

Fengler, Matthias R.; Härdle, Wolfgang Karl; Villa, Christophe

**Working Paper**

## The dynamics of implied volatilities: A common principal components approach

SFB 373 Discussion Paper, No. 2001,38

**Provided in Cooperation with:**

Collaborative Research Center 373: Quantification and Simulation of Economic Processes, Humboldt University Berlin

*Suggested Citation:* Fengler, Matthias R.; Härdle, Wolfgang Karl; Villa, Christophe (2001) : The dynamics of implied volatilities: A common principal components approach, SFB 373 Discussion Paper, No. 2001,38, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin, <https://nbn-resolving.de/urn:nbn:de:kobv:11-10049795>

This Version is available at:

<https://hdl.handle.net/10419/62716>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The Dynamics of Implied Volatilities: A Common Principal Components Approach\*

Matthias R. FENGLER<sup>a</sup>, Wolfgang K. HÄRDLE<sup>a</sup>, Christophe VILLA<sup>b</sup>

<sup>a</sup>*Institute for Statistics and Econometrics, Department of Business Administration and Economics,  
Humboldt-Universität zu Berlin, Spandauer Straße 1, 10187 Berlin, Deutschland*

<sup>b</sup>*University of Rennes 1 - IGR, CREREG-Axe Finance and CREST-ENSAI,  
11 rue Jean Mace, F-35019 Rennes, France*

June 2001

---

## Abstract

It is common practice to identify the number and sources of shocks that move implied volatilities across space and time by applying Principal Components Analysis (PCA) to pooled covariance matrices of changes in implied volatilities. This approach, however, is likely to result in a loss of information, since the surface structure of implied volatilities in the maturities and moneyness dimension is neglected. In this paper we propose to estimate the implied volatility surface at each point in time nonparametrically and to analyze the implied volatility surface slice by slice with a common principal components analysis (CPCA). As opposed to traditional PCA, the basic assumption of CPCA is that the space spanned by the eigenvectors is identical across groups, whereas variances associated with the components are allowed to vary. This allows us to study a  $p$  variate random vector of  $k$  groups, say the "volatility smile" at  $p$  different grid points of moneyness for  $k$  maturities, simultaneously. Our evidence suggests that surface dynamics can indeed be traced back to a common eigenstructure between covariance matrices of the surface "slices", which allow for the usual shift, slope, and twist interpretation of shocks to implied volatilities. This insight is a suitable starting point for VaR Monte Carlo Simulations of delta-gamma neutral, vega sensitive option portfolios.

---

\*We gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 373 "Quantifikation und Simulation ökonomischer Prozesse". We would like to thank the participants of the Seminar in Mathematical Statistics, WIAS Berlin, the Conference of the German Finance Association, the HEC Finance Seminar (Paris), the CMAP Seminar (Ecole Polytechnique) for comments and helpful discussions.

Corresponding authors: fengler@wiwi.hu-berlin.de, haerdle@wiwi.hu-berlin.de, christophe.villa@univ-rennes1.fr.

**Keywords:** Common Principal Component Analysis, Implied Volatility Surface, Principal Component Analysis, Smile  
**JEL Classification Codes:** C13, C49, G13

## 1 Introduction

Understanding volatility of financial assets has become a first rank issue in modern financial theory and practice: Whether in risk management, portfolio hedging, or option pricing, a precise notion of the market's assessment and expectation of volatility is clearly inevitable. Much research has been spent on *realized historic volatilities*, (Roll, 1977, and references therein). However, since it seems unsettling to draw conclusions from past to expected behavior of volatility, recently, the focus shifted to *implied volatilities* (Dumas, Fleming, and Whaley, 1998). To derive implied volatilities the Black and Scholes formula is solved for the volatility parameter  $\sigma$  of the underlying asset price dynamics using observed option prices. Inferring volatility directly from observed option prices is more natural as the option value is decisively determined by the market's assessment of current and future volatility. Hence implied volatility may be interpreted as the market's expectation of average volatility over the remaining lifetime of the option.

As is well known, the volatilities implied by observed market prices exhibit a pattern that is far different from that actually assumed for deriving the Black Scholes formula: instead of being constant across strikes and time to maturity, implied volatility appears to be non flat, a stylized fact which has been called "smile" effect. Especially, options far out of the money have higher implied volatility than those with an exercise price at the money, though having the same features otherwise.

Various attempts have been made in order to explain this departure from the Black and Scholes model, the most prominent of which is to assume stochastic volatility (Hull and White, 1987; Johnson and Shanno, 1987; Scott, 1987; Wiggins, 1987; Härdle and Hafner, 2000). Recently, this interpretation has been supported by the work of Dumas, Fleming, and Whaley (1998) who present evidence against the deterministic volatility function hypothesis.

However, unlike Hull and White (1987) and others, the approach taken in this paper is not to explain deformations or deviations of implied volatilities from the common Black and Scholes model, but to regard implied volatilities observed in the markets as yet another financial variable interesting by itself. Clearly, this notion seems to be unsatisfactory from a theoretical standpoint, however, it is not without merit: First, Black Scholes implied volatilities play an important rôle

for practitioners, since they serve a valuable one-to-one mapping from the spaces of option prices, strike prices, interest rates, and maturities to the (positive) real line. This reduces considerably the amount of information to be taken into account when taking investment decisions. Moreover, with growing liquidity of organized option markets, traders are interested in investing in "volatility" itself, i.e. in setting up portfolios that only have a vega sensitivity. This volatility again may be measured in terms of implied volatilities, such as in the VDAX, a product of the EUREX. Second, from a theoretical point of view, new attention has been drawn to implied volatilities with the emergence of the *market models of volatility*. Originally inspired by the market models of the term structure of interest rates (Miltersen, Sandmann, and Sondermann, 1997; Jamshidian, 1997), they were set into the current framework by Ledoit and Santa-Clara (1998) and Schönbucher (1999) amongst others. Market models of volatility assume the existence of a sufficient number of traded plain vanilla options which are used as input variables to price exotic and illiquid options consistently with the smile.

Yet, practically, to set up a model that truly reflects the dynamics of implied volatilities – be it for trading, pricing or risk management – one has to identify the number and shapes of the shocks that move the volatility surface across space and time. Borrowing from the literature of the term structure of interest rates (Rebonato, 1998; Bliss, 1997), the most common technique employed is principal component analysis (PCA). When applied to the term structure of implied volatilities of ATM options (Avellaneda and Zhu, 1997; Härdle and Schmidt, 2000; Sylla and Villa, 2000) or to a smile at a given maturity (Alexander, 2001), this approach carries directly over from the interest rate literature. However, there is an important difference between principal component analysis of interest rates and implied volatilities: Implied volatilities have both a strike *and* a term structure dimension, and hence essentially constitute a three dimensional data set, a fact that has to our knowledge been only accounted for by stacking and grouping of variables. Derman and Kamal (1997) analyze changes in implied volatilities by stacking the surface into a vector with a dimension equal to the number of grid points in the strike dimension (here a delta metric) times the number of grid points in the term structure of the options. Thus, surface dynamics are given by a multi-variate time series and standard PCA can be applied. However, this approach neglects the natural group structure in maturities. The grouping approach is taken by Skiadopoulos, Hodges, and Clewlow (1999) who form three large "maturity buckets" in the surface, average implied volatilities of options whose maturities fall into them and apply PCA to each bucket covariance matrix separately. Besides neglecting the surface structure, this approach may result in a hybrid set of principle components disturbed by both within and between group variation and are hence difficult to interpret (Basilevsky, 1994, p. 313). Most importantly, however, both approaches

fail to separate between *common* and *specific* factors that drive the implied volatility surface, a task which is at the heart of our study.

We depart from the research mentioned before in two important ways: First, we estimate nonparametrically the implied volatility surface day by day with a procedure laid out in Härdle and Vieu (1992); Härdle and Tsybakov (1997); Aït-Sahalia and Lo (1998, 2000). In smoothing the volatility surfaces, we recover the time series of implied volatilities  $\{\hat{\sigma}_t(\kappa, \tau)\}_{t=1}^T$  on a given grid of moneyness  $\kappa$  and maturity  $\tau$ . Second, we apply an old technique of multi-variate analysis to the implied volatility surface: *common* principle component analysis. This method seems to be the preferred modeling approach of implied volatilities, as it exploits a natural group structure in the data and does not rely on a pooling methodology. Third, in the maximum likelihood framework we use, we may test the specification of our models.

Common principle component analysis (CPCA) is based on the assumption that covariance matrices share a common structure across groups. The basic idea of CPCA is that the space spanned by the eigenvectors is identical across groups, whereas variances associated with the components are allowed to vary. With CPCA, we may analyze a  $p$  variate random vector of  $k$  groups, say  $k$  maturities of implied volatilities *simultaneously*. Hypothesis testing for this structure has been derived and allows us to test for the validity of the CPCA specification versus other possible structures. More precisely, we may test for the hypothesis of only  $q \leq p - 2$  common eigenvectors, while the remaining  $p - q$  eigenvectors are specific in each group. This model has been called *partial* common principle components model of order  $q$ , pCPC( $q$ ). The decisive advantage of the common principal components framework is that it is mathematically appealing, and empirically parsimonious.

Our evidence suggests that the CPC model is appropriate for analyzing the implied volatility surface, especially for the shorter maturities. We interpret our findings as supporting our notion that pooling can distort estimates. As generally found, the shift, slope, and twist interpretation of implied volatilities, known from the literature of the term structure of interest rates is also revealed. Moreover, the common factors extracted from the volatility surface are suitable a starting point for VaR Monte Carlo Simulations of delta-gamma neutral, vega sensitive option portfolios.

In the following section, we give a graphical motivation that suggests naturally the hypothesis of a CPC model to be natural. Section 3 describes the CPC and the partial CPC model, parameter estimation, and a hierarchy of covariance matrices linking the various models. In Section 4 the empirical results are presented, and Section 5 concludes.

## 2 An Intuitive Motivation of Common Principal Components

Principal component analysis, as originated by Hotelling (1933) and Pearson (1901) is a one-group method. In many applications, however, the data obviously fall into groups in which the same variables are measured. Since for options at each point in time only a very limited number of maturities exist, whereas in the strike dimension a large number of options is available, one may consider the implied volatility surface as falling into maturity groups. In this case it is natural to assume that the covariance structure between groups should be similar to some extent. The CPC model hypothesizes that the space spanned by the eigenvectors from group covariance matrices is common to all groups, while eigenvalues are different. In this section, we will demonstrate how this assumption arises from maturity groups of implied volatilities.

Denote  $\{\hat{\sigma}_t(\kappa, \tau)\}_{t=1}^T$  the time series of implied volatilities at time  $t$  as a function of moneyness  $\kappa$  and maturity  $\tau$  measured in months. Moneyness is defined as  $\kappa = \frac{K}{F_t}$ , where  $K$  is the option's strike price, and  $F_t = S_t e^{r\tau}$  the implied future price. As we recover implied volatilities on a constant grid of moneyness  $\kappa_i$  and moneyness  $\tau_j$  (see appendix), the times series of volatility returns, i.e. log-differences of implied volatilities,  $\{\Delta \ln \hat{\sigma}_t(\kappa_i, \tau_j)\}_{t=2}^T$ , are well defined. In Figure 1, we show a scatter plot of  $\{\Delta \ln \hat{\sigma}_t(0.90, 1)\}_{t=2}^T$  against  $\{\Delta \ln \hat{\sigma}_t(1.10, 1)\}_{t=2}^T$  in the left panel, and in the right panel a scatter plot of  $\{\Delta \ln \hat{\sigma}_t(0.90, 3)\}_{t=2}^T$  against  $\{\Delta \ln \hat{\sigma}_t(1.10, 3)\}_{t=2}^T$ , i.e. for two different maturity groups, one and three months respectively, we plot the volatility returns of two different moneyness against each other. Additionally we display the principle axes and the ellipses of constant standard distance, the so called square root of the Mahalanobis distance, (Mardia, Kent, and Bibby, 1992), from the respective mean vector, i.e.

$$\left\{ (x - \bar{x}_j)^T S_j^{-1} (x - \bar{x}_j) \right\}^{-\frac{1}{2}} = 2 \quad j = 1, 2$$

where  $x = \Delta \ln \hat{\sigma}_t(\kappa, \tau)$ .  $S_j$  is the sample covariance matrix and  $\bar{x}_j$  the mean vector in group  $j$ . The ellipse may be interpreted as the approximate 95% confidence region if our sample were drawn from a bivariate normal distribution.

As can readily be seen, the principle axes or the eigenvectors  $\Gamma_j = (\gamma_{j1}, \gamma_{j2})$ , are almost parallel (Figure 1; Table 1). Variability, however, either reflected in the size of the eigenvalues  $\lambda_{ji}$  or in the size of the ellipses, is different across groups. This illustrates the well-known fact that volatility returns for short maturities are more volatile and hence more spread out in space than for long maturities. Given these insights, it seems natural to assume a model where differences of the eigenvectors are attributed to sampling variability and are restricted to be common, while

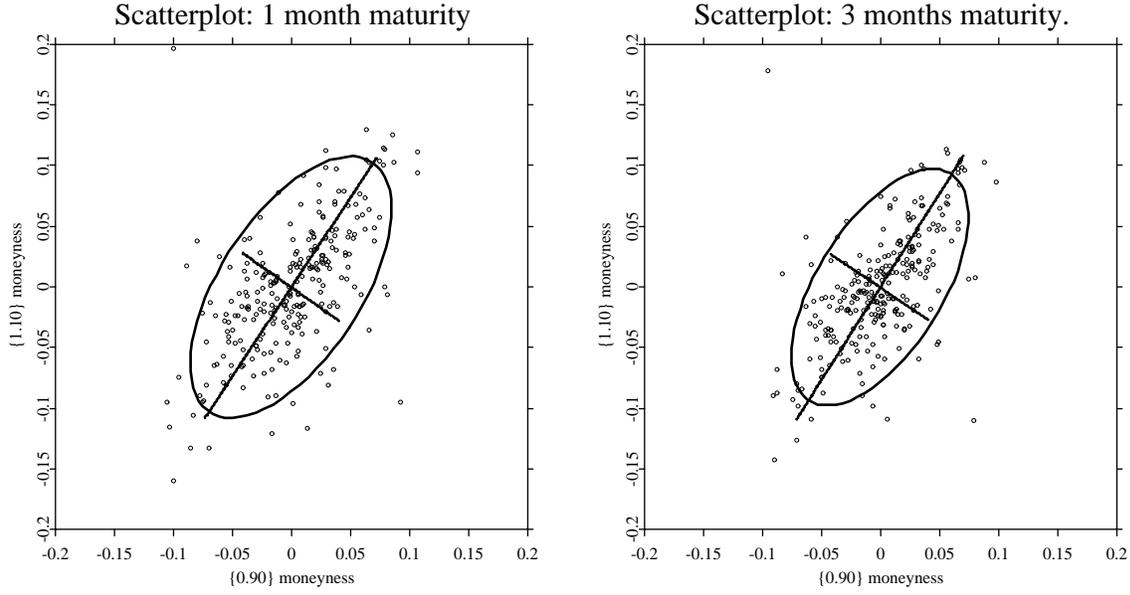


Figure 1: *Principle axes obtained by separate PCA for groups of 1 month and 3 months maturity; moneyness is  $\kappa = 0.90$  against  $\kappa = 1.10$ , ODAX 1999.*

eigenvalues are allowed to be different. This is shown in Figure 2, which illustrates the same scatter plots with principle axes and ellipses estimated under the restriction of a common transformation of both groups, i.e. by assuming the model

$$\begin{aligned} \Psi_1 &= \Gamma \Lambda_1 \Gamma^T \\ \text{and } \Psi_2 &= \Gamma \Lambda_2 \Gamma^T \quad \Lambda_i = \text{diag}(\lambda_{i1}, \lambda_{i2}), \end{aligned}$$

where  $\Gamma$  is an orthogonal transformation matrix and  $\Lambda_i = \text{diag}(\lambda_{i1}, \lambda_{i2})$  the matrices of eigenvalues in group  $i = 1, 2$ . Results are displayed in Table 2. This is exactly the CPC model for the dimension  $p = 2$ ; the formal presentation of the  $p$ -dimensional case is delayed until Section 3.1.

Since higher dimensional illustrations are limited to the dimension of three, consider the parallel coordinate plots in Figures 3 to 5. Here, we display the  $p = 6$  coordinates of the eigenvectors associated with the three largest eigenvalues. They were obtained by applying a PCA to each covariance matrix of implied volatility returns separately in the maturity groups of one, two, and three months. Each group contains the full set of the moneyness grid,  $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$ , which

Maturity group	1 months	3 months
	$S_i, \hat{\lambda}_i, \bar{x}_i \times 10^2$	
Sample covariance	$S_1 = \begin{pmatrix} 0.182 & 0.137 \\ 0.137 & 0.290 \end{pmatrix}$	$S_2 = \begin{pmatrix} 0.141 & 0.109 \\ 0.109 & 0.238 \end{pmatrix}$
Mean	$\bar{x}_1 = \begin{pmatrix} -0.052 & -0.063 \end{pmatrix}$	$\bar{x}_2 = \begin{pmatrix} -0.054 & -0.078 \end{pmatrix}$
Matrix of eigenvectors	$\hat{\Gamma}_1 = \begin{pmatrix} 0.563 & 0.827 \\ 0.827 & -0.563 \end{pmatrix}$	$\hat{\Gamma}_2 = \begin{pmatrix} 0.546 & 0.837 \\ 0.837 & -0.546 \end{pmatrix}$
Characteristic roots	$\hat{\lambda}_1 = \begin{pmatrix} 0.383 & 0.089 \end{pmatrix}^T$	$\hat{\lambda}_2 = \begin{pmatrix} 0.309 & 0.070 \end{pmatrix}^T$

Table 1: *PCA applied separately to the groups of 1 and 3 months time to maturity; each group contains time series of  $\{\Delta \ln \hat{\sigma}_t(0.90, \cdot)\}_{t=2}^{T=254}$  and  $\{\Delta \ln \hat{\sigma}_t(1.10, \cdot)\}_{t=2}^{T=254}$ , ODAX 1999.*

Maturity group	1 month	3 months
	$\hat{\Psi}_i, \hat{\lambda}_i \times 10^2$	
Estimated Covariance	$\hat{\Psi}_1 = \begin{pmatrix} 0.179 & 0.136 \\ 0.136 & 0.293 \end{pmatrix}$	$\hat{\Psi}_2 = \begin{pmatrix} 0.144 & 0.110 \\ 0.110 & 0.235 \end{pmatrix}$
Matrix of eigenvectors	$\hat{\Gamma}_{\text{CPC}} = \begin{pmatrix} 0.555 & 0.832 \\ 0.832 & -0.555 \end{pmatrix}$	
Characteristic roots	$\hat{\lambda}_2 = \begin{pmatrix} 0.383 & 0.089 \end{pmatrix}^T$	$\hat{\lambda}_2 = \begin{pmatrix} 0.309 & 0.070 \end{pmatrix}^T$

Table 2: *CPCA applied jointly to the groups of 1 and 3 month time to maturity; each group contains time series of  $\{\Delta \ln \hat{\sigma}_t(0.90, \cdot)\}_{t=2}^{T=254}$  and  $\{\Delta \ln \hat{\sigma}_t(1.10, \cdot)\}_{t=2}^{T=254}$ , ODAX 1999.*

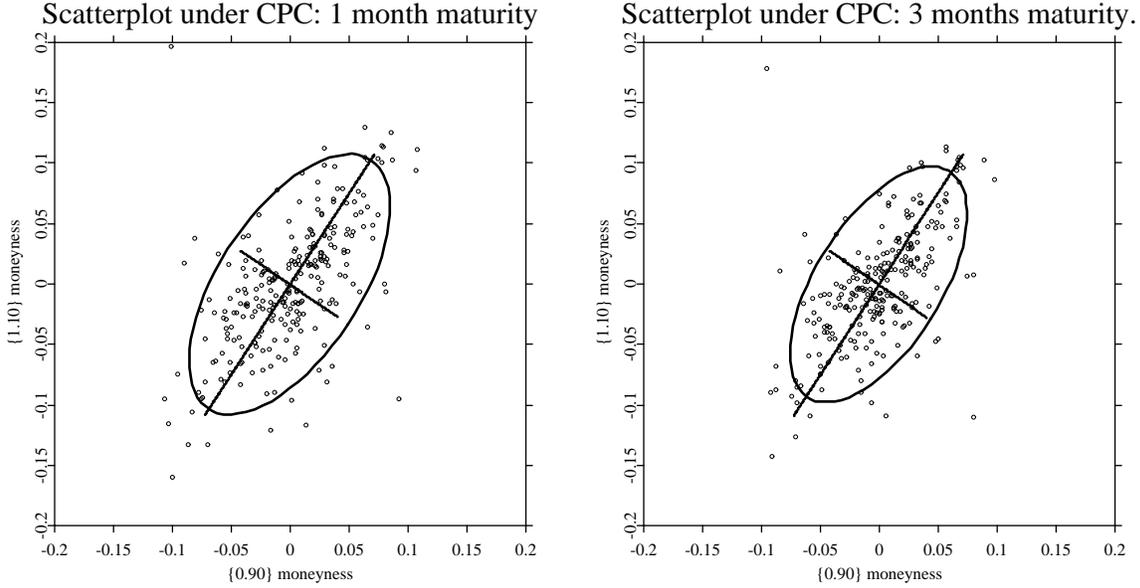


Figure 2: *Principle axes obtained under the CPC model for groups of 1 month and 3 months maturity; moneyness is  $\kappa = 0.90$  against  $\kappa = 1.10$ , ODAX 1999.*

was taken on a constant scale. (A log-scale would also be a valid choice.) For better comparison the scale on the axes is the same in all plots. Again, if the assumption of a common eigenstructure holds, we should expect the plots to deliver almost parallel lines. Clearly, this is the case.

The eigenvectors allow for a distinct interpretation: The first factor loadings are all of the same sign and are hence interpreted as a shift. The second eigenvectors exhibit a *Z*-shaped slope, centered at the money (corresponding to index 4), and the third characteristic vector displays a twist formation, again centered at the money. We will discuss these findings and their implications in greater detail in Section 4.

The main advantages of the CPC approach are: First CPC allows for jointly estimating common factors, i.e. it takes advantage of the *principle of parsimony* (Dempster, 1972) which says that whenever two competing models fit the data about equally well, the one involving the smaller number of parameters to be estimated should be preferred. Or, to put it differently: if the information from several samples can be combined to estimate the same parameters, then the estimates thus obtained are in general more stable, that is, they have smaller variability. Second, one avoids

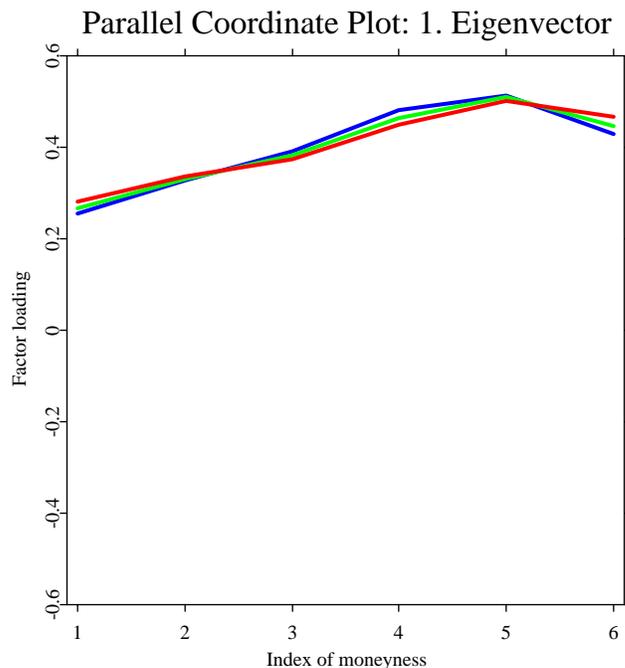


Figure 3: *Set of first eigenvectors obtained by separate PCA for 1 month (blue), 2 months (green), 3 months (red) maturity; index 1 to 6 is corresponding to moneyness  $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$ , ODAX 1999.*

the consequences of pooling samples from different populations. Recall that one may either pool samples directly as Skiadopoulos, Hodges, and Clewlow (1999) or the sample covariance matrices of each group (Thorpe, 1983). In the first approach directions of the principle components are determined by between-group as well as by within-group variability. The latter avoids this shortcoming of mixing between-group and within-group variability. However, Airolidi and Flury (1988, p. 31) out that *"... pooling the variance-covariance matrices is not appropriate unless all populations are assumed to have identical variability"*. Otherwise, the group with the highest variability will determine largely the directions of the extracted components. In the context of implied volatility this can be an issue, since variability in front contracts is higher than for the longer maturities. Both approaches, however, are likely to result in a loss of information, if populations are different.

Notably, a very intriguing feature of the CPC approach is the models allow for being ordered in a hierarchical fashion, as will be stated more precisely in Section 3.3. From this point of view,

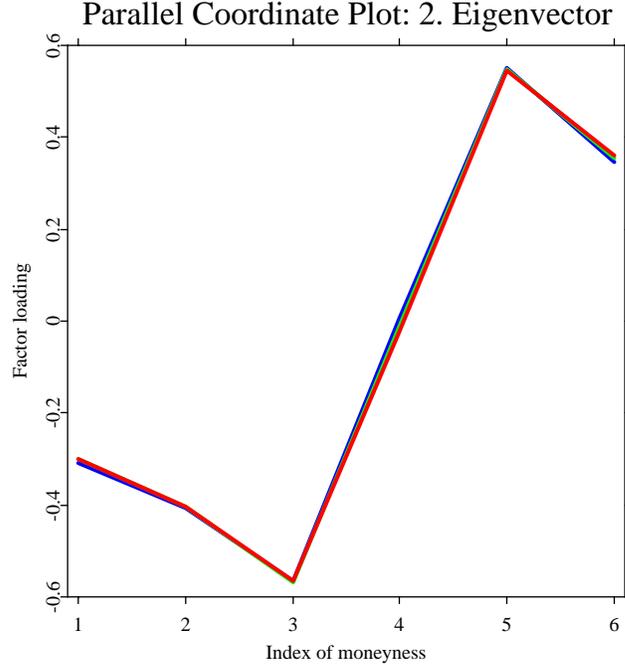


Figure 4: Set of second eigenvectors obtained by separate PCA for 1 month (blue), 2 months (green), 3 months (red) maturity; index 1 to 6 is corresponding to moneyness  $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$ , ODAX 1999.

a well organized analysis of covariance matrices can be accomplished. This helps understand the relationship between various subsets of the time series of the implied volatility surface.

### 3 Common Principle Components Analysis of Implied Volatilities

As above, denote  $\hat{\sigma}_t(\kappa, \tau)$  the implied volatility at time  $t$  recovered on a given grid of moneyness  $\kappa_j \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$  and maturity  $\tau_i \in \{1, 2, 3, 6, 9, 12\}$  [in months] by a Nadaraya-Watson estimator (see appendix for details). Next we form maturity groups indexed by  $\tau_i$  and stack implied volatilities to multiple time series of the smile  $X_{\tau_i}$ . The smile is observed at each  $\kappa_j$ . Hence in our empirical framework, we can compute  $k = 6$  sample covariance  $S_i$  each belonging to a different maturity group  $i$ . The hypothesis of a common principle component model for the

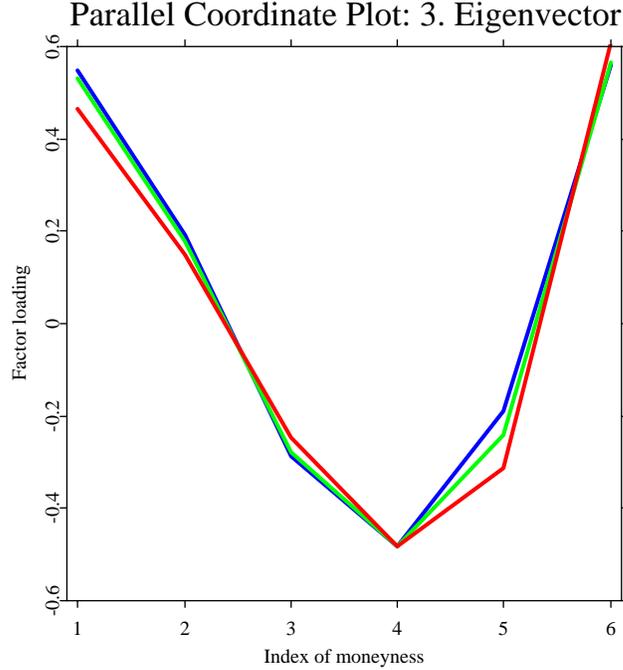


Figure 5: Set of third eigenvectors obtained by separate PCA for 1 month (blue), 2 months (green), 3 months (red) maturity; index 1 to 6 is corresponding to moneyness  $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$ , ODAX 1999.

population covariance matrices  $\Psi_i$  can formally be stated as

$$(1) \quad H_{\text{CPC}} : \Psi_i = \Gamma \Lambda_i \Gamma^T, \quad i = 1, \dots, k.$$

$\Psi_i$  are positive definite  $p \times p$  population covariance matrices,  $\Gamma$  is an orthogonal  $p \times p$  matrix and  $\Lambda_i = \text{diag}(\lambda_{i1}, \dots, \lambda_{ip})$  is the matrix of eigenvalues. Assume that all CPCs are well defined and all  $\lambda_i$  are distinct.

### 3.1 Estimating a common eigenstructure

Let  $S_i$  be the (unbiased) sample covariance matrix of implied volatilities, which are assumed to stem from an underlying  $p$ -variate normal distribution  $N_p(\mu, \Psi_i)$ . Sample size is  $n_i (> p)$ . Then the distribution of  $S_i$  is a generalization of the chi-squared variate, the Wishart distribution, with

$(n_i - 1)$  degrees of freedom (Muirhead, 1982, p.86), denoted by

$$n_i S_i \sim \mathcal{W}_p(\Psi_i, n_i - 1).$$

For the  $k$  Wishart matrices  $S_i$  the likelihood function is given by

$$(2) \quad L(\Psi_1, \dots, \Psi_k) = C \prod_{i=1}^k \exp \left\{ \text{tr} \left( -\frac{1}{2} (n_i - 1) \Psi_i^{-1} S_i \right) \right\} |\Psi_i|^{-\frac{1}{2}(n_i - 1)}$$

where  $C$  is a constant not depending on the parameters. Maximizing the likelihood is equivalent to minimizing the function

$$g(\Psi_1, \dots, \Psi_k) = \sum_{i=1}^k (n_i - 1) \left\{ \ln |\Psi_i| + \text{tr}(\Psi_i^{-1} S_i) \right\}.$$

Assuming that  $H_{\text{CPC}}$  in equation (1) holds, yields

$$g(\Gamma, \Lambda_1, \dots, \Lambda_k) = \sum_{i=1}^k (n_i - 1) \sum_{j=1}^p \left( \ln \lambda_{ij} + \frac{\gamma_j^T S_i \gamma_j}{\lambda_{ij}} \right).$$

Next, the orthogonality constraints of  $\Gamma$  have to be imposed. This is achieved by the Lagrange method, where we denote  $\mu_j$  the Lagrange multiplier of the  $p$  constraints  $\gamma_j^T \gamma_j = 1$ , and  $\mu_{hj}$  the Lagrange multiplier for the  $p(p-1)/2$  constraints  $\gamma_h^T \gamma_j = 0$  ( $h \neq j$ ). Hence the function to be minimized is given by

$$(3) \quad g^*(\Gamma, \Lambda_1, \dots, \Lambda_k) = g(\cdot) - \sum_{j=1}^p \mu_j (\gamma_j^T \gamma_j - 1) - 2 \sum_{h < j} \mu_{hj} \gamma_h^T \gamma_j.$$

Taking partial derivatives with respect to all  $\lambda_{im}$  and  $\gamma_m$ , it can be shown that the solution of the CPC model can be written as the generalized system of characteristic equations

$$(4) \quad \gamma_m^T \left( \sum_{i=1}^k (n_i - 1) \frac{\lambda_{im} - \lambda_{ij}}{\lambda_{im} \lambda_{ij}} S_i \right) \gamma_j = 0, \quad m, j = 1, \dots, p, \quad m \neq j,$$

which needs to be solved using

$$\lambda_{im} = \gamma_m^T S_i \gamma_m, \quad i = 1, \dots, k, \quad m = 1, \dots, p$$

and the constraints

$$\gamma_m^T \gamma_j = \begin{cases} 0 & m \neq j \\ 1 & m = j \end{cases}.$$

Flury (1988) proves existence and uniqueness of the maximum of the likelihood function, and Flury and Gautschi (1986) provide a numerical algorithm, which has been implemented in XploRe (<http://www.i-xplo.de>).

The maximum likelihood estimates of  $\Psi_i$  are given by  $\hat{\Psi}_i = \hat{\Gamma} \hat{\Lambda}_i \hat{\Gamma}^T$ ,  $i = 1, \dots, k$ . For the asymptotic distribution theory of  $\hat{\gamma}_{ij}$  and  $\hat{\lambda}_{ij}$ , we refer the reader to Flury (1988). Sample common principle components of the maturity groups are given by  $Y_{\tau_i} = \hat{\Gamma}^T X_{\tau_i}$  where  $X_{\tau_i} \in \mathbb{R}^6$  is the multiple time series of the implied volatility smile at maturity  $\tau_i$ .

### 3.2 Partial Common Principle Components

There is an obvious generalization of the CPC model: Instead of assuming the transformation matrix to be the same for all groups, one could hypothesize that only some  $q < p$  vectors are common, while others are specific. In terms of the analysis of implied volatilities this means that e.g. one source of the shocks, say the shift, is common to the whole surface, thereby pushing the surface up and down by a common degree, while slope or twist shocks differ across maturity groups. Since for shorter contracts the smile is more concave than for longer maturities, this hypothesis might apply to implied volatilities. Formally, for a partial CPC model of order  $q$ , pCPC( $q$ ), the hypothesis is given by

$$H_{\text{pCPC}(q)} : \Psi_i = \Gamma^{(i)} \Lambda_i \Gamma^{(i)T}, \quad i = 1, \dots, k \quad ,$$

where the  $\Psi_i$  are positive definite population covariance matrices, and  $\Lambda_i = \text{diag}(\lambda_{i1}, \dots, \lambda_{ip})$ , again the matrix of eigenvalues.  $\Gamma^{(i)} = (\Gamma_c, \Gamma_s^{(i)})$  are orthogonal  $p \times p$  matrices, where  $\Gamma_c$  is  $p \times q$  and denotes the matrix of eigenvectors common to all groups, and  $\Gamma_s^{(i)}$  the  $p \times (p - q)$  matrix of eigenvectors that are specific. By orthogonality the pCPC( $p - 1$ ) model implies the pCPC( $p$ ) model, i.e. ordinary CPC. Therefore,  $q$  is restricted to the range  $1 \leq q \leq p - 2$ , meaning that a minimum dimension of  $p = 3$  for any pCPC model is required. When comparing a higher pCPC( $q$ ) with a lower pCPC( $q-1$ ) model, we assume throughout that the analysis follows the order given by the size of the corresponding characteristic roots. Of course, one cannot be sure that any other eigenvectors is common, with the first few being specific. Nevertheless, it is a very natural way to start the analysis, especially when one is only interested in the first few principal components.

Establishing the likelihood function essentially follows the same lines as in the CPC model aside from respecting the additional orthogonality constraints of the specific factors  $\gamma_s^{(i)}$ . For the common

eigenvectors  $m, j \leq q$ , one obtains the same system of equations as in (4), however, a more intricate second equation links common and specific components, making a solution laborious (Flury, 1988).

### 3.3 Testing Relationships between Covariance Matrices

The maximum likelihood delivers a convenient framework for testing the specification assumed. For example, the log-likelihood ratio statistic for testing the  $H_{\text{CPC}}$  against the unrestricted model (unrelatedness between covariance matrices) is given by

$$T_{\text{CPC}} = -2 \ln \frac{L(\hat{\Psi}_1, \dots, \hat{\Psi}_k)}{L(S_1, \dots, S_k)} = \sum_{i=1}^k (n_i - 1) \ln \frac{|\hat{\Psi}_i|}{|S_i|},$$

where  $L(S_1, \dots, S_k)$  denotes the unrestricted maximum of the log-likelihood. Since the number of parameters estimated in the CPC model is  $p(p-1)/2$  (for the orthogonal matrix  $\Gamma$ ) plus  $kp$  (for the eigenvalues  $\Lambda_i$ ), and the number of parameters in the unrelated case is given by  $kp(p-1)/2 + kp$ , the statistic  $T_{\text{CPC}}$  is asymptotically  $\chi^2$  with  $(k-1)p(p-1)/2$  degrees of freedom as  $\min(n_i) \rightarrow \infty$  (Rao, 1973).

It is important to notice that the CPC and pCPC( $q$ ) models presented so far can be ordered in a hierarchical fashion, which allows a detailed analysis of the involved covariance matrices of different maturity groups. This hierarchy highlights the relationships between various subsets of the implied volatility surface, and whether there is at all a potential gain in inspecting the surface as a whole. The highest level of similarity would be to assume equality between covariance matrices of different maturity groups  $\Psi_i$ . In this case the number of parameters to be estimated is  $p(p+1)/2$ , and one may obtain the parameters by one single PCA applied to one pooled sample covariance matrix of all  $k$  groups. From a modeling perspective this would lead to a model using only *one* maturity or, equivalently, *one pooled* maturity bucket across moneyness. A second possible relationship is proportionality of all  $\Psi_i$ , i.e.  $\Psi_i = \rho_i \Psi_1$ ,  $i = 2, \dots, k$ . This model may be regarded as an offspring of the standard CPC model obtained by imposing additional constraints on the eigenvalues, i.e. by imposing  $\lambda_{ij} = \rho_i \lambda_{1j}$ ,  $\rho_i > 0$  for  $i = 2, \dots, k$  and  $j = 1, \dots, p$ . The number of parameters here is  $p(p+1)/2 + (k-1)$ . In this case again, one set of implied volatilities together with the proportionality constants  $\rho_i$  characterize the dynamics. It is first in the third level, the CPC level, i.e.  $\Psi_i = \Gamma \Lambda_i \Gamma^T$ ,  $i = 2, \dots, k$ , that the idea of a 'surface' is properly captured: for, though groups obey to a common transformation, they have their own sources of risk, between which no relation is assumed a priori. The following levels in the hierarchy are given by the pCPC( $q$ ) models starting

higher Model	lower Model	Degrees of freedom
Equality	Proportionality	$k - 1$
Proportionality	CPC	$(p - 1)(k - 1)$
CPC	pCPC( $q$ ) ( $1 \leq q \leq p - 2$ )	$\frac{1}{2}(k - 1)(p - q)(p - q - 1)$
pCPC(1)	Arbitrary covariance matrices	$(p - 1)(k - 1)$

Table 3: *Testing sequentially against the next lower model in the hierarchy and corresponding degrees of freedom.*

from  $q = p - 2$  down to  $q = 1$ . The relations between different groups disappear subsequently, until the last level: the  $\Psi_1, \dots, \Psi_k$  do not share any common eigenstructure.

One may appreciate the hierarchical structure of this set of assumptions by realizing that if two matrices share two principal components in common, then they necessarily share one component. Moreover, if two matrices are proportional, then they satisfy the CPC model, and all of the pCPC( $q$ ) models. From the nested structure, it is possible to decompose the total chi square into partial chi squares as

$$\begin{aligned}
T_{total} = & T(\text{inequality of proportionality constants} \mid \text{proportionality}) \\
& + T(\text{deviation from proportionality} \mid \text{CPC}) \\
& + T(\text{nonequality of last } p\text{-}q \text{ components} \mid \text{pCPC}(q)) \\
& + T(\text{nonequality of the first } q \text{ components}).
\end{aligned}$$

The decomposition of the log-likelihood and the corresponding degrees of freedom are displayed in Table 3 (Flury, 1988, p. 151).

In the step-up procedure, which is suggested by this decomposition, one tests sequentially one model against the next lower model in the hierarchy. One stops as soon as the higher model is not rejected against the lower one. However, one should be cautious, since these sequential tests are not independent from each other. This is why one should apply a model selection approach based on the Akaike or Schwarz Information Criterion (AIC, SIC, Section 3.4). Since from a hypothesis testing point of view, the most natural way to proceed is to test directly against the unrelated model, we performed these tests as well. By the summation property a test against any lower model is simply given by adding up the chi square test statistics and the degrees of freedom in between.

### 3.4 Alternative Model Selection: Akaike Information Criterion and Schwarz Information Criterion

Alternatively the Akaike information criterion (Akaike, 1973) may be used for model selection. The AIC is defined by

$$AIC = -2 (\text{maximum of log-likelihood}) + 2 (\text{number of parameters estimated}).$$

According to this criterion, the model with the lowest AIC should be chosen. Following Flury (1988), we use a modified Akaike information criterion. Let us assume that we have  $I$  hierarchically ordered models to compare, with  $r_1 < r_i < \dots < r_I$  ( $i = 1, \dots, I$ ) parameters in model  $i$ . We define the AIC to be

$$AIC(i) = -2(L_i - L_I) + 2(r_i - r_1)$$

where  $L_i$  is the maximum of the log-likelihood function of model  $i$ . Then selecting the model with the lowest AIC is equivalent to selecting the model with the lowest  $AIC(i)$ . Moreover we have

$$AIC(I) = 2(r_I - r_1) \quad \text{and} \quad AIC(1) = -2(L_1 - L_I)$$

such that  $AIC(I)$  is twice the difference of the number of parameters of the two extreme models, whereas  $AIC(1)$  is equal to the chi-square test statistic for comparing these two models.

Another model selection approach is the Schwarz Information Criterion (Schwarz, 1978, SIC). It is defined as

$$SIC = -(\text{maximum of log-likelihood}) \\ + (\text{number of parameters estimated}) \times \ln(\text{number of observations}).$$

The SIC, which is derived from Bayesian principles, gives a more severe complexity penalty as the AIC. Following the line of thoughts developed above, we define our modified SIC as

$$SIC(i) = -2(L_i - L_I) + 2(r_i - r_1) \ln(N),$$

where  $N = \sum_{i=1}^k n_i$  denotes the overall sum of observations across the  $k$  groups. According to this definition of the SIC, the model with the lowest SIC is the best fitting one.

	Mean	Standard deviation	Minimum	Maximum
Implied volatility	30.27%	8.41%	11.51%	79.97%
Moneyiness	0.94	0.17	0.37	1.50
Time to maturity (years)	0.48	0.44	0.03	2.02

Table 4: *Summary statistics of the ODAX 1999 implied volatility data. Source: EUREX, Thomson Datastream; own calculations.*

## 4 Empirical Results

### 4.1 Presentation of the Database

Our data set contains daily data on DAX options from the German-Swiss Futures Exchange, EUREX for the entire year 1999 (254 trading days). Options are European style. Interest rate data and spot prices were provided by Thomson Financial Datastream.

For data preparation, the following procedure is applied to the 120,401 observations of the initial database (Skiadopoulos, Hodges, and Clewlow, 1999): First, we replace the prices of all in-the-money options, whose prices might possibly contain a liquidity premium, with the corresponding prices implied by the put-call parity. Specifically, we replace the price of each in-the-money call option by its out-of-the-money put price. After this procedure, all the information contained in liquid put prices is extracted and preserved in the corresponding call prices via the put-call parity. Put prices may now be discarded without any loss of reliable information. Second, we omit options quoted less than one tick ( $\text{€}1/10$ ), those with an implied volatility greater than 80% (extreme far out of the money options) and those with a time to maturity of less than 10 days because of their sensitivity to small errors. This filtering method leaves us with a final sample of 57,702 observations, i.e. around 230 per day. Table 4 describes the main features of our data set. Since the options are of European style, we calculate the moneyiness metric as the strike price  $K$  divided by the (implied) futures price  $F$ , i.e.  $\frac{K_i}{F_i}$ . Finally, we are using the daily closing notation of the DAX from the German Stock Exchange and the term structure of the EURIBOR interest rate on a daily basis. After approximating the riskless rate of a given maturity by linear interpolation, we calculate implied volatilities by solving the Black Scholes formula for the volatility parameter  $\sigma$  using the market prices. We apply a nonparametric smoothing method (Härdle and Vieu, 1992; Härdle and Tsybakov, 1997; Ait-Sahalia and Lo, 1998, 2000) to obtain for each day the surface values of implied volatilities on a given grid on moneyiness and maturity (see appendix). For the

Maturity groups	Mean	Standard deviation	Skewness	Kurtosis	Minimum	Maximum
1	-0.00032	0.051	0.026	3.259	-0.212	0.273
2	-0.00040	0.047	0.024	3.232	-0.212	0.261
3	-0.00052	0.044	0.010	3.248	-0.207	0.242
6	-0.00070	0.038	-0.002	3.167	-0.176	0.197
9	-0.00082	0.034	-0.027	3.175	-0.160	0.168
12	-0.00100	0.029	-0.046	3.228	-0.138	0.128

Table 5: *Summary statistics of log-differences of the ODAX 1999 implied volatility data. Mean, standard deviation, skewness, and kurtosis averaged across moneyness except for minimum and maximum. Source: EUREX, Thomson Datastream; own calculations.*

moneyness  $\kappa$ , we chose a grid of  $\{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$  and for maturity  $\tau$  a grid of  $\{1, 2, 3, 6, 9, 12\}$  months. Due to the absence of data for a couple of observation dates, we did not extend the smile to  $\kappa = 1.15$ . With this choice of a grid, the procedure leaves us with 36 time series of the implied volatility surface. From the perspective of our CPC model, this time series may be regarded as 6 multi-variate time series of the smile or the changes of the smile for a *given* maturity. This data set is stored in the financial data base MD\*Base located at the Center for Applied Statistics and Economics at Humboldt University, Germany.

Summary statistics of the log differences used for estimation are given in Table 5. Note that there seems not to be a large deviation from normality: skewness is positive for the short and negative for the long maturities, however quite close to zero, and kurtosis is close to 3. We also performed Jarque-Bera-Tests on each single time series to test for normality in the data (Bera and Jarque, 1982). In only 6 out of 36 cases the hypothesis of normality was rejected at the 10% level of significance, which justifies the maximum likelihood estimation technique.

## 4.2 The Common Principal Components Analysis

Consider first the data already presented in the separate PCA in Section 2. Covariance matrices were computed from log-differences of the multi-variate time series of implied volatilities. Table 8 to 10 (in appendix) display the three samples and estimated covariance matrices. The common eigenvectors and eigenvalues are given in Table 11. In Table 6 we display the model selection procedure. The Step-up approach suggests a model, when the higher model is not rejected against

Model		Chi. Sqr	df	p-val	AIC	SIC
higher	lower					
Equality	Proport.	237.0	2	0.00	352.0	352.0
Proport	CPC	82.7	10	0.00	118.0	141.0*
CPC	pCPC(4)	7.1	2	0.03	55.7	191.0
pCPC(4)	pCPC(3)	0.2	4	1.00	52.6*	210.0
pCPC(3)	pCPC(2)	8.1	6	0.23	60.4	263.0
pCPC(2)	pCPC(1)	4.5	8	0.81	64.4	335.0
pCPC(1)	Unrelated	11.9	10	0.29	75.9	436.0
Unrelated					84.0	557.0

Table 6: *Decomposition of the chi square test statistic (step-up  $\mathcal{E}$  model building approaches) - 1, 2, and 3 months maturity*

the lower one, which is a pCPC(4) model here. The AIC is lowest for this model, but also very close to the AIC of the CPC model. The SIC, penalizing complexity more strongly as the AIC, even favors a proportional model, but a CPC model still appears to to very well. Testing a CPC model directly against a the unrelated model yields  $T_{\text{CPC}} = 31.8$ , which corresponds to a  $p$ -value  $p = 0.38$  of the  $\chi^2$  at 30 degrees of freedom. Interestingly, the equality and the proportional model are rejected against the unrestricted model. To sum-up, each selection criterion advises us to accept a model which imposes a *common* eigenstructure for implied volatility dynamics. The criteria only vary in the degree of additional constraints which one can assume, ranging from a proportional, over a CPC or a pCPC(4) model.

As is seen from Table 11 and in Figure 6, where we present the parallel coordinate plot for the three eigenvectors associated with the three biggest characteristic roots under CPC, the first components exhibit the shift,  $Z$ -shaped slope and twist structure we mentioned earlier. The first factor loadings are all of the same sign, giving biggest weight to the 1.05 moneyness (corresponding to index 5). The second eigenvectors exhibits a  $Z$ -shaped slope, as factor loadings have opposite sign at each side of the smile dynamics. This is in line with Skiadopoulos, Hodges, and Clewlow (1999); Härdle and Schmidt (2000); Sylla and Villa (2000), but interestingly ours is quite symmetric and well centered at the money (index 4), which bears a weight of almost zero. The third characteristic vector displays the twist formation, giving a large weight to at the money implied volatility returns and large weights of opposite sign to the outer parts of the smile dynamics.

CPC Coordinate Plot: First three Eigenvectors

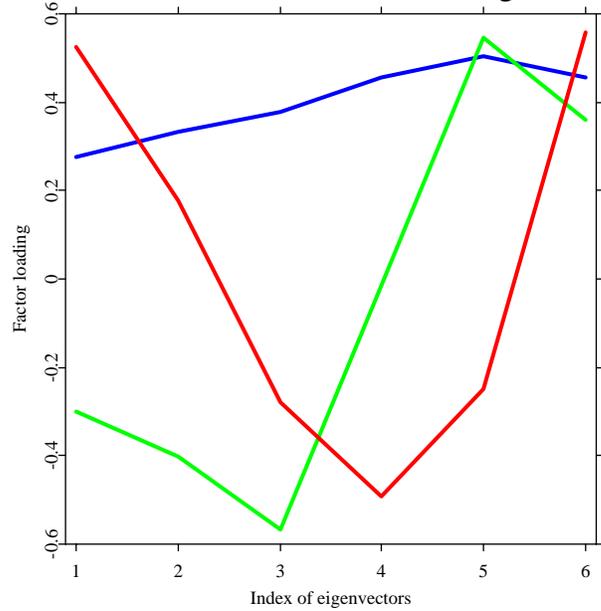


Figure 6: *First three eigenvectors as obtained by CPC for 1 month (blue), 2 months (green), 3 months (red) maturity; index 1 to 6 is corresponding to moneyness  $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$ , ODAX 1999.*

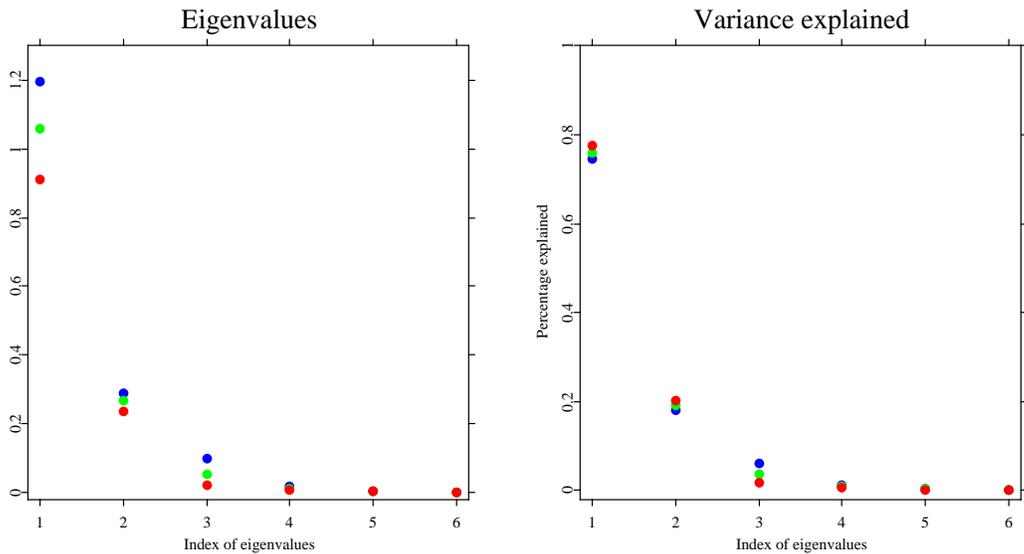


Figure 7: *Eigenvalues and the variance explained as obtained in the CPC model*

Eigenvalues differ in size dropping from the shortest maturity to the biggest. To evaluate the the percentage of variance explained up to the  $j$ th component in the  $k$ th group we computed  $\frac{\sum_{i=1}^j \lambda_{ki}}{\sum_{i=1}^p \lambda_{ki}}$  as a measure of variance explained. Altogether the first three eigenvalues,  $j = 3$ , explain between 98.5% (one month maturity), 98.9% (two months maturity), and 99.2% (three months maturity) of the dynamics of the implied volatility surface. The relative proportion explained increases slightly the longer the maturity, as can be seen in the right panel of Figure 7: Whereas the first PC explains 74.5% of all variability of the one-month maturity, already 77.5% are explained in the three-month maturity. At closer inspection, this holds true for the first two eigenvalues in all groups, meaning that while variability in PCs drops, the longer the maturity, the portion of variance explained by the first PCs increases. Beginning from the third eigenvalue this seems to be reversed. This is one of the reasons why the proportional model or even the equality model for the implied volatility surface cannot hold true, leaving us with the more general CPC model.

Naturally, one asks how the CPC model extends to longer maturities. However, as we learned, when adding the six months and higher maturities to the short ones, the CPC models were rejected highly significantly. This is why we analyzed the long maturities in a separate model with the covariance matrices of the six, nine and twelve months. Model selection is presented in Table 7.

The step-up approach suggests a pCPC(4) or even CPC model at a 5% level. However, when testing directly against the unrelated model they are both rejected. AIC is lowest for the pCPC(1) and testing the pCPC(1) against the unrelated model, yields a  $T_{pCPC(1)} = 19.7$  at 10 degrees of freedom and a  $p$ -value of 0.03. Hence one may accept a pCPC(1) model for the long maturities. Indeed, at closer inspection of the eigenvectors obtained by a separate PCA, it is apparent that only the "shift" vector is similar. See Table 12 for the transformation matrix, where  $\hat{\gamma}_1^c$  is the first common eigenvector, and  $\hat{\gamma}_2^{(i)}$  and  $\hat{\gamma}_3^{(i)}$  denote the specific ones. As can be seen, the latter differ. Possibly, the increased distance between maturities (three months as opposed to one month as before) might "loosen" the common structure inherent in shorter maturities to some extent. However, following the SIC, which again favors more simple models, a proportional or CPC model fits best.

Overall, we again find strong evidence for the common eigenstructure in the implied volatility surface dynamics. To check for robustness, we divided the sample in three non-overlapping sub-samples and estimated the CPC models in each of them. As factor loadings turned out to be of negligible difference and model selection stayed the same in each sample, we believe that the structure revealed remains stable across a considerable amount of time (here the entire year 1999).

Model		Chi. Sqr	df	p-val	AIC	SIC
higher	lower					
Equality	Proport.	251.0	2	0.00	486.0	486.0
Proport	CPC	81.0	10	0.00	239.0	262.0*
CPC	pCPC(4)	5.1	2	0.08	178.0	313.0
pCPC(4)	pCPC(3)	3.2	4	0.52	177.0	335.0
pCPC(3)	pCPC(2)	104.0	6	0.00	182.0	385.0
pCPC(2)	pCPC(1)	21.9	8	0.01	89.4	360.0
pCPC(1)	Unrelated	19.6	10	0.03	83.6*	444.0
Unrelated					84.0	557.0

Table 7: *Decomposition of the chi square test statistic (step-up & model building approaches) - 6, 9 and 12 months maturity*

This is an important result as the common structure can be a reliable starting point for modeling and simulating implied volatility dynamics.

## 5 Conclusions

In this paper, we present a model that is capable of modeling the implied volatility surface dynamics, by reducing its dimension to a small number of factors *common* to several maturity groups. This is accomplished by working with a multi-variate principal components technique that is designed for the multi-group case: *common* principle components. The common principal component analysis (CPCA) exploits a group structure given by the data, and allows for jointly estimating a common eigenstructure across groups. This allows us to model the implied volatility surface returns simultaneously for different maturity groups. In a CPCA framework several models with varying degrees of similarity in the eigenstructure can be compared and tested. The time series of implied volatility returns we use are obtained by a nonparametric kernel smoothing procedure from German DAX option data from 1999.

Our findings are summarized as follows: The CPC model or a pCPC(4) model, where only four eigenvectors are shared across groups, is well justified for modeling the short maturities across the smile dynamics of implied volatilities. In line with earlier literature, the shocks driving the implied

volatility surface can largely be attributed to a shift,  $Z$ -shaped slope and twist shock. For the longer maturities a pCPC(1) can be a valid choice. Hence, the preferred modeling strategy is to assume that implied volatility surface dynamics obey a common eigenstructure or, equivalently, are driven by a small number of common factors. Results prove to be stable across different sub-samples of the time period considered.

The results have direct implications for the market models of volatility that aim at pricing illiquid or exotic options with observed implied volatilities. By our evidence, the dimension of these models can be considerably reduced, because a common eigenstructure applies across different maturities. Moreover our evidence suggests that variances of the random sources are not to be linked in a proportional, but in a nonlinear, though monotone fashion. The remaining randomness could be modeled with a white noise process.

The power to reduce the dimension of implied volatility surface dynamics is also a suitable starting point for VaR calculations in Monte Carlo based methods. Developing a Monte Carlo simulation for delta-gamma-neutral, vega sensitive portfolios along the lines of Jamshidian and Zhu (1997) is straight forward: Delta-gamma-neutral, vega sensitive portfolio changes can be approximated by a first order Taylor expansion of Black and Scholes option prices. For two (or three at the most) sources of risk simulate Brownian motions in each group, and apply the estimated transformation matrix to recover volatility changes, from which possible portfolio changes can be obtained by the first order approximation. Thus the evidence presented provides valuable insights for risk modeling.

## A Appendix

### A.1 Implied Volatility Surface Smoothing and Bandwidth Choice

We shortly present the kernel regression procedures we employed in order to obtain our implied volatility surface time series on a given grid of maturities  $\{\tau_i\}$  and moneyness  $\{\kappa_j\}$ .

For a partition of explanatory variables  $(x_1, x_2) = (\kappa, \tau)$ , i.e. of moneyness and maturity, a two-dimensional Nadaraya-Watson kernel estimator of  $\hat{\sigma}$  is given by (Härdle, 1990; Härdle, Müller, Sperlich, and Werwatz, 2001)

$$\hat{\sigma}(x_1, x_2) = \frac{\sum_{i=1}^n K_1\left(\frac{x_1 - x_{1i}}{h_1}\right) K_2\left(\frac{x_2 - x_{2i}}{h_2}\right) \hat{\sigma}_i}{\sum_{i=1}^n K_1\left(\frac{x_1 - x_{1i}}{h_1}\right) K_2\left(\frac{x_2 - x_{2i}}{h_2}\right)},$$

where  $\hat{\sigma}_i$  is the Black Scholes volatility implied by observed options prices.  $K_1(u)$  and  $K_2(u)$  are univariate kernel functions, and  $h_1$  and  $h_2$  are the bandwidths chosen. We use quartic kernels of order 2, i.e.

$$K_i(u) = \frac{15}{16} (1 - u^2)^2 I(|u| \leq 1).$$

As from an empirical point of view, the choice of the kernel function has little influence on the results (Härdle, 1990), this choice was lead by reasoning that the quartic kernel generally behaves well in practical applications.

For bandwidth selection we employed the following procedure: As a starting point of the optimal bandwidths we computed a cross validation for each of the 254 observation dates by using a penalizing functions approach, i.e. we minimized the prediction error

$$p(h_1, h_2) = N^{-1} \sum_{j=1}^N \{\sigma_j - \hat{\sigma}_{h_1, h_2}(X_j)\}^2 \times \Xi(N^{-1} h_1^{-1} h_2^{-1} K_1(0) K_2(0)),$$

where  $\Xi(u) = \exp(2u)$  is the Akaike penalizing function (Akaike, 1970). Since possible choices of penalizing functions have the same first order expansion, they asymptotically behave in the same way. Next we averaged the bandwidths across observation dates, which yielded  $h_1^* = 0.042$  for moneyness and  $h_2^* = 0.092$  for maturity, at a standard deviation of  $s(h_1) = 0.012$  and  $s(h_2) = 0.010$  respectively. As the standard deviation of estimated bandwidths proved to be quite low, we used one single bandwidth for all estimation dates. However, bandwidths were still insufficient in the maturity dimension due to the special structure of the data: the data points behave like "pearls in a necklace" in the three-dimensional space of the volatility surface. However, penalizing approaches, as other cross validation procedures, evaluate the quality of the estimations right at actually observed data points. As estimates at the grid points deviate from the actual observations, the bandwidth in maturity dimension obtained by cross validation can be too small. Hence in maturity dimension, we chose to oversmooth. In finally using  $h_2^* = 0.35$  for the surface up to the three months horizon, and  $h_2^* = 1.1$  for the long maturities, we also accounted for the fact that observations are closer in the front contracts than in the long term contracts (1 month distance compared to 3 months).

## A.2 Tables

---



---


$$S_1 = \begin{pmatrix} 0.1402 & 0.1460 & 0.1505 & 0.1217 & 0.1002 & 0.1221 \\ 0.1460 & 0.1822 & 0.2145 & 0.1784 & 0.1347 & 0.1366 \\ 0.1500 & 0.2145 & 0.2863 & 0.2360 & 0.1532 & 0.1328 \\ 0.1218 & 0.1784 & 0.2360 & 0.3000 & 0.3033 & 0.2208 \\ 0.1002 & 0.1347 & 0.1532 & 0.3033 & 0.4086 & 0.3023 \\ 0.1221 & 0.1367 & 0.1328 & 0.2208 & 0.3023 & 0.2898 \end{pmatrix}$$

$$\hat{\Psi}_1 = \begin{pmatrix} 0.1505 & 0.1539 & 0.1554 & 0.1276 & 0.1096 & 0.1408 \\ 0.1539 & 0.1851 & 0.2121 & 0.175 & 0.1351 & 0.1485 \\ 0.1554 & 0.2121 & 0.2756 & 0.2218 & 0.1434 & 0.1363 \\ 0.1276 & 0.175 & 0.2218 & 0.2754 & 0.2839 & 0.2205 \\ 0.1096 & 0.1351 & 0.1434 & 0.2839 & 0.3996 & 0.3127 \\ 0.1408 & 0.1485 & 0.1363 & 0.2205 & 0.3127 & 0.3211 \end{pmatrix}$$


---



---

Table 8: *Sample covariance  $S \times 10^2$  and estimated population covariance  $\hat{\Psi} \times 10^2$ : 1 month maturity*

---



---


$$S_2 = \begin{pmatrix} 0.1183 & 0.1305 & 0.1428 & 0.1194 & 0.0959 & 0.1093 \\ 0.1305 & 0.1621 & 0.1919 & 0.1588 & 0.1183 & 0.1230 \\ 0.1428 & 0.1919 & 0.2465 & 0.1954 & 0.1247 & 0.1212 \\ 0.1195 & 0.1588 & 0.1954 & 0.2411 & 0.2531 & 0.2038 \\ 0.0959 & 0.1183 & 0.1247 & 0.2531 & 0.3581 & 0.2813 \\ 0.1093 & 0.1230 & 0.1212 & 0.2038 & 0.2813 & 0.2642 \end{pmatrix}$$

$$\hat{\Psi}_2 = \begin{pmatrix} 0.1228 & 0.1340 & 0.145 & 0.1213 & 0.0987 & 0.1149 \\ 0.134 & 0.1638 & 0.1916 & 0.1578 & 0.1183 & 0.1267 \\ 0.145 & 0.1916 & 0.2434 & 0.1916 & 0.1217 & 0.1227 \\ 0.1213 & 0.1578 & 0.1916 & 0.2346 & 0.2468 & 0.2041 \\ 0.0987 & 0.1183 & 0.1217 & 0.2468 & 0.353 & 0.2841 \\ 0.1149 & 0.1267 & 0.1227 & 0.2041 & 0.2841 & 0.2729 \end{pmatrix}$$


---



---

Table 9: *Sample Covariance  $S \times 10^2$  and estimated Covariance  $\hat{\Psi} \times 10^2$ : 2 months maturity*

---



---

$S_3 =$	$\begin{pmatrix} 0.1010 & 0.1162 & 0.1321 & 0.1123 & 0.0886 & 0.0974 \\ 0.1162 & 0.1415 & 0.1667 & 0.1367 & 0.1007 & 0.1092 \\ 0.1321 & 0.1667 & 0.2050 & 0.1572 & 0.0985 & 0.1086 \\ 0.1123 & 0.1369 & 0.1572 & 0.1879 & 0.2033 & 0.1825 \\ 0.0886 & 0.1007 & 0.09849 & 0.2037 & 0.3018 & 0.2540 \\ 0.0974 & 0.1092 & 0.1086 & 0.1825 & 0.2540 & 0.2376 \end{pmatrix}$
$\hat{\Psi}_3 =$	$\begin{pmatrix} 0.0981 & 0.1139 & 0.1306 & 0.1107 & 0.0863 & 0.0920 \\ 0.1139 & 0.1411 & 0.1679 & 0.1382 & 0.101 & 0.1058 \\ 0.1306 & 0.1679 & 0.2094 & 0.1617 & 0.1012 & 0.1072 \\ 0.1107 & 0.1382 & 0.1617 & 0.1955 & 0.2095 & 0.1825 \\ 0.0863 & 0.1010 & 0.1012 & 0.2095 & 0.3038 & 0.2503 \\ 0.0920 & 0.1058 & 0.1072 & 0.1825 & 0.2503 & 0.2269 \end{pmatrix}$

---



---

Table 10: *Sample Covariance  $S \times 10^2$  and estimated Covariance  $\hat{\Psi} \times 10^2$ : 3 month maturity*

---



---

$\hat{\Gamma}_{CPC}$	=	$\begin{pmatrix} 0.275 & -0.299 & 0.525 & -0.601 & 0.318 & 0.31 \\ (0.0085) & (0.0119) & (0.0153) & (0.0149) & (0.0174) & (0.0114) \\ 0.334 & -0.402 & 0.176 & -0.091 & -0.451 & -0.696 \\ (0.0099) & (0.0089) & (0.0091) & (0.0139) & (0.0201) & (0.013) \\ 0.379 & -0.567 & -0.278 & 0.374 & -0.207 & 0.524 \\ (0.014) & (0.0107) & (0.0134) & (0.0108) & (0.0177) & (0.0073) \\ 0.457 & -0.014 & -0.493 & -0.03 & 0.665 & -0.324 \\ (0.0038) & (0.0139) & (0.0072) & (0.0197) & (0.0104) & (0.0192) \\ 0.504 & 0.546 & -0.25 & -0.39 & -0.44 & 0.199 \\ (0.0135) & (0.0133) & (0.0139) & (0.0129) & (0.0112) & (0.0133) \\ 0.455 & 0.359 & 0.557 & 0.581 & 0.13 & -0.008 \\ (0.0099) & (0.015) & (0.0151) & (0.0137) & (0.0146) & (0.0077) \end{pmatrix}$
$\hat{\lambda}_1$	=	$\begin{pmatrix} 1.197 & 0.289 & 0.097 & 0.018 & 0.004 & 0.001 \\ (0.1064) & (0.0257) & (0.0086) & (0.0016) & (0.0004) & (0.0001) \end{pmatrix}^T$
$\hat{\lambda}_2$	=	$\begin{pmatrix} 1.057 & 0.266 & 0.052 & 0.012 & 0.003 & 0.001 \\ (0.0940) & (0.0236) & (0.0046) & (0.0011) & (0.0003) & (0.0001) \end{pmatrix}^T$
$\hat{\lambda}_3$	=	$\begin{pmatrix} 0.910 & 0.236 & 0.020 & 0.007 & 0.002 & 0.001 \\ (0.0809) & (0.0210) & (0.0018) & (0.0006) & (0.0001) & (0.0001) \end{pmatrix}^T$

---



---

Table 11: *Common transformation  $\hat{\Gamma}$  and eigenvalues  $\lambda_i \times 10^2$  from a CPC with groups of 1, 2, and 3 months maturity; standard errors in parenthesis.*

---



---

Maturity group $i$		1	2	3	1	2	3	
$\hat{\Gamma}_{CPC}^{(i)}$	=	0.3127	0.3128	0.3150	0.3295	0.5379	0.4904	-0.1298
		0.3425	0.3980	0.4000	0.4056	0.1411	0.1300	-0.0873
		0.3783	0.5429	0.5395	0.5225	-0.3112	-0.2655	-0.0463
		0.4399	0.0079	0.0114	0.02141	-0.4812	-0.4855	0.5133
		0.4869	-0.5526	-0.5451	-0.5270	-0.2738	-0.3111	0.4111
		0.4594	-0.3787	-0.3901	-0.4190	0.5359	0.5826	-0.7354
$\hat{\lambda}_1$	=	0.675	0.177	0.024	0.005	0.001	0.0005	)
		(0.0401)	(0.0105)	(0.0014)	(0.0003)	(0.00006)	(0.00003)	)
$\hat{\lambda}_2$	=	0.548	0.148	0.011	0.003	0.001	0.0004	)
		(0.0326)	(0.0088)	(0.0007)	(0.0002)	(0.00004)	(0.00002)	)
$\hat{\lambda}_3$	=	0.389	0.113	0.004	0.001	0.001	0.0002	)
		(0.0231)	(0.0067)	(0.0003)	(0.0001)	(0.00003)	(0.00001)	)

---



---

Table 12: *Estimated eigenstructure  $\hat{\Gamma}_{CPC}^{(i)} = (\gamma_1^c, \gamma_2^{(i)}, \gamma_3^{(i)})$  and eigenvalues of the pCPC(1) model for 6, 9, and 12 months maturity ( $i = 1, 2, 3$ ) – common and first three specific eigenvectors of each groups. Remaining specific vectors omitted for sake of clarity.*

## References

- AIROLIDI, J.-P., AND B. D. FLURY (1988): “An application of common principal component analysis to cranial morphometry of *Microtus californicus* and *M. ochrogaster* (Mammalia, Rodentia),” *J. Zool., Lond.*, 216, 21–36.
- AÏT-SAHALIA, Y., AND A. LO (1998): “Nonparametric estimation of state-price densities implicit in financial asset prices,” *Journal of Finance*, 53, 499–548.
- (2000): “Nonparametric risk management and implied risk-aversion,” *Journal of Econometrics*, 94, 9–51.
- AKAIKE, H. (1970): “Statistical predictor identification,” *Annals of the Institute for Statistical Mathematics*, 22, 203–217.
- (1973): “Information Theory and an Extension of the Maximum Likelihood Principle,” in *2nd International Symposium on Information Theory*. Akademiai Kiado, Budapest.
- ALEXANDER, C. (2001): “Principles of the Skew,” *Risk - Equity Risk Special Report*, January, s29–s32.
- AVELLANEDA, M., AND Y. ZHU (1997): “E-ARCH Model for the Term-Structure of Implied Volatility of FX Options,” Discussion paper, Courant Institute of Mathematical Sciences, submitted to Applied Mathematical Finance.
- BASILEVSKY, A. (1994): *Statistical Factor Analysis and Related Methods: Theory and Applications*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons.
- BERA, A., AND C. JARQUE (1982): “Model Specification Tests: a Simultaneous Approach,” *Journal of Econometrics*, 20, 59–82.
- BLISS, R. (1997): “Movements in the Term Structure of Interest Rates,” *Economic Review*.
- DEMPSTER, A. P. (1972): “Covariance Selection,” *Biometrics*, 28, 157–175.
- DERMAN, E., AND M. KAMAL (1997): “The patterns of change in implied index volatilities,” Quantitative strategies research notes, Goldman Sachs.
- DUMAS, B., J. FLEMING, AND R. E. WHALEY (1998): “Implied Volatility Functions: Empirical Tests,” *Journal of Finance*, LIII(6), 2059–2106.

- FLURY, B. (1988): *Common Principal Components and Related Multivariate Models*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Son, New York.
- FLURY, B., AND W. GAUTSCHI (1986): “An Algorithm for Simultaneous Orthogonal Transformations of Several Positive Definite Matrices to nearly Diagonal Form,” *Journal on Scientific and Statistical Computing*, 7, 169–184.
- HÄRDLE, W. (1990): *Applied Nonparametric Regression*. Cambridge University Press, Cambridge, UK.
- HÄRDLE, W., AND C. HAFNER (2000): “Discrete Time Option pricing with flexible volatility Estimation,” *Finance and Stochastics*, 4(2), 189–207.
- HÄRDLE, W., M. MÜLLER, S. SPERLICH, AND A. WERWATZ (2001): *Non - and Semiparametric Modelling*. Springer Verlag, <http://www.i-xplore.de/>.
- HÄRDLE, W., AND P. SCHMIDT (2000): “Common Factors Governing VDAX Movements and the Maximum Loss,” Discussion Paper 97/00, Sfb373, Humboldt-University, Berlin.
- HÄRDLE, W., AND A. TSYBAKOV (1997): “Local polynomial estimators of the volatility function in nonparametric autoregression.,” *Journal of Econometrics*, 81, 223–242.
- HÄRDLE, W., AND P. VIEU (1992): “Kernel Regression Smoothing of Time Series.,” *Journal Time Series Analysis*, 13, 209–232.
- HOTELLING, H. (1933): “Analysis of a complex of statistical variables into principal components,” *Journal of Educational Psychology*, 24, 417–441.
- HULL, J., AND A. WHITE (1987): “The Pricing of Options on Assets with Stochastic Volatilities,” *Journal of Finance*, 42, 281–300.
- JAMSHIDIAN, F. (1997): “LIBOR and swap market models and measures,” *Finance and Stochastics*, 1, 291–328.
- JAMSHIDIAN, F., AND Y. ZHU (1997): “Scenario Simulation: Theory and Methodology,” *Finance and Stochastics*, 1, 43–67.
- JOHNSON, H., AND D. SHANNO (1987): “Option Pricing when the Variance is Changing,” *Journal of Financial and Quantitative Analysis*, 22, 143–151.
- LEDOIT, O., AND P. SANTA-CLARA (1998): “Relative Option Pricing With Stochastic Volatility,” Working paper, UCLA, Los Angeles, USA.

- MARDIA, K. V., J. T. KENT, AND J. M. BIBBY (1992): *Multivariate Analysis*. Academic Press, 8th edn.
- MILTERSEN, K., K. SANDMANN, AND D. SONDERMANN (1997): “Closed form solutions for term structure derivatives with log-normal interest rates,” *The Journal of Finance*, 1(52), 409–430.
- MUIRHEAD, R. (1982): *Aspects of Multivariate Statistical Theory*. Wiley, Wiley, New York.
- PEARSON, K. (1901): “On lines and planes of closest fit to systems of points in space,” *Philosophical Magazine*, 2(6), 559–572.
- RAO, C. R. (1973): *Linear Statistical Inference and Its Applications*. Wiley, New York, 2nd edn.
- REBONATO, R. (1998): *Interest-Rate Option Models: Understanding, Analyzing and Using Models for Exotic Interest-Rate Options*, Wiley Series in Financial Engineering. John Wiley & Son Ltd., 2nd edn.
- ROLL, R. (1977): “A Critique of the Asset Pricing Theory’s Tests: Part I,” *Journal of Financial Economics*, 4, 129–176.
- SCHARZ, G. (1978): “Estimating the Dimension of a Model,” *Annals of Statistics*, 6, 461–464.
- SCHÖNBUCHER, P. J. (1999): “A Market Model for Stochastic Implied Volatility,” Working paper, Department of Statistics, Bonn University.
- SCOTT, L. (1987): “Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application,” *Journal of Financial and Quantitative Analysis*, 22, 419–37.
- SKIADOPOULOS, G., S. HODGES, AND L. CLEWLOW (1999): “The Dynamics of the S&P 500 Implied Volatility Surface,” *Review of Derivative Research*, 3, 263–282.
- SYLLA, A., AND C. VILLA (2000): “Measuring Implied Volatility Surface Risk using Principal Components Analysis,” in *Measuring Risk in Complex Stochastic Systems*, no. 147 in Lecture Notes in Statistics. Springer Verlag.
- THORPE, R. (1983): “A Review of the Numerical Methods for Recognising and Analysing Racial Differentiation,” in *Numerical Taxonomy*, ed. by J. Felsenstein, pp. 404–23. Springer, New York.
- WIGGINS, J. B. (1987): “Option Values under Stochastic Volatility,” *The Journal of Financial Economics*, 19, 351–72.