portfolios for the six-month maturity. Protecting options portfolios against the two primary risk factors affecting risk-neutral densities can undeniably enhance passive portfolio management strategies.

Table 8. – Options portfolio hedging

Comparison of the volatility of a delta-gamma neutral options portfolio (DG portfolio) with an options portfolio immune to the two primary risk factors affecting risk-neutral densities (RND portfolio). The portfolios were built to hedge a long position on an at-the-money option ($m = 1$) with options that are out-of-the-money ($m = 1.1$) and in-the-money ($m = 0.9$). The options have maturities of one-month, three-months, and six-months, which correspond to the immunization period. The volatility of each portfolio was calculated for the entire data sample. The numbers below are the average volatilities for each immunization period.

Immunization period	DG portfolio	RND portfolio
1 month	0.376	0.136
3 months	2.865	1.660
6 months	2.989	1.343

7. Conclusion

In our study we investigated the dynamics of risk-neutral density functions implied in the prices of options on the CAC 40 index. We calculated risk-neutral densities using a moneyness metric for maturities ranging from one month to six months, and carried out a Principal Component Analysis (PCA) to determine the shock factors affecting risk-neutral density curves.

We used moneyness measures ranging from 0.5 to 1.5 and confirmed that the curve's integral over this range equals one, meaning that

the interval covers the entire risk-neutral density curve. The PCA revealed two significant risk factors. The first factor explains on average 60% of the total variance, and has strong correlations in opposite directions at the density curve centre versus the ends. This factor seems to be linked to the second density moment. The effect of the second factor varies according to maturity. This factor appears to be consistently linked to the third and fourth density moments, and explains on average 20% of the total variance.

We then developed two mean-reversion models for risk-neutral density dynamics, one with a jump component and one without, based on the shock factors determined in our PCA. We then compared the forecasts given by these models for several maturities. The addition of a jump component significantly improved the model's forecasting power for the first PCA factor (which accounts for 60% of the total variance). This factor has the highest correlation with the variance of the risk-neutral density of the underlying asset return. It is linked to historic volatility, and subsequently to implied volatility. Our finding is then consistent with other research on implied volatility dynamics. For the second PCA factor, the addition of a jump component led to a smaller improvement in the model's forecasting power. This improvement varied with maturity; it was significant for the one-month maturity, but minor for the three-month and six-month maturities. Indeed, neither model gave adequate results for the second factor for the three-month and six-month maturities. While it is important to account for discontinuities in the modelling process, and a jump process is a good method for doing so, other processes should be tested on other data samples. In addition, because the presence of jumps limits the use of a PCA as a data reduction tool, other avenues should be explored in this area as well. Semi-parametric and non-parametric methods could be possible options, especially since they could be used to analyze the joint evolution of density curves over several maturities, and therefore simultaneously evaluate options of different maturities.

Studying risk-neutral density curve dynamics offers many benefits, especially for risk management and options portfolio valuation. In this paper we give an example of how our findings can be used in options portfolio hedging. A possible extension of our work could be to develop a valuation model based on the shock factors we isolated by adjusting the factors so as to eliminate arbitrage opportunities.

REFERENCES

[1] AÏT-SAHALIA Y. and A.W. LO (1998), Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance*, Vol. 53(2), 499-547.

[2] AÏT-SAHALIA Y. and A.W. LO (2000), Nonparametric risk management and implied risk aversion. *Journal of Econometrics*, Vol. 94, 9-51.

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

[4] BEBER A. and M. W. BRANDT (2006), The Effect of Macroeconomic News on Beliefs and Preferences: Evidence from the Options Market. *Journal of Monetary Economics*, Vol. 53(8), 1997-2039.

[5] BLACK F. and M. SCHOLES (1973), The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, Vol. 81, 637-654.

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

[7] BLISS R.R. and N. PANIGIRTZOGLOU (2004), Option-Implied Risk Aversion Estimates. *Journal of Finance*, Vol. 59(1), 407-446.

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

[9] CLEWS R., N. PANIGIRTZOGLOU and J. PROUDMAN (2000), Recent developments in extracting information from options markets. *Bank of England Quarterly Bulletin 40*, 50-60.

[10] CONT R. (1997), Beyond Implied Volatility: Extracting Information from Options Prices. In J. Kertesz and I. Kondor, eds., *Econophysics*. Dordrecht: Kluwer.

[11] CONT R. and J. DA FONSECA (2002), Dynamics of implied volatility surfaces. *Quantitative Finance*, Vol. 2, N 1, 45-60.

[12] CONT R. and P. TANKOV (2004), Financial Modelling With Jump Processes. CRC Press, 306p.

[13] COUTANT S. (1999), Implied Risk Aversion in Options Prices Using Hermite Polynomials. *Working paper*, Banque de France.

[14] COUTANT S., E. JONDEAU and M. ROCKINGER (2001), Reading PIBOR futures options smiles: The 1997 snap election. *Journal of Banking and Finance*, Vol. 25, 1957-1987.

[15] COX J. and S. ROSS (1976), The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, Vol. 3, 145-166.

[16] FENGLER M., W. HÄRDLE and E. MAMMEN (2007), A Semiparametric Factor Model for Implied Volatility Surface Dynamics. *Journal of Financial Econometrics*, Vol. 5(2), 189-218.

[17] GEMMILL G. and A. SAFLEKOS (2000), How Useful are Implied Distributions? Evidence from Stock Index Options. *Journal of Derivatives*, Vol. 7(3), 83-98.

[18] GIACOMINI E., W. HÄRDLE and V. KRÄTSCHMER (2008), Dynamic Semiparametric Factor Models in Risk Neutral Density Estimation. *SFB 649 Discussion Paper* 2008-038.

[19] HARVEY C.R. and R.E. WHALEY (1991), S&P 100 Index Option Volatility. *Journal of Finance*, Vol. 46(4).

[20] HARVEY C.R. and R.E. WHALEY (1992), Market volatility prediction and the efficiency of the S&P 100 index option market. *Journal of Financial Economics*, Vol. 31, 43-73.

[21] HULL J.C. (2006), Options, Futures and Other Derivatives. 6th edition, *Pearson Prentice Hall*, 789p.

[22] JACKWERTH J.C. (1999), Option-Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review. *Journal of Derivatives*, Vol. 7(2), 66-82.

[23] JACKWERTH J.C. and M. RUBINSTEIN (1996), Recovering Probability Distributions from Option Prices. *Journal of Finance*, Vol. 51(5), 1611-1631.

[24] JONDEAU E. and M. ROCKINGER (1998), Estimating Gram-Charlier Expansions under Positivity Constraints. Les Cahiers de recherche du groupe HEC.

[25] JONES C.S. (2006), Nonlinear Factor Analysis of S&P500 Index Options Returns. *Journal of Finance*, Vol. 61(5), 2325-2363.

[26] KERMICHE L. (2008), Une modélisation de la surface de volatilité implicite par processus sauts. *Finance*, Vol. 29(2), 57-101.

[27] LATANE H. and R. RENDLEMAN (1976), Standard Deviations of Stock Price Ratios Implied in Option Prices. *Journal of Finance*, Vol. 31(2), 369-381.

[28] LYNCH D. and N. PANIGIRTZOGLOU (2003), *Summary Statistics of Option-Implied Probability Density Functions and their Properties*. Working paper, Bank of England.

[29] MALZ A.M. (1996), Using option prices to estimate reslignment probabilities in the European Monetary System: the case of sterling-mark. *Journal of International Money and Finance*, Vol. 15, 717-748.

[30] MANDLER M. (2002), *Comparing risk-neutral probability density functions implied by option prices - market uncertainty and ECB-council meetings*. Working paper.

[31] McManus D.J. (1999), *The information content of interest rate futures options*. Working paper, Bank of Canada.

[32] Melick W.R. and C.P. Thomas (1997), Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis. *Journal of Financial and Quantitative Analysis*, Vol. 32, 91-115.

[33] Merton R. (1973), Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, Vol. 4, 141-183.

[34] Merton R. (1976), Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics*, 125-144.

[35] Panigirtzoglou N. and G. Skiadopoulos (2004), A new approach to modeling the dynamics of implied distributions: Theory and evidence from the S&P 500 options. *Journal of Banking and Finance*, Vol. 28, 1499-1520.

[36] Pérignon C. and C. Villa (2002), Extracting Information from Options Markets : Smiles, State-Price Densities and Risk Aversion. *European Financial Management*, Vol. 8(4), 495-513.

[37] Shimko D. (1993), Bounds of probability. *Risk*, Vol. 6(4), 33-37.

[38] Skiadopoulos G. and S. Hodges (2001), Simulating the Evolution of the Implied Distribution. *European Financial Management*, Vol. 7(4), 497-521.