FINANCIAL OPTIONS RESEARCH CENTRE

University of Warwick

The Dynamics of Smiles

George Skiadopoulos
Stewart Hodges
and
Les Clewlow

March 1998

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry
CV4 7AL
Phone: (01203) 524118

FORC Preprint:1998/85

The Dynamics of Smiles*

George Skiadopoulos† Stewart Hodges‡ Les Clewlow§

Abstract

This empirical study is motivated by Dupire (1992) and Derman and Kani (1998) models. We investigate the number and shape of shocks that move implied volatility smiles and subsequently we look at their correlation with changes in the underlying asset. Principal Components Analysis is applied to the changes of implied volatilities over time, for fixed ranges of days to maturity. Two different metrics are used: the strike and the moneyness metric. In contrast to earlier papers in the interest rate literature, we decide on the number of components by using Velicer's non-parametric criterion. Subsequently, a "Procrustes" type rotation is performed in order to interpret the retained components. Similar results are found in both metrics. Two principal components explain the dynamics of smiles. After the rotation the first one is interpreted as a shift and the second has a Z-shape. The correlations for the first principal component depend on the metric, while for the second are positive under both metrics.

JEL Classification: G13

Keywords: Dynamics of Implied Volatilities, Implied Volatility Smiles, Principal Components Analysis, Procrustes Rotation.

1 Introduction

The Black-Scholes formula (B-S) [4] is used widely to price and hedge options due to its tractability. It is based on several assumptions, such as costless trading takes place in continuous time, the short term interest rate is constant, the underlying asset pays no dividends and that the underlying asset follows a geometric Brownian motion dB with constant volatility σ , i.e.

*This is a revised version of the "Dynamics of Smiles", Financial Options Research Centre 1998/85 Pre-Print. We would like to thank Russell Grimwood, Wojtek Krzanowski, Joao Pedro Nunes, Juan Carlos Mehia Perez, and the participants at the AFFI 1998 (Lille) conference for helpful discussions and comments. Part of this paper was written while the first author benefited as a FORC Research Fellow from the Human Capital & Mobility Programme of the European Commission. Financial support from the Corporate Members of FORC is also gratefully acknowledged. Any remaining errors are our responsibility alone.

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

‡Director of the Financial Options Research Centre, Warwick Business School, University of Warwick, Coventry CV4 7AL, UK; Tel: 01203-523606, e-mail: S.D.Hodges@warwick.ac.uk

§Associate Research Fellow at the Financial Options Research Centre; e-mail: Les_Clewlow@compuserve.com

$$\frac{dS}{S} = \mu dt + \sigma dB \tag{1}$$

Reparameterizing option values by using the B-S model yields the implied volatilities. The empirical evidence shows that at any given date and for any given maturity, implied volatilities vary across strikes by exhibiting smiles and skews (see for example, Rubinstein [21], and Sheikh [23]). An implied volatility smile is in contrast to the B-S model prediction of a constant implied volatility, and it implies that one or more of the B-S assumptions are violated.

One of the potential explanations for the existence of smiles is offered by stochastic volatility models (see among others, Hull and White [14], Johnson and Shanno [16], Scott [22], Wiggins [28]). These models are motivated by the empirical evidence that shows that the variance of the underlying asset changes randomly over time (see Kon [17], Scott [22], Bodurtha and Courtadon [5], Hull and White [14]). However, the limitation of these models is that they do not fit the data well. Furthermore, they rely on the specification of the market price of risk of volatility, and hence they allow only for equilibrium pricing that is of no practical use.

As a response, Dupire [10] and Derman and Kani [9] proposed models that allow for arbitrage pricing with stochastic volatility, with no need for any volatility risk premium to be specified (see Skiadopoulos [24] for review of these models). This is achieved by using a methodology that is similar to the Heath, Jarrow, Morton [12] approach. More specifically, they assume a process for the forward (rather than the instantaneous) variance V_T of the form

$$\frac{dV_T(t,S)}{V_T(t,S)} = a(t,S)dt + \sum_{i=1}^{n} b_i(t,S)dW_i \tag{2}$$

where dW_i is the ith Brownian motion.

Using this process either implicitly (Dupire), or explicitly (Derman and Kani), they can price and hedge standard and exotic options by Monte Carlo simulations. Their models seem to be promising, but there are three issues that have to be dealt with for implementation purposes. These are (a) the specification of the number n of stochastic shocks appearing in equation (2), (b) the estimation/interpretation of the volatility parameters $b(.)$, and (c) the estimation of correlation of the shock in the asset process with each one of the shocks in the forward variance process. This is necessary for the joint simulation of the underlying asset price and the forward variance.

In this paper we answer (a) and (b) by applying Principal Components Analysis (PCA) on sets of the differences of implied volatilities. Then, we use the results from the PCA to answer (c)[1]. One of the contributions of the paper is that we use a non-parametric criterion (and not rules of thumb as in other studies) in order to answer the first question and we develop a method which enables us to answer the second question and subsequently the third one.

The remainder of the paper is structured as follows. In the second section, we describe the data set that we use and the way that we filter it. In the third section, we describe the

[1] By studying implied volatilities, we implicitly assume that they provide us with all the required information to specify the forward volatility process. This is a reasonable assumption because the implied volatility is just the expected average of the forward volatilities (see Derman and Kani [9]).

PCA technique and how this can be applied to the analysis of smiles. In the fourth and fifth sections, we apply the technique to the differences of implied volatilities when these are classified into two different sets of variables (metrics) for given maturity buckets: the strike and moneyness metric. We decide on the number of components to retain by following several steps which include Velicer's criterion. Next, we try to interpret them by using the idea of a Procrustes type rotation method. In the sixth section, we calculate the correlations between the changes of the principal components and the underlying asset price. The last section concludes.

2 The Data Set and the Filtering Method

2.1 The Data Set

The data that we use is daily closing prices of futures options on the Standard and Poor index (S&P 500) obtained from the Chicago Mercantile Exchange (CME) for the years 1992-95. The primary source database for this study is the transaction report "Stats Database", compiled daily by CME. This database contains the following information for each option traded during a day: the date, the style (call or put), the options and futures expiration month, the exercise price, the number of contracts traded, the opening, closing, low and high future's and option's price, the opening, closing, low and high bid-ask future's and option's price and the settlement price. We are going to use the closing options and futures prices, since we regard those as containing more information than the opening, low, or high ones. In addition we use London Euro-Currency interest rates (middle-rates) on the US dollar, collected from Datastream in order to approximate the riskless rate in our option pricing model. We collected daily interest rates for 7-days, one-month, three months, six months and one year and for other maturities we interpolated linearly.

Given that futures options on the S&P 500 are American-style options, implied volatilities σ_{imp} were calculated by the Barone-Adesi, Whaley [3] quadratic approximation which explicitly takes into account the early exercise premium. However, even if the market uses the Barone-Adesi Whaley model, the calculation of implied volatilities will be subject to other sources of errors. These are described in the following section.

2.2 Filtering the Data

Harvey and Whaley [11] and Roll [20] have shown that in the presence of non-synchroneity and bid-ask spread effects, respectively, spurious negative serial correlation is induced in the volatility changes if closing option prices are used. The trading hours for the S&P 500 Index are 8.30 a.m. till 3.15 p.m. (Chicago time), while those for the S&P 500 Index Options on futures are 8.30 a.m. till 3.15 p.m., but the option can be exercised until 7.00 p.m. on any business day the option is traded. Furthermore, deep in-the money and out-of the money options are rather illiquid and they trade less frequently than those nearer to the-money. This implies that for those options the option market has closed some time before the asset's market. Therefore, there are measurement errors in the calculation of our implied volatilities that need to be filtered out.

In the first stage a preliminary screening of the data is done by excluding the data that violate the arbitrage boundary conditions. The arbitrage boundary conditions are given by:

$$C > F - K \tag{3}$$

$$P \geq K - F \tag{4}$$

where C (P) is the American call (put) option price, F is the futures price and K is the strike price. We also exclude options having a price of less than 0.01 cents, and we eliminate short term options with less than 10 days to maturity because they are very sensitive to small errors in the option price.

In a second stage we exclude from our data sets in-the-money (ITM) calls and puts and we construct our smiles by using for the left part of the smile out-of-the-money (OTM) puts and for the right part OTM calls. We do so because the delta of ITM options is high. Hence, their implied volatilities are prone to measurement errors. This is not going to bias the results of the subsequent analysis because the implied volatilities for both calls and puts should be the same if the put-call parity is to be respected. Moreover, we only keep the implied volatilities calculated from options having a vega bigger than eight. The idea is that the vega evaluated at the quoted implied volatilities is equal to the ratio of the measurement errors in the option price over the errors in the implied volatilities. Hence, the interpretation of our *vega constraint* (as we will call it hereafter) is that if the given measurement error ΔC is required to produce an error in the calculation of σ_{imp} less than $\Delta\sigma_{imp}$, then vega has to be greater than a certain cut-off point.

The advantages of the vega-constraint over the constraint on $\frac{F}{K}$ that other researchers have used so far, are that (a) the cut-off point is not determined exogenously, but endogenously as we will explain in a while[2] and (b) the ratio $\frac{F}{K}$ remains constant regardless of the days to maturity of the option, while the calculation of the vega-constraint takes explicitly into account the days to maturity.

The cut-off point of eight was chosen by looking at the trade-off between accuracy and number of observations that we throw away. The more we want to constrain the errors $\Delta\sigma_{imp}$ the greater vega has to be. On the other hand,the greater vega is the more information is thrown away. Therefore, we had to experiment with different values of vega, in order to see the number of observations that we were excluding and how noisy our data was. The chosen cut-off point retained a sufficient amount of data, while the amount of noise was checked by what we call the "accuracy graph"[3].

An accuracy graph is constructed as follows: a specific expiry date contract is chosen and it is followed from the point that it has 30 days to expire down to 27 days to expire. Hence, four implied volatilities are obtained for each strike. In the case that the data is not

[2]Rubinstein [21] eliminates options with a ratio (S/X) less than 0.75. Sheikh [23] replicating Rubinstein's paper applied the same criterion. Canina and Figlewski [7] eliminate options that are more than 20 points in or out-of-the money. Xu and Taylor [29] eliminate options when $X < 0.8S$ or $X > 1.2S$. However, no justification is given for this range of values.

[3]The retained amount of data for a given level of vega differs across maturities and years and between calls and puts. Setting the vega constraint to eight retains 52%-30% of the call observations and 90%-50% of the put observations.

noisy, then for each strike the implied volatilities should not differ a lot over such a short time interval, i.e. for each strike the standard deviation should be small. Therefore, for each strike the average of those four implied volatilities and their standard deviation are calculated. Then, the average and plus, minus one standard deviation are plotted.

Figure 1 shows such a graph for a call contract that expires in 9309 when no filtering has been applied. It is obvious that the implied volatilities of ITM calls have very big standard deviations, i.e. they are very noisy something which was expected because of the non-synchroneity effects. Similar figures were constructed for both calls and puts for the case that only the vega constraint was imposed (i.e., ITM calls and puts were not eliminated). The figures were examined for values of vega $v = 1, 2, 4, 8, 16$. We found that the vega constraint trims the extreme strikes, i.e. the deep ITM and OTM options, but the large amount of noise for the ITM options remained. This justifies the elimination of the ITM options as we have already explained. Moreover, such figures were constructed for the case that vega equals 16. There is of course a further reduction in noise, but it is not that significant compared to the amount of informaton that is discarded as vega increases and to the fact that there is still noise for ITM options despite the value of vega. Hence, the data dictate that the cut-off point for the vega constraint should be eight. Figure 2 shows a recovered from OTM options implied volatility skew. It confirms that the filtering procedure is satisfactory.

3 Principal Components Analysis and Smiles

The natural technique to identify the number and the interpretation of stochastic shocks that affect implied volatility smiles, is Principal Components Analysis (PCA). We describe it briefly to explain why it is appropriate and to introduce the notation we will use.

PCA. has been constructed to answer the following question: How can we explain the systematic behavior of the observed variables by means of a smaller set of computed but unobserved latent random variables? From a purely mathematical viewpoint the purpose of a population principal component (PC) model is to transform p correlated random variables to an orthogonal set which reproduces the original variance-covariance structure. This is equivalent to rotating a p-dimensional quadratic form to achieve independence between the variables.

Consider n observations on p random variables represented by the $(n \times 1)$ vectors $x_1, x_2, \ldots,$ x_p, where for simplicity we assume that the means of these vectors are zero, i.e. $\overline{x_1} = \overline{x_2} = \ldots = \overline{x_p} = 0$ and an estimator S of the variance covariance matrix is given by $S = X'X$ where X is a $(n \times p)$ matrix. Construct now linear transformations Z_i for $i = 1, 2 \ldots, p$ so that

$$Z = XP \tag{5}$$

where Z is a $(n \times p)$ matrix and P is a $(p \times p)$ matrix with the *ith* column the $\underline{P_i'}$ vector $\underline{P_i'} = (p_{1i}, p_{2i}, \ldots p_{pi})$, for $i = 1, 2, \ldots, p$. The *ith* column Z_i is the *ith principal component* (PC hereafter). The elements of the vector $\underline{P_i'}$ are the coefficients of the *ith* PC and they are called the *loadings*.

The variance of the Z_i PC is given by $Var(Z_i) = \underline{P_i'} X'X \underline{P_i}$ and the PCs are constructed such that $Cov(Z_i, Z_j) = 0$ for $Z_i \neq Z_j$.

The purpose of the PCA is achieved by determining the unknown fixed loadings, so as to maximize sequentially the variance of the PC's, starting from the first one up to the *pth*, under the constraint that $P'P = I$ where I is the identity matrix. The first order condition of this maximization problem results to

$$(X'X - lI)P = 0 \tag{6}$$

where l_i are the lagrange multipliers.

From equation (6) it is evident that the PCA boils down to the calculation of the eigenvalues l_i and the eigenvectors of the variance-covariance matrix $X'X$. S is now represented as

$$S = P'LP \tag{7}$$

where $L{=}diag(l_1, l_2, \ldots, l_p)$ and the sum of the variances of the PCs equals the sum of the variances of the X variables. Moreover, it can be proved that if the covariance matrix is diagonal, then there is no gain in performing the PCA.

As Basilevsky explains (Basilevsky [1] page 143), it is better to work with the standardized variables of Xs and therefore with the correlation matrix, rather than the covariance matrix of Xs. For the purposes of our study, standardized variables and PCs are used.

From Equation (5) standardizing the PCs to unit length yields a new matrix $Z^* = ZL^{-\frac{1}{2}} = XPL^{-\frac{1}{2}}$. Hence,

$$X = Z^*A' \tag{8}$$

where

$$A' = L^{\frac{1}{2}}P' \tag{9}$$

When both variables and components are standardized to unit length, the elements a_{ij} of A' are correlations between the variables and PCs and they are called *correlation loadings*. If $r \prec p$ PCs are retained then

$$X = Z^*_{(r)}A'_{(r)} + \varepsilon_{(r)} \tag{10}$$

where $\varepsilon_{(r)}$ is a $(n \times p)$ matrix of residuals and the other matrices are defined as before having r rather than p columns. The percentage of variance of X_i $(i = 1, 2, \ldots, p)$ that is explained by the retained PCs (communality of X_i) is calculated from the correlation loadings. Hence, after retaining $r < p$ components that explain a sufficient amount of the original variance-covariance structure, we use equation (10) to check the magnitude of communalities and the interpretation of the retained components.

Since we want to determine the number n of shocks that appear in the stochastic differential equation (2) we apply the PCA to the *first differences of implied volatilities*. This is the natural thing to look at, if we work with a discretized version of equation (2).

The variables that we apply the PCA to will be measured in two different ways (metrics): (a) we will consider as variables the first differences of implied volatilities classified across strikes (strike metric) which is the natural metric to look at, and (b) we will consider as variables the first differences of implied volatilities classified across moneyness i.e. $\frac{K-F}{F} * 100$ (moneyness metric). The reason for choosing this metric is that there are theoretical reasons that support that smiles are a function of moneyness (see Heynen [13] Taylor and Xu [25]).

Moreover, since the purpose of this paper is to analyze only the smile dynamics (and not the dynamics of the whole implied volatility surface) our analysis is applied to fixed buckets of days to maturity that are as fine as possible, so as to isolate the maturity effect. We fix six such intervals: 30-10, 60-30, 90-60, 150-90, 240-150 and 360-240 days to maturity. The maturity buckets were chosen, so as to cope with two constraints: (a) getting a sufficient amount of data for each range in order to perform the PCA, and (b) treating the missing observations that occur due to the filtering procedure that has been applied. The literature on treating missing observations is vast, but as Anderson et al. [6] note "The only real cure for missing data is to not have any". Therefore, the missing values are not replaced since we do not know whether this will bias the results. Instead, listwise deletion is applied, i.e. the whole day for which at least the observation for one variable (strike or moneyness) is missing, is deleted.

4 PCA on the Strike Metric

4.1 Number of Retained Principal Components and a First Interpretation

The variables for each year and within a given range, on which we will perform the PCA were chosen so as to get (a) a sufficient number of strikes (say not less than 7 ideally) so as to examine a wide range of the smiles, (b) a sufficient amount of observations (say no less than 100), and (c) a satisfactory correlation between them, since the higher the correlations between the variables, the greater the gain from the PCA. The overall level of correlation was measured by the Kaiser-Meyer-Olkin (KMO) measure. It was found to be between 0.7 and 0.9. The fact that correlations are high, is not only encouraging for the application of the PCA, but it is also a necessary condition for the selection of a linear PCA model. In the case that there are nonlinearities, a set of highly related random variables can exhibit low correlation, unless the nonlinearities are taken explicitly into account (see Basilevsky [1], page 162).

Next, we decide on the number of components to be retained and their interpretation. In order to do so, researchers usually use rules of thumb (see for instance, Litterman and Scheinkman [18]). For example, they keep the components that explain 90% of the total variance, or they omit the correlation loadings which are smaller than 0.20 (for a description and discussion of the several rule of thumbs see Jackson [15]). As Basilevsky 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 applying Velicer's [27] non-parametric criterion, looking at the communalities that the retained, according to the criterion, components explain, and examining sequentially the noisiness of the correlation loadings. If any PC appears to be just noise, then we prefer to reject it.

Velicer proposed 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, based on the partial correlations of the residuals of the PCs model, after $r < p$ components have been extracted[4]. The idea of the criterion can be described as follows: Let

$$X = Z^*_{(r)}A'_{(r)} + \varepsilon_{(r)} \tag{11}$$

Basilevsky [1] shows (theorem 3.13, page 132) that $X'X = AA'$. Hence, the variance-covariance matrix of the residuals $\varepsilon_{(r)}$ is given by the following expression

$$\varepsilon'_{(r)}\varepsilon_{(r)} = X'X - A_{(r)}A'_{(r)} \tag{12}$$

Let $D = diag(\varepsilon'_{(r)}\varepsilon_{(r)})$. Then, $R^* = D^{-\frac{1}{2}}\varepsilon'_{(r)}\varepsilon_{(r)}D^{-\frac{1}{2}}$ 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

$$f_r = \sum_{i \neq j}\sum \frac{r^{2*}_{ij}}{p(p-1)} = \sum_{i=j}\sum \frac{r^{2*}_{ij} - p}{p(p-1)} \tag{13}$$

and lies in the interval 0 to 1.

One might expect that as we retain more PCs, $f_r \rightarrow 0$. However, this is not the case. f_r decreases until a number r^* and then it increases again (see Velicer [27] for the explanation of this behavior). Velicer suggests that $r = r^*$ should be the number of components to be retained. The idea behind this is that as long as f_r declines, there is still space for the additional components to capture part of the covariance of the residuals. Since it is obvious from the formula of the partial correlation that f_r declines as long as the partial covariances decline faster than the residual variances, Velicer's procedure terminates when, on the average, additional PCs explain more of the residual variances than their covariances, i.e. when they explain unsystematic rather than systematic behavior.

Table 1 shows that the number of retained PCs according to Velicer's criterion varies across the maturity ranges. In general, we should not accept more than two components. This is in contrast to the mean eigenvalue rule of thumb $\bar{l}$ which retains the PCs having eigenvalues greater than the mean eigenvalue. We can see that the mean eigenvalue rule keeps in many cases one PC more than Velicer's criterion. We also show f_0, a second summary statistic that is useful for comparative purposes. f_0 is the usual Velicer's statistic when no PCs have been retained. If $f_1 > f_0$ then no components would be extracted and the explanation to that lies in the behavior of Velicer's criterion. Table 1 shows that $f_1 < f_0$ in all the cases. Hence, the PCA can extract a limited number of components[5].

[4] Application of the Bera-Jarque test showed that the null-hypothesis of univariate and hence of multivariate normality was rejected. This necessitates the use of a non-parametric test. Another aspect of rejecting multivariate normality, apart from not being able to apply parametric tests for the PCA analysis, is that we can not use any of the existing tests for eliminating outliers (for a review of the techniques about dealing with outliers, see Barnett and Lewis [2]). PCA is very sensitive to their presence (see Basilevsky [1]). This makes necessary the use of filters on the data set, as we have already demonstrated.

[5] Reddon [19] evaluated the type-I error (reject the null while it is true) rates of this test. Towards this end, he specified the null hypothesis of Velicer's test as the reduction in dimensionality. Then, he carried

	Year	f_o	f_1	f_2	f_3	r*	l	1st PC	2nd PC	3rd PC
30-10	92	0.2863	0.2835	0.2843	0.2847	1	2	57.60	16.70	5.80
	93	0.3156	0.3131	0.3138	0.3143	1	1	61.30	16.30	6.30
	94	0.2586	0.2557	0.2563	0.2567	1	2	55.90	14.00	8.70
	95	0.3125	0.3074	0.3088	0.3096	1	1	60.00	17.30	6.40
60-30	92	0.1383	0.1377	0.1377	0.1377	2	3	39.30	18.20	9.30
	93	0.1210	0.1205	0.1204	0.1204	2	2	34.90	24.00	7.70
	94	0.0999	0.0995	0.0994	0.0994	2	3	28.00	23.40	14.60
	95	0.3323	0.3297	0.3304	0.3310	1	1	64.40	11.40	8.10
90-60	92	0.1736	0.1718	0.1720	0.1723	1	2	46.40	15.80	9.90
	93	0.1529	0.1516	0.1516	0.1518	2	2	43.20	18.80	10.70
	94	0.1266	0.1253	0.1252	0.1254	2	2	36.90	17.70	10.60
	95	0.1044	0.1024	0.1022	0.1027	2	2	39.20	21.10	60.30
150-90	92	0.2296	0.2283	0.2285	0.2286	1	4	50.30	8.60	7.30
	93	0.1528	0.1519	0.1519	0.1520	2	3	41.20	14.70	8.80
	94	0.1845	0.1835	0.1835	0.1836	2	4	44.00	14.00	7.60
	95	0.1910	0.1896	0.1898	0.1900	1	2	46.60	15.50	7.00
240-150	92	0.3452	0.3441	0.3442	0.3444	1	2	61.80	9.70	7.50
	93	0.3063	0.3053	0.3054	0.3055	1	2	58.10	12.10	6.50
	94	0.2886	0.2877	0.2878	0.2879	1	2	56.70	10.80	6.90
	95	0.1551	0.1543	0.1544	0.1545	1	2	42.30	21.40	10.60
360-240	92	0.3177	0.3168	0.3169	0.3172	1	2	61.90	18.70	9.60
	93	0.3712	0.3696	0.3698	0.3701	1	2	64.90	13.20	8.40
	94	0.4062	0.4054	0.4053	0.4055	2	2	66.30	18.50	7.00
	95	0.2234	0.2217	0.2222	0.2226	1	2	56.30	20.10	11.50

Table 1: Principal Components in the Strike Metric: r* = the number of components retained under Velicer's criterion (minimum of f0,...f3), l = number of components retained under rule of thumb, with percentage of variance explained by components 1-3.

Table 1 also shows the percentage of variance that the first three PCs explain, despite of the fact that Velicer's criterion does not retain the third PC. This is so that we can compare it to the corresponding component found in the interest rate literature (see Litterman and Scheinkman [18]). However, the decision on the number of retained components there, has not been justified by any testing procedure, but it was based only on the grounds of the amount of variance explained by the retained components.

From Table 1 the average variance across the maturity buckets can be calculated for each year. We find that the first PC explains on average 41%-62% of the total variance, and the second PC explains on average 13%-19%. The explained average cumulative variance lies between 59% and 80%[6]. The first PC approaches its upper variance limit in the shortest and longest maturity options. The second PC approaches its upper limit for the first three ranges, then it dies down, but it comes up again for the longest range. In general, the first two PCs together explain most of the variance in the shortest and longest maturity options.

We consider Velicer's procedure as the first step in order to decide how many components we should retain. The second step is to look at the interpretation of the first three PCs[7]. The interpretation of the PCs can be revealed by looking at the correlation loadings A' in equation (8). Hence, the first three columns of A' reveal the impact on implied volatilities of the first three PCs, respectively.

Plots of the correlation loadings show that the first PC looks roughly as a parallel shift with an attenuation at the edges in the range 30-10 days to maturity and for all the years (the terminology "shift" and "slope" is used in the same spirit as Litterman and Scheinkman have already established). In the range 60-30 and for the years 93 and 95 the first PC still moves implied volatilities to the same direction. However, for all the other ranges and the years 92, 93, and 94 the first PC does not have the interpretation of a shift, because it has both negative and positive correlation loadings. In fact, it looks more like a slope. For the year 95 in all the ranges the first PC is a shift.

Regarding the interpretation of the second PC the figures reveal that for the range 30-10 over all the years the second PC has the interpretation of a slope (in fact it is like a Z-shape). For the range 60-30 the same is true with the exception of year 94, while for the rest of the ranges and for the years 92, 93, and 94 the second PC seems to have a triangular shape. However, for the year 95 and for all the ranges to maturity, the second PC seems to have the slope interpretation.

---

out simulations using data generated from populations having unit variances and zero covariances and he found out that Velicer's test made excessive type-I errors in the cases where the number of observations did not exceed two times the number of variables. As the number of observations increased beyond two times the number of variables, the type-I error rates rapidly became zero. Since the number of observations that we use is far more than twice the number of variables, we feel comfortable with the results from comparing f_o with f_1.

[6] These results are very different from the interest rate literature corresponding ones. For example, Litterman and Scheinkman [18] had found that on average the first PC explained 89.5% of the total variance, the second PC explained 8.5% and the third 2%. In total, the three retained PCs explained there, 98.4% on average.

[7] In fact, we looked also at the communalities that the retained (according to Velicer's test) PCs explain. This is because we would like the number of retained components to explain a sufficient amount of the variability of the implied volatilities dynamics. The results showed that two PCs provide a satisfactory communality. The inclusion of a third PC increased the explained variance significantly, only in a limited number of cases.

Figure 3 shows the correlation loadings of the first and second PCs for the years 92 and 93 over the different maturity buckets (for each range, interpolation has been performed across the missing strikes). In general, the figures for the third PCs' correlation loadings reveal that for all the years and for all the ranges, the third PC does not have a typical pattern, and therefore it can be treated as noise. This is consistent with Velicer's criterion results. Hence, we decide to retain two PCs in the strike metric.

However, even though we were able to get the shift and slope interpretation for the shorter maturity options in some years and over all the ranges in year 95, we would like to have this *simple* interpretation for the remaining ranges and years. This would help to implement equation (2)[8].

4.2 The Rotation Method

In the case that interpretation of the resulting PCs is not straightforward, people usually apply the standard technique of rotation, i.e. the cartesian coordinates are moved to such an angle so as a simple interpretation to be possible. Effectively, a rotation (this is the second rotation after the PCA) is performed on the characteristic vectors, producing some new "components" that may be useful, although they are obtained by a different criterion from PCA.

Given that we want to get a simple interpretation of the retained PCs and since the intuition coming from the Taylor series expansion tells us that the first PC should have a level (shift) interpretation and the second PC a slope interpretation, we use a "Procrustes" type rotation[9]. The question that a "Procrustes" type rotation addresses is the following: Let two $(p \times r)$ matrices, A and B. What $(r \times r)$ transformation matrix T will best transform A into B? In other words, what matrix T will make AT most like B[10]? B is called the target matrix, and in our case it will be the $(p \times 2)$ matrix of rotated loadings which will give the parallel and the slope character for the first and second PC, respectively. Therefore, the task is to determine T in such a way, as to get the desired result.

The way that we construct our rotation method is by using the general properties of orthogonal rotations as outlined in Basilevsky [1]. Then, we regress a vector of constants on the vector of loadings of the first two PCs. The latter determines the elements of the first row of the matrix T and combining this with the former we get the elements of the second

[8] In our case, in order to implement equation (2) it does not matter whether we have a simple (or even any) interpretation of the retained PCs, since the correlation loadings are the estimates of the b_i's coefficients. However, in order to estimate the b_i's by another econometric technique, their functional form should be specified. A simple interpretation of the PCs will help us in achieving this. Intuitively thinking, it is possible to interpret the first PC as a level and the second as a slope. This is because any well-behaved function can be approximated by a Taylor series expansion of first order, where the zero order expansion is the level, and the first order expansion is the slope.

[9] Before trying to apply "Procrustes" type of rotation, we applied the most popular rotation methods, i.e. the varimax, the quartimax and the oblique method. However, none of these methods produced the desired interpretation because of the way that they are constructed. For more details on these methods, see Jackson [15].

[10] In Greek mythology, when Thesus was cleaning up Greece's highways of criminals, one of those he killed was Procrustes, the Stretcher. Procrustes had an iron bed on which he tied any traveller who fell into his hands. If the victim was shorter than the bed, he or she was stretched out to fit; if too long, Procrustes chopped off what ever was necessary.

row[11]. Doing so, we force the first rotated PC to look like a shift but we do not know a priori how the second rotated PC is going to look like.

Figures 4 and 5 show the first two rotated PCs loadings for the years 92 up to 95. These graphs have been produced by working with unstandardized variables and PCs, and consequently we do not look at correlation loadings, but just at loadings. We can see that for the intervals 30-10, 60-30 and 90-60 and for all the years the first rotated PC can be interpreted as a shift, while the second rotated PC is interpreted as a slope (it has a Z-shape with large loadings for the implied volatilities in the small strikes and small loadings for the implied volatilities in the big strikes). For the intervals 150-90, and especially for 240-150 and 360-240 the rotation was not successful for the first PC as it does not have a consistent shape over the years. However, it was successful for the second PC as it brought up a Z-shape, for all the ranges of days to maturity and for all the years.

In Table 2 we show the percentage of variance that the first and second *rotated* PCs explain. We also show the percentage of the variance that the first unrotated PC explains and the cumulative percentage variance. From the properties of the orthogonal rotation, the cumulative variance explained by the first two PCs remains the same as before the rotation; however, the percentage of the total variance that each rotated PC explains has changed. The first rotated PC explains on average 26%-59% of the total variance, while the second PC explains 16%-36%. These indicate that the first rotated PC explains less variance than the unrotated one, while the second rotated PC explains more. In fact, once a rotation is performed, the second rotated PC may explain more variance than the first one (e.g., see range 90-60). However, it is still true that the percentage variance that the first rotated PC explains, is bigger for the shortest and longest maturity options, as it was for the unrotated first PC, but this is no longer true for the second PC.

Another way of interpreting Figures 4 and 5 is by examining whether the effect of the retained components is greater on the shorter, or the longer maturity options' implied volatilities. These figures show that the first rotated PC affects more the longer maturity options' implied volatilities (e.g., 90-60 days), than the shorter maturity ones (e.g., 30-10). Regarding the effect of the second rotated PC over the different maturity buckets, it seems as if the effect of the second rotated PC over the maturity buckets is the same, and this picture is repeated for all the years.

5 PCA on the Moneyness Metric

5.1 Construction of the Moneyness Metric

In the strike metric, PCA is performed on fixed variables that are dictated by the option contract specification. However, in the moneyness metric the variables $\frac{K_i - F_t}{F_t} * 100$ (we

[11] Notice that we apply the rotation to P and not to A', because of the property that the rotated eigenvectors remain orthogonal that is used to construct our method. This property is not valid for rotated correlation loadings (see Basilevsky [1]).

	Year	Unrot. 1st PC	1st PC	2nd PC	Cumulative
30-10	92	57.60%	57.50%	16.90%	74.40%
	93	61.30%	61.10%	16.60%	77.70%
	94	55.90%	55.90%	14.00%	69.90%
	95	60.00%	59.90%	17.40%	77.30%
	Average	58.70%	58.60%	16.20%	74.80%
60-30	92	39.30%	26.80%	30.70%	57.40%
	93	34.90%	33.60%	25.40%	59.00%
	94	28.00%	23.50%	27.90%	51.40%
	95	64.40%	64.40%	11.40%	75.80%
	Average	41.65%	37.10%	23.90%	60.90%
90-60	92	46.40%	16.00%	46.20%	62.20%
	93	43.20%	24.70%	37.20%	61.90%
	94	36.90%	28.10%	26.50%	54.60%
	95	39.20%	35.70%	24.60%	60.30%
	Average	41.43%	26.10%	33.60%	59.80%
150-90	92	50.30%	16.40%	42.50%	58.90%
	93	41.20%	30.80%	25.10%	55.90%
	94	44.00%	43.00%	15.00%	58.00%
	95	46.60%	46.00%	16.10%	62.10%
	Average	45.53%	34.10%	24.70%	58.70%
240-150	92	61.80%	59.50%	12.10%	71.60%
	93	58.10%	13.30%	56.90%	70.20%
	94	56.70%	24.70%	42.90%	67.60%
	95	42.30%	42.20%	21.40%	63.70%
	Average	54.73%	34.90%	33.30%	68.30%
360-240	92	61.90%	61.30%	19.40%	80.70%
	93	64.90%	31.80%	46.30%	78.10%
	94	66.30%	36.60%	48.20%	84.80%
	95	56.30%	46.50%	29.90%	76.40%
	Average	62.35%	44.10%	36.00%	80.00%

Table 2: Percentage of Variance Explained by the Unrotated first PC and by the Rotated PCs in the Strike Metric.

will call this the "natural" moneyness metric) for $i = 1, 2, \ldots, s$ where s is the number of strikes and t measures the calendar time, are different from day t to day $t+1$, as the futures price changes from F_t to F_{t+1}. This means that for every day we have different variables, and therefore we can not apply the PCA. The only solution to this is to fix the moneyness variables before starting the analysis (we will call these fixed moneyness variables the "artificial" moneyness metric). Then, for each day, we interpolate across the implied volatilities for these "fixed" variables.

The spacing between the variables of the "artificial" moneyness metric (step-size) was set so that between any two consecutive variables of the "natural" moneyness metric, there will be only one variable of the "artificial" moneyness metric. Otherwise, the implied volatilities created by the interpolation of both of the "fixed" variables, would depend on the *same* two consecutive values of the "natural" moneyness metric and consequently they would exhibit spurious dependence, something which of course would distort the results of our subsequent analysis[12].

It still holds, as in the strike metric, that the variables that we are going to use, should give us a sufficient amount of observations after the listwise deletion, and a satisfactory KMO correlation measure. KMO ranges between 70% and 90%, just as it was the case with the strike metric.

5.2 Number of Retained Principal Components and a First Interpretation

In order to see how many PCs we retain, we use once more Velicer's criterion.

Table 3 shows that Velicer's criterion keeps either one or two PCs depending on the maturity range that we look at, but in any case, it does not keep more than two PCs[13]. Again in all the cases, $f_0 > f_1$ something which confirms that we can reduce the dimensionality of the variables that we work with. We show also the percentage of variance that the first three PCs explain.

From Table 3 we can calculate for each year, the average variance that each PC explains across the maturity buckets and subsequently we can look at the range that the average variance lies in across the years. The first PC explains on average 45%-80% and the second

[12]The step-size was calculated by tracing the minimum futures price for each year, and looking at the spacing between strikes for each maturity bucket. However, the bigger the moneyness-step size, the fewer are the variables to perform PCA on. Hence, before deciding on the step-size, we had to check the number of variables that correspond to any given step-size. We dealt with this by finding for every day the minimum and maximum moneyness and then we constructed distribution graphs for them (the moneyness variables were classified in bins with a spread of 1%, since this should be the spacing between most of the variables in the natural moneyness metric). Subsequently, the range was set by taking the right tail from the distribution of the minimum moneyness and by taking the left tail from the distribution of the maximum moneyness.

[13]Regarding the comparison between Velicer's test and the mean eigenvalue rule of thumb, the tests give identical results for the ranges 30-10, 90-60 and 360-240 and they agree for most of the years in the range 60-30. In the remaining ranges, the mean eigenvalue rule retains one more PC than Velicer's, as it was the case with the strike metric.

	Year	f_0	f_1	f_2	f_3	r*	l	1st PC	2nd PC	3rd PC
30-10	92	0.4818	0.4792	0.4801	0.4807	1	1	73.30	11.00	5.50
	93	0.5521	0.5491	0.5500	0.5508	1	1	77.80	8.10	6.00
	94	0.6959	0.6936	0.6946	0.6949	1	1	85.70	6.40	2.30
	95	0.4764	0.4735	0.4745	0.4752	1	1	73.00	11.10	5.90
60-30	92	0.1816	0.1812	0.1810	0.1811	2	2	39.60	30.30	8.10
	93	0.2065	0.2058	0.2058	0.2059	2	2	44.60	28.70	7.70
	94	0.2761	0.2747	0.2751	0.2753	1	2	53.60	22.90	5.30
	95	0.2189	0.2183	0.2183	0.2185	2	2	47.60	25.50	8.70
90-60	92	0.2291	0.2276	0.2274	0.2278	2	2	48.80	27.30	9.00
	93	0.2164	0.2133	0.2132	0.2137	2	2	46.80	30.60	9.70
	94	0.2168	0.2153	0.2152	0.2154	2	2	44.20	29.60	7.40
	95	0.1395	0.1383	0.1381	0.1384	2	2	39.70	25.50	9.50
150-90	92	0.2326	0.2313	0.2316	0.2317	1	2	51.50	16.80	7.50
	93	0.1989	0.1969	0.1972	0.1975	1	2	48.10	16.90	8.90
	94	0.2578	0.2560	0.2563	0.2565	1	2	54.20	18.00	8.00
	95	0.2008	0.1985	0.1988	0.1991	1	2	48.80	18.10	8.50
240-150	92	0.4314	0.4302	0.4303	0.4307	1	1	69.20	11.90	7.60
	93	0.4334	0.4322	0.4324	0.4327	1	2	68.70	10.40	6.50
	94	0.3595	0.3585	0.3588	0.3590	1	2	63.40	16.90	7.40
	95	0.3224	0.3211	0.3213	0.3215	1	2	60.70	12.80	7.30
360-240	92	0.5668	0.5653	0.5657	0.5663	1	1	77.90	14.10	4.40
	93	0.4933	0.4917	0.4922	0.4928	1	1	75.10	16.90	4.10
	94	0.6852	0.6826	0.6839	0.6844	1	1	85.10	7.80	2.50
	95	0.5932	0.5874	0.5894	0.5911	1	1	80.90	6.80	5.80

Table 3: Principal Components in the Moneyness Metric: r* = number of components retained under Velicer's criterion (minimum f0,...f3), l = number of components retained under rule of thumb with percentage variance explained by components 1-3.

PC explains 9%-28%. In total, the first two PCs explain 68%-91%. Hence, the retained PCs in the moneyness metric, explain about 11% more of the total variance than in the strike metric. Moreover, similarly to the results that we got for the strike metric, the first PC approaches its upper explained variance limit in the shortest and longest maturity options. The second PC approaches its upper limit for the first three ranges and then it decreases (this was not the case for the strike metric). The first two PCs together explain most of the variance in the shortest and longest maturity options, as it was the case in the strike metric. However, as we did in the strike metric, we consider Velicer's procedure only as the first step of deciding on the number of PCs to retain. The second step is to look at the interpretation of the first three PCs[14].

The plots of the correlation loadings revealed that for the ranges 30-10 and for all the years the first PC is like a parallel shift with an attenuation at the edges. In the range 60-30, it has positive correlation loadings for all the years but 92. In the other ranges and for all the years, it has both positive and negative correlation loadings (it is like a Z-shape), especially for the two longest ranges. This is the interpretation that we would like to give to the second PC (but not the first one), where the shape of the shock varied across ranges and across years within each range.

Regarding the shape of the second PC, in the ranges 30-10 and 60-30 and for all the years, it is like a Z-shape. However, in the range 90-60 it has positive correlation loadings for 92, 93, 94 (that is the interpretation that we would like to give to the first PC, but not to the second one). In the year 95, it has both positive and negative loadings. In the range 150-90, it has positive correlation loadings in all the years but 94. In the range 240-150, the shape is like a pyramid, with very big correlation loadings for the ATM options, while for the range 360-240 it is noisy, having positive correlation loadings for all the years but 93.

Finally, the figures for the third PC revealed that it was just noise, something which is consistent with Velicer's procedure results. Given that Velicer's criterion retains at most two PCs, and looking at the interpretation of the third PC, we decide to keep two PCs in the moneyness metric, just as we did in the strike metric. Although we were able to get the shift and slope interpretation for the ranges 30-10 and 60-30, we would like to get the same interpretation for the other ranges and for all the years if possible. Hence, we will have to go through the "Procrustes" type rotation method, in order to get the desired interpretation of the two retained PCs.

5.3 The Rotation Method

The results for the first and second rotated PC, for the years 92 and 93, appear in Figure 6 and for the other two years appear in Figure 7. We can see that the rotation of the first PC was successful since it has brought for all the maturity buckets and for all the years, positive loadings (apart from 240-150 for the year 92). In the range 30-10, the parallel shift with the attenuation at the edges has been maintained, while in the range 60-30 there are positive loadings for all the years, with a much smoother shape than the unrotated ones. The shape

[14] We also looked at the communalities explained by one and two PCs. In general, the number of retained PCs did better in terms of communalities in the moneyness metric, than they did on the strike metric. One PC seems to need the assistance of the second, but not the assistance of the third, since the latter does not increase the explained communalities significantly.

is not far away from indicating a parallel move. In the ranges 90-60 and 150-90 there are positive loadings for all the years. In the ranges 240-150 and 360-240 the pyramid shape is preserved. This is expected, since the rotated first PC is a linear combination of the two unrotated PCs.

For the second PC, the rotation was equally successful, since it produced the Z-shape in the cases where there was not before. So, in the ranges 30-10 and 60-30, it has preserved the Z-shape, while in the remaining ranges it has revealed it.

The interesting point is that in the ranges 150-90, 240-150 and 360-240, it seems as if after the rotation there is a kind of "changing" names between the unrotated components (we will call this the PCs' swap effect); before the rotation the first PC had a Z-shape and the second had positive loadings. After the rotation the first PC has the shift and the second has the Z-shape interpretation. In order to verify the swap effect, we check the percentage of the total variance that each one of the rotated PCs explain. This is compared with the percentage of the total variance that the unrotated PCs explained, so as to see if there is a swap effect in the explained variances, as well.

Table 4 shows the percentage of the variance that the first and second *rotated* PCs explain on the moneyness metric. The first rotated PC explains on average 20%-77.5%, and the second rotated PC explains 9%-64%. These results indicate that the first rotated PC explains less variance than the unrotated one, while the second rotated PC explains more, a result which is similar to that for the strike metric.

Regarding the swap effect, comparing Table 4 to Table 3, we can see that for the ranges 150-90 and 240-150 and for all the years, but 94 in the former bucket and 95 in the latter bucket, the first rotated PC explains roughly the variance that the second unrotated PC explained, while the second rotated PC explains the variance that the first unrotated PC explained. However, this is not the case for 360-240. Therefore, the PC swap effect is confirmed for the ranges 150-90 and 240-150.

From Figures 6 and 7 we can also see whether the retained components affect more the shorter, or the longer maturity options' implied volatilities. We examine this effect, for the ranges 30-10, 60-30, 90-60, and 150-90, over the years 92, 93, 94, and 95, since in the remaining ranges the first rotated PC had the noisy pyramid wise shape. These figures show that the effect of the first PC is pronounced in the 30-10 days range implied volatilities, then it dies down for the 60-30 range and then it comes back for the 90-60 range, while it seems that the influence falls again in the 150-90 range. This is different from the results from the strike metric analysis, where the first PC affected more the longer maturity options' implied volatilities. In contrast to the effect of the first component, it seems as if the effect of the second rotated PC is the same for all the maturity buckets. This pattern is repeated for all the years, just as it was the case in the strike metric.

	Year	Unrot. 1st PC	1st PC	2nd PC	Cumulative
30-10	92	73.30%	73.30%	11.10%	84.40%
	93	77.80%	77.80%	8.10%	85.80%
	94	85.70%	85.70%	6.40%	92.10%
	95	73.00%	73.00%	11.10%	84.10%
	average	77.45%	77.50%	9.20%	86.60%
60-30	92	39.60%	36.00%	33.90%	69.90%
	93	44.60%	44.40%	28.90%	73.30%
	94	53.60%	53.20%	23.30%	76.50%
	95	47.60%	44.70%	28.30%	73.10%
	average	46.35%	44.60%	28.60%	73.20%
90-60	92	48.80%	40.00%	36.10%	76.10%
	93	46.80%	33.50%	43.90%	77.40%
	94	44.20%	37.90%	35.90%	73.90%
	95	39.70%	31.80%	33.50%	65.20%
	average	44.88%	35.80%	37.40%	73.20%
150-90	92	51.50%	19.90%	48.40%	68.30%
	93	48.10%	19.60%	45.40%	65.00%
	94	54.20%	34.00%	38.30%	72.30%
	95	48.80%	25.50%	41.30%	66.90%
	average	50.65%	24.80%	43.40%	68.10%
240-150	92	69.20%	22.20%	59.00%	81.20%
	93	68.70%	15.80%	63.20%	79.00%
	94	63.40%	17.20%	63.10%	80.30%
	95	60.70%	25.10%	48.40%	73.50%
	average	65.50%	20.10%	58.40%	78.50%
360-240	92	77.90%	14.90%	77.10%	92.00%
	93	75.10%	50.70%	41.30%	92.00%
	94	85.10%	24.50%	68.50%	92.90%
	95	80.90%	16.80%	70.90%	87.70%
	average	79.75%	26.70%	64.50%	91.20%

Table 4: Percentage of Variance Explained by the Unrotated first PC and by the Rotated PCs in the Moneyness Metric.

6 Correlations between the Futures Price and the Principal Components

So far, we have investigated the number and shape of shocks appearing in equation (2). However, in order to implement the models proposed by Dupire [10] and Derman and Kani [9] (e.g., by Monte Carlo simulation) it is necessary to know the sign and the size of the correlations, between the Brownian motions of the processes for the underlying asset and the forward volatility. In this section, we calculate the correlations between the changes in the future price and the changes of the principal components.

The correlations are measured by using the Pearson coefficient. Since the desired interpretation comes (most of the time) from the rotated PCs, we look in Table 5 at the correlations between the changes of the futures price with the changes of each one of the first two *rotated* principal components, under the strike and moneyness metric, respectively. We calculated also the correlations by using the non-parametric Spearman's coefficient in order to capture any non-linear association. However, the results were the same as with Pearson's coefficient, and hence we prefer not to report them. One asterisk is displayed when the coefficient is significant at 5% significance level, and two asterisks are displayed when the coefficient is significant at 10% significance level.

We can see that in the strike metric the correlations between the changes of the rotated first principal component with the changes of the futures price are positive, in most of the cases. Negative correlations occur in the range 150-90 for the year 94, and in the ranges 240-150 and 360-240 for the year 92. On the other hand, in the moneyness metric, the correlations between the changes of the rotated first principal component with the changes of the futures price are negative. The only exception occurs in the range 360-240 for the years 93, 94, and 95. This could be justified as a leverage effect (see Christie [8]). Notice that the sign of the correlations between the changes of the first unrotated PC with the changes of the futures price depends on the metric[15]. In the strike metric, the correlations between the changes of the rotated second principal component with the changes of the futures price are positive, in most of the cases. Exceptions occur in the range 240-150 for the year 95 and in the range 360-240 for the year 92. In the moneyness metric the correlations are always positive.

However, the size of the correlation coefficients changes over years and over ranges of days to maturity. This is not surprising because the correlations depend on the variances. Therefore, in the case that variances follow a stochastic process, the correlations should vary stochastically, as well.

[15] Since the correlation has the same sign as the covariance, it is easy to show that the correlation depends on the metric that we work on, by looking at the covariance between $\Delta\sigma_t = \sigma_{t+1} - \sigma_t$ and ΔF under both metrics. Say that the covariance in the strike metric is $Cov_{strike} = Cov(\Delta\sigma_t(K), \Delta F)$. Then, in the moneyness metric, for a given moneyness level we have $Cov_{mon} = Cov[\sigma_{t+1}(K+\Delta F) - \sigma_t(K), \Delta F]$. Expanding $\sigma_{t+1}(K+\Delta F)$ as a Taylor series of order one around a point K yields $Cov_{mon} = Cov[\sigma_{t+1}(K) + \Delta F\sigma'_{t+1} - \sigma_t(K), \Delta F]$ $= Cov_{strike} + \sigma'_{t+1} Var(\Delta F)$. Therefore, the correlation sign depends on the slope of the skew, the point around which the expansion is performed, and the variance of ΔF.

	Coefficient		92	93	94	95
30-10	Strike	$\Delta PC1$	0.04	-0.06	-0.18	-0.01
		$\Delta PC2$	0.38**	0.34**	0.48**	0.00
	Moneyness	$\Delta PC1$	-0.40**	-0.52**	-0.65**	-0.42**
		$\Delta PC2$	0.12	0.01	0.05	-0.06
60-30	Strike	$\Delta PC1$	0.23**	0.11	-0.10	0.19*
		$\Delta PC2$	0.31**	0.31**	0.36**	0.06
	Moneyness	$\Delta PC1$	-0.35**	-0.38**	-0.58**	-0.41**
		$\Delta PC2$	0.24**	0.22**	0.25**	0.11
90-60	Strike	$\Delta PC1$	-0.15	0.20*	0.10	0.29**
		$\Delta PC2$	0.32**	0.27**	0.39**	0.15
	Moneyness	$\Delta PC1$	-0.36**	-0.49**	-0.49**	-0.37**
		$\Delta PC2$	0.25*	0.17	0.33**	0.37**
150-90	Strike	$\Delta PC1$	0.30**	0.23**	-0.35**	0.28**
		$\Delta PC2$	0.36**	0.36*	0.24**	0.28**
	Moneyness	$\Delta PC1$	-0.28**	-0.31**	-0.19*	-0.32**
		$\Delta PC2$	0.33**	0.26**	0.37**	0.16
240-150	Strike	$\Delta PC1$	-0.31**	0.12	0.30**	0.34**
		$\Delta PC2$	0.15*	0.38**	0.38**	-0.17*
	Moneyness	$\Delta PC1$	0.08	0.04	-0.28**	0.08
		$\Delta PC2$	0.28**	0.36**	0.37*	0.38**
360-240	Strike	$\Delta PC1$	-0.38**	0.33**	0.33**	0.05
		$\Delta PC2$	0.04	0.40**	0.41**	0.27**
	Moneyness	$\Delta PC1$	-0.45**	0.33**	0.14	0.27*
		$\Delta PC2$	0.31**	0.35**	0.47**	0.52**

Table 5: Correlations between Changes of the Futures Price with Changes of the Rotated PCs in the Strike and Moneyness Metrics.

7 Conclusions

We applied Principal Components Analysis (PCA) to the first differences of implied volatilities for six fixed ranges of days to maturity for the years 1992-1995. The results from this study contribute to implementing of Dupire's and Derman and Kani's models. In particular, we investigated the number of shocks that drive the forward volatility process, their interpretation and their correlation with the changes in the underlying asset.

Since PCA is sensitive to the choice of the metric that it is performed on and since there are strong theoretical reasons that suggest that smiles are a function of moneyness we applied the technique to two metrics: the strike and the moneyness. As a first stage to the determination of the number of PCs to be retained, we did not rely on rules of thumb. Instead, we used Velicer's non-parametric criterion. This retained in both metrics, either one or two PCs, depending on the maturity bucket under scrutiny. As a second stage, we looked at the interpretation of the first three PCs. We found that the third PC was just noise in both metrics. Combining Velicer's criterion with the results from the explained communalities and the interpretation of PCs we decided to keep two PCs for both metrics. This is in contrast to the three PCs that researchers kept in the interest rate literature. The two retained PCs explain across the four years 60%-80% (70%-90%) of the total variance of the changes of implied volatilities in the strike (moneyness) metric on average, across the various maturity buckets.

Since a simple interpretation of the retained PCs would help in the implementing the above mentioned models, we would like to interpret the first PC as a (parallel) shift and the second one as a slope. Intuitively this is what a Taylor's expansion tells us should be the case. However, the results from the interpretation of the first two PCs, showed that we have the expected shape only in the shortest maturities. Therefore, a rotation of the first two PCs had to be performed in order to get the desired simple interpretation. It turned out that the "Procrustes" type of rotation that we developed was appropriate for such a purpose. The rotation method was successful in that it delivered the desired interpretation for both PCs in the moneyness metric and for the second PC in the strike metric. It delivered the desired interpretation for the first PC in the strike metric for the first three ranges.

Finally, we looked at the correlations between the changes of the futures price and each one of the first two rotated PCs in the strike and the moneyness metric. We found that the correlation for the first rotated PC is positive in the strike metric, while it becomes negative in the moneyness metric. The correlation for the second rotated PC is positive in both metrics.

A question that should be answered by future research is whether our results about the dynamics of implied volatilities, are robust for other data sets, since it is well documented that the magnitude of smiles (or skews) depends on the underlying asset (see for example Tompkins [26]).

References

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

[2] Barnett, V. and Lewis, T. (1994): Outliers in Statistical Data, 3rd Edition, Wiley Series in Probability and Mathematical Statistics.

[3] Barone-Adesi, G. and Whaley, R. (1987): "Efficient Analytic Approximation of American Option Values", *Journal of Finance*, 42, 301-320.

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

[5] Bodurtha, J. and Courtadon, G. (1987): "Tests of an American Option Pricing Model on the Foreign Currency Options Markets", *Journal of Financial and Quantitative Analysis*, 22, 153-167.

[6] Anderson, A., Basilevsky, A. and Hum A. (1983). "Missing Data: A Review of the Literature" in Rossi, P., Wright, J., and Anderson, A., Handbook of Survey Research, Academic Press.

[7] Canina, L. and Figlewski, S. (1993): "The Informational Content of Implied Volatility", *Review of Financial Studies*, 6, 659-681.

[8] Christie, A. (1982): "The Stochastic Behavior of Common Stock Variances, Value, Leverage and Interest Rate Effects", *Journal of Financial Economics*, 10, 407-432.

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

[10] Dupire, B. (1992): "Arbitrage Pricing with Stochastic Volatility", Working Paper, Societe Generale Division Options, Paris.

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

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

[13] Heynen, R. (1993): "An Empirical Investigation of Observed Smile Patterns", Working paper, Tinberger Institute, Erasmus University Rotterdam.

[14] Hull, J. and White, A. (1987): "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance* 3, 281-300.

[15] Jackson, E. (1991): A User's Guide to Principal Components, Wiley Series in Probability and Mathematical Statistics.

[16] Johnson, H. and Shanno, D. (1987): "Option Pricing when the Variance is Changing", *Journal of Financial and Quantitative Analysis* 22, 143-151.

[17] Kon, S. (1984): "Models of Stock Returns: A Comparison", *Journal of Finance* 39, 147-166.

[18] Litterman, R. and Scheinkman, J. (1988): "Common Factors Affecting Bond Returns", Goldman Sachs, Financial Strategies Group.

[19] Reddon, J. (1985): "Monte Carlo type I Error Rates for Velicer's Partial Correlation test for the Number of Principal Components", *Criminometrica*, 1, 13-23.

[20] Roll, R. (1984): "A Simple Implicit Measure of the Effective Bid-Ask Spread", *Journal of Finance*, 39, 1127-1139.

[21] Rubinstein, M. (1985): "Non-Parametric Tests of Alternative Option Pricing Models", *Journal of Finance*, 40, 455-480.

[22] Scott, L. (1987): "Option Pricing when the Variance Changes Randomly: Theory, Estimation and an Application", *Journal of Financial and Quantitative Analysis* 22, 419-438.

[23] Sheikh, A. (1991): "Transaction Data Tests of S&P 100 Call Option Pricing", *Journal of Financial and Quantitative Analysis*, 26, 459-475.

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

[25] Taylor, S. and Xu, X. (1993): "The Magnitude of Implied Volatility Smiles: Theory and Empirical Evidence for Exchange Rates", Working Paper, Financial Options Research Centre.

[26] Tompkins, R. (1998): "Implied Volatility Surfaces: Uncovering Regularities for Options on Financial Futures", Working paper, Financial Options Research Centre, University of Warwick.

[27] Velicer, W. (1976): "Determining the Number of Components from the Matrix of Partial Correlations", *Psychometrica*, 41, 321-327.

[28] Wiggins, J. (1987): "Option Values under Stochastic Volatility", *Journal of Financial Economics* 19, 351-372.

[29] Xu, X. and Taylor, S. (1994): "The Term Structure of Volatility Implied by Foreign Exchange Options", *Journal of Financial and Quantitative Analysis*, 29, 57-74.

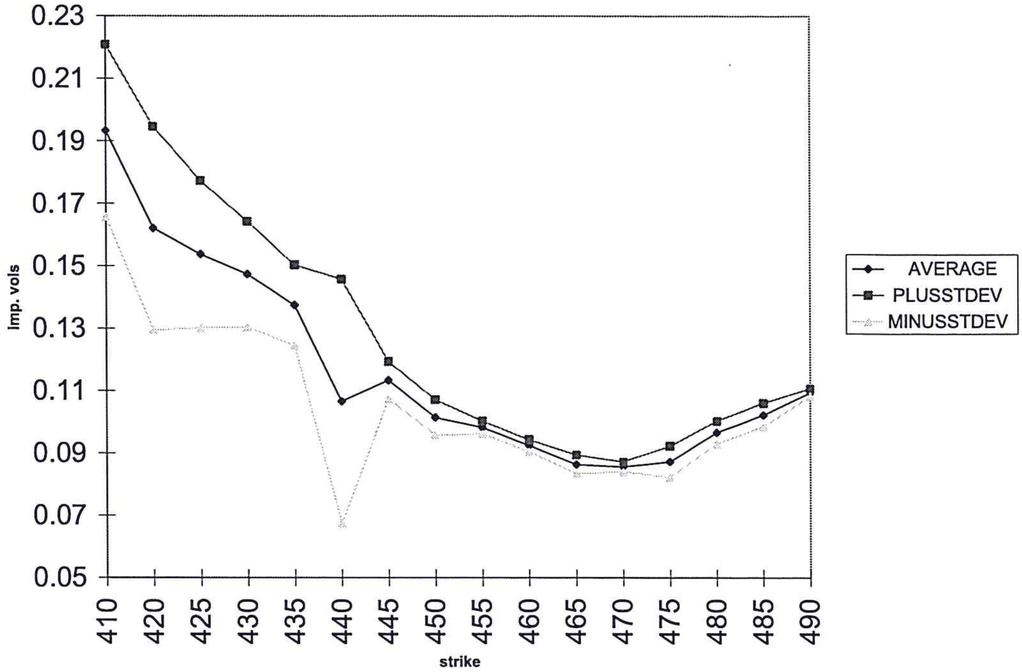

Figure 1: Unfiltered Data for a Call Contract with 9309 expiry date.

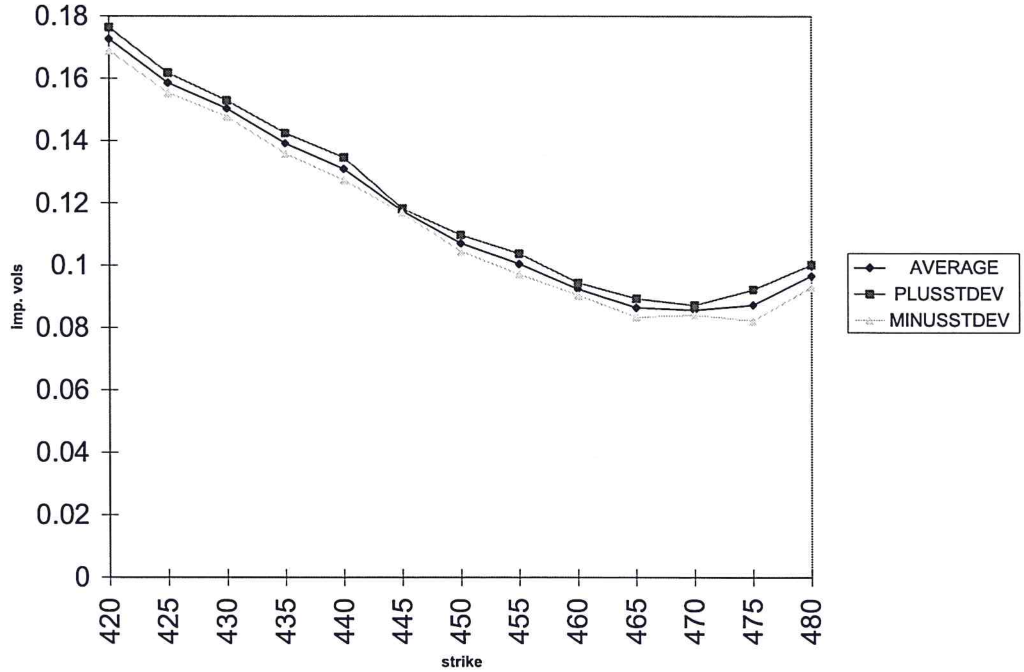

Figure 2: Recovered Smile From Calls and Puts that expire in 9309.

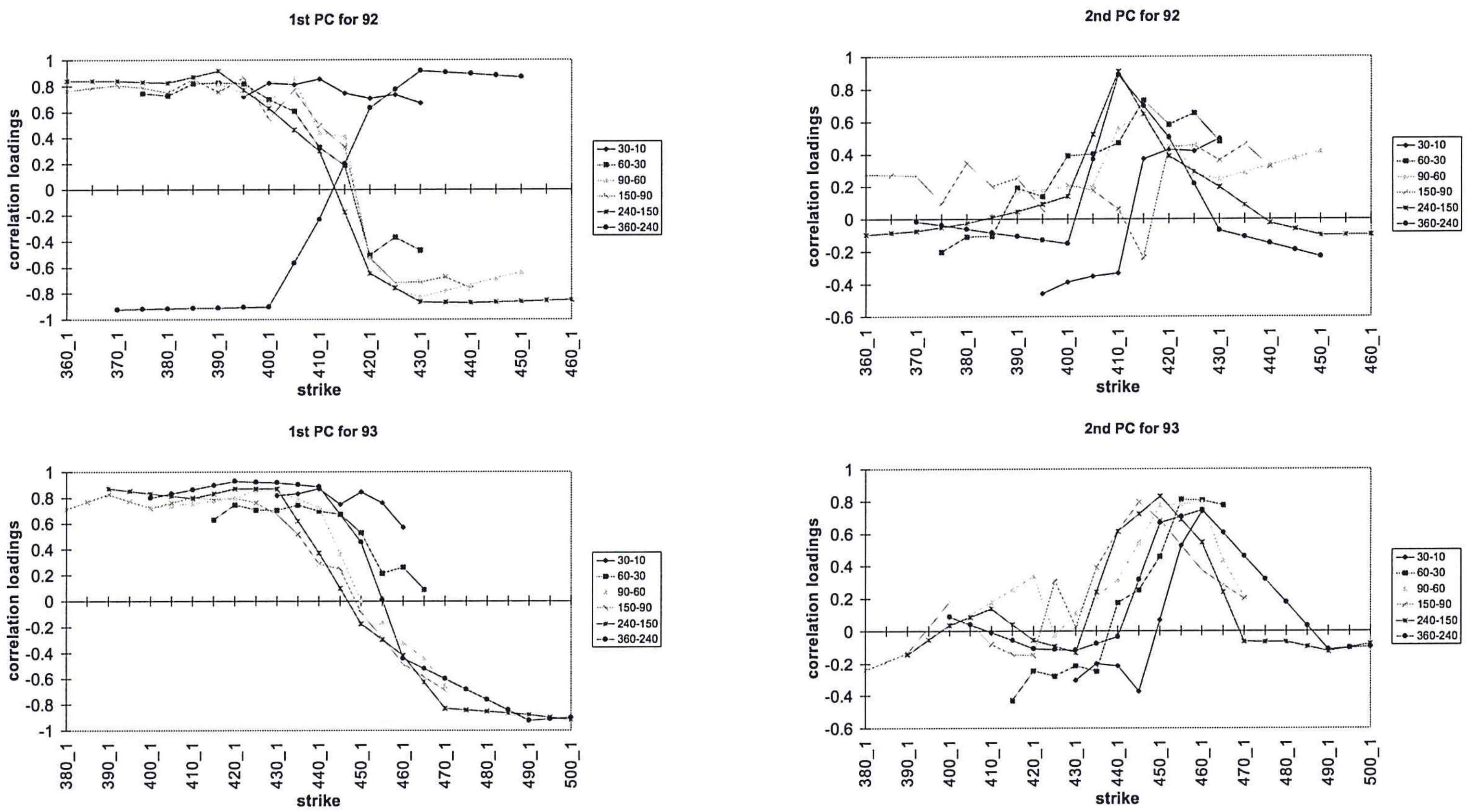

Figure 3: First and Second PCs for Years 92 and 93.

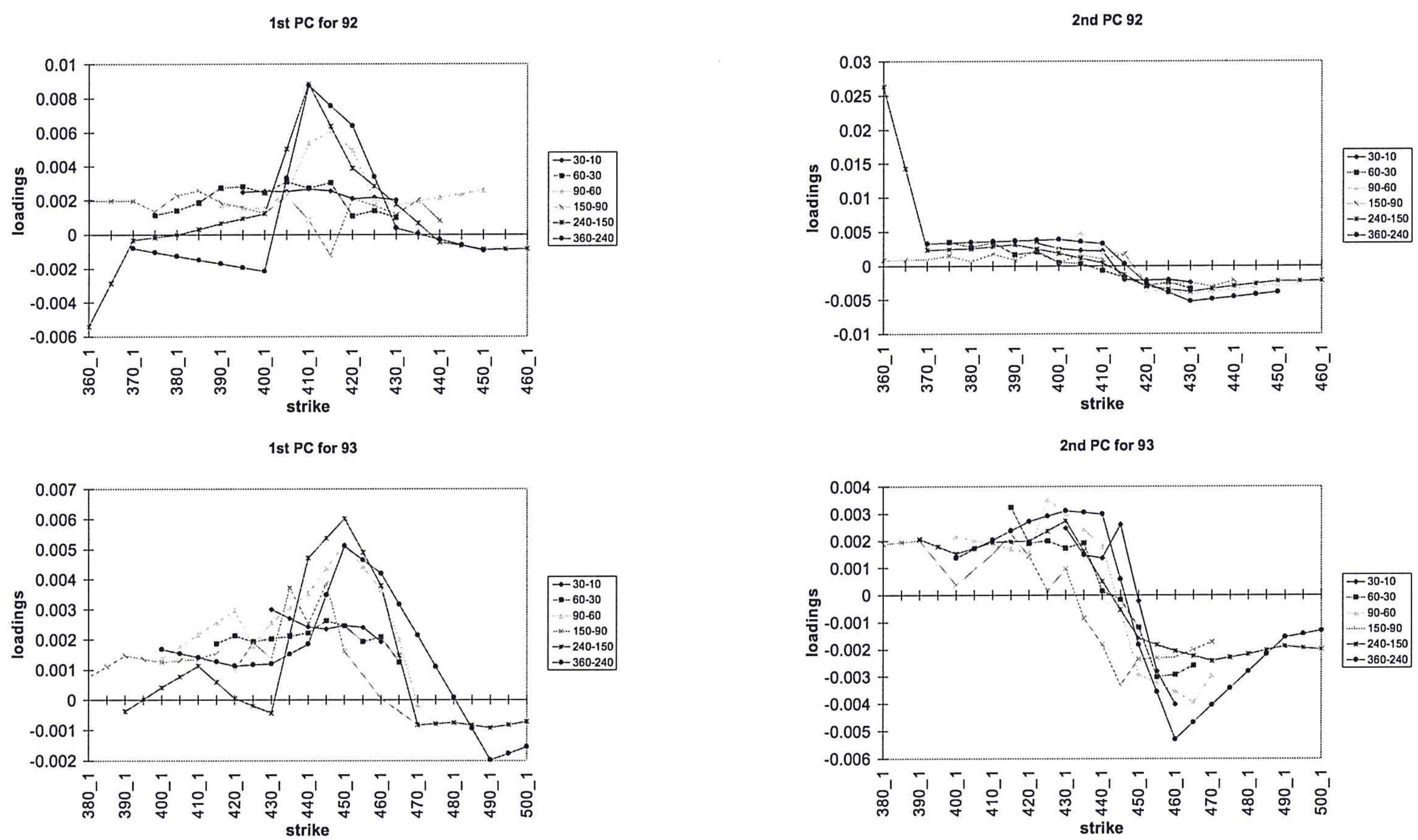

Figure 4: First and Second Rotated PCs for Years 92 and 93.

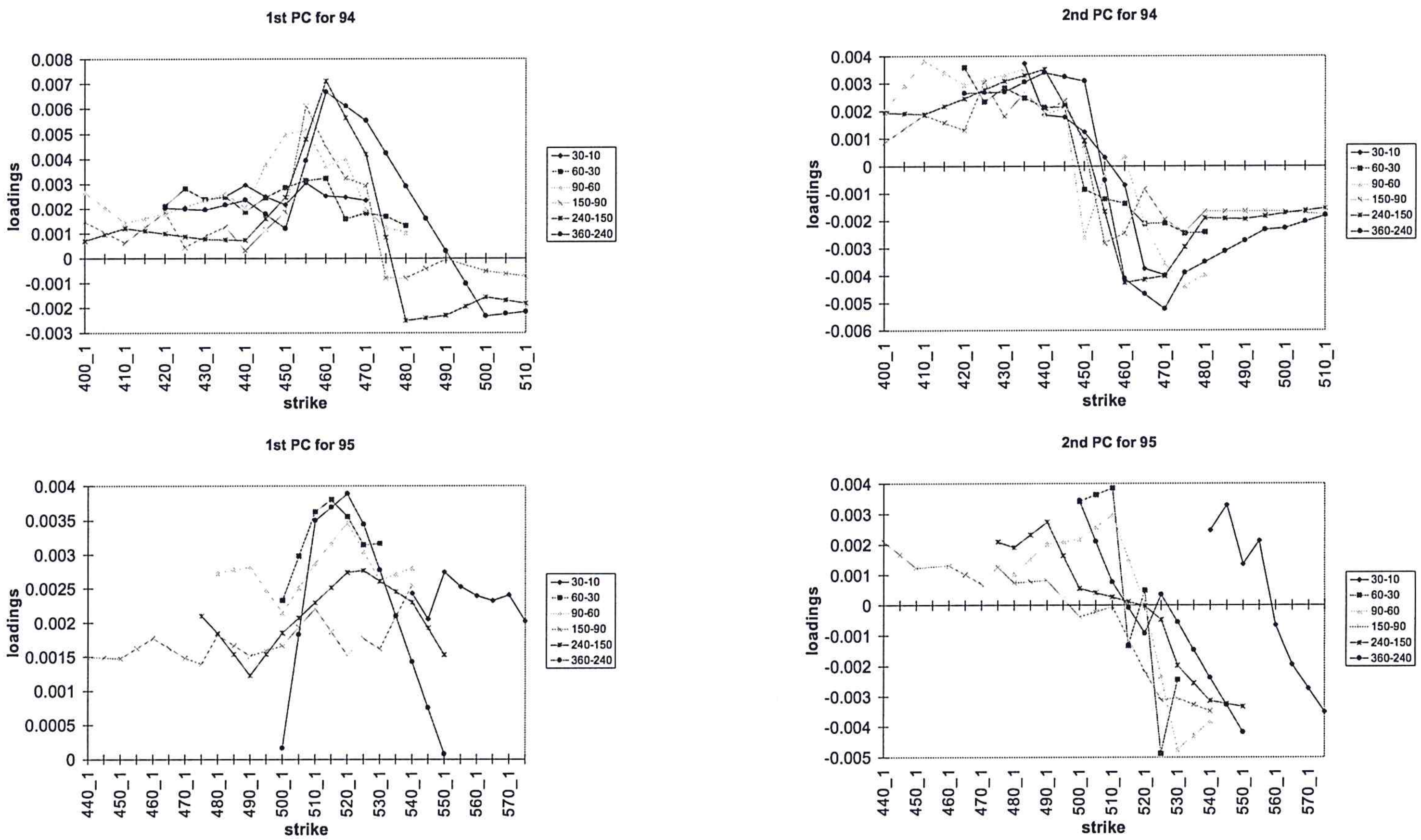

Figure 5: First and Second Rotated PCs for Years 94 and 95.

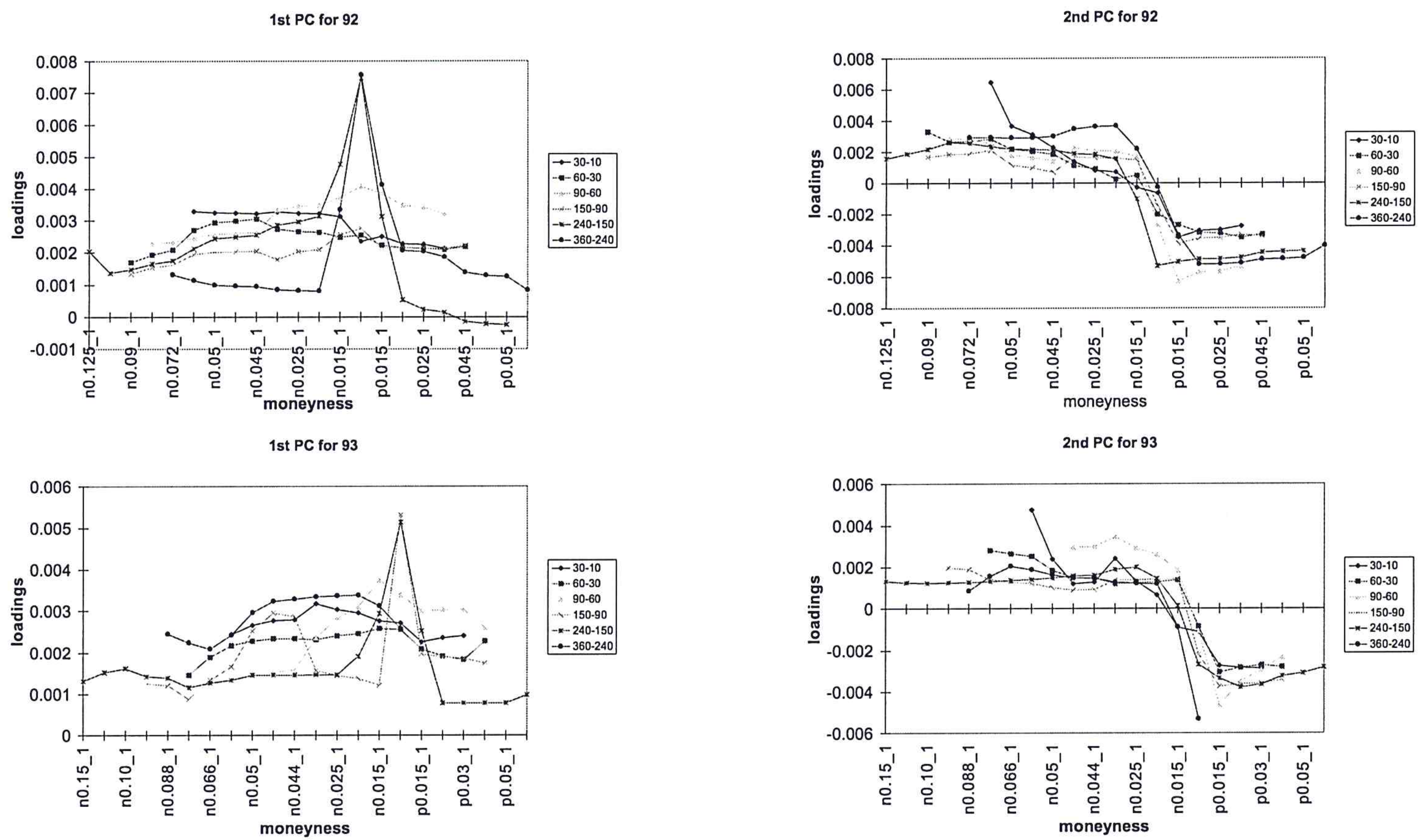

Figure 6: First and Second Rotated PCs for Years 92 and 93.

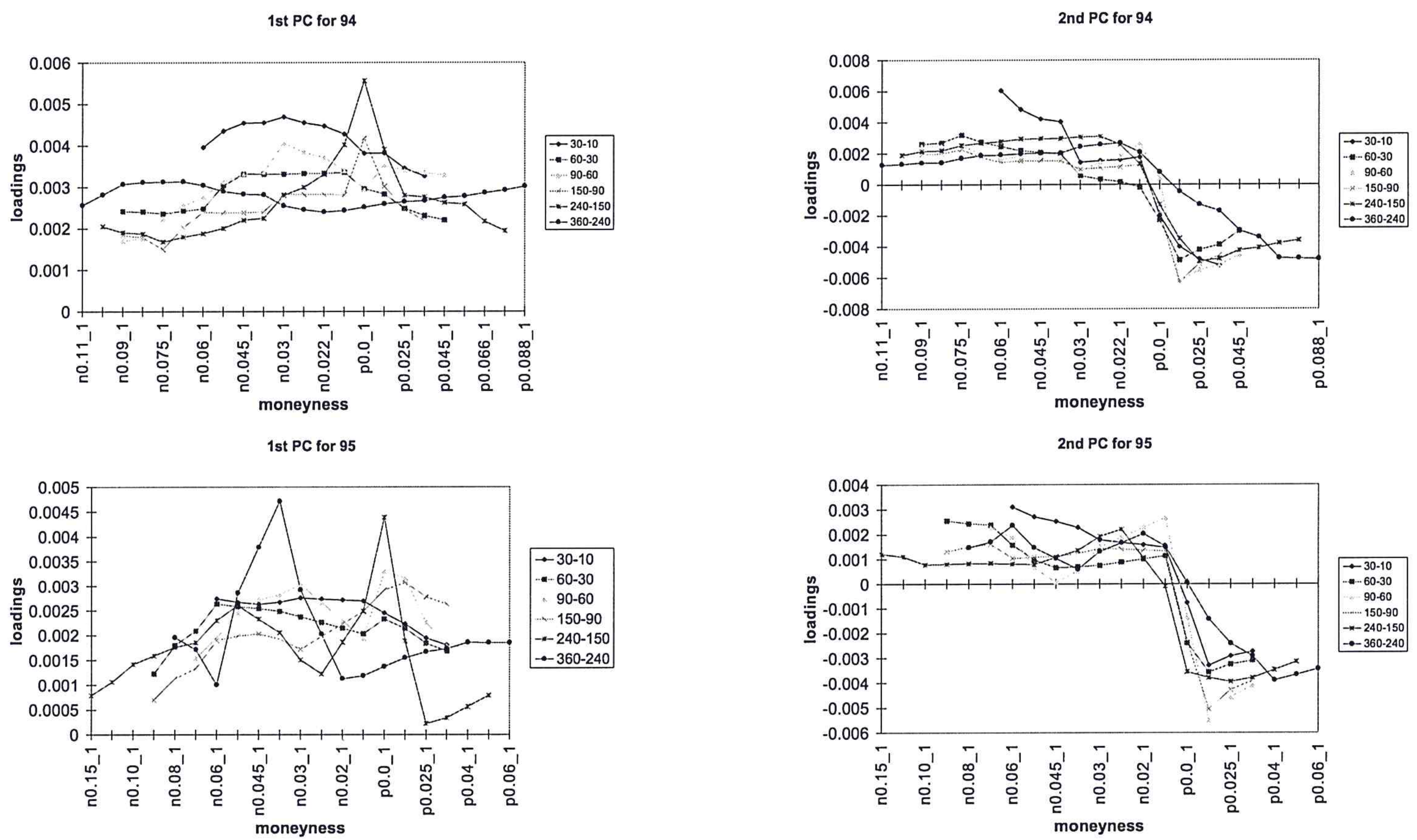

Figure 7: First and Second Rotated PCs for Years 94 and 95.