FINANCIAL OPTIONS RESEARCH CENTRE

University of Warwick

A New Approach to Modeling the Dynamics of Implied Distributions: Theory and Evidence from the S&P 500 Options

Nikolaos Panigirtzoglou
and
George Skiadopoulos

September 2002

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry
CV4 7AL
Phone: (0)24 76 524118

FORC Preprint: 2002/120

A New Approach to Modeling the Dynamics of Implied Distributions: Theory and Evidence from the S&P 500 Options*

Nikolaos Panigirtzoglou† and George Skiadopoulos‡

This Version: 14/10/2002

Abstract

This paper presents a new approach to modeling the dynamics of implied distributions. First, we obtain a parsimonious description of the dynamics of the S&P 500 implied cumulative distribution functions (CDFs) by applying Principal Components Analysis. Subsequently, we develop new arbitrage-free Monte-Carlo simulation methods that model the evolution of the whole distribution through time as a diffusion process. Our approach generalizes the conventional approaches of modeling only the first two moments as diffusion processes, and it has important implications for "smile-consistent" option pricing and for risk management. The out-of-sample performance within a Value-at-Risk framework is examined.

JEL Classification: G11, G12, G13.

Keywords: Implied cumulative distribution function, Monte Carlo Simulation, Option Pricing, Principal Components Analysis, Value-at-Risk.

1 Introduction

Cross sections of option prices can be used to infer (imply) the risk-neutral probability density function (PDF- or the equivalent cumulative distribution function-CDF) for the future values

*We are particularly grateful to Stewart Hodges for the inspiration he provided. We would also like to thank Karim Abadir, Robert Bliss, Peter Carr, George Constandinides, Jens Jackwerth, George Kapetanios, Anastassios Malliaris, Anthony Neuberger, Joao Pedro Nunes, Olivier Renault, Michael Rockinger, Pedro Santa-Clara, Phillip Schonbucher, Robert Tompkins, Dimitris Vayanos, Stavros Zenios, William Ziemba, and the participants at the 2002 FEES Conference (Athens), the 2002 Bachelier World Congess (Crete), the 2002 FORC Conference (Warwick), the 2002 EIR Conference (LSE), the Athens University of Economics and Business, and the University of Piraeus seminars for many helpful discussions and comments. The views expressed herein are those of the authors and do not necessarily reflect those of the Bank of England. Any remaining errors are our responsibility alone.

†Monetary Instruments and Markets Division, Bank of England. Threadneedle Street, London EC2R 8AH, England. nikolaos.panigirtzoglou@bankofengland.co.uk

‡University of Piraeus, Department of Banking and Financial Management, Financial Options Research Centre, Warwick Business School, University of Warwick, and Athens University of Economics and Business. gskiadopoulos@hotmail.com

of the underlying asset. There is an extensive literature on methods to extract the implied PDF at a *single point* in time. The objective of this paper is to investigate the *dynamics* of implied distributions, and to provide algorithms that can apply the empirical results to areas like option pricing, risk management and economic policy issues.

Over the last twenty-five years, there has been a considerable interest among academics, market participants, and policy makers to extract information from option prices regarding the market participants expectations. Breeden and Litzenberger (1978) showed that the second derivative of a standard European call option with respect to the strike price delivers the risk-neutral PDF. Ever since, many techniques have been developed to estimate the risk-neutral PDF [see Bahra (1997), Cont (1997), Jackwerth (1999), and Mayhew (1995), among others for extensive surveys].

The extracted PDFs are used as a guide to economic policy issues such as assessing monetary credibility, revealing the public's expectations about future interest rates, and evaluating the timing of monetary policy actions. They are also used to test whether specific events (e.g., the '87 crash, or the outcome of elections) were expected [see Bates (1991), Gemmill and Saflekos (2000), Coutant *et al.* (2001)].

Even though the estimation of the implied distributions at a *single* point in time has been studied thoroughly, we are not aware of any attempt to empirically explore *and* model *explicitly the stochastic evolution of the whole PDF through time*[1]. So far, researchers in Finance have been modeling the first two moments of the distribution, either in a static framework [e.g. a single factor standard Capital Asset Pricing Model, (CAPM)], or in a dynamic one e.g. Black and Scholes (1973) model, stochastic volatility option pricing models]. However, the empirical evidence shows that higher-order moments should be modeled, as well; Bates (1991, 1996a, 1996c), and Lynch and Panigirtzoglou (2001), find that the skewness and kurtosis of implied distributions change considerably over time. Recent studies [e.g., Fama and French (1992)] indicate that the CAPM cannot explain the cross-asset variation in expected returns, leaving space for introducing additional variables. Towards this direction, Harvey and Siddique (2000) develop an asset pricing model that incorporates the (conditional) skewness. Modeling the entire distribution is the most general approach one can take, so as to capture the effect of higher-order moments, as well.

Understanding and modeling the dynamics of implied distributions may allow us to address a number of issues in financial economics such as market microstructure questions, option pricing and risk management problems, and the effects of economic policy decisions. In particular, the evolution of implied PDFs can be viewed as a result of changes in the market participants information set. Hence, the implications of various microstructure models can be examined, e.g., Kyle's (1985) model implication that more information is revealed in

[1]So far, the stochastic evolution of implied distributions has been modeled only indirectly [see Bates (2002) for a discussion]. This is done via the equivalent methods of modeling implied volatilities [Ledoit and Santa-Clara (1998), Schönbucher (1999)], or local volatilities [Dupire (1992), Derman and Kani (1998)], or option prices [Britten-Jones and Neuberger (2000)]. Empirical studies such as Kamal and Derman (1997), Skiadopoulos *et al.* (1999), Alexander (2001), Ane´ and Labidi (2001), Fengler *et al.* (2001), and Cont and da Fonseca (2002) have also investigated the dynamics of implied distributions implicitly, by examining the properties of implied volatilities. On the other hand, Skiadopoulos and Hodges (2001) simulate the evolution of the implied distribution explicitly, without relying on any empirical evidence. Lynch and Panigirtzoglou (2001) study only the time series properties of the implied PDFs summary statistics.

the market as the volume increases.

Developing a theory about the dynamics of implied distributions can be helpful for "smile-consistent" option pricing and hedging when the moments of the distribution change in a *stochastic* way. "Smile-consistent" models (also called "implied models") reverse the traditional approach taken in option pricing [see Bates (2002) for a survey of the development of the approaches taken in option pricing, so far]. The traditional approach specifies a process for the underlying asset in advance, and then it develops the option pricing model. Models falling within this approach include the Black-Scholes (1973) model, stochastic volatility models, e.g., Hull and White (1987), jump diffusion models, e.g., Merton (1976), Bates (1996b), or a combination of stochastic volatility and jumps, e.g., Scott (1997), Bates (1996a). On the other hand, the implied approach takes the prices of traded plain vanilla options as given, and it uses them to price exotic and illiquid options. Implied distributions reflect the same information with the implied volatility smile. In addition, option prices can be obtained by integrating their payoffs over the risk-neutral distribution [Harrison and Kreps (1979), Harrison and Pliska (1981)]. Therefore, starting from today's distribution, the stochastic evolution of the implied distribution is equivalent to the consistent with the smile *stochastic evolution* of option prices[2].

Knowledge of the evolution of implied distributions is also useful for risk management purposes such as Value-at-Risk (VaR), and for economic policy decisions. VaR is the loss that corresponds to a given probability over a pre-specified time horizon. Implied distributions reveal the VaR that (risk-neutral) market agents expect to be realized over the distribution's horizon. Ait-Sahalia and Lo (2000) define as Economic-VaR (E-VaR) the calculated from the implied distribution VaR. They have proposed it as a more general measure of risk compared to the conventional VaR measures calculated from the actual distribution. This is because E-VaR incorporates investor's preferences. Therefore, it can account for the fact that two investors evaluate in a different way the same dollar loss, depending on their wealth [see also Lo (1999) for a discussion of the merits of such a measure]. Consequently, the dynamics of implied distributions reveal the variation of (risk-neutral) VaR over time. In the same spirit, the changes in implied distributions indicate the change in market consensus about the credibility of economic policy announcements, or the evolution of market participants expectations regarding the future level of economic variables such as interest rates [Söderlind and Svensson (1997)].

In this paper, initially we investigate the dynamics of three *constant*-maturity S&P 500 implied CDFs by applying Principal Components Analysis (PCA). We use a variety of criteria in order to determine the number of factors that explain the evolution of implied distributions through time. We obtain a parsimonious description of the dynamics of implied distributions

[2]Studies by Bakshi *et al.* (1997), and Das and Sundaram (1999) have shown that models falling within the traditional approach cannot account for the empirically observed smiles, nor for their evolution over time. As a response, the "smile-consistent" models have emerged. These fall within two categories: deterministic, and stochastic volatility "smile-consistent" models. The former category assumes that volatility is a deterministic function of the asset price and time. The latter assumes that volatility evolves stochastically. Dumas *et al.* (1998), and Buraschi and Jackwerth (2001) have shown that "smile-consistent" deterministic volatility models cannot forecast accurately the stochastic evolution of option prices over time. Hence, "smile-consistent" stochastic volatility models is the next natural candidate to capture the stochastic evolution of option prices over time [see Jackwerth (1999), and Skiadopoulos (2001) for a review of "smile-consistent" models].

since we identify two components that explain 89-92% of the variance depending on the maturity. Using a rotation technique we are able to interpret the two components in an intuitive way. The first component is interpreted as a shock to the location of the distribution. The second component affects the variance, and possibly the skewness and the kurtosis of the distribution.

Subsequently, the implied distribution is modeled as a diffusion process. We develop two new Monte Carlo (MC) simulation algorithms that model the evolution of the implied PDF and of the implied CDF over time, respectively. The presented methods are *arbitrage-free*. Our modeling approach is new. It can be regarded as the generalization of the conventional approaches of modeling only the first two moments of the implied distribution as diffusion processes. It differs from the approach suggested by Skiadopoulos and Hodges (2001) who first performed MC simulation of implied distributions. Their algorithm simulates the implied distribution by using mixture of distributions as a building block. They assume that sequential shocks affect the first two moments of the distribution. A mixture of distributions that captures the effect of a shock to a given moment is specified. The mixture is used as a model for today's distribution. Then, the simulated distribution emerges as a mapping from the future to the original distribution. In principle, Skiadopoulos and Hodges' model can be generalized to include shocks to higher-order moments of the distribution, as well. However, this comes at the cost of specifying additional and more complex forms of mixtures. In contrast, the use of a diffusion process to model implied distributions is a more flexible, and hence a more operational approach. Shocks to higher-order moments can be incorporated by simply adding extra factors to the process. In addition, the number and form of factors can be easily identified by PCA. This is not the case with a mixture of distributions model where it is not obvious how the assumed specification of the mixture can be tested empirically.

The proposed algorithms require as inputs only the initial implied distribution and the volatility shocks of the process (i.e. the PCA results). Modeling the dynamics of implied distributions by using the PCA results is analogous to implementing Heath-Jarrow-Morton (1992) type of models using PCA. In the context of smile-consistent option pricing with stochastic volatility, the simulation of the implied distribution is preferable to the simulation of local, or implied volatilities. This is because it avoids the problems associated with the calculation of the drift of the process of local volatilities [as it is the case in Derman and Kani's (1998) model], or those associated with the estimation of the implied volatility process [e.g. in Ledoit and Santa-Clara's (1998) model]. On the other hand, the algorithms can not be used for the pricing of American options [as it is the case with the conventional Monte Carlo simulation, Boyle (1979)], and for the simultaneous pricing of options having different expiries.

Finally, we provide an out-of-sample application of the algorithm that simulates the implied CDF, in a E-VaR framework. Using the historically estimated PCA factors, the algorithm is applied to forecast the variability of an extreme initial implied probability that corresponds to a given horizon's E-VaR level (probability E-VaR). The variability of a given S&P 500 probability VaR level over the next fourteen days in year 2001 is calculated, and 95% confidence intervals are constructed. We find that the constructed intervals forecast accurately the range within which the future probability E-VaR will lie. The proposed application can be viewed as a test of the out-of-sample performance of our model and it has important implications for the economics of option pricing. So far, the out-of-sample

performance of only deterministic volatility "smile-consistent" models has been examined by Dumas *et al.* (1998).

The remainder of the paper is structured as follows. First, we present the data set used, and explain the method we chose to extract the implied distributions. The metric under which implied distributions are measured, is discussed. In Section 3 we describe the PCA technique including why it was chosen, and how it will be applied to analyze the data. Next, the dynamics of the implied CDFs are explored. The correlations of the rotated components with the PDF's moments are calculated. In Sections 6 and 7, we develop the methods for the simulation of the implied PDF and CDFs, respectively. In Section 8, the out-of-sample application for VaR purposes is explained, and results are presented. In the last Section, we conclude with a brief summary, certain implications of this paper are discussed, and we suggest directions for future research.

2 The Data Set

2.1 Source Data Description

The raw dataset contains the daily settlement prices of the S&P 500 futures options and of the underlying futures contracts traded in the Chicago Mercantile Exchange (CME) for the period from 19/06/98 to 29/12/2001. We extracted from the option prices the daily-implied probabilities of the underlying future prices.

The CME S&P 500 options contract is an American style futures option; the underlying future is the CME S&P 500 futures contract. S&P 500 options trade with expiries on the same expiry dates as the underlying futures contract. These trade out to one year with expiries in March, June, September and December. In addition, there are monthly serial option contracts out to one quarter. Options expire on the 3rd Friday of the expiry month, as do the futures contracts in their expiry months. Options quotations used to compute PDFs are the settlement prices, and the associated price of the underlying is the settlement price of the S&P 500 futures contract maturing on, or just after, the option expiry date.

2.2 Extracting the Implied Distributions

We estimate the implied PDFs using the non-parametric method developed by Panigirtzoglou in previously unpublished work at the Bank of England and presented in Bliss and Panigirtzoglou (2002). The technique makes use of Breeden and Litzenberger (1978) non-parametric result, and it uses a natural spline to fit implied volatilities as a function of the deltas of the options in the sample. This method is chosen because it generates PDFs that are robust to quite significant measurement errors in the quoted option prices, as shown in Bliss and Panigirtzoglou (2002).

The development of the method is motivated by some earlier papers in the literature of estimating implied probability distributions. Breeden and Litzenberger (1978) showed that assuming that we observe option prices for a continuum of strikes, the second derivative of a European call price with respect to the strike price delivers the risk neutral PDF $\Phi_T(K)$,

i.e.

$$\frac{\partial^2 C}{\partial K^2} = e^{-r(T-t)}\Phi_T(K), \tag{1}$$

where C is the European call option price, and K is the strike price.

However, in practice, available option quotes do not provide a continuous call price function. To construct such a function we must fit a smoothing function to the available data. Following Shimko (1993), the interpolation is done in the implied volatility space. Implied volatilities are estimated from option prices by using the analytical quadratic approximation of Barone-Adesi and Whaley (BAW, 1987). This is an accurate and computationally efficient modification of the Black-Scholes formula that captures the early exercise premium of the American-style S&P 500 futures options. In addition, the implied volatility calculated via the BAW formula can be inserted in the Black-Scholes (1973) formula to calculate the European option prices [see BAW (1987) for a discussion]. Hence, equation (1) that was derived for European options can also be applied to our American option data set.

For the purposes of calculating implieds, the standard filtering constraints were imposed. Only at-the-money and out-of-the-money options were used because they are more liquid than in-the-money. Hence, measurement errors in the calculation of implied volatilities due to non-synchroneity [Harvey and Whaley, (1991)] and bid-ask spreads are less likely to occur because their delta is smaller than that of in-the-money options. Option prices that violate the monotonicity and convexity no-arbitrage property of the call/put pricing function [Merton (1973)] were discarded. Option prices with less than five working days to maturity were also excluded; these prices are excessively volatile as market participants close their positions. The three-month LIBOR interest rates taken from Bloomberg, are used for the calculation of implied volatilities. The choice of the interest rate does not introduce any noise in the calculation of implied volatilities because rho is very small for out-of-the-money options. Finally, we discarded options for which it was impossible to compute an implied volatility.

Panigirtzoglou method follows Malz (1997) in smoothing in implied volatility/delta space and Campa *et al.* (1997) in using a natural spline to fit a smoothing function to the transformed raw data. Using a delta rather than a strike metric groups the away-from-the-money implied volatilities more closely together than near-the-money implied volatilities; the delta of away-from-the-money options does not change significantly, as the strike price changes. This has the effect of permitting greater "shape" near the center of the distribution where the data is more reliable (frequently traded). The use of a natural spline rather than of a low order polynomial, as in Shimko (1993), permits the user to control the smoothness of the fitted function. The spline is also less restrictive of the shape the fitted function can assume.

The delta metric is constructed by converting strikes into their corresponding call deltas by using the at-the-money implied volatility[3]. Hence, a set of implied volatilities and cor-

[3]A small delta corresponds to a high strike (i.e. out-of-the-money calls), while a large delta corresponds to a low strike (i.e. in-the-money calls). The formula for transforming each strike into a delta is the Black's (1976) model delta

$$\delta(X) = \exp(-r\tau) \cdot \Phi((\ln(S/X) + 0.5\sigma_{atm}^2\tau)/(\sigma_{atm}\sqrt{\tau})),$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution. We use the at-the-money implied volatility so

responding deltas is constructed for each available contract. Implied volatilities of deltas greater than 0.99 or less than 0.01, are discarded. These volatilities correspond to far out-of-the-money call and put prices, which have generally low liquidity. An implied volatility smile is constructed if there are at least three implied volatilities, with the lowest delta being less than or equal to 0.25 and the highest delta being greater than or equal to 0.75. This ensures that the available strikes cover a wide range of the PDF available outcomes. In the case that the range of strikes does not spread along the required interval, no PDF is extracted.

The natural spline used is a weighted piece-wise cubic polynomial. It is determined by minimizing the following function

$$\min_{\Theta}\{\sum_{i=1}^{N}[w_i(IV_i - g(\delta_i, \Theta))]^2 + (1-\lambda)\int_0^1 g''(\delta;\Theta)^2 d\delta\}, \tag{2}$$

where N is the number of strike prices and IV_i is the observed implied volatility of the ith option in the cross-section. $g(\delta_i, \Theta)$ is the fitted to implied volatilities smoothing spline which is a function of the ith option's delta δ_i, and of the set of parameters Θ that define the smoothing spline; w_i is the weight applied to the ith option's fitted implied volatility error. The weight is set equal to the ith option's vega so as most weight is given to the fitting errors that occur for the (liquid) close to the-money options. The piecewise function $g(\delta, \Theta)$ has knot points at the observed deltas; it is constrained so as it has continuous first and second derivatives at the knot points.

The parameter λ $(0 \leq \lambda \leq 1)$ is a smoothing parameter that controls the trade-off between the goodness of fit of the fitted spline and its smoothness. The latter is measured by the integrated second derivative of the fitted implied volatility function. The smoothing parameter is set equal to 0.99 for all the contracts in our sample. This value of λ is chosen among others because it accommodates the trade-off between goodness-of-fit and smoothness of the estimated PDFs. The spline is linear outside the range of available deltas. We choose to extrapolate the spline horizontally outside the range of available deltas to prevent unreasonably high or low implied volatilities. The estimated implied PDFs are not sensitive to the horizontal extrapolation because the vega of far away from-the-money options is small.

Once the spline $g(x;\Theta)$ is fitted, 20,000 points along the function are converted back to option price/strike space using BAW approximation. The 20,000 points are selected to produce equally spaced strikes over the range where the PDF is significantly different from zero. This range varies every day with each cross-section, primarily as the price level of the underlying changes. Finally, we use the 20,000 call price/strike data points to numerically differentiate the call price function to obtain the estimated PDF for the cross-section.

2.3 Constructing Constant Horizon Implied Distributions

The S&P 500 option contracts have fixed expiry dates, that is, the time to maturity changes with time. This poses a problem to the purposes of our analysis. As the option contract approaches its expiry, the implied distribution will degenerate. We want to exclude this

as the ordering of deltas is the same as that of the strikes. Using the implied volatilities that correspond to each strike could change the ordering in the delta space, in cases where steep volatility skews are observed. This would result in generating volatility smiles with artificially created kinks.

"time-to-maturity" effect from studying the dynamics of implied PDFs. Hence, we construct *constant-horizon* PDFs by using the extracted implied PDFs [see also Coutant *et al.* (2001), for a similar approach].

The construction of constant-horizon probabilities is done as described in Clews, Panigirtzoglou and Proudman (2000). It is based on the non-parametric technique for estimating fixed expiry-date PDFs. We use the cubic spline method to construct an implied volatility smile as a function of delta, for each available contract. For each delta, we use the same cubic smoothing spline interpolation technique to interpolate across the implied volatilities of the different option futures contracts, in maturity space. From this interpolation we pick the value of the implied volatility that corresponds to the PDF's desired maturity (e.g. 3 months). In the case that the desired maturity does not fall within the range of traded maturities, we have to extrapolate. No extrapolation is performed if the desired maturity is more than 1.25 months apart from the shortest/longest traded maturities. This process is repeated for nine different deltas from 0.1 to 0.9. Hence, we construct an implied volatility smile as a function of delta for hypothetical options with the desired time-to-maturity. We then use the constructed implied volatility smile to generate the constant-horizon PDF using the same non-parametric method we used for extracting the fixed expiry-date PDFs.

We construct implied PDFs and CDFs for fixed maturities τ=1, 3 and 6 months. For the purposes of our analysis, implied PDFs and CDFs are measured in a call delta metric; this is constructed as described above. Hence, the state probabilities correspond to the cumulative probabilities $\pi_t \equiv \Pr ob[S\tau \leq X\tau(\delta)]$, where $X_T(\delta)$ corresponds to a strike of a particular delta δ with τ time-to-maturity. We have calculated cumulative probabilities $\pi_t(T, \delta)$ for strikes $X\tau(\delta)$ of 21 different deltas δ (0.01, 0.059,..., 0.5,...,0.941, 0.99).

The advantage of measuring implied distributions on a (call) delta metric is that the delta variables are constrained in the interval $[0, 1]$. Therefore, we can capture the entire area of the distribution by using a small number of variables, as opposed to using an alternative metric (e.g. a futures Profit/Loss metric). The reduction in the number of variables will increase the effectiveness of our subsequent Principal Components Analysis. In addition, the delta metric adjusts the changes in probabilities of the states $S\tau \leq X\tau(\delta)$ to changes in the level of the underlying asset, to changes in the volatility, and to changes in the time to maturity. Due to its latter characteristic, Kamal and Derman (1997) have employed the delta metric to analyze the dynamics of implied volatilities. For the same reason, currency option traders also quote option prices in terms of implied volatilities measured as a function of delta [see Malz (1997)].

Our chosen metric does not affect probabilities. Probabilities are unaffected by the choice of metric since any variable transformation leaves probabilities unchanged[4]. However, the analysis of the dynamics of implied distributions does depend on the chosen metric.

[4]Let the density $f_r(R)$ as measured on metric R, and the density $f_y(Y)$ as measured on a transformed metric $Y = g(R)$. The two measured on different metrics densities are related via the Jacobian, i.e.

$$f_y(Y) = f_r(R) \cdot \left| \frac{dR}{dg(R)} \right| \Leftrightarrow f_R(R) \cdot dR = f_y(Y) \cdot dY.$$

Therefore, densities change to ensure that probabilities $f_R(R) \cdot dR$ and $f_y(Y) \cdot dY$ are equal.

3 Principal Components Analysis and Implied Distributions

We outline the Principal Components Analysis (PCA) methodology used in this paper. PCA is used to explain the systematic behavior of observed variables, by means of a smaller set of unobserved latent random variables. Its purpose is to transform p correlated variables to an orthogonal set which reproduces the original variance-covariance structure (or correlation matrix).

PCA has been used in the interest rate literature to explore the dynamics of the yield curve and to provide alternative hedging schemes to the traditional duration hedge [see Litterman and Scheinkman (1991), Knez *et al.* (1994), among others]. Kamal and Derman (1997), Skiadopoulos *et al.* (1999), Alexander (2001), Ane´ and Labidi (2001), Fengler *et al.* (2001), and Cont and da Fonseca (2002) have applied PCA to investigate the dynamics of implied volatilities.

We apply PCA to decompose the correlation structure of first differences of implied CDFs. To achieve this, we measure the daily differences of implied CDFs across different levels of delta for each one of the three different fixed maturities. For example, one of our variables provides a time series of the first differences of implied CDFs that correspond to a delta of 0.5 for one-month maturity[5].

Denote time by $t = 1, \ldots, K$ and let p be the number of variables. Such a variable is a $(K \times 1)$ vector $\mathbf{x}$. The purpose of the PCA is to construct Principal Components (PCs hereafter) as linear combinations of the vectors $\mathbf{x}$, orthogonal to each other, which reproduce the original variance-covariance structure. The first PC is constructed to explain as much of the variance of the original p variables, as possible (maximization problem). The second PC is constructed to explain as much of the remaining variance as possible, and so on. The coefficients with which these linear combinations are formed are called the *loadings*. In matrix notation

$$\mathbf{Z} = \mathbf{XA}, \tag{3}$$

where $\mathbf{X}$ is a $(K \times p)$ matrix, $\mathbf{Z}$ is a $(T \times p)$ matrix of PCs, and $\mathbf{A}$ is a $(p \times p)$ matrix of loadings. The first order condition of this maximization problem results to

$$(\mathbf{X}'\mathbf{X} - l\mathbf{I})\mathbf{A} = \mathbf{0}, \tag{4}$$

where l_i are the Lagrange multipliers. From equation (4) it is evident that the PCA is simply the calculation of the eigenvalues l_i and the eigenvectors of the variance-covariance matrix $\mathbf{S} = \mathbf{X}'\mathbf{X}$. Furthermore, the variance of the *ith* PC is given by the *ith* eigenvalue, and the sum of the variances of the PCs equals the sum of the variances of the $\mathbf{X}$ variables.

When both variables and components are standardized to unit length, the elements of $\mathbf{A}'$ are correlations between the variables and PCs and they are called *correlation loadings* [see Basilevsky (1994) for more details]. If we retain $r < p$ PCs then

[5]We applied Principal Components Analysis to the "percentage futures Profit/Loss" metric $P/L{=}\frac{F_T - F_t}{F_t}$, where F_t is the futures price observed at time t, and S_T is the futures price at the option's maturity. However, the PCA on P/L gave rather noisy results compared to the delta metric analysis. Therefore, we only report the results obtained from applying PCA to the delta metric.

$$\mathbf{X} = \mathbf{Z}_{(r)}\mathbf{A}'_{(r)} + \boldsymbol{\varepsilon}_{(r)}, \tag{5}$$

where $\boldsymbol{\varepsilon}_{(r)}$ is a $(K \times p)$ matrix of residuals and the other matrices are defined as before having r rather than p columns. The percentage of variance of $\mathbf{x}$ which is explained by the retained PCs (*communality* of $\mathbf{x}$) is calculated from the correlation loadings. After retaining $r < p$ components, we look at equation (5) to examine the size of the communalities, and the meaning of the retained components.

PCA is a natural and parsimonious technique to identify the number and the interpretation of stochastic shocks that move implied distributions. It enables us to simplify the complex dynamics of implied distributions, by identifying their most important components. We investigate the dynamics of implied distributions by grouping them in three different constant maturities. This will shed light on whether the dynamics of distributions depend on the maturity we examine.

Next, we perform a separate PCA on the first differences of implied CDFs for each maturity. This is because we tested for unit-roots by applying standard Dickey-Fuller tests to time series of the levels of CDFs for a given constant-maturity and delta. We find that the null hypothesis of the existence of unit root cannot be rejected. In addition, we show in Table 1 the means (and their t-statistics) of the daily CDF differences across different deltas and maturities. We can see that the average constant-maturity CDF differences are not significantly different from zero. Hence, implied CDFs need to be differenced so as to remove the non-stationarity; Frachot *et al.* (1992) have shown that PCA is misleading when applied to non-stationary variables. We perform PCA on a data set over the period 19/06/98 to 29/12/2000. The data for year 2001 will be used for an out-of-sample application of the PCA (see Section 8).

To summarize, in our framework, each variable $\mathbf{x}$ is a time series collection of differences in implied CDFs for a given delta level and within a constant maturity. The first PC is the linear combination of these variables that contributes the most to explaining the overall variance of the changes in the implied distribution. The second PC is the linear combination (of the same variables) orthogonal to the previous one, which explains as much as possible of the remaining variance, and so on.

4 PCA on Implied CDFs

In this section, we investigate the dynamics of implied distributions by applying PCA to implied cumulative distribution functions. Working with CDFs rather than with probability density functions has the advantage that we can capture the entire PDF through a mapping from $[0, 1] \rightarrow [0, 1]$.

We decide on the number of components to be retained and we look at their interpretation. Earlier researchers have used a variety of rules of thumb to determine the number of components to be retained. For example, they keep the components corresponding to eigenvalues larger than the mean of all the eigenvalues (mean eigenvalue rule of thumb), or they

	1 month	*3 months*	*6 months*
CDF delta	**Mean**	**Mean**	**Mean**
0.01	-2.19E-04 (-0.156)	-1.26E-04 (-0.340)	-5.78E-05 (-0.252)
0.059	-2.38E-03 (-0.239)	-1.56E-03 (-0.379)	-8.00E-04 (-0.230)
0.108	-3.44E-03 (-0.242)	-2.31E-03 (-0.317)	-1.23E-03 (-0.187)
0.157	-2.71E-03 (-0.214)	-1.80E-03 (-0.222)	-9.67E-04 (-0.133)
0.206	-7.80E-04 (-0.099)	-3.76E-04 (-0.051)	-2.53E-04 (-0.039)
0.255	1.60E-03 (0.214)	1.38 E-03 (0.187)	7.01E-04 (0.099)
0.304	3.94E-03 (0.304)	3.12E-03 (0.351)	1.65E-03 (0.177)
0.353	5.79E-03 (0.319)	4.56E-03 (0.422)	2.43E-03 (0.208)
0.402	7.10E-03 (0.328)	5.61E-03 (0.455)	3.02E-03 (0.226)
0.451	7.86E-03 (0.337)	6.26E-03 (0.473)	3.39E-03 (0.237)
0.5	8.13E-03 (0.345)	6.51E-03 (0.478)	3.49E-03 (0.241)
0.549	7.93E-03 (0.347)	6.37E-03 (0.468)	3.39E-03 (0.237)
0.598	7.43E-03 (0.347)	5.98E-03 (0.448)	3.14E-03 (0.230)
0.647	6.60E-03 (0.342)	5.32E-03 (0.416)	2.72E-03 (0.213)
0.696	5.51E-03 (0.332)	4.44E-03 (0.376)	2.15E-03 (0.186)
0.745	4.22E-03 (0.319)	3.39E-03 (0.324)	1.51E-03 (0.149)
0.794	2.74E-03 (0.288)	2.14E-03 (0.242)	7.47E-04 (0.089)
0.843	1.12E-03 (0.177)	7.56E-04 (0.105)	-9.85E-05 (-0.014)
0.892	-6.00E-04 (-0.097)	-7.25E-04 (-0.119)	-1.03E-03 (-0.170)
0.941	-2.22E-03 (-0.248)	-2.17E-03 (-0.374)	-2.12E-03 (-0.343)
0.99	-2.48E-03 (-0.339)		

Table 1: Means and their t-statistics of the CDF differences across different deltas. Results are reported for One, Three, and Six-Months constant-maturities.

keep the components which explain 90% of the total variance[6]. As Basilevsky (1994) notes "-such practice is statistically arbitrary, and seems to be prompted more by intuitive concepts of practicality and "parsimony", than by probabilistic requirements of sample-population inference."

We determine the number of components to be retained by looking at a range of criteria. First, we apply Velicer's (1976) non-parametric criterion[7]. Next, working with components retained under this criterion, we look at the communalities. Finally, we look at the interpretation of the PCs. If any component appears to be noise then we prefer to reject it.

4.1 Number of Retained Principal Components and their Interpretation

Velicer proposes a non-parametric method for selecting nontrivial PCs, i.e. components which have not arisen as a result of random sampling, measurement error, or individual variation. His method is based on the partial correlations of the residuals of the PCs model, after $r < p$ components have been extracted. The criterion can be described as follows: The variance-covariance matrix of the residuals $\varepsilon_{(r)}$ in equation (5) is given by

$$\varepsilon'_{(r)}\varepsilon_{(r)} = \mathbf{X}'\mathbf{X} - \mathbf{A}_{(r)}\mathbf{A}'_{(r)}. \tag{6}$$

Let $\mathbf{D} = diag(\varepsilon'_{(r)}\varepsilon_{(r)})$. Then, $\mathbf{R}^* = \mathbf{D}^{-\frac{1}{2}}\varepsilon'_{(r)}\varepsilon_{(r)}\mathbf{D}^{-\frac{1}{2}}$ is the matrix of partial correlations of the residuals. If r^*_{ij} represents the ith row, jth column element of R^*, then the Velicer statistic is given by

$$f_r = \sum\sum_{i \neq j} \frac{r^{2*}_{ij}}{p(p-1)} = \sum\sum_{i=j} \frac{r^{2*}_{ij} - p}{p(p-1)}, \tag{7}$$

and lies in the interval 0 to 1. The behavior of f_r is that it is decreasing until a number r^* and then it increases again. Velicer suggests that $r = r^*$ should be the number of components to be retained.

Table 2 shows the retained PCs under Velicer's criterion when the PCA has been performed on implied CDFs. We can see that according to this criterion one component is retained for each maturity. The first two PCs explain on average 91% of the total variation of implied CDFs, across the three constant maturities.

We checked the communalities that the first PC itself explains. We compare these to the communalities that the first two components explain. We find that the first PC alone give

[6] Litterman and Scheinkman (1991), and Knez *et al.* (1994) apply PCA to the yield curve and they retain three components because these are explaining about 98% of the total variance. However, as Jackson (1991, page 44) notes "this procedure is not recommended. There is nothing sacred about any fixed proportion". For a description and discussion of the several rules of thumb, see Jackson (1991).

[7] Most of the tests used for determining the number of PCs to be retained, are parametric based on the assumption of multivariate normality [For a review of these tests, see Basilevsky (1994)]. However, application of Q-Q plots to our dataset, showed that the null-hypothesis of univariate, and hence of multivariate normality, was rejected. Therefore, we decided to use a non-parametric method.

	f_0	f_1	f_2	f_3	f_4	r^*	**1st PC**	**2nd PC**	**3rd PC**
1 month	0.46218	0.46192	0.46203	0.46214	0.46217	1	65.17%	23.70%	7.95%
3 months	0.45711	0.45696	0.45702	0.45708	0.45710	1	62.51%	30.12%	4.74%
6 months	0.44224	0.44211	0.44214	0.44221	0.44222	1	60.95%	30.87%	5.44%

Table 2: Principal Components from the PCA on Implied Cumulative Distributions: r* = the number of components retained under Velicer's criterion (minimum of f0,...f4).

rise to low communalities (almost zero) for some deltas. Including the second PC increases the explained communalities to more than 90% in most of the deltas. The communality criterion indicates that two components should be retained.

As a final step in our methodology for the number of components to retain, we look at the interpretation of the first three components. Figure 1 shows the correlation loadings of the first three PCs across the three maturities as these have been obtained from the PCA on implied CDFs. For middle deltas, the first PC has almost equal size correlation loadings. For very low and high deltas the size of the correlation loadings decreases. This attenuation is stronger for the nearest contract. These results imply that the first component moves the central part of the CDF in a parallel fashion. Hence, if the first shock has a positive sign, then it will accumulate more probability mass in the center of the CDF and less mass at the edges.

The second PC does not move the CDF is a parallel way; assuming that the shock has a positive sign, it accumulates more mass in the lower part of the CDF. In the middle part the mass reduces and then it increases again. The third PC has a rather noisy behavior. The correlation loadings of the first three PCs confirm that two components drive the dynamics of the implied CDFs.

Skiadopoulos *et al.* (1999) had explored the dynamics of the CME S&P 500 implied volatilities for 1992-1995. They found that two components move implied volatilities, as well. So far, it is well understood that at any *given point in time*, implied distributions, implied volatilities, and option prices convey the same information [see among others, Bahra (1997), Bates (1996b, 1996c), and Rubinstein (1985)]. "Fat" tailed implied distributions are mapped to implied volatility smiles; equivalently stated, the Black-Scholes (1973) option model underprices away from-the-money options, and overprices close to-the-money options. Our results show that the *evolution* of implied distributions and that of implied volatilities is affected by the same number of factors. However, the interpretation of the CDF obtained components is not intuitive. To obtain a simple interpretation we resort to a rotation.

4.2 Rotation of the Principal Components and their Interpretation

We use a rotation technique, so as to achieve a simple and intuitive interpretation of the retained components. This is a common procedure followed by researchers in the context of PCA, when the interpretation of the retained PCs is not apparent. Knez *et al.* (1994) and Skiadopoulos *et al.* (1999) apply various rotation techniques to obtain a clear interpretation of the factors driving money market returns, and implied volatilities, respectively. In this

section, first we describe the idea of a rotation. Then, we apply an appropriate for the purposes of our study rotation.

The first two components and implied distributions are related via

$$\mathbf{X} = \mathbf{ZP}',$$

where

$$\mathbf{P}' = \mathbf{L}^{-\frac{1}{2}}\mathbf{A}',$$

L being a diagonal matrix containing the eigenvalues that correspond to the retained components.

Hence, $\mathbf{P}$ is the matrix of loadings corresponding to unstandardized variables and components. Rotating the r retained orthogonal PCs to a new orthogonal position (orthogonal rotation), is equivalent to multiplying the loadings and PCs by an $(r \times r)$ orthogonal matrix $\mathbf{T}$ (transformation matrix) and its inverse $\mathbf{T}^{-1}$, i.e.

$$\mathbf{X} = \mathbf{Z}_{(r)}\mathbf{T}\mathbf{T}^{-1}\mathbf{P}'_{(r)} + \varepsilon = \mathbf{V}_{(r)}\mathbf{G}'_{(r)} + \varepsilon, \tag{8}$$

where

$$\mathbf{V}_{(r)} = \mathbf{Z}_{(r)}\mathbf{T}, \tag{9}$$

and

$$\mathbf{G}_{(r)} = \mathbf{P}_{(r)}\mathbf{T}. \tag{10}$$

Equations (9), and (10) deliver the new rotated PCs and eigenvectors, respectively. The rotated components still explain the same total amount of variance, as the unrotated ones, but the total variance may have been re-distributed between the two. Moreover, as a property of orthogonal rotations (rotations that maintain the axes at right angles) applied to loadings, the rotated eigenvectors remain orthogonal, but the rotated principal components do not [see Basilevsky (1994), Theorem 5.2, page 259].

The matrix $\mathbf{T}$ is not unique, since an infinite number of orthogonal rotations is possible. An additional criterion has to be introduced to fix the location of the axes. Depending on the criterion several rotation methods have been developed. We choose a "Procrustes" type rotation [see Jackson (1991)]. In a "Procrustes" type rotation, $\mathbf{T}$ is chosen so as the initially obtained correlation loadings fit a targeted interpretation.

Intuitively, we would expect that changes in implied distributions stem from changes in their moments [Skiadopoulos and Hodges (2001)]. Positive changes in the mean would shift distributions to the left (right). Changes in the variance and kurtosis would shrink distributions and make them leptokurtic. If the first PC had correlation loadings of equal size and of the same sign, then it would move the initial PDF to the left (or to the right, if the shock was negative). In such a case, the first PC could be interpreted as a shock to the mean. We could interpret the second PC as a shock to the variance, or to the kurtosis, if it's correlation loadings were increasing and then declining.

	Unrot. 1st PC	1st Rot. PC	2nd Rot. PC	Cumulative
1 month	65.17%	30.53%	58.34%	88.87%
3 months	62.51%	57.70%	34.93%	92.64%
6 months	60.95%	52.45%	39.37%	91.82%
Average	62.88%	46.89%	44.21%	91.11%

Table 3: Percentage of variance explained by the unrotated first Principal Component (PC), and by the rotated PCs in the CDF analysis.

Therefore, we use the Procrustes rotation to make the first PC to be close to a parallel shift, as possible[8]. We accomplish this by using the Procrustes type rotation developed by Skiadopoulos *et al.* (1999-see Appendix A for the construction of the technique). The technique is based on a regression that is performed so as to find the orthogonal rotation which minimizes the least squares distance between the loadings of the first PC and a vector of constants. We have no freedom on choosing how the second PC will look like. For the purposes of the rotation method, the loadings matrix **P** will be used (and not the correlation loadings matrix **A**-see Appendix A).

Figure 2 shows the loadings of the first and second rotated PC. The first PC has positive loadings of almost equal size with attenuation at the edges, being interpreted as a (almost) parallel shift of the CDF. This is more evident in the longer maturities. Hence, the first component can be associated with shocks to the mean of the distribution. The second PC can be described by a concave quadratic function: it has positive loadings for middle deltas and negative loadings for either end in the delta space. Therefore, the second shock affects the kurtosis and/or the variance of the distribution.

In Table 3 we show the percentage of variance that the first and second rotated PCs explain, as well as the percentage for the original first component. We see that the first shock dominates in the longest expiries, while the second factor explains more of the variance in the one month horizon.

5 Correlations of the rotated PCs with the PDF's moments

We calculate the correlations between the rotated PCs obtained from the PCA analysis on CDFs and the first four moments of the PDF. This could shed some additional light on the PCs interpretation. The moments have been calculated by measuring the implied PDF on the delta metric. We measure the skewness and kurtosis of the PDF using the risk-reversal and strangle measure, respectively, instead of using the "conventional" measures of

[8]The first PC cannot move CDFs in an exact parallel fashion, since this would make the perturbed PDF to integrate to more than one. Other rotation methods such as the varimax, the quartimax and the oblique, can not give us the desired intuitive interpretation [see Jackson (1991)]. This is the reason we choose to apply a Procrustes rotation.

	1 month		*3 months*		*6 months*	
	1st PC	**2nd PC**	**1st PC**	**2nd PC**	**1st PC**	**2nd PC**
mean	0.17**	0.12**	0.18**	0.06	0.18**	0.06
variance	0.10*	0.12**	0.10*	0.07	0.16**	0.10*
atm-vol	0.16**	-0.03	0.14**	-0.02	0.11**	-0.01
risk reversal	-0.05	0.12**	-0.06	0.05	-0.04	0.05
strangle	-0.05	0.05	-0.04	0.07	0.05	0.08*

Table 4: Correlations between the rotated Principal Components (PCs) obtained from PCA on CDFs, and the moments of the PDF.

the asymmetry and "fatness" of tails[9]. This is because the risk-reversal and strangle measures have been calculated on a delta metric, by construction [see also Malz (1997) for an extensive discussion of risk-reversal and strangle measures]. We also calculate the correlations between the rotated PCs and the at-the-money implied volatility (atm-vol).

The correlations are calculated using the Pearson and the non-parametric Spearman correlation coefficients. The results obtained are similar for both coefficients. Therefore, we report only the Pearson coefficient's results in Table 4. One and two asterisks indicates that the correlation is significant at 5% and 1%, respectively (2-tailed test)[10].

The first rotated PC is significantly positively correlated with the mean. An increase in the first component increases the mean of the PDF, that is to say it shifts the PDF to the right. The size and the sign of the correlation remains the same across the three maturities. This is in accordance with the interpretation of the first PC as a shock to the mean of the distribution. The first PC is also positively correlated with the at-the-money implied volatility and with the variance. This could indicate that the CDF's first factor is correlated with the first factor affecting implied volatilities. Studies on the dynamics of implied volatilities have shown that two or three factors govern the evolution of implied volatilities [see Kamal and Derman (1997), Skiadopoulos *et al.* (1999), Alexander (2001), Ane´ and Labidi (2001), Fengler *et al.* (2001), and Cont and da Fonseca (2002)]; the first factor moves implieds in a parallel fashion. An alternative explanation of the positive correlation between the first PC

[9]The risk-reversal provides a further measure of asymmetry that is independent of the shape of the tails of the implied PDF. It is calculated as the difference between the 25-delta call and 75-delta call implied volatilities. Hence, it reflects the slope of the volatility smile. A lognormal PDF has a risk reversal equal to zero. Therefore, the risk reversal shows the asymmetry of the implied PDF in excess of the benchmark lognormal PDF.

The strangle is a measure of the degree of "fatness" of tails of the implied PDF. It is independent of the shape of the tails of the PDF. It is obtained by taking the difference between the average of the 25- and 75-delta call implied volatilities and the at-the-money implied volatility. As such, it provides a measure of the degree of curvature of the volatility smile. A lognormal PDF has a strangle equal to zero. Therefore, the strangle shows the degree of "fatness" of tails of the implied PDF in excess of the benchmark lognormal PDF.

[10]Given that the rotated components deliver a clear interpretation, we use those in order to calculate the correlations. We calculate the rotated components using equation (9). We derive the unstandardized components $\mathbf{Z}$ from the standardized components $\mathbf{Z}^*$, used so far, via the relationship $\mathbf{Z}^* = \mathbf{Z}\mathbf{L}^{-\frac{1}{2}}$.

and the implied volatility/variance is due to the use of the delta metric. The mean of the PDF is measured in the delta space. Delta depends on volatility. Hence, the first PC that captures changes in the mean, has to be correlated with volatility. The correlation of the first PC with the risk reversal and with the strangle is insignificant.

The second rotated PC is significantly correlated only with some of the moments in the nearest and longest maturity. So, in the first month contract it is positively correlated with the mean, the variance, and with the risk reversal. In the six month maturity it is correlated positively with the variance and with the strangle. There is no significant correlation with any of the moments in the three months contract.

6 Simulating the Implied PDF: The Algorithm

In this section, we present a new algorithm for Monte-Carlo simulation of the implied PDF. The algorithm can be implemented easily by using the PCA results. In contrast to models that simulate the evolution of implied [e.g., Ledoit and Santa-Clara (1998)], or local volatilities [Derman and Kani (1998)], the proposed algorithm avoids the problem of estimating the drift of the process of the implied distribution. This is because in a no-arbitrage environment, implied probabilities have to follow a driftless process. In addition, the simulation of the implied distribution captures the effect from stochastic shocks to the higher-order moments. The algorithm for the Monte-Carlo simulation of the implied CDF is discussed in Section 7.

6.1 The General Idea

Let $\pi_t(S_T)$ be the risk-neutral probability as formed at current time t that the asset price S_T will be reached at time T, where $t < T$. Breeden and Litzenberger (1978) first showed that this probability can be extracted from a continuum, across strikes, of market call option prices observed at time t, $\forall S_T$.

If we knew what the implied risk neutral probability $\pi_{t+1}(S_T)$ would be in the next time step $t+1$ for $\forall S_T$, then we would be able to answer questions like the smile-consistent pricing and hedging of European style options, and how to forecast future VaR levels. We would also be able to address economic policy issues, such as the evolution of economic participants expectations over time.

Monte Carlo (MC) simulation appears as the natural tool for the implementation of the above idea. MC can be performed on the implied PDF, or equivalently on the implied CDF. However, the simulation has to be subject to certain constraints, depending on whether the simulation is performed on implied PDFs, or on implied CDFs. In the case that one chooses to simulate the implied PDF, the MC simulation needs to be performed in a way so as to guarantee (a) an arbitrage-free evolution of the implied probabilities, (b) that the simulated implied density integrates to one at every time step of the simulation, and (c) the simulated $\pi_t(S_T)$ is non-negative. In the case that one chooses to simulate the implied CDF, the simulation algorithm must ensure (a) an arbitrage-free evolution of the implied CDF, and (b) that every simulated point of the implied CDF is non-negative and less than one.

The suggested simulation method is easy to implement because it does not require calculating the drift term of the process. It requires as inputs only a starting value, and the number and form of shocks (volatility coefficients) of the process. The starting point is the extracted from the observed option prices implied distribution. The volatility coefficients of the process can be estimated by performing PCA on implied distributions.

6.2 The Arbitrage-Free Evolution of the Probabilities

Let the probabilities evolve obeying the widely used diffusion process

$$\frac{d\pi_t(X_T,T)}{\pi_t(X_T,T)} = a(X,T)dt + \sum_{j=1}^{r} f_{j,T}(X)dB_{j,t}, \tag{11}$$

where X is the metric that we work under, r is the number of factors, T is the PDF's expiry date, and B_j is the jth Brownian motion with $Cor(dB_i, dB_j) = 0$, for $i \neq j$. The notation indicates that the jth factor depends on T. Hence, in our framework $\pi_t(X_T,T)$ is the formed at time t probability that corresponds to a particular delta, $f_{j,T}(X)$ is the PCA correlation loading estimated for a particular delta and expiry date, and B_j is the jth retained principal component ($j = 1, 2$).

The simulation has to be carried out by requiring that the simulated probabilities are martingales under the risk-neutral measure. Hence, when their expectation is taken with respect to the equivalent martingale measure conditional on the information at time t, the following property holds

$$E_t^*[\pi_s(X_T,T)] = \pi_t(X_T,T) \tag{12}$$

for $t < s$, i.e. today's 30-day PDF/CDF is an unbiased predictor of tomorrow's 29-day PDF/CDF under the risk-neutral measure.

The martingale condition follows from the law of iterated expectations [see Skiadopoulos and Hodges (2001)]. Intuitively, we expect risk-neutral probabilities to behave as martingales since the forward price of any asset is a martingale under the risk-neutral measure; implied probabilities can be thought of as a forward Arrow-Debreu price. The martingale property ensures that any arbitrage opportunities are excluded [Harrison and Kreps (1979)]. The arbitrage-free property is particularly important in the case that the proposed algorithms are used for option pricing purposes; it ensures that the Merton's (1973) no-arbitrage results hold.

It follows that equation (11) has to be modified so as probabilities to evolve as martingales. The following Proposition shows the appropriate transformation.

Proposition 1 *The process*

$$\frac{d\pi_t(X_T,T)}{\pi_t(X_T)} = \sum_{j=1}^{r} f_{j,T}(X)dB_{j,t} \tag{13}$$

ensures that the simulated PDF evolves as a martingale.

Proof. (See Appendix B).

Proposition 1 delivers a driftless process (13) that can be used (in principle) to simulate the implied PDF. The simulation can be performed easily because we do not have to estimate the drift. The drift has to vanish so as to ensure that no-arbitrage opportunities arise. However, equation (13) has to be modified further so as to exclude the occurrence of negative simulated probabilities.

Notice that Proposition 1 holds only for *fixed expiry date* distributions. This is because the maturity date T needs to be fixed for PDFs/CDFs to follow a martingale. In contrast, constant-maturity distributions do not need to follow a martingale. In this case, the driftless process in equation (13) cannot be used. However, the empirical evidence in Section 3, shows that constant-maturity distributions evolve as martingales and therefore Proposition 1 can also be applied to constant-maturity CDFs. In this case, T corresponds to a particular constant horizon and $f_{j,T}(X)$ is the PCA correlation loading estimated for a particular delta and constant horizon.

6.3 Integration of Probabilities to One

The simulation algorithm has to ensure that the simulated PDF integrates to one at every time step of the simulation. The initial density at time $t = 0$, integrates to one[11]. However, this is not necessarily true for the simulated PDF for future times s $(0 < s < T)$. The random numbers used to perturb the initial density may produce a new density that integrates to more or less than one. We can accommodate the issue of the integration of the simulated PDF to one due to the linearity of equation (13).

Let a continuum of k probabilities $(k = 1, ..., +\infty)$ each one associated with the corresponding variable X_k. Then, we can discretize equation (13) as follows,

$$\pi_{t+1}(X,T) = \pi_t(X,T)(1 + \sum_{j=1}^{r} f_{j,T}(X) dB_{j,t}). \tag{14}$$

We need to ensure that $\sum_{k=1}^{\infty} \pi_{t+1}(X_k, T) = 1$, $\forall t$. This is equivalent to requiring the sum over k of the right hand side of equation (14) to be equal to one, i.e.

$$\sum_{k=1}^{\infty} \pi_t(X_k,T) + \sum_{k=1}^{\infty} \sum_{j=1}^{r} \pi_t(X_k,T) f_{j,T}(X_k) dB_{j,t} = 1. \tag{15}$$

[11] Differentiation of a call option price with respect to the strike price K yields $\frac{\partial C}{\partial K} |_{k=0} = -e^{-r\tau}$ and $\frac{\partial C}{\partial K} |_{k=\infty} = 0$. Since the implied risk-neutral density function $\Phi(S_T)$ is given by $\Phi(K) = e^{r\tau} \frac{\partial^2 C}{\partial K^2}$ we get that

$$\begin{aligned} \int_0^{\infty} \Phi(K) dK &= e^{r\tau} \int_0^{\infty} \frac{\partial^2 C}{\partial K^2} dK = e^{r\tau} [\frac{\partial C}{\partial K}]_0^{\infty} \\ &= e^{r\tau} [0 - (- - e^{-r\tau})] = 1 \end{aligned}$$

Since $\sum_{k=1}^{\infty} \pi_t(X_k, T) = 1$ (the initial density integrates to one), a sufficient condition to ensure that equation (15) is satisfied is

$$\sum_{k=1}^{\infty} \pi_t(X_k, T) f_{j,T}(X_k) = 0, \tag{16}$$

for $j = 1, ..., r$.

Equation (16) can be satisfied for appropriate choices of $f_{j,T}(X)$. However, we have no freedom in choosing $f_{j,T}(X)$. These are dictated by the data and the used estimation technique (e.g. PCA). We circumvent this problem by taking $f_{j,T}(X)$ as it is estimated. Then, equation (16) will not necessarily be satisfied, i.e.

$$\sum_{k=1}^{\infty} \pi_t(X_k, T) f_{j,T}(X_k) = \mu_j, \tag{17}$$

for $j = 1, ..., r$. μ_j is a random term since it depends on the simulated probabilities which are generated by random numbers.

We make the simulated PDF to integrate to one by re-adjusting the volatility coefficients using the μ_j. The following Proposition shows how to do this.

Proposition 2 *Given that the probability distribution integrates to one at the current time t it will also integrate to one in every time step of the simulation s, $t < s < T$, if we use the process*

$$\pi_{t+1}(X, T) = \pi_t(X, T)\{1 + \sum_{j=1}^{r} [f_{j,T}(X) - \mu_j] dB_{j,t}\}. \tag{18}$$

Proof. (See Appendix B).

The main advantage of the normalization depicted in equation (18) is that it preserves the martingale property[12]. If instead of the chosen normalization we had applied the obvious one $\frac{\pi_{t+1}(X,T)}{\sum_{k=1}^{\infty} \pi_{t+1}(X,T)}$, then the martingale property would not be preserved any longer. This stems from equations (15), and (16) which show that

$$\sum_{k=1}^{\infty} \pi_{t+1}(X_k, T) = 1 + \sum_{j=1}^{r} \mu_j,$$

where μ_j is a random term. Therefore, the denominator of the obvious normalization is not a constant. As a result, the rescaled probabilities would not necessarily be martingales[13].

Corollary 1 *In order to ensure non-negative probabilities we must have*

$$\sum_{j=1}^{r} [f_{j,T}(X, T) - \mu_j] dB_{j,t} \geqslant -1, \forall X. \tag{19}$$

[12]The preservation of the martingale property can be shown very easily by applying Ito's Lemma on the logarithm of $\pi_t(K)$.

[13]If a variable X is a martingale, and a is a constant number, then the variable $Y = X/a$ is a martingale, as well. However, if a is a random term, then the variable Y is not necessarily a martingale.

Proof. It follows directly from equation (18). ■

Using the *linear* stochastic differential equation (18) allows us to force the simulated density to integrate to one, and to accommodate for non-negative probabilities. This would not be the case if we had used the exponential equation (42)

However, Corollary 1 poses numerical problems to the implementation of the simulation. For every X, appropriate random numbers have to be generated so that Corollary 1 is satisfied. This would mean that different shocks occur *at the same time*, each one affecting different point of the PDF. However, in our analysis, we consider that at a certain point in time *one shock* occurs; it's impact on every point of the PDF is different, and it is quantified by the corresponding correlation loading. We circumvent this obstacle by performing the simulation using the solution of equation (18), i.e.

$$\pi_{t+\Delta t} = \pi_t \times \exp\{-\frac{1}{2}\sum_{j=1}^{r}[f_{j,T}(X) - \mu_j]^2\Delta t + \sum_{j=1}^{r}[f_{j,T}(X) - \mu_j]\Delta B_{j,t}\}. \tag{20}$$

The exponential term of equation (20) ensures that the simulated probabilities will be always positive.

7 Simulating the Implied CDF: The Algorithm

In this section, we develop a new algorithm for Monte Carlo simulation of the implied CDF. The simulation algorithm has to ensure that the simulated state cumulative probabilities lie between zero and one. In addition, the simulated probabilities have to evolve as a martingale as it was the case with the simulated implied PDF. On the other hand, the simulated CDF does not have to integrate to one.

The initial probabilities at time $t = 0$ fall between zero and one, by construction. However, this is not necessarily true for the simulated probabilities for future times s $(0 < s < T)$. To ensure that the simulated probabilities will lie in the interval $[0, 1]$ we simulate first a non-linear transformation of the probabilities, and then we switch to the corresponding simulated probability.

In particular, we define a new variable $y_t \equiv \Phi^{-1}(\pi_t)$, where $\Phi(\cdot)$ is the cumulative standard normal distribution function (the arguments T and X have been suppressed for simplicity). The variable y_t can be thought of as the strike price of a hypothetical normal distribution, corresponding to a given delta and maturity. Irrespective of the value that y_t gets, π_t will be between zero and one. Hence, we simulate the defined variable y_t and then we convert it to the corresponding value of π_t. This ensures that the simulated cumulative probabilities will lie between zero and one. The following Proposition derives the no-arbitrage process for y_t.

Proposition 3 *The variable y_t follows the process*

$$dy_t = \frac{1}{2}y_t\frac{\Phi(y_t)^2}{\phi(y_t)^2}(\sum_{j=1}^{r} f_{j,T}^2)dt + \frac{\Phi(y_t)}{\phi(y_t)}\sum_{j=1}^{r} f_{j,T}dB_{j,t}, \tag{21}$$

where $\phi(y_t)$ is the standard normal probability density function of y_t, and $\Phi(y_t)$ is the cumulative standard normal distribution.

Proof. (See Appendix B).

8 A Value-at-Risk Application

In this section, we provide an application of the CDF simulation algorithm in the framework of Value-at-Risk (VaR). We calculate VaR using the Economic-VaR (E-VaR) measure proposed by Ait-Sahalia and Lo (2000). E-VaR is defined similarly to the conventional VaR (Ait-Sahalia and Lo call the latter Statistical VaR, S-VaR); it is the percentile (loss) that corresponds to a given probability and time horizon. However, E-VaR is calculated from the risk-neutral distribution, while S-VaR is estimated from the actual distribution. E-VaR is a more general measure of the risks facing investors since it incorporates the investor's risk preferences, the demand-supply effects and the probabilities that correspond to extreme losses. Hence, it is determined by the three P's (Preferences, Prices, and Probabilities) that any complete risk management system should take into account [see Lo (1999)]. On the other hand, S-VaR is just a statistical evaluation of uncertainty. The ratio of S-VaR and E-VaR measures the market participants' risk aversion. S-VaR and E-VaR are identical in the special case that investors are risk-neutral.

Given an initial constant-maturity E-VaR, we use the CDF algorithm and the PCA results to forecast the VaR's variability over time by constructing confidence intervals. This is of great importance to a fund manager, or to a bank that wants to plan its future reserve deposits for VaR purposes.

Bliss and Panigirtzoglou (2001) have found that the risk-neutral PDF variability (e.g. E-VaR variability) is similar to that of the subjective PDF (e.g. S-VaR variability); their results are robust to different maturities and to different specifications of the utility function. Since our application focuses on VaR variability, we expect that the distinction between E-VaR and S-VaR is not important.

8.1 The Application

Conventionally, any measure of VaR is defined in terms of the loss associated with a particular probability level (*dollar VaR*). However, since our data set contains probabilities, we specify E-VaR in the *equivalent reverse* way, i.e. in terms of the probability that corresponds to a particular loss (e.g. the loss level that corresponds to a call delta of 0.941- *probability VaR*).

The application is structured as follows: let a sample of K observations of an extreme probability $\pi_t(\delta)$, $t = 0, 1, \ldots, K$. We start from an initially observed extreme probability and we simulate it for the next n days (n-simulation horizon) using equation (21) and the corresponding PCA constant-maturity correlation loadings as volatility coefficients. We store the *nth* simulated extreme probability. Then, we take as an initial point, the extreme $(n+1)$ probability of our sample and we repeat the simulation for the next n days; we store the $(2n+1)$ simulated probability. The simulation terminates at the $(K-n)$ observation. At the end, we have stored $(K-n)/n$ simulated extreme probabilities. We repeat the simulation exercise 1,000 times (runs). Finally, for each one of the stored simulated extreme probabilities, we calculate the standard deviation across the 1,000 runs.

The mean of each one of the stored probabilities across the 1,000 runs is the corresponding initial (today's) probability VaR, by construction of the algorithm (simulated probabilities evolve as martingales). In addition, the simulated probability E-VaR is distributed normally since we use Brownian shocks. Then, for a level of significance $a\%$, we can construct $(1-a)\%$

confidence intervals for the future probability VaR by adding/subtracting from the mean the appropriate number m of standard deviations, as follows

$$\pi_i(\delta) - m\sigma_{i+n} < \pi_{i+n}(\delta) < \pi_i(\delta) + m\sigma_{i+n}, \tag{22}$$

where $i = 0, n, \ldots, K-n$, σ_{i+n} is the standard deviation of the $i+n$ date probability E-VaR, and $\pi_{i+n}(\delta)$ is the $(i+n)$ date realized probability E-VaR, as a function of delta. For example, the algorithm predicts that the probability E-VaR to be realized after 14 days will fall between $\pm$ two standard deviations from the current probability E-VaR with probability 95%. In this respect, the calculated (simulated) standard deviation measures the probability E-VaR's variability.

In the case that the $(1-a)\%$ of the $(K-n)/n$ realized probabilities falls between the upper and lower confidence limits, the algorithm will forecast accurately the variation of tomorrow's E-VaR. These conclusions are based both on the historically estimated correlations loadings and on the way the algorithm is constructed.

8.2 Implementation Issues

We perform the application using the S&P 500 implied CDF data for the year 2001 (out-of-sample application). For each one of the three constant-horizon CDFs, we simulate the extreme probability of delta equal to 0.941 for a 14-days simulation horizon. Hence, the fund manager forecasts the variability of the probability E-VaR over the next fourteen days, every fourteen days in year 2001. The number of observations in our sample is 238. Therefore, for each simulation run we store sixteen values.

A potential limitation to assessing the implications of our application is the lack of traded away from-the-money strikes; this may hamper the accurate estimation of the left tail of the implied distribution to which the 0.941 call delta corresponds. However, in our analysis we have estimated the left tail of implied distributions using out-of-the-money puts. These options are traded even for deep out-of-the-money strikes. Figure 3 shows the average highest tradable delta strike for all the quarterly contracts with expiries from December 1998 to March 2001. The average is calculated over a six-month period prior to the expiry of each contract. We can see that our data set allows the accurate estimation of the distribution's left tail, since there are traded strikes even in the region between 0.97 and 0.99 delta.

In principle, the simulation is carried out via equation (21). However, implementation-wise, there are three issues we should take into account.

First, the factors $f_{j,t}^{PCA}(T,\delta)$ are extracted from applying PCA to the *changes* of cumulative probabilities. Therefore, they are related to the factors $f_{j,t}(T,\delta)$ in equation (13) via

$$f_{j,t}^{PCA}(T,\delta) = f_{j,t}(T,\delta)\pi_t(T,\delta) \Longrightarrow f_{j,t}^{PCA}(T,\delta) = f_{j,t}(T,\delta)\Phi(y_t).$$

Hence, equation (21) should be used as

$$dy_t = \frac{1}{2}y_t \frac{1}{\phi(y_t)^2}\left(\sum_{j=1}^{r}(f_{j,t}^{PCA})^2\right)dt + \frac{1}{\phi(y_t)}\sum_{j=1}^{r} f_{j,t}^{PCA} dB_{j,t}. \tag{23}$$

The simulation of the implied CDF can be performed by discretizing equation (23) using a standard Euler discretization scheme, i.e.

$$y_{t+1}(T,\delta) - y_t(T,\delta) = \frac{1}{2} y_t(T,\delta) \frac{1}{\phi(y_t(T,\delta))^2} \left(\sum_{j=1}^{r} (f_{j,t}^{PCA}(T,\delta))^2\right)\Delta t + \frac{1}{\phi(y_t(T,\delta))} \sum_{j=1}^{r} f_{j,t}^{PCA}(T,\delta)\Delta B_{j,t}. \quad (24)$$

The second issue is that the Brownian motions in equation (23) are orthogonal to each other. Therefore, the *unrotated* components (and not the rotated) should be used for the simulation. This is because the rotated components are not orthogonal, by construction of the Procrustes rotation (see Appendix A). Hence, the simulation should be performed using the correlation loadings (and not loadings). This brings forward the third issue. The correlation loadings have been estimated by performing PCA on the standardized variables $d\pi_t(\delta)$. However, the simulation refers to unstandardized variables. Therefore, in equation (24) the factors $f_{j,t}^{PCA}(T,\delta)$ should be scaled by multiplying them with the standard deviations $\sigma(\delta,T)$ of the corresponding variable $\delta(T)$.

8.3 Discussion of the Results

Figure 4 shows the standard deviations across the 1,000 simulation runs, for each one of the sixteen stored values. Results are shown for each one of the three constant-horizons. We can see that in general, the standard deviations are small. In addition, the standard deviation does not seem to vary between the sixteen points. We confirm this by testing the null hypothesis that the mean of the changes of standard deviations is equal to zero. Application of a standard t-test reveals that we cannot reject the null hypothesis. Notice that the one-month standard deviation is greater than the longer maturities deviation. This reflects the pattern of standard deviations observed historically in the in-sample analysis.

Next, we translate the probability E-VaR confidence interval to a dollar E-VaR confidence interval. For any given delta, we map the lower and upper confidence limit of the confidence interval (equation (22)) to the corresponding (strike) E-VaR level. The transformation is done via Black's (1976) delta. Figure 5 shows the spread in strike space of the S&P 500 probability E-VaR 95% confidence interval (i.e. upper confidence limit minus lower confidence limit) across the sixteen monitoring points, for each one of the three maturities. We can see that the six-months confidence interval is wider than the shorter horizon intervals. This reflects the increased dispersion of the implied probability density functions for longer maturities. However, the width of the confidence interval is similar for the one and three-months maturities. This indicates that the increased dispersion of the three-months densities relative to the one month density is offset by the higher standard deviation of the 0.941 delta probability for the one-month.

Figures 6, 7, and 8 provide a visual test of the predictive power of the constructed probability E-VaR confidence intervals for each one of the three maturities. The figures show the realized 0.941 delta probability E-VaR, and the simulation-based upper and lower

confidence limits of the 95% confidence interval across the monitoring dates (probabilities have been multiplied by 100).

We can see that the realized probability VaR falls within the constructed confidence interval for all but one the monitoring dates. The exception is the realized probability on the 17th September 2001 for the one and three-month's maturity. Inspection of the data reveals that this is due to the 11th of September crash that made the subsequent short horizon implied distributions more negatively skewed (the 17th September simulated confidence interval has as a starting point the 21st August implied distribution). However, the implied distribution for the six-months horizon was not affected by the crash, as much as the shorter horizon distributions. This reflects the (risk-neutral) investors' perception that the probability of a further decline after the crash was higher only for shorter horizons.

The fact that the simulated intervals "miss" one observation is expected. This is because these are 95% confidence intervals; we expect that almost one out of sixteen observations will lie out of these intervals. Therefore, the simulated confidence interval forecast accurately the range within the S&P 500 probability E-VaR will lie in year 2001. This holds regardless of the starting point.

9 Conclusions and Implications of the Research

We proposed and implemented a new approach to model the evolution of implied distributions over time. First, we investigated the dynamics of the S&P 500 implied cumulative distribution functions (CDFs), by applying Principal Components Analysis (PCA). The Bliss and Panigirtzoglou (2001) method has been used in order to extract the implied distributions. Our study has been performed for one, three, and six-month constant horizons implied distributions. The results provide important insights into the evolution of implied distributions and allow it's parsimonious modeling as a diffusion process. The second contribution of the paper is the development of methods to model the implied distribution as a diffusion process. Two new Monte Carlo simulation methods have been presented. These simulate the evolution of the implied PDF and that of the implied CDF over time, respectively. The algorithms forecast accurately the future S&P 500 Value-at-Risk (VaR) variability.

We considered three alternative criteria (Velicer's criterion, communalities, interpretation) to decide on the number of components driving implied distributions. We found that two are the components that account for the CDF evolution over time in each one of the three maturities. These explain on average about 90% of the CDF's variance across the three maturities. To obtain a clear interpretation of the effect of the two shocks on implied CDFs, we applied a "Procrustes"-type rotation. The application of the rotation on implied CDFs revealed that the first component affects the location, while the second component affects the variance and possibly the kurtosis/skewness of implied distributions.

The proposed simulation methods are arbitrage-free, and they can be implemented easily. They are based on a diffusion process that describes the evolution of the implied distribution. They require as inputs only the initial implied distribution, and the correlation loadings obtained from the PCA.

Finally, we have provided an application for VaR purposes by using the Economic-VaR (E-VaR) measure proposed by Ait-Sahalia and Lo (2000). E-VaR is specified in terms of

the probability that corresponds to a particular loss (probability E-VaR). Then, using the historically estimated PCA correlation loadings, we performed out-of-sample simulations to forecast the variability of today's S&P 500 probability E-VaR, over the next fourteen days in year 2001. The results show that the algorithm forecasts accurately the interval within which the S&P 500 probability E-VaR will lie in year 2001. These findings are robust to the choice of the VaR horizon, and to the choice of the initial probability VaR, as a starting point for the simulation. Furthermore, we expect that the choice of the VaR measure (E-VaR versus the traditional statistical VaR) doesn't affect the constructed probability bounds [see Bliss and Panigirtzoglou (2001)].

Our study has further implications in terms of both the empirical analysis results and the presented simulation methods. The PCA analysis supports Skiadopoulos and Hodges (2001) model that assumes that implied distributions change as a result of sequential shocks to its moments. Moreover, the number of shocks driving the S&P 500 implied CDFs is the same as the number of shocks driving the S&P 500 implied volatilities [Skiadopoulos *et al.* (1999)]. So far, it was well known that at any given point in time, implied distributions and implied volatilities are two equivalent ways to describe option pricing biases relative to the Black-Scholes prices. Our empirical analysis implies that the *dynamics* of implied distributions and those of implied volatilities are governed by the same number of factors.

In terms of the presented simulation methods, the simulation of the risk-neutral distributions is the natural choice for "smile-consistent" option pricing in continuous time. This is because it avoids the problems associated with estimating the drift of the process under consideration, encountered in using a local, or implied volatility process. In an arbitrage-free setting, the risk-neutral distribution has to evolve as a martingale, i.e. it is driftless. The accuracy of the algorithms in forecasting the VaR's variability indicates that the simulation of implied distributions may be a very useful tool to risk managers.

Finally, the proposed application can be regarded as a test of the out-of-sample performance of our model. Dumas *et al.* (1998) have examined the out-of-sample performance of "smile-consistent" deterministic volatility models. They performed their study by considering the stability of implied deterministic volatility functions for the S&P 500 index. They found that *deterministic* volatility models cannot predict the evolution of option prices. Our results indicate that a more general "smile-consistent" model where all the moments change *stochastically* may be the appropriate approach to capture the dynamics of option prices.

This paper creates three strands for future research. First, the empirical issue of the dynamics of implied distributions should be explored further by using different data sets and by examining different time periods. It may be the case that the number and shape of shocks that affect implied distributions, depend on the underlying asset, just as it is the case with the dynamics of implied volatilities. For example, Kamal and Derman (1997) found that three principal components affect the dynamics of the S&P 500 index options, and the dynamics of the Nikkei 225 index options implied volatilities, over the period 1994-1997. On the other hand, Skiadopoulos *et al.* (1999) found that two principal components drive the dynamics of the S&P 500 futures options implied volatilities, over the period 1992-95. Alternative variants of PCA models may also be employed. The decomposition assumed in this paper is based on the assumption of a linear relationship between probabilities, and the set of principal components. The application of non-linear PCA models deserves to become a topic for future research.

Second, the economic implications of the empirical evidence on the dynamics of implied distributions should be investigated. For example, the issue of complete markets can be addressed by examining whether the two components are consistent with the complete markets hypothesis. Buraschi and Jackwerth (2001) have tested formally whether the S&P 500 market is complete over the period 1986-1995. Their analysis is performed on the returns of at-the-money and away-from-the-money options. They found that the market was incomplete with the evidence being stronger in the post-crash period. Their results imply that stochastic rather than deterministic volatility models should be used for option pricing and risk management purposes. Similar tests can also be applied to implied distributions. We would expect that the implied distribution analysis results would also reject the completeness hypothesis for the S&P 500 market. This is because option prices and implied distributions convey the same information.

Finally, researchers should look into the way that the suggested algorithms can be used to various applications. For economic policy purposes, the application of the algorithms requires converting the simulated risk-neutral distribution to the corresponding subjective distribution, i.e. it requires the estimation of risk-aversion. Ait-Sahalia and Lo (2000), Bliss and Panigirtzoglou (2001), Jackwerth (2000), and Rosenberg and Engle (2002) have suggested alternative methodologies for such an estimation. For option-pricing purposes, future research should assess the forecasting performance of the algorithms, i.e. whether they predict accurately the evolution of option prices over time. By construction, the algorithm predicts that on average future option prices should be equal to today's option price. However, the algorithm allows the construction of (arbitrage) bounds option prices. This can be particularly useful for the construction of bounds to the prices of exotic options, complementing recent work by Hodges and Neuberger (2001). Also, trading strategies that aim to profit from changes in higher-order moments (e.g. skewness and kurtosis trades) can be designed [see Ait-Sahalia *et al.* (2001) for such strategies]. Alternative processes such as a jump-diffusion process [e.g., Merton (1976), Bates (1996b)] for the evolution of implied distributions should also be considered. This will be necessary so as to capture any discontinuities in the underlying asset price.

References

[1] Ané, T., and C. Labidi, 2001, "Implied Volatility Surfaces and Market Activity over Time," *Journal of Economics and Finance*, 25, 259-275.

[2] Alexander, C., 2001, "Principal Component Analysis of Volatility Smiles and Skews," working paper, ISMA, University of Reading.

[3] Ait-Sahalia, Y and A. W. Lo, 2000, "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics* 94, 9-51.

[4] Ait-Sahalia, Y., Y. Wang, and F. Yared, 2001, "Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset?," *Journal of Econometrics,* 102, 67-110.

[5] Bahra, B., 1997, "Implied Risk-Neutral Probability Density Functions From Option Prices: Theory and Application," Working Paper 66, Bank of England, London.

[6] Bakshi, G., C. Cao, and Z. Chen, 1997, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003-2049.

[7] Barone-Adesi, G., and R. E. Whaley, 1987, "Efficient Analytical Approximation of American Option Values," *Journal of Finance*, 42, 301-320.

[8] Basilevsky, A., 1994, *Statistical Factor Analysis and Related Methods: Theory and Applications*, Wiley Series in Probability and Mathematical Statistics.

[9] Bates, D. S., 1991, "The Crash of '87: Was it Expected? The Evidence from Options Market," *Journal of Finance*, 46, 1009-1044.

[10] Bates, D. S., 1996a, "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69-107.

[11] Bates, D. S., 1996b, "Dollar Jump Fears, 1984-1992: Distributional Abnormalities Implicit in Currency Futures Options," *Journal of International Money and Finance*, 15, 65-93.

[12] Bates, D. S., 1996c, "Testing Option Pricing Models," in G.S., Maddala, and C. R. Rao, (ed.) *Statistical Methods in Finance, Handbook of Statistics*, 14, 567-611, Elsevier, Amsterdam .

[13] Bates, D. S., 2002, "Empirical Option Pricing: A Retrospection," Working Paper, University of Iowa; forthcoming in *Journal of Econometrics.*

[14] Black, F., and M. Scholes., 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81, 637-654.

[15] Black, F., 1976, "The Pricing of Commodity Contracts," *Journal of Financial Economics*, 3, 167-179.

[16] Bliss, R., and N. Panigirtzoglou, 2002, "Testing the Stability of Implied Probability Density Functions," *Journal of Banking and Finance,* 26, 381-422.

[17] Bliss, R., and N. Panigirtzoglou, 2001, "Recovering Risk Aversion from Options," Working Paper WP-2001-15, Federal Reserve Bank of Chicago.

[18] Boyle, P. P. (1977). "Options: a Monte Carlo approach", *Journal of Financial Economics*, 4, 323-338.

[19] Breeden, D., and R. Litzenberger, 1978, "Prices of State-Contingent Claims Implicit in Option prices," *Journal of Business*, 51, 621-651.

[20] Britten-Jones, M., and A. Neuberger, 2000, "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 55, 839-866.

[21] Buraschi, A., and J. C. Jackwerth, 2001, "The Price of a Smile: Hedging and Spanning in Option Markets," *Review of Financial Studies*, 14, 495-527.

[22] Campa, J. M., P. H. K. Chang, and R. Reider, 1998, "Implied Exchange Rate Distributions: Evidence from OTC Option Markets," *Journal of International Money and Finance*, 17, 117-160.

[23] Clews, R., N. Panigirtzoglou, and J. Proudman, 2000, "Recent Developments in Extracting Information form Option Markets," *Bank of England Quarterly Bulletin*, 40, 50-60.

[24] Cont, R., 1997, "Beyond Implied Volatility," in J., Kertesz, and I. Kondor (ed.) *Econphysics*, Dordrecht, Kluwer.

[25] Cont, R., and J., da Fonseca, 2002, "Dynamics of Implied Volatility Surfaces," *Quantitative Finance*, 2, 45-60.

[26] Coutant, S., E. Jondeau, and M. Rockinger, 2001, "Reading PIBOR Futures Options Smiles: The 1997 Snap Election," *Journal of Banking and Finance*, 25, 1957-1987.

[27] Das, S., and R. Sundaram, 1999, "Of Smiles and Smirks: A Term-Structure Perspective," *Journal of Financial and Quantitative Analysis*, 34, 211-230.

[28] Dumas, B., J. Fleming, and R. E. Whaley, 1998, "Implied Volatility Functions: Empirical Tests," *Journal of Finance*, 53, 2059-2106.

[29] Derman, E., and I. Kani, 1998, "Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility," *International Journal of Theoretical and Applied Finance*, 1, 61-110.

[30] Dupire, B., 1992, "Arbitrage Pricing with Stochastic Volatility," working paper, Societe Generale Division Options, Paris.

[31] Fama, E., and K. French, 1992, "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427-465.

[32] Fengler, M. R., W. K. Härdle, and C. Villa, 2001, "The Dynamics of Implied Volatilities: A Common Principal Components Approach," Working Paper 38, Humboldt University, Berlin, June.

[33] Gemmill, G., and A. Saflekos, 2000, "How Useful are Implied Distributions? Evidence from Stock-Index Options," *Journal of Derivatives*, 7, 83-98.

[34] Harrison, M. J., and D. M. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20, 381-408.

[35] Harrison, M. J., and S. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Applications*, 11, 215-260.

[36] Harvey, C. R., and R. E. Whaley, 1991, "S&P 100 Index Option Volatility," *Journal of Finance*, 46, 1551-1561.

[37] Harvey, C. R., and A. Siddique, 2000, "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 1263-1295.

[38] Heath, D., Jarrow, R., and A. Morton, 1992, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology For Contingent Claims Valuation", *Econometrica*, 60, 77-105.

[39] Hodges, S. D., and A. Neuberger, 2001, "Rational Bounds on Exotic Options," working paper, Financial Options Research Centre, University of Warwick.

[40] Hull, J. C., and A. White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance*, 42, 281-300.

[41] Jackson, E., 1991, *A User's Guide to Principal Components*, Wiley Series in Probability and Mathematical Statistics.

[42] Jackwerth, J. C., 1999, "Implied Binomial Trees: A Literature Review," *Journal of Derivatives* 7, 66-82.

[43] Jackwerth, J. C., 2000, "Recovering Risk Aversion from Option Prices and Realized Returns," *Review of Financial Studies* 13, 433-451.

[44] Kamal, M., and E. Derman, 1997, " The Patterns of Change in Implied Index Volatilities," Quantitative Strategies Research Notes, Goldman Sachs, New York.

[45] Knez, P. J., R. Litterman, and J. Scheinkman, 1994, "Explorations into Factors Explaining Money Market Returns," *Journal of Finance* 49, 1861-1882.

[46] Kyle, A. S., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 51, 1315-1335.

[47] Ledoit, O., and P. Santa-Clara, 1998, "Relative Pricing of Options with Stochastic Volatility," working paper, University of California, Los Angeles.

[48] Litterman, R., and J. Scheinkman, 1991, "Common Factors Affecting Bond Returns," *Journal of Fixed Income* 1, 54-61.

[49] Lo, A. W., 1999, "The Three P's of Total Risk Management," *Financial Analysts Journal*, 55, 13-26.

[50] Malz, A., 1997, "Estimating the Probability Distribution of Future Exchange Rates from Option Prices," *Journal of Derivatives* 5, 18-36.

[51] Mayhew, S., 1995, "On Estimating The Risk-Neutral Probability Distribution Implied by Option Prices," working paper, University of California, Berkeley.

[52] Merton, R. C., 1973, "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141-183.

[53] Merton, R. C., 1976, "Option Pricing when Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3, 125-144.

[54] Rosenberg, J., and R. F. Engle, 2002, "Empirical Pricing Kernels," Working Paper Fin-99-014, Leonard N. Stern School of Business, New York University; forthcoming in *Journal of Financial Economics*, 64.

[55] Rubinstein, M., 1985, "Non-Parametric Tests of Alternative Option Pricing Models using all Reported Trades and Quotes on the 30 most Active CBOE Option Classes from August 23, 1976 through August 31, 1978," *Journal of Finance* 40, 455-480.

[56] Schönbucher, P. J., 1999, "A Market Model of Stochastic Implied Volatility," *Philosophical Transactions of the Royal Society*, Series A, 357, 2071-2092.

[57] Scott, L. O., 1997, "Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Application of Fourier Inversion Methods," *Mathematical Finance* 7, 413-426.

[58] Shimko, D., 1993, "Bounds of Probability," *Risk*, 6, 33-37.

[59] Skiadopoulos, G., S.D. Hodges, and L. Clewlow, 1999, "The Dynamics of the S&P 500 Implied Volatility Surface," *Review of Derivatives Research*, 3, 263-282.

[60] Skiadopoulos, G., 2001, "Volatility Smile Consistent Option Models: A Survey," *International Journal of Theoretical and Applied Finance* 4, 403-437.

[61] Skiadopoulos, G., and S. D. Hodges, 2001, "Simulating the Evolution of the Implied Distribution," *European Financial Management Journal* 7, 497-521.

[62] Söderlind, P., and L. Svensson, 1997, "New Techniques to Extract Market Expectations from Financial Instruments," *Journal of Monetary Economics*, 40, 383-429.

Appendix A: Construction of the Procrustes Rotation Method

Denote by $\mathbf{g}_1, \mathbf{g}_2$, the $(p \times 1)$ vectors of rotated *loadings* of the first and second retained PCs, respectively. Let the transformation matrix $\mathbf{T} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$. Then, equation (10) delivers the rotated loadings as linear combinations of the unrotated ones, i.e.

$$\mathbf{g}_1 = a_1\mathbf{p}_1 + b_1\mathbf{p}_2, \tag{25}$$

and

$$\mathbf{g}_2 = a_2\mathbf{p}_1 + b_2\mathbf{p}_2, \tag{26}$$

where $\mathbf{p}_i$ is the $(p \times 1)$ vector of loadings of each component, for $i = 1, 2$. The inner product $< \mathbf{g}_1, \mathbf{g}_2 >$ is given by

$$\begin{aligned} & < \mathbf{g}_1, \mathbf{g}_2 > = < a_1\mathbf{p}_1 + b_1\mathbf{p}_2, a_2\mathbf{p}_1 + b_2\mathbf{p}_2 > = \\ a_1a_2 \; & < \mathbf{p}_1, \mathbf{p}_1 . > + a_1b_2 < \mathbf{p}_1, \mathbf{p}_2 > + b_1a_2 < \mathbf{p}_2, \mathbf{p}_1 > + b_1b_2 < \mathbf{p}_2, \mathbf{p}_2 > . \end{aligned} \tag{27}$$

Given that the normalization constraint $\mathbf{P}'\mathbf{P} = \mathbf{I}$ has been imposed on the construction of the PCA , we have that

$$< \mathbf{p}_1, \mathbf{p}_1 > = < \mathbf{p}_2, \mathbf{p}_2 > = 1, \tag{28}$$

and

$$< \mathbf{p}_1, \mathbf{p}_2 > = < \mathbf{p}_2, \mathbf{p}_1 > = 0, \tag{29}$$

[Basilevsky (1994), Theorem 3.1, page 103]. Hence, equation (27) combined with equations (28) and (29) yields

$$< \mathbf{g}_1, \mathbf{g}_2 > = a_1a_2 + b_1b_2. \tag{30}$$

In the case that an orthogonal rotation is applied to loadings, rotated eigenvectors remain orthogonal [Basilevsky (1994), Theorem 5.2, page 259]. This is not the case if the rotation is applied to correlation loadings, since rotated correlation loadings are not orthogonal [Basilevsky (1994), Theorem 5.3, page 261]. Thus, $< \mathbf{g}_1, \mathbf{g}_2 > = 0$. Then, equation (30) yields

$$a_1a_2 + b_1b_2 = 0. \tag{31}$$

A Procrustes type rotation allows choosing the coefficients forming the linear combinations in equations (25) and (26), so as to achieve a pre-defined interpretation. We choose them, so as to get the parallel shift interpretation for the first PC; our method does not guarantee a priori the shape of the second PC.

We perform an Ordinary Least Squares (OLS) regression of a vector of constants on $\mathbf{p}_1$, and $\mathbf{p}_2$ so as to get the parallel shift interpretation for the first PC. This regression will deliver a_1 and b_1. To calculate a_2 and b_2 from equation (31), we need one more equation; the following Proposition presents it.

Proposition 4 *A sufficient condition that the cumulative variance of the unrotated PCs equals the cumulative variance of the rotated PCs is that*

$$a_2^2 + b_2^2 = 1, \tag{32}$$

and the estimated by the OLS regression a_1 and b_1 to be standardized, to a_1^* and b_1^*

where

$$a_1^* = \frac{a_1}{\sqrt{a_1^2 + b_1^2}},$$

and

$$b_1^* = \frac{b_1}{\sqrt{a_1^2 + b_1^2}},$$

so as $a_1^{*2} + b_1^{*2} = 1$.

Proof.

Let $\mathbf{z}_i$, and $\mathbf{v}_i$ be the $(T \times 1)$ vectors of the unrotated and rotated PCs, respectively, for $i = 1, 2$. Then, from equation (9) we have that

$$\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{z}_1 & \mathbf{z}_2 \end{pmatrix} \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = \begin{pmatrix} \mathbf{z}_1 a_1 + \mathbf{z}_2 b_1 & \mathbf{z}_1 a_2 + \mathbf{z}_2 b_2 \end{pmatrix}.$$

Since, by the construction of the PCA, the unrotated PCs are orthogonal, we get that the variance of $\mathbf{v}_i$ is given by

$$Var(\mathbf{v}_1) = Var(\mathbf{z}_1)a_1^2 + Var(\mathbf{z}_2)b_1^2, \tag{33}$$

and

$$Var(\mathbf{v}_2) = Var(\mathbf{z}_1)a_2^2 + Var(\mathbf{z}_2)b_2^2. \tag{34}$$

Since $Var(\mathbf{z}_i) = l_i$, the cumulative variance of the unrotated PCs will equal the cumulative variance of the rotated PCs, if

$$(a_1^2 + a_2^2)l_1 + (b_1^2 + b_2^2)l_2 = l_1 + l_2. \tag{35}$$

Equation (35) is true if $a_1^2 + b_1^2 = 1$ and $a_2^2 + b_2^2 = 1$.

■

Solving equations (31) and (32) yields

$$a_2 = \pm\sqrt{b_1^{*2}}, \tag{36}$$

and

$$b_2 = \pm\sqrt{a_1^{*2}}, \tag{37}$$

where a_2, and b_2 are chosen, so that the condition which results from equation (31), is respected, i.e.

$$b_2 = \frac{-a_1 a_2}{b_1}. \tag{38}$$

Appendix B: Proof of Propositions

Proof of Proposition 1.

The solution of the standard stochastic differential equation (11) is

$$\pi_t(X_T) = \pi_0(X_T) \exp\{[a - \frac{1}{2}\sum_{j=1}^{r} f_{j,T}(X)^2]t + \sum_{j=1}^{r} f_{j,T}(X)B_{j,t}\}, \tag{39}$$

where the argument T has been suppressed for simplicity.

For $\pi_t(X_T)$ to be a martingale, the exponent term in equation (39) has to be equal to one, i.e.

$$\exp\{[a - \frac{1}{2}\sum_{j=1}^{r} f_{j,T}(X)^2]t + \sum_{j=1}^{r} f_{j,T}(X)B_{j,t}\} = 1. \tag{40}$$

Equation (40) holds if and only if its left hand side is a martingale. In general, in order to make an exponent term a martingale we need to find a function $g(t)$, such that

$$L = \exp[Y_t + g(t)]$$

becomes a martingale, where

$$Y_t = [a - \frac{1}{2}\sum_{j=1}^{r} f_{j,T}(X)^2]t + \sum_{j=1}^{r} f_{j,T}(X)B_{j,t}.$$

Applying Ito's lemma on $L = L(Y_t, t)$ yields

$$dL = [g^{'}(t) + a]Ldt + \sum_{j=1}^{r} f_{j,T}(X)dB_{j,t}L. \tag{41}$$

Equation (41) shows that L becomes a martingale if and only if $g^{'}(t) = -a$, i.e. $g(t) = -at+c$ (c being a constant). Consequently, $\pi_t(X_T)$ is a martingale if and only if

$$\pi_t(X_T) = \pi_0(X_T)\exp\{[a - \frac{1}{2}\sum_{j=1}^{r} f_{j,T}(X)^2]t + \sum_{j=1}^{r} f_{j,T}(X)dB_{j,t} - at + c\},$$

which simplifies to

$$\pi_t(X_T) = \pi_0(X_T)\exp\{-\frac{1}{2}\sum_{j=1}^{r} f_{j,T}(X)^2 t + \sum_{j=1}^{r} f_{j,T}(X)dB_{j,t}\}, \tag{42}$$

assuming that $g(0) = 0$. It can be verified that equation (42) is the solution of equation (13). ■

Proof of Proposition 2.

Applying the summation operator in both sides of equation (18) across the continuum of the k variables, and using equation (17) we get

$$\sum_{k=1}^{\infty}\pi_{t+1}(X_k, T) = \sum_{k=1}^{\infty}\pi_t(X_k, T) + \sum_{j=1}^{r}[\sum_{k=1}^{\infty}\pi_t(X_k, T)f_{j,T}(X_k) - \mu_j\sum_{k=1}^{\infty}\pi_t(X_k, T)]dB_{j,t} = 1.$$

■

Proof of Proposition 3.

Applying Ito's lemma to $y_t = \Phi^{-1}(\pi_t)$ and using equation (13) yields

$$dy_t = \frac{dy_t}{d\pi_t}\pi_t \sum_{j=1}^{r} f_{j,T} dB_{j,t} + \frac{1}{2}\frac{d^2 y_t}{d\pi_t^2}\pi_t^2 (\sum_{j=1}^{r} f_{j,t}^2)dt, \tag{43}$$

where $\pi_t = \Phi(y_t)$.

Next, we derive closed-form expressions for $\frac{dy_t}{d\pi_t}$ and $\frac{d^2 y_t}{d\pi_t^2}$ by differentiating the transformation equation $\pi_t = \Phi(y_t)$ with respect to π_t, i.e.

$$1 = \phi(y_t)\frac{dy_t}{d\pi_t} \Longrightarrow \frac{dy_t}{d\pi_t} = \frac{1}{\phi(y_t)}, \tag{44}$$

where $\phi(\cdot)$ is the standard normal density.

Differentiating equation $1 = \phi(y_t)\frac{dy_t}{d\pi_t}$ with respect to π_t yields

$$0 = \frac{d\phi(y_t)}{dy_t}\frac{dy_t}{d\pi_t}\frac{dy_t}{d\pi_t} + \phi(y_t)\frac{d^2 y_t}{d\pi_t^2} \Longrightarrow 0 = -\phi(y_t)y_t(\frac{dy_t}{d\pi_t})^2 + \phi(y_t)\frac{d^2 y_t}{d\pi_t^2} \Longrightarrow$$

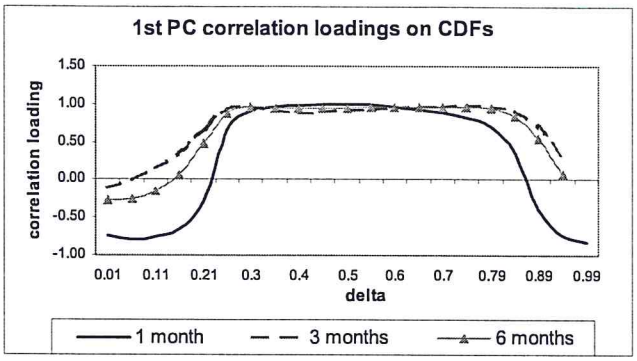
$$\frac{d^2 y_t}{d\pi_t^2} = y_t(\frac{dy_t}{d\pi_t})^2 \Longrightarrow$$

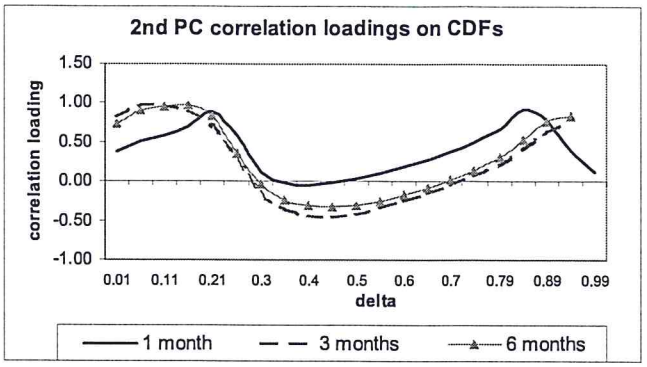
$$\frac{d^2 y_t}{d\pi_t^2} = y_t\frac{1}{\phi(y_t)^2}. \tag{45}$$

Using equations (44), and (45), equation (43) can be modified as follows:

$$dy_t = \frac{1}{2}y_t\frac{\Phi(y_t)^2}{\phi(y_t)^2}(\sum_{j=1}^{r} f_{j,T}^2)dt + \frac{\Phi(y_t)}{\phi(y_t)}\sum_{j=1}^{r} f_{j,T} dB_{j,t}. \tag{46}$$

■





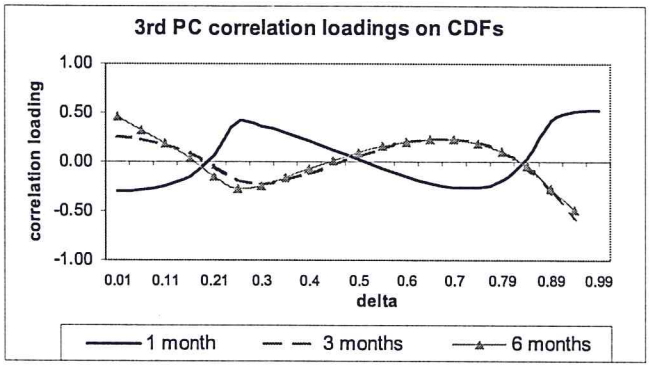
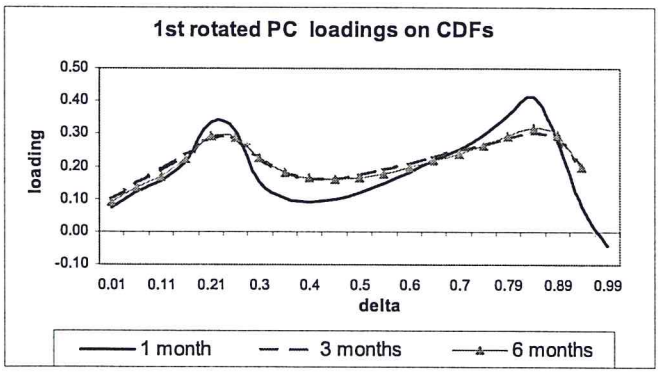


Figure 1: Principal Components Analysis (PCA) on Implied CDFs: Correlation loadings of the 1st, 2nd, and 3rd PC across the One, Three, and Six-months constant maturities.



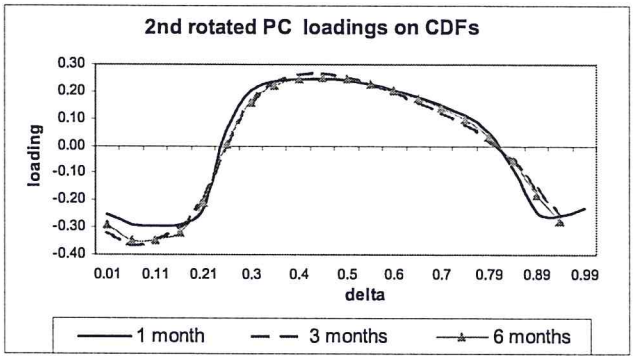


Figure 2: Principal Components Analysis (PCA) on Implied CDFs: Loadings of the 1st and 2nd rotated PCs across the One, Three, and Six-months constant maturities.

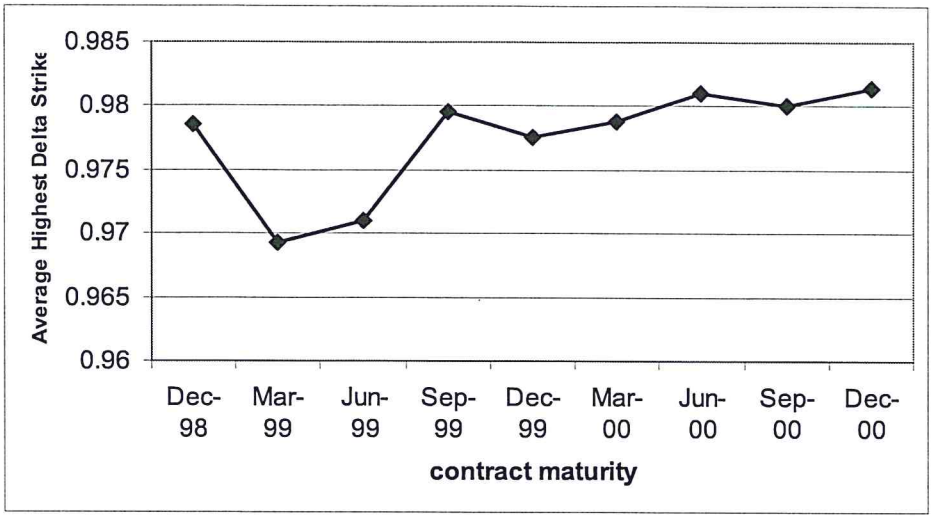


Figure 3: Average Highest Traded Delta Strike Over the Six-Month Period Prior to the Maturity of Each Contract.

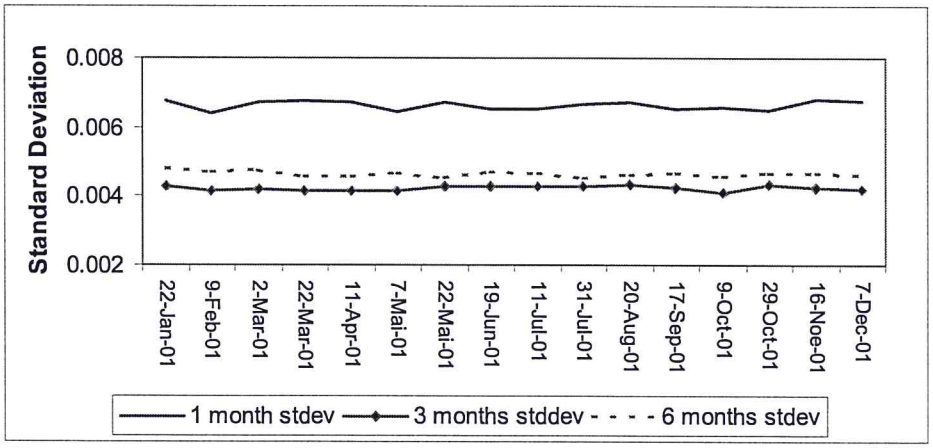


Figure 4: Standard Deviation of the Simulated 0.941 Delta Probability monitored every 14-days in the Year 2001.

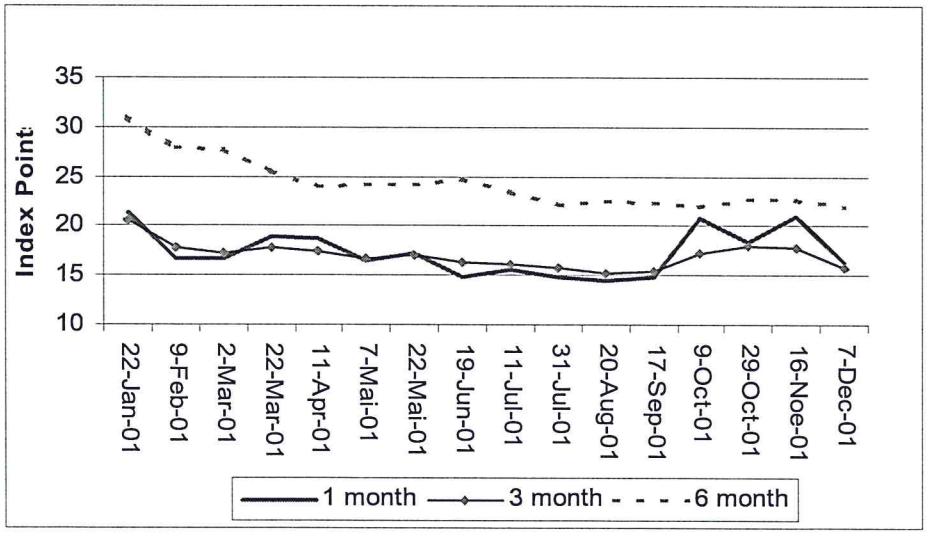


Figure 5: Width of the Confidence Interval in Strike Space around the 0.941 Delta Space.

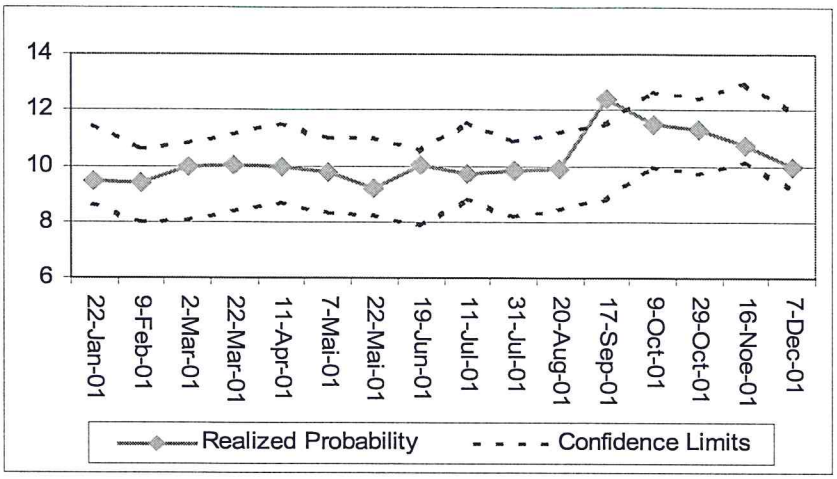


Figure 6: One-Month Realized 0.941 Delta Probability and Simulated 95% Confidence Intervals.

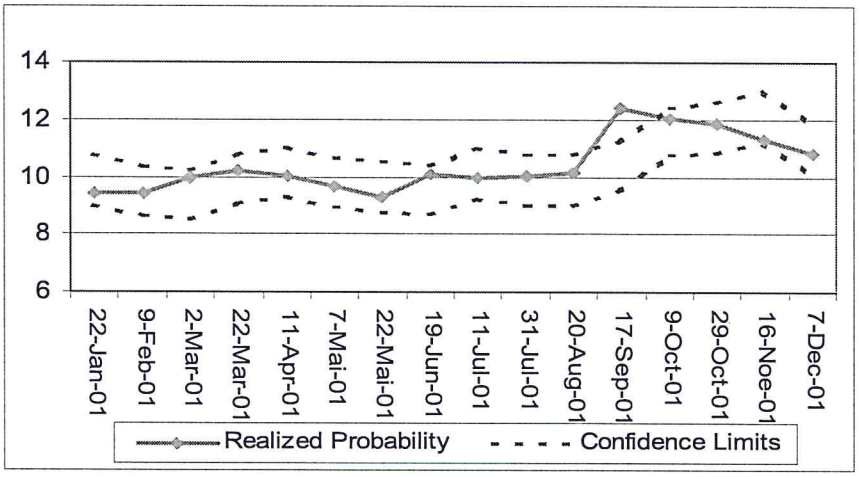


Figure 7: Three-Months Realized 0.941 Delta Probability and Simulated 95% Confidence Intervals.

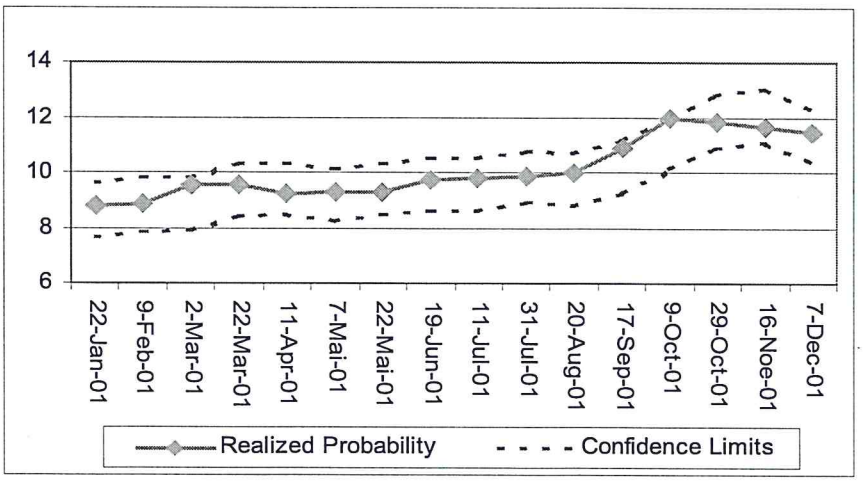


Figure 8: Six-Months Realized 0.941 Delta Probability and Simulated 95% Confidence Intervals.