

Abschlußarbeit im Rahmen des Masterstudienganges Statistik

# The Phenomenology of Implied Volatility Surfaces

Humboldt Universität zu Berlin

eingereicht von	Matthias R. Fengler
Matrikel	164707
geb.	am 28.10.1973
Betreuer:	Professor Dr. Wolfgang Härdle

Berlin, den 16. Oktober 2002

Erklärung:

Ich erkläre, daß ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 16.Oktober 2002

---

Matthias R. Fengler

# Contents

- 1 Introduction** **1**
  
- 2 Implied Volatilities** **4**
  - 2.1 The Black and Scholes formula . . . . . 4
  - 2.2 A quick literature overview . . . . . 5
  
- 3 Descriptive statistics of implied volatility data** **7**
  - 3.1 Data description and preparation . . . . . 7
  - 3.2 Estimation of the implied volatility surface . . . . . 10
  - 3.3 General features of implied volatility surfaces: 1995 to 2001 . . . . . 11
  
- 4 Common principal components models** **16**
  - 4.1 General overview . . . . . 16
  - 4.2 Estimating common eigenstructures . . . . . 18
  - 4.3 Hypotheses testing . . . . . 19
  
- 5 Empirical results** **25**
  - 5.1 The entire sample period . . . . . 25
  - 5.2 Model selection . . . . . 32
  - 5.3 Stability analysis . . . . . 34
    - 5.3.1 Stability of eigenvectors . . . . . 34

5.3.2	Stability of eigenvalues . . . . .	36
6	Time series models	39
7	Conclusions	51
A	Additional implied volatility surface plots	55
B	Asymptotic theory	63

# 1 Introduction

Expected volatility is a cornerstone in modern financial theory: The rational investor balances the tradeoff between the risk he bears and the return he expects. [Markowitz \(1959\)](#), followed by [Sharpe \(1964\)](#) and [Lintner \(1965\)](#), were among the first to quantify the naïve idea of the simple equation ‘more risk means higher return’ in terms of equilibrium models. Since then the analysis of volatility and price fluctuations has sparked a huge number of studies in both theoretical and quantitative finance refining and extending these early models.

A decisive advance in the research of volatility was possible, when [Black and Scholes \(1973\)](#) published their seminal work on pricing European put and call options: Since the input parameters of the fundamental option pricing formula, i.e. current asset value, strike price, a riskless interest rate, and option maturity, are known except for the diffusion coefficient of the asset price dynamics, an obvious question became: Which volatility is *implied* in option prices observed on markets, given the Black and Scholes (BS) model were a true description of market conditions? As options are bets on the future development of the underlying asset, the key advantage of this *option implied volatility* is the fact that it is – unlike volatility measures based on historical data – a forward looking variable by nature and thus should reflect market expectations on volatility over the remaining live time of the option. Consequently, the information content of option implied volatility with respect to actually realized price fluctuations was of primary concern in the first studies approaching this new topic.

Nowadays, implied volatility is ubiquitous in any application of market participants: It serves as a convenient way of quoting options, ‘volatility trading’ is common practice on trading floors, market models incorporate the risk from changing implied volatility for hedge strategies, and risk management tools, which are approved by bank regulators to steer the allocation of economic capital, include models of implied volatility.

Implied volatility enjoys unravelled popularity despite the evidence that it actually con-

### IVS Ticks of 20000502

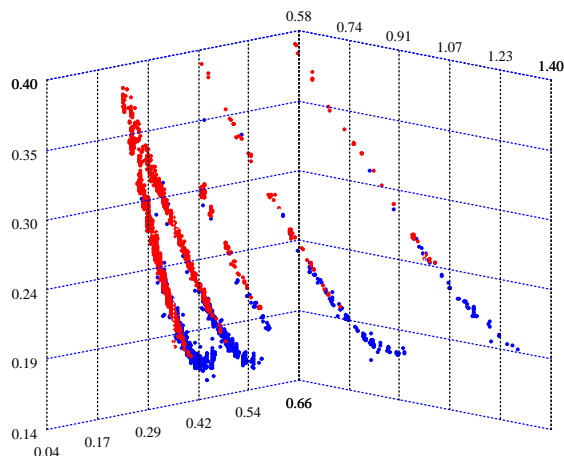


Figure 1.1: *Call (blue) and put (red) implied volatilities observed on 2nd May, 2000, ODAX.*

tradicts the model it is estimated from: According to [Black and Scholes \(1973\)](#), implied volatility is a constant: As a matter of fact however, implied volatilities display distinct, time dependent functional patterns across strikes  $K$  and term structure  $\tau$ . This dependence, captured by the function  $\hat{\sigma}_t : (K, \tau) \rightarrow \hat{\sigma}_t(K, \tau)$ , is called implied volatility surface (IVS). A typical picture is given in Figure (1.1) showing a pronounced curvature across strikes (here, in the scale of moneyness) on 2nd May, 2000. For a given maturity  $\tau$ , this function has been named ‘smile’ or, if it is skewed, ‘smirk’.

There is a considerable amount of literature which aims at reconciling this empirical finding with financial theory. Generally, this may be achieved by including another degree of freedom into the option pricing models. Well known examples are stochastic volatility models, models with jump diffusions, or models relying on general Lévy processes, such as the generalized hyperbolic. However, these models generate smile and term structure phenomena that only partially capture the richness of functional patterns of the IVS and the complexity of its dynamics ([Tompkins, 2001](#)), and parameter estimates seem to suffer from time inconsistency

(Dumas et al., 1998) making practical applications difficult.

A new interpretation of the inaccuracy of these models is nourished from the observation that due to their depth and liquidity option markets behave more and more self-governing: Apart from the underlying asset price dynamics, they are additionally driven by demand and supply conditions inherent to options markets, Bakshi et al. (2000); Cont and da Fonseca (2002). From this point of view, one interprets the IVS as a ‘state variable’, i.e. an additional financial indicator in option markets.

The aim of the present thesis is to describe the phenomenology of the IVS derived from German option data for the DAX index from January 1995 to May 2001 and investigate its statistical properties. The analysis is focussed on two salient features that characterize IVS dynamics: Firstly, the instantaneous profile of the IVS is subject to changes, while most shocks tend to move it into the same direction. Secondly, the size of the shocks decreases with the option’s maturity. This leads to high spatial correlation between contemporaneous surface values, while at the same time ‘volatility’ of implied volatility is highest for short maturity contracts. It is due to the insight by Fengler et al. (2001) that these two observations imply that implied volatilities of different maturity groups obey a *common* eigenstructure. The so called *common principle components* (CPC) models exactly feature this structure. Thus, IVS dynamics can be generated by a small number of uncorrelated factors from a low-dimensional space spanned by the eigenvectors of a common transformation matrix. Here, we extend the study by Fengler et al. (2001) to a larger data set, and investigate the statistical properties of these factors. Contrary to other studies in this field, particular concern will be put on the question of model stability in sub-samples and time series modelling.

This work is organized as follows: In a first step, we introduce the IVS formally and give an overview on related literature. Section (3) describes the data, explains data preparation and the estimation of the IVS. In Section (4) we present the CPC models and its estimation. Necessary tests are derived. Section (5) provides the static results of CPC models and the stability analysis, while Section (6) is devoted to time series analysis. Section (7) summarizes the results obtained and concludes.

## 2 Implied Volatilities

### 2.1 The Black and Scholes formula

European style calls and puts are contingent claims on an asset  $S_t$  (for simplicity, paying no dividends, here), which yield as pay-off  $\max(S_T - K, 0)$  and  $\max(K - S_T, 0)$ , respectively, at a given expiry day  $T$ .  $K$  is called strike price. In the traditional BS model it is assumed that the asset price process  $S_t$  follows a geometric Brownian motion with a constant diffusion coefficient  $\sigma$ . Under no arbitrage, the BS option pricing formula for calls is given by (Black and Scholes, 1973):

$$C_t^{BS}(S_t, K, \tau, r, \sigma) = S_t \Phi(d_1) - K e^{-r\tau} \Phi(d_2) \quad (2.1)$$
$$d_1 = \frac{\ln(S_t/K) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} \quad d_2 = d_1 - \sigma\sqrt{\tau} \quad ,$$

where  $\Phi(u)$  denotes the cumulative distribution function of the standard normal distribution,  $\tau = T - t$  time to maturity of the option,  $r$  a riskless interest rate over the option's life time. Put prices  $P_t$  are obtained via the put-call-parity  $C_t - P_t = S_t - e^{-r\tau} K$ .

Implied volatility  $\hat{\sigma}$  is defined as that  $\sigma$  which equates observed market prices  $\tilde{C}_t$  with the theoretical BS option price:

$$\hat{\sigma} : \quad C_t^{BS}(S_t, K, \tau, r, \hat{\sigma}) - \tilde{C}_t = 0, \quad (2.2)$$

i.e. which is *implied* by the data, given the model were true. Due to monotonicity of the BS formula in the volatility parameter  $\sigma$  there exists a unique solution  $\hat{\sigma} > 0$ .

From the BS assumptions, implied volatility  $\hat{\sigma}$  as a function of  $t, K, \tau$  is constant. Empirically however, implied volatility is a time-dependent function of  $K$  and  $\tau$ . Hence it can be interpreted as the random surface  $\hat{\sigma}_t : (K, \tau) \rightarrow \hat{\sigma}_t(K, \tau)$ . To center the surface around the most interesting and most liquid contracts, one rescales the strike dimension to moneyness  $\kappa = K/S_t$ , and studies  $\hat{\sigma}_t : (\kappa, \tau) \rightarrow \hat{\sigma}_t(\kappa, \tau)$ , Figure (1.1). Often, out-of-the-money (OTM) put options (i.e.  $\kappa < 1$ ) have higher implied volatility than corresponding in-the-money (ITM) options. This phenomenon is usually referred to as 'smile' or 'smirk'.

## 2.2 A quick literature overview

Since the IVS is a complex, high-dimensional random object, a first aim of the empirical literature is reducing the dimensionality of the surface, usually by means of principal component analysis (PCA). A general finding of this literature is that a small number of factors, usually two or three, is sufficient to explain up to 90% of the IVS dynamics. Studies of this type include [Skiadopoulos et al. \(1999\)](#) on the smile of given maturity buckets and [Fengler et al. \(2002c\)](#) for the term structure of VDAX sub-indices within an application to Maximum Loss. [Alexander \(2001\)](#) analyzes fixed strike deviations of implied volatility  $\hat{\sigma}_K - \hat{\sigma}_{ATM}$ . This choice allows for a natural way of modeling different volatility regimes as identified by [Derman \(1999\)](#). A more comprehensive surface perspective on IVS dynamics is adopted in [Fengler et al. \(2001\)](#) in the common CPC framework, which allows for a joint decomposition of maturity groups of the IVS. [Cont and da Fonseca \(2002\)](#) use a functional PCA to identify relevant eigenfunctions of the IVS.

A procedure specifically tailored to the discrete design of the IVS is recently proposed by [Fengler et al. \(2002a\)](#). As is obvious from [Figure \(1.1\)](#), implied volatilities only occur in a discrete string structure owed to the fact that on institutionalized exchanges, option contracts are only traded for a limited number of maturities. Taking this particularity into account, the authors devise a procedure motivated from generalized linear models and backfitting that only exploits information in the neighborhood of the observed design points to estimate the relevant basis functions. They thus avoid global, possibly biased fits.

[Hafner and Wallmeier \(2001\)](#) provide an ample time series analysis of ‘smile measures’ of implied volatility. However, it seems to be difficult to find significant and stable explanatory variables for the smile dynamics. Using a more parsimonious model approach, [Rosenberg \(2000\)](#) develops a time dependent volatility model and show its superiority to standard models. Closest to this study is [Avellaneda and Zhu \(1997\)](#) who employ an EARCH model for PCA-extracted factors of FX-implied volatilities of various markets.

The recent interest in implied volatility is also due to the development of market models of volatility, ([Ledoit and Santa-Clara, 1998](#); [Schönbucher, 1999](#), among others). Starting point of these models is the presence of a sufficient number of European options with varying

strike prices and maturities. Necessarily, certain conditions on the processes have to be identified in order to avoid arbitrage. In a similar spirit, [Derman and Kani \(1998\)](#) construct a stochastic implied tree of local volatilities. All these models employ the current deformation of the IVS as input parameter to price other illiquid and exotic options on the underlying asset. However, the empirical performance of these kinds of models seems still to be under investigation. Especially, the typical lack of model stability as pointed out in the study by [Dumas et al. \(1998\)](#) is an open question.

## 3 Descriptive statistics of implied volatility data

### 3.1 Data description and preparation

The data set under consideration contains tick statistics on the DAX future and DAX index options and was provided by the German-Swiss Futures Exchange EUREX for the period from 1995 to May 2001. Both future and option data are contract based data, i.e. each single contract is registered together with its price, contract size, and time of settlement up to a hundredth second. Interest rate data in daily frequency, i.e. 1, 3, 6, 12 months FIBOR rates for the years 1995–1999 and EURIBOR rates for the period 2000–2001, were obtained from *Thomson Financial Datastream*. Interest rate data was linearly interpolated to approximate a ‘riskless’ interest rate for the option specific time to maturity. Option raw data has undergone a complex preparation scheme which is due to [Hafner and Wallmeier \(2001\)](#) and described in the following.

In first step, we recover the DAX index values. To this end, we group to each option price observation  $H_t$  the future price  $F_t$  of the nearest available future contract, which was traded within a one minute interval around the observed option. The future price observation was taken from the most heavily traded futures contract on the particular day, usually the 3 months contract. The no-arbitrage price of the underlying index in a frictionless market without dividends is given by  $S_t = F_t e^{-r_{T_F,t}(T_F-t)}$ , where  $S_t$  and  $F_t$  denote the index and the future price respectively,  $T_F$  the future’s maturity date, and  $r_{T,t}$  the interest rate with maturity  $T - t$ . Implied volatilities will be measured in terms of moneyness  $\kappa = \frac{K}{F_t}$ , where  $K$  is the option’s strike price.

The DAX index is a capital weighted performance index ([Deutsche Börse, 2002](#)), i.e. dividends less corporate tax are reinvested into the index. Therefore, at a first glance, dividend payments should have no or almost little impact on index options. However, when only the interest rate discounted future is used to recover implied volatilities from an inversion of the Black and Scholes formula, implied volatilities of calls and puts can differ significantly. This may produce difficulties for the estimation of the IVS. This discrepancy is especially large during spring when most of the 30 companies listed in the DAX distribute dividends. The point is best visible in [Figure \(3.1\)](#) and [Figure \(3.2\)](#), both on April 4, 2000: As can be

## Implied Volatility Surface Ticks

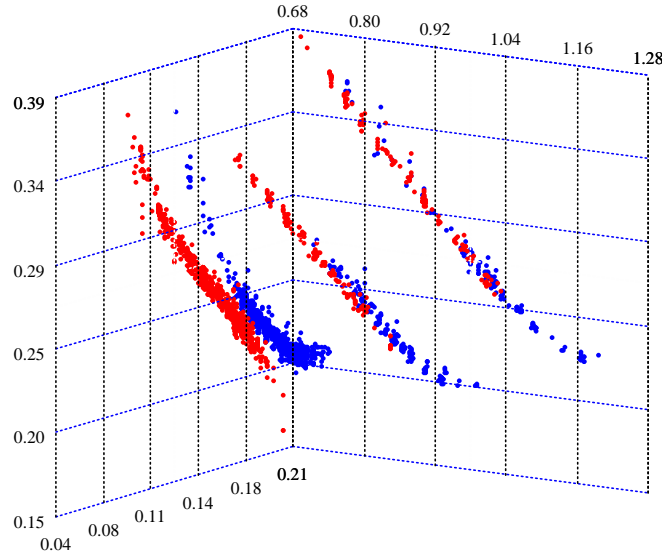


Figure 3.1: *Implied volatility surface ticks on April 4, 2000, derived by interest rate discounted futures price only; put implied volatility red, call implied volatility blue.*

seen in Figure (3.1) implied volatilities of calls (blue points) and puts (red) fall apart, thus violating the put-call-parity and general market efficiency considerations. After correcting the discount factor of the futures price with an implied dividend  $\Delta \hat{D}_t = (10.3, 5.0, 1.9)^\top$ , where the first entry refers to 16 days, the second to 45 days and the third to 73 days to maturity, implied volatilities of puts and calls converge on the same levels, Figure (3.2). The effect of the concave put volatility smile is remedied, too.

Hafner and Wallmeier (2001) argue that the marginal investor's individual tax scheme is different from the one actually assumed to compute the DAX index. If this is the case, the net dividend for this investor can be higher or lower than the one used for the index computation. This discrepancy, which is called 'difference dividend', has the same impact as

## Implied Volatility Surface Ticks

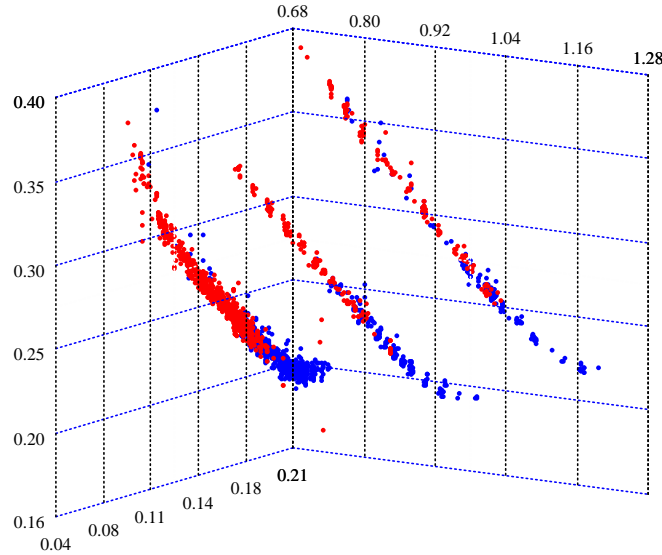


Figure 3.2: *Implied volatility surface ticks on April 4, 2000, derived by interest rate and implied dividend yield discounted futures price; put implied volatility red, call implied volatility blue.*

a dividend payment for an unprotected option, i.e. it drives a wedge into the option prices and hence into implied volatilities. Denote by  $\Delta D_{t,T}$  the time  $T$  value of this difference dividend incurred between  $t$  and  $T$ . Then, to overcome these consequences, the authors observe that both the dividend adjusted futures pricing formula,

$$F_t = S_t e^{r_F(T_F-t)} - \Delta D_{t,T_F}, \quad (3.1)$$

and the dividend adjusted put-call parity,

$$C_t - P_t = S_t - \Delta D_{t,T_H} e^{-r_H(T_H-t)} - K e^{-r_H(T_H-t)} \quad (3.2)$$

(with  $T_H$  denoting the call's  $C_t$  and the put's  $P_t$  maturity date), can be exploited to derive

the difference dividend. Inserting equation (3.1) into (3.2) yields

$$C_t - P_t = F_t e^{-r_F(T_F-t)} + \Delta D_{t,T_H,T_F} - K e^{-r_H(T_H-t)}, \quad (3.3)$$

where  $\Delta D_{t,T_H,T_F} \equiv \Delta D_{t,T_F} e^{-r_F(T_F-t)} - \Delta D_{t,T_H} e^{-r_H(T_H-t)}$  is the desired difference dividend.

An ‘adjusted’ the index level

$$\tilde{S}_t = F_t e^{-r_F(T_F-t)} + \Delta D_{t,T_H,T_F} \quad (3.4)$$

is the index level, which ties put and call implied volatilities exactly to the same levels when used in the inversion of the Black-Scholes formula.

For an estimate of  $\Delta \hat{D}_{t,T_H,T_F}$ , pairs of puts and calls of the same maturity are identified provided they were traded within a five minutes interval. For each pair the  $\Delta D_{t,T_H,T_F}$  is derived from equation (3.4). To ensure robustness  $\Delta \hat{D}_{t,T_H,T_F}$  is estimated by the median of all  $\Delta D_{t,T_H,T_F}$  of the pairs of the given maturity at day  $t$ . Implied volatilities are recovered by inverting the Black and Scholes formula using the corrected index value  $\tilde{S}_t = F_t e^{-r_F(T_F-t)} + \Delta \hat{D}_{t,T_H,T_F}$ . Note that  $\Delta D_{t,T_H,T_F} = 0$ , when  $T_H = T_F$ . Indeed, when calculated also in this case,  $\Delta \hat{D}_{t,T_H,T_F}$  proved to be close to zero, which supports the validity of this approach. The described procedure is applied on a daily basis throughout the entire data set from 1995–2001.

## 3.2 Estimation of the implied volatility surface

To avoid any parametric mis-specification of the IVS, estimation is performed by means of nonparametric smoothing, more specifically, by local linear estimation, [Härdle \(1990\)](#); [Härdle et al. \(2002\)](#). Suppose the model is given by

$$Y_i = m(X_i) + \zeta \varepsilon_i, \quad i = 1, \dots, n, \quad (3.5)$$

where  $X_i = (x_{i1}, x_{i2})$  is the 2-dimensional vector of independent variables, i.e. time to maturity  $\tau = T - t$  and moneyness  $\kappa$ .  $Y_i$  denotes implied volatility  $\hat{\sigma}$  here.  $m$  is a regression function which can locally approximated by polynomial of degree one, and  $\zeta$  a constant scaling the conditional variance of  $Y_i$  given  $X_i$ .  $\varepsilon_i$  are iid distributed with  $E[\varepsilon_i] = 0$ , and  $E[\varepsilon_i^2] = 1$ .

By expanding equation (3.5) in a Taylor series in the neighborhood of  $x$

$$m(x) \approx m(x_0) + \sum_{i=1}^2 \left( \frac{\partial m}{\partial x_i} \Big|_{x_0} (x_i - x_{0i}) \right)$$

and realizing the neighborhood of  $x$  by kernel weights, an estimator of  $m(x)$  can be formulated in terms of the quadratic minimization problem

$$\min_{\beta} \sum_{i=1}^N \left\{ Y_i - \beta_0 - \beta_{11}(X_{i1} - x_{i1}) - \beta_{12}(X_{i2} - x_{i2}) \right\}^2 K_h(X_i - x_0). \quad (3.6)$$

Here  $\beta = (\beta_0, \beta_{11}, \beta_{12})^\top$  denotes the vector of coefficients,  $N$  is the number of observations, and  $K_h$  is a two-dimensional kernel function with bandwidth  $h$  controlling the neighborhood in each direction.

The minimization problem (3.6) can be interpreted as a weighted least squares estimator and be solved accordingly. For details, the asymptotic behavior of the estimates etc. see [Härdle et al. \(2002\)](#). We estimate the IVS daily on the moneyness grid  $\kappa \in \{0.9250, 0.9500, 0.9750, 1.0000, 1.025, 1.05\}$  and a maturity grid  $\tau = \{0.0625, 0.125, 0.1875, 0.25\}$  months. As kernel function we choose a product kernel of the univariate quartic kernels:  $K(u) = \frac{15}{16}(1 - u^2)^2 I(|u| \leq 1)$ . Bandwidth selection is based on the experience of [Fengler et al. \(2001\)](#), who use an Akaike penalizing technique to determine bandwidths. For robustness, implied volatilities with maturity of less than 10 days were excluded from the estimation.

The entire data set of the option data, the IVS's, and the time series of  $\Delta \hat{D}_t$  is stored in the financial data base [MD\\*base](#) located at the Center for Applied Statistics and Economics (CASE), Humboldt Universität zu Berlin, Germany.

### 3.3 General features of implied volatility surfaces: 1995 to 2001

In this section we present a number of plots of the data set under investigation and give a short description of their behavior. For a first general overview inspect Figures (3.3) and (3.4) which present the 3 months ATM implied volatility and the DAX index from 1995-05/2001 and, for details, 3 months ATM implied volatility only.

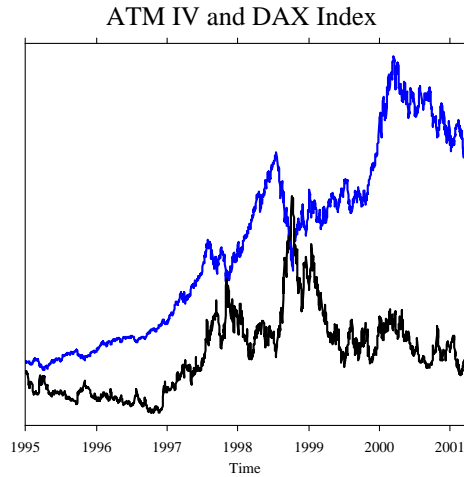


Figure 3.3: *German DAX  $\times 10^{-4}$  (blue) and 3 months ATM implied volatility levels (black).*

The DAX index started 1995 from around 2100 points and increased only moderately till the beginning of 1997. During this time ATM implied volatility was below 20% and gradually fell till 1997 towards 14%. In the terminology introduced by [Derman \(1999\)](#) and [Alexander \(2001\)](#) the market regime of this period is usually referred to as a ‘range-bounded’ market.

Beginning from the end of 1996 the DAX commenced a steady and smooth increase till mid 1998, which was shortly interrupted by the Asian crisis in the second half 1997. The entire ascent of the DAX was accompanied by steadily increasing implied volatility levels rising as high as 35%, when DAX fell sharply at the peak of the market turmoil. From then, implied volatilities gradually declined, but remained – also relatively volatile – at historically high levels, while the index rose again. The situation had changed to ‘stable trending market’, [Alexander \(2001\)](#). Between summer 1998 the DAX fell again about 2000 points, followed by a sharp increase of implied volatilities with peaks up to 50%.

During the recovery of the index between 1999 and 2000, implied volatilities again fell to the levels recorded before the late 1998 increase. Although increasingly volatile, they remained at these levels, when the DAX began its gradual decline from the post war peak of 8000 points in March 2000.

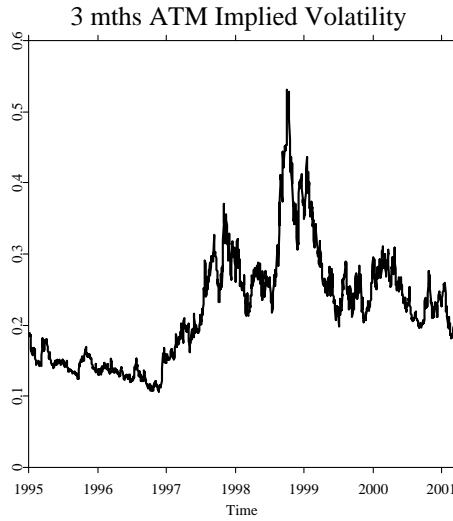


Figure 3.4: *3 months ATM implied volatility levels of DAX Index options.*

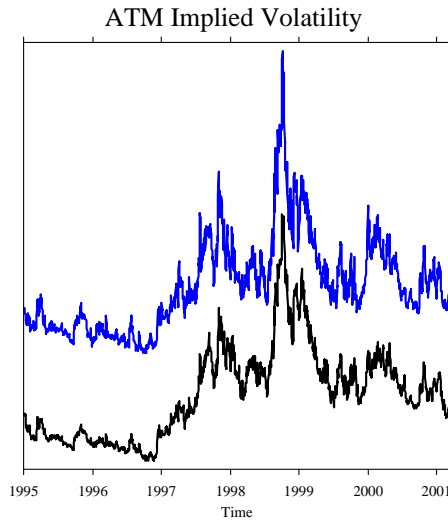


Figure 3.5: *22 days (blue) and 3 months implied volatilities (black), shifted.*

The general term structure features during this time are visible in Figure (3.5), which plots 3 month implied volatility and 22 days implied volatility (shifted scaled): It is seen that generally all effects, increases and drops, appear more pronounced for short implied volatilities. Figure (3.6) displays average term structures. In 1995 and 1996 they moderately increase, 1998 and 1999 they are slightly steeper. 2000 and 2001 have a ‘humped’ appearance with lower mid term volatilities and higher short and long implied volatility levels.

Annual standard deviation of implied volatility can be inferred from Figure (3.7). As already observed short run volatilities are more subject to daily variation than long term volatilities, which is reflected in the downward sloping functions. 1998 was the period of the highest volatility, followed by 1997 and 1999.

For clarity, a number of additional three-dimensional plots are collected in the appendix Section (A). They mainly reiterate Figure (3.6) and (3.7), but additionally allow conclusions about the smile structure of implied volatilities. A general feature, e.g. visible in Figure (A.3) and (A.4) is a skew in the smile function: for moneyness  $\kappa < 1$  implied volatilities tend to be higher than for  $\kappa > 1$ .  $\kappa < 1$  is the OTM region for puts, and obviously in crisis times, increased demand for OTM puts pushes up put implied volatility. As is visible from these plots, the IVS is subject to considerable change over time, the dynamics of which will be studied in Section (6).

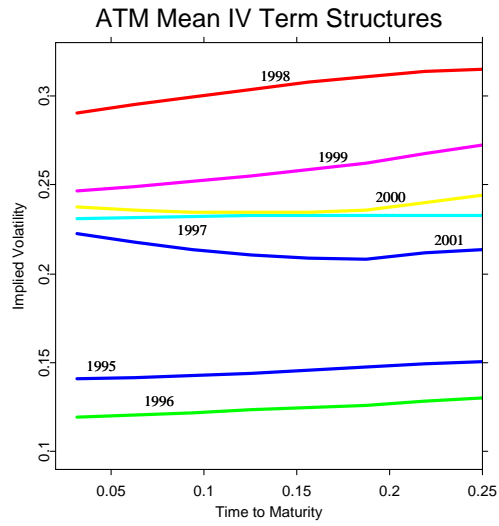


Figure 3.6: Mean ATM implied volatility surface from 1995 to (May) 2001, ODAX.

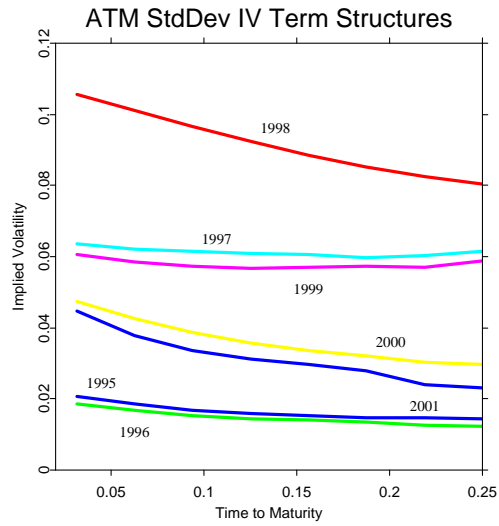


Figure 3.7: Standard deviation of ATM implied volatility surfaces from 1995 to (May) 2001, ODAX.

## 4 Common principal components models

### 4.1 General overview

Common principal components (CPC) models are models about certain properties of the eigenstructure of correlated variables appearing in groups. They aim at finding spaces spanned by common directions of maximal variance of these variables, while these variances may have different size in each group. By projecting into a lower dimensional subspace, one obtains factors which can be interpreted as determining the general structure of the original variables. For our analysis of implied volatilities this means that we can uncover sources of shocks driving the IVS dynamics.

Due to their flexibility and scope CPC models provide a number of advantages compared with ordinary principal component analysis (PCA): First, instead of estimating separately in the groups, CPCs are estimated jointly in all groups which results in efficiency gains. Secondly, estimates are not contaminated by differing variances in each group, as would be the case in a pooled PCA. Finally, and this will be the most important aspect in terms of pragmatic applications, CPC models constitute a parsimonious framework within which careful model selection and model comparison can be accomplished, and which allows for a substantial reduction of dimensionality of complex variables.

As noted in the introduction there are two key properties of the IVS which motivate CPC modeling: Shocks move the surface into the same direction, and variance of these shocks decreases with the option's maturity. This phenomenon implies that implied volatilities of different maturity groups obey a common eigenstructure. As above, denote  $\hat{\sigma}_i(\kappa, \tau)$  the implied volatility function at time  $t$  recovered on a given grid of moneyness  $\kappa$  in  $k$  maturity groups by means of a local linear estimator. Next group the observations in maturity groups indexed by  $i$  and stack implied volatilities to multiple time series of the smile  $\mathbf{X}_i$ . After taking log-differences, we obtain  $k$  sample covariance matrices  $\mathbf{S}_i$  of implied volatility returns, which belong to different maturity groups  $i$ .

The hypothesis of a CPC model for population covariance matrices  $\Psi_1, \dots, \Psi_k$  can formally

be stated as

$$H_{\text{CPC}} : \mathbf{\Psi}_i = \mathbf{\Gamma} \mathbf{\Lambda}_i \mathbf{\Gamma}^\top, \quad i = 1, \dots, k. \quad (4.1)$$

$\mathbf{\Psi}_i$  are positive definite  $p \times p$  population covariance matrices (of implied volatilities),  $\mathbf{\Gamma}$  is an orthogonal  $p \times p$  matrix and  $\mathbf{\Lambda}_i = \text{diag}(\lambda_{i1}, \dots, \lambda_{ip})$  is the matrix of eigenvalues. Assume that all CPCs are well defined and that there is at least one population  $i$  for each  $j \in \{1, \dots, p\}$  in which the eigenvalue  $\lambda_{ij}$  is distinct.

A particular strength of CPC models is that they enclose a whole family of models with varying degrees of flexibility in the eigenstructure. All these models are nested, which makes model selection convenient. The *proportional* model, e.g., puts additional constraints on the matrix of eigenvalues  $\mathbf{\Lambda}_i$  by imposing that  $\lambda_{ij} = \rho_i \lambda_{1j}$ , where  $\rho_i > 0$  are unknown constants. This is equivalent to writing

$$H_{\text{prop}} : \mathbf{\Psi}_i = \rho_i \mathbf{\Psi}_1, \quad i = 1, \dots, k. \quad (4.2)$$

In terms of implied volatility modeling this means that variances of the components increase with the same proportional rate as we move between different maturity groups. Leaving the assumptions on the eigenvalues as in hypothesis (4.1), one may also ease the restrictions on the transformation matrix  $\mathbf{\Gamma}$ : This leads to *partial* CPC models, pCPC( $q$ ), where  $q$  denotes the order of common eigenvectors in  $\mathbf{\Gamma}$ . Formally, the hypothesis of the pCPC( $q$ ) model is

$$H_{\text{CPC}} : \mathbf{\Psi}_i = \mathbf{\Gamma}^{(i)} \mathbf{\Lambda}_i \mathbf{\Gamma}^{(i)\top}, \quad i = 1, \dots, k, \quad (4.3)$$

where  $\mathbf{\Lambda}_i$  as in (4.1) and  $\mathbf{\Gamma}^{(i)} = \begin{pmatrix} \mathbf{\Gamma}_c & \mathbf{\Gamma}_s^{(i)} \end{pmatrix}$ . Here, the  $p \times q$  matrix  $\mathbf{\Gamma}_c$  contains the  $q$  common eigenvectors, while  $\mathbf{\Gamma}_s^{(i)}$  of dimension  $p \times (p-q)$  holds the  $p-q$  group specific eigenvectors. The  $\mathbf{\Gamma}^{(i)}$  are still orthogonal matrices. This implies that the necessary dimension  $p$  to estimate a pCPC(1) model is at least 3. When all possible pCPC( $q$ ) are to be estimated sequentially moving from  $p-q$  to 1, it is open to the modeling approach in which order the constraints on  $\gamma_i$  are to be eased. A natural way to proceed is to allow in each step for group specific eigenvectors in the ‘least important’ case, where ‘least important’ is measured in terms of size of the corresponding eigenvalues.

For implied volatility modeling the pCPC( $q$ ) hypothesis implies that in each maturity group a slightly different projection of implied volatilities is optimal, while the transformation of the most important components stays the same for the entire IVS. This is sensible since the prevailing number of shocks hitting the IVS move it primarily into the same direction.

## 4.2 Estimating common eigenstructures

In the foregoing section we presented a number of admissible hypotheses with respect to the IVS. Here, we want to turn to estimation of the common eigenstructures. However, due to its similarity with the models (4.2) and (4.3) and its practical importance, we only focus on the ordinary CPC model as given in (4.1). For all other models as well as proofs the reader is referred to Flury (1988). Most important results on the asymptotic distributions of the estimates necessary for the derivations of the tests in Section (4.3) will be contained in the Appendix (B).

Let  $\mathbf{S}_i$  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  $\mathbf{S}_i$  is a generalization of the chi-squared variate, the Wishart distribution, with  $(n_i - 1)$  degrees of freedom (Härdle and Simar, 2003), denoted by

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

For the  $k$  Wishart matrices  $\mathbf{S}_i$  the likelihood function is given by

$$L(\Psi_1, \dots, \Psi_k) = c \prod_{i=1}^k \exp \left\{ \text{tr} \left( -\frac{1}{2} (n_i - 1) \Psi_i^{-1} \mathbf{S}_i \right) \right\} |\Psi_i|^{-\frac{1}{2}(n_i - 1)} \quad (4.4)$$

where  $c$  is a constant not depending on the parameters.

**Proposition 4.1 (CPC).** *The solution to the common principle components hypothesis for the  $j = 1, 2, \dots, p$  variables in  $i = 1, 2, \dots, k$  groups can be written as the generalized system of characteristic equations*

$$\gamma_m^\top \left( \sum_{i=1}^k (n_i - 1) \frac{\lambda_{im} - \lambda_{ij}}{\lambda_{im} \lambda_{ij}} \mathbf{S}_i \right) \gamma_j = 0, \quad m, j = 1, \dots, p, \quad m \neq j, \quad (4.5)$$

to be solved using

$$\lambda_{im} = \gamma_m^\top \mathbf{S}_i \gamma_m, \quad i = 1, \dots, k, \quad m = 1, \dots, p \quad (4.6)$$

and the constraints

$$\gamma_m^\top \gamma_j = \begin{cases} 0 & m \neq j \\ 1 & m = j \end{cases}. \quad (4.7)$$

**Proof 4.1.** *Flury (1988), Fengler et al. (2001).*

The system of equations (4.5) is basic to the CPC analysis. An algorithm for solving (4.5) under the constraints (4.6) and (4.7) is provided in [Flury and Gautschi \(1986\)](#). They also prove existence and uniqueness of the maximum of the likelihood function. The numerical algorithm is implemented in [XploRe](#).

Maximum likelihood estimates of  $\hat{\Psi}_i$  are given by  $\hat{\Psi}_i = \hat{\Gamma} \hat{\Lambda}_i \hat{\Gamma}^\top$ ,  $i = 1, \dots, k$ . Sample common principle components of the maturity groups are given by  $\mathbf{Y}_i = \hat{\Gamma}^\top \mathbf{X}_i$  where  $\mathbf{X}_i \in \mathbb{R}^p$  is the multiple time series of the implied volatility smile returns at maturity  $\tau_i$ .

### 4.3 Hypotheses testing

With these results we are in the shape of conducting hypothesis tests about the eigenvalues  $\hat{\lambda}_{ij}$  and the eigenvectors  $\hat{\gamma}_j$  in the multi-sample framework. Suppose we estimate a CPC model in  $R$  sub-samples and wish to test the hypothesis of equality of the  $j$ th eigenvalue  $\hat{\lambda}_{ij}^{(r)}$  in the  $i$ th group across  $R$  sub-samples:

$$H_0 : \hat{\lambda}_{ij}^{(1)} = \dots = \hat{\lambda}_{ij}^{(r)} = \dots = \hat{\lambda}_{ij}^{(R)}$$

against the alternative  $H_1 : \exists \left( \hat{\lambda}_{ij}^{(r_1)}, \hat{\lambda}_{ij}^{(r_2)} \right)$  such that  $\hat{\lambda}_{ij}^{(r_1)} \neq \hat{\lambda}_{ij}^{(r_2)} \forall r_1, r_2$ .  $H_0$  can be written as

$$H_0 : \begin{array}{l} \hat{\lambda}_{ij}^{(1)} - \hat{\lambda}_{ij}^{(2)} = 0 \\ \vdots \\ \hat{\lambda}_{ij}^{(1)} - \hat{\lambda}_{ij}^{(r)} = 0, \\ \vdots \\ \hat{\lambda}_{ij}^{(1)} - \hat{\lambda}_{ij}^{(R)} = 0 \end{array} \quad (4.8)$$

which allows us to formulate the test statistic. It is useful to define a contrast matrix  $\mathbf{C}$ :

**Definition 4.1 (Contrast matrix).**  $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_l, \dots, \mathbf{c}_L)$  is called a contrast matrix if  $\sum_{l=1}^L \mathbf{c}_l = \mathbf{0}$ , and its rows are linearly independent, [Johnson and Wichern \(1998\)](#).

We define by  $\mathbf{C}_1$  the special  $(R - 1) \times R$  contrast matrix

$$\mathbf{C}_1 = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{pmatrix}. \quad (4.9)$$

A test on the equality of the  $j$ th eigenvalue in the  $i$ th group in  $R$  sub-samples, drawing on Proposition (B.1) in the appendix, is given in the following

**Proposition 4.2 (Equality of the  $j$ th eigenvalue in the  $i$ th group in  $R$  sub-samples).**

Denote by  $\tilde{\boldsymbol{\lambda}}_{ij}$  the  $R \times 1$  stacked vector of  $\hat{\lambda}_{ij}^{(r)}$ , in  $r = 1, \dots, R$  sub-samples and its variance by the  $R \times R$  matrix

$$V_{\tilde{\boldsymbol{\lambda}}} = 2 \left\{ \text{diag} \left( \frac{\lambda_{ij}^{(1)2}}{n_{i1} - 1}, \dots, \frac{\lambda_{ij}^{(r)2}}{n_{ir} - 1}, \dots, \frac{\lambda_{ij}^{(R)2}}{n_{iR} - 1} \right) \right\},$$

where  $n_{ir}$  is the sample size of group  $i$  and sub-sample  $r$ . A test for (4.8) is given by

$$T_{\text{equ}} = (\mathbf{C}_1 \tilde{\boldsymbol{\lambda}}_{ij})^\top (\mathbf{C}_1 \mathbf{V}_{\tilde{\boldsymbol{\lambda}}} \mathbf{C}_1^\top)^{-1} (\mathbf{C}_1 \tilde{\boldsymbol{\lambda}}_{ij}), \quad (4.10)$$

and  $\chi^2$  distributed with  $(R - 1)$  degrees of freedom.

**Proof 4.2.**  $T_{\text{equ}}$  is a sum of  $(R - 1)$  squares of independently, asymptotically standard normally distributed random variables.

In practice  $\mathbf{V}_{\tilde{\boldsymbol{\lambda}}}$  is unknown, but it can be estimated consistently given the consistent estimates of  $\hat{\lambda}_{ij}$ . The asymptotic distribution of  $T$  will stay asymptotically the same. A number similar tests, such as testing the  $j$ th eigenvalue across the  $r = 1, \dots, R$  samples jointly for all groups can be constructed from this accordingly.

There are a lot of ways of formulating the aforementioned hypothesis by a different choice of the contrast matrix  $\mathbf{C}$ . However, it turns out that  $T$  does not depend on this particular choice. For example, an equivalent formulation of the hypothesis using the contrast matrix

$$\mathbf{C}_2 = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

would be

$$H_0 : \begin{array}{l} \hat{\lambda}_{ij}^{(2)} - \hat{\lambda}_{ij}^{(1)} = 0 \\ \hat{\lambda}_{ij}^{(3)} - \hat{\lambda}_{ij}^{(2)} = 0 \\ \vdots \\ \hat{\lambda}_{ij}^{(R)} - \hat{\lambda}_{ij}^{(R-1)} = 0 \end{array} .$$

The equivalence of the tests is due to the fact that any pair of contrast matrices is related by a nonsingular matrix  $\mathbf{A}$  such that  $\mathbf{C}_1 = \mathbf{A}\mathbf{C}_2$ . Inserting  $\mathbf{A}\mathbf{C}_2$  into  $T$  yields

$$\begin{aligned} T &= (\mathbf{C}_1 \tilde{\boldsymbol{\lambda}})^\top (\mathbf{C}_1 \mathbf{V} \mathbf{C}_1^\top)^{-1} (\mathbf{C}_1 \tilde{\boldsymbol{\lambda}}) \\ &= (\mathbf{A}\mathbf{C}_2 \tilde{\boldsymbol{\lambda}})^\top (\mathbf{A}\mathbf{C}_2 \mathbf{V} \mathbf{C}_2^\top \mathbf{A}^\top)^{-1} (\mathbf{A}\mathbf{C}_2 \tilde{\boldsymbol{\lambda}}) \\ &= (\mathbf{C}_2 \tilde{\boldsymbol{\lambda}})^\top (\mathbf{C}_2 \mathbf{V} \mathbf{C}_2^\top)^{-1} (\mathbf{C}_2 \tilde{\boldsymbol{\lambda}}), \end{aligned}$$

which is the same as before.

Developing tests for the eigenvectors one faces the difficulty that the covariance matrix  $\mathbf{V}_{\hat{\Gamma}}$  given in equation (B.1) is singular. Therefore we adopt a strategy proposed by Anderson (1963) and Flury (1988). As tests will be carried out on submatrices of  $\mathbf{V}_{\hat{\Gamma}}$  note first of all that the  $qp \times qp$  submatrix  $\mathbf{V}_{\hat{\Gamma}}(q)$  for  $q < p$  given by

$$\mathbf{V}_{\hat{\Gamma}}(q) = \begin{pmatrix} \sum_{\substack{j=1 \\ j \neq 1}}^p \theta_{1j} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^\top & -\theta_{12} \boldsymbol{\gamma}_2 \boldsymbol{\gamma}_1^\top & \cdots & -\theta_{1q} \boldsymbol{\gamma}_q \boldsymbol{\gamma}_1^\top \\ -\theta_{21} \boldsymbol{\gamma}_1 \boldsymbol{\gamma}_2^\top & \sum_{\substack{j=1 \\ j \neq 2}}^p \theta_{1j} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^\top & \cdots & -\theta_{2q} \boldsymbol{\gamma}_q \boldsymbol{\gamma}_1^\top \\ \vdots & \vdots & \ddots & \vdots \\ -\theta_{q1} \boldsymbol{\gamma}_1 \boldsymbol{\gamma}_q^\top & -\theta_{q2} \boldsymbol{\gamma}_2 \boldsymbol{\gamma}_q^\top & \cdots & \sum_{\substack{j=1 \\ j \neq p}}^p \theta_{qj} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^\top \end{pmatrix} \quad (4.11)$$

has a particular spectral decomposition which is stated in Proposition (4.3). For convenience, call the  $p$  elements in rows  $(j-1)p+1$  to  $jp$  of a characteristic vector of  $\mathbf{V}_{\hat{\Gamma}}(q)$  the ‘ $j$ th position’.

**Proposition 4.3.** *The matrix  $\mathbf{V}_{\hat{\Gamma}}(q)$ ,  $q \leq p$ , has the following spectral decomposition:*

1.  $q(q-1)/2$  eigenvectors (one for each pair of  $j, h$  such that  $1 \leq j < h \leq q$ ) having  $\gamma_h/\sqrt{2}$  in position  $j$  and  $\gamma_j/\sqrt{2}$  in position  $h$  with all other positions being zero. Associated roots are  $2\theta_{jh}$ .
2.  $q(p-q)$  eigenvectors (one for each pair of  $j, h$  such that  $1 \leq j \leq q < h \leq p$ ) having  $\gamma_h$  in position  $j$  and zero elsewhere. Associated roots are  $\theta_{jh}$ .
3.  $q(q-1)/2$  eigenvectors (one for each pair of  $j, h$  such that  $1 \leq j < h \leq q$ ) which can be chosen to have  $\gamma_h/\sqrt{2}$  in position  $j$  and  $\gamma_j/\sqrt{2}$  in position  $h$  with all other positions being zero. Associated roots are zero.
4.  $q$  eigenvectors (one for each  $j$  from 1 to  $q$ ) which may be chosen to have  $\gamma_j$  in position  $j$  and zero elsewhere, again associated roots are zero.

**Proof 4.3.** *By direct verification.*

Thus  $\mathbf{V}_{\hat{\Gamma}}(q)$  has rank  $\frac{q(q-1)}{2} + q(p-q)$ . After these preparations we are able to form a test for  $\hat{\Gamma}$  in  $R$  sub-samples. Here we formulate the statistic for testing the first eigenvector only, i.e.

$$H_0 : \gamma_1^{(1)} = \dots = \gamma_1^{(r)} = \dots = \gamma_1^{(R)}$$

against the alternative  $H_1 : \exists \left( \gamma_1^{(r_1)}, \gamma_1^{(r_2)} \right)$  such that  $\gamma_1^{(r_1)} \neq \gamma_1^{(r_2)} \forall r_1, r_2$ . Testing for any other of the  $j$  eigenvectors can be accomplished by permutation of the estimated matrix of eigenvectors. Moreover, equivalence of more than one eigenvector, such as

$$H'_0 : \left( \gamma_1^{(1)}, \gamma_2^{(1)}, \gamma_3^{(1)} \right) = \left( \gamma_1^{(r)}, \gamma_2^{(r)}, \gamma_3^{(r)} \right) \quad \forall r = 2, \dots, R$$

can also be derived from the result of Proposition (4.3).

Again we rewrite  $H_0$  as

$$H_0 : \begin{array}{l} \gamma_1^{(1)} - \gamma_1^{(2)} = 0 \\ \gamma_1^{(1)} - \gamma_1^{(3)} = 0 \\ \vdots \\ \gamma_1^{(1)} - \gamma_1^{(r)} = 0 \\ \vdots \\ \gamma_1^{(1)} - \gamma_1^{(R)} = 0 \end{array} , \quad (4.12)$$

by using a  $p(R - 1) \times pR$  contrast matrix  $\mathbf{C}_2$  as defined in (4.9), with the ones replaced by the  $p \times p$  identity matrix  $\mathbf{I}_p$ , and the zeros replaced by  $p \times p$  matrices of zeros. A test for (4.12), whose derivation draws on Proposition (B.2), is stated in the following

**Proposition 4.4 (Test of equality of the first eigenvector in  $R$  sub-samples).** *Let  $\mathbf{C}_2$  be a  $p(R - 1) \times pR$  the contrast matrix. Denote by  $\tilde{\boldsymbol{\gamma}}_1$  the  $pR \times 1$  vector of stacked eigenvectors  $\hat{\boldsymbol{\gamma}}_1^{(r)}$ . Suppose the  $R$  sub-samples are independently drawn, and define  $\mathbf{V}_{\tilde{\boldsymbol{\gamma}}}$  as the  $pR \times pR$  block-diagonal matrix*

$$\mathbf{V}_{\tilde{\boldsymbol{\gamma}}} = \begin{pmatrix} \mathbf{V}_{\hat{\Gamma}}^{(1)}(1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{V}_{\hat{\Gamma}}^{(R)}(1) \end{pmatrix}, \quad (4.13)$$

where  $\mathbf{V}_{\hat{\Gamma}}^{(r)}(1)$  the first  $p \times p$  block out of  $\mathbf{V}_{\hat{\Gamma}}$  in equation (B.1) in sub-sample  $r$ . Collect all the  $p - 1$  nonzero characteristic roots of  $\mathbf{V}_{\hat{\Gamma}}^{(r)}(1)$  in the  $R(p - 1) \times R(p - 1)$  matrix

$$\boldsymbol{\Theta} = \text{diag} \left( \frac{\theta_{12}^{(1)}}{n_1 - k}, \dots, \frac{\theta_{1h}^{(r)}}{n_r - k}, \dots, \frac{\theta_{1p}^{(R)}}{n_R - k} \right),$$

where  $h = 2, \dots, p$ ,  $n_r$  the number of observations in sub-sample  $r$  and  $k$  the number of groups. Next collect all associated characteristic vectors in the  $Rp \times R(p - 1)$  matrix

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}^{(1)} & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & \cdots & \mathbf{B}^{(R)} \end{pmatrix}.$$

Then an  $\alpha$ -level test of equality

$$H_0 : \mathbf{C}_2 \tilde{\boldsymbol{\gamma}}_1 = \mathbf{0}$$

for the  $R$  first eigenvectors against  $H_1 : \mathbf{C}_2 \tilde{\boldsymbol{\gamma}}_1 \neq \mathbf{0}$  is given by

$$T_{\text{equ}} = (\mathbf{C}_2 \tilde{\boldsymbol{\gamma}}_1)^\top (\mathbf{C}_2 \mathbf{B} \boldsymbol{\Theta}^{-1} \mathbf{B}^\top \mathbf{C}_2^\top) (\mathbf{C}_2 \tilde{\boldsymbol{\gamma}}_1). \quad (4.14)$$

Reject  $H_0$ , if  $T_{\text{equ}} > \chi^2(1 - \alpha; (R - 1)(p - 1))$ , where  $\chi^2(\alpha; \nu)$  is the  $\alpha$ -quantile of the chi-squared distribution with  $\nu$  degrees of freedom.

**Proof 4.4.** Note first that  $\mathbf{V}_{\tilde{\boldsymbol{\gamma}}}$  has rank  $(p - 1)(R - 1)$ . A  $g$ -inverse of  $\mathbf{V}_{\tilde{\boldsymbol{\gamma}}}$  is given by

$$\mathbf{V}_{\tilde{\boldsymbol{\gamma}}}^- = \mathbf{B} \boldsymbol{\Theta}^{-1} \mathbf{B}^\top.$$

Thus,  $\mathbf{Z} = \mathbf{\Theta}^{-1/2}(\mathbf{C}_2\mathbf{B})^\top(\mathbf{C}_2\tilde{\gamma})$  has a limiting normal distribution with mean zero and variance  $\mathbf{I}_{(R-1)(p-1)}$ . Then  $T_{\text{equ}} = \mathbf{Z}^\top\mathbf{Z}$  is a sum  $(R-1)(p-1)$  of independently normally distributed random variables.

## 5 Empirical results

The empirical part of this study will be devoted to an analysis of the entire sample period. We will study the CPC model and the related models pCPC( $q$ ).

### 5.1 The entire sample period

The estimation results for the entire sample period exhibit the same stylized facts as documented in Fengler et al. (2001) for the year 1999 for daily settlement prices. In Figure (5.1) we display the results for the first three eigenvectors, Table (5.2) reports estimation results of  $\hat{\Gamma}$ , numbers given in parenthesis are asymptotic standard errors. The factor loadings of the first eigenvector, the blue line in Figure (5.1), are of the same sign throughout (eigenvectors are unique up to sign), and give approximately the same weight to each volatility shock across the smile. We hence interpret this factor as a *common shift* factor. In Figure (5.2) we present the projection  $y_1$  of the longest implied volatility maturity group (3 months maturity); the upper panel shows the principal component, the lower the integrated process. The shift interpretation of the first component is also visible from general structure of this process: It exhibits almost the same patterns as the implied volatility process itself, i.e. especially the rising levels in implied volatility beginning from 1997s. A measure to gauge the fraction of variance explained by the given factor is the fraction of the corresponding eigenvalue to the sum of all eigenvalues. For the the first component this amounts to 88% in the longest maturity group.

Component	Variance explained	Standard deviation	Skewness	Kurtosis	Mean reversion	Correlation with underlying
1	0.88	0.078	0.34	4.12	227.7	-0.48
2	0.06	0.020	0.30	6.54	36.5	0.08
3	0.03	0.015	0.22	7.30	2.2	-0.03

Table 5.1: *Descriptive statistics of principal components (daily); ODAX.*

Common Coordinate Plot: First three Eigenvectors

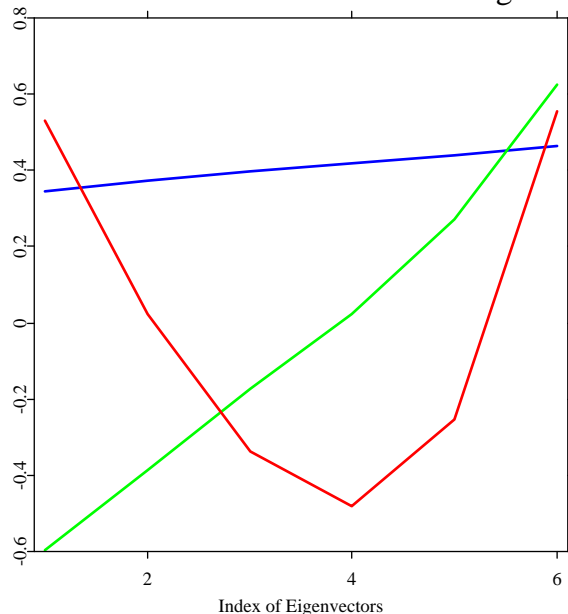


Figure 5.1: *CPC model for the entire sample period 1995 - May 2001, ODAX. Compare with Fengler et al. (2001).*

The second eigenvector, the green line, switches its sign at-the-money, and gives opposite weights to implied volatility shocks. Thus we may interpret the second type of shocks as *common slope* shocks. Figure (5.3) displays this component. The integrated second PC has a stable downward trend, which seems to revert around 1999. The third eigenvector may be interpreted as a *common twist* factor, hitting the curvature of the surface, as the sign of the factor loadings switch for the near-the-money region. Again the projections on this component and integrated process is shown in Figure (5.4). These components account for only 6% and 3% of the variance. The interpretations of the factor loadings in terms of shift, slope and twist shocks are similar to those reported in PCA studies on interest and forward rates, Rebonato (1998).

Table (5.1) summarizes the descriptive statistics of the PCs. The results are grossly similar to the findings of Cont and da Fonseca (2002) reported on the S&P 500 and the FTSE 100 except for the mean reversion. Whereas skewness is close zero for the three PCs, there is

evidence for excess kurtosis especially for the the second and third PC. The mean reversion of the integrated first PC is found to be 230 days, i.e. almost a year, while the second PC exhibits are more short-lived mean reversion of 36 days. The third PC has a mean reversion of 2 days. To our experience however, estimates of the mean reversion tend to be very sensitive to the sample size chosen, and change significantly in annual sub-samples. Thus, the estimates of the mean reversion coefficient should be taken with caution.

Correlation with the returns of underlying is around -0.5 for the first component. This is in line with the so called ‘leverage effect’: According to this argument (implied) volatility rises, when there is a negative shock in the market value of the firm, since this results to an increase in the debt-equity ratio. For the the second the third component correlation is negligible.

$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$	$\hat{\gamma}_6$
0.344	-0.598	0.530	0.472	-0.129	0.055
(0.0021)	(0.0095)	(0.0113)	(0.0070)	(0.0085)	(0.0037)
0.373	-0.385	0.022	-0.614	0.502	-0.288
(0.0014)	(0.0044)	(0.0096)	(0.0086)	(0.0105)	(0.0068)
0.397	-0.173	-0.339	-0.326	-0.457	0.618
(0.0010)	(0.0065)	(0.0055)	(0.0090)	(0.0090)	(0.0056)
0.419	0.024	-0.482	0.250	-0.337	-0.644
(0.0011)	(0.0085)	(0.0038)	(0.0083)	(0.0088)	(0.0043)
0.440	0.270	-0.252	0.432	0.610	0.334
(0.0012)	(0.0057)	(0.0074)	(0.0105)	(0.0081)	(0.0074)
0.463	0.625	0.554	-0.213	-0.191	-0.073
(0.0022)	(0.0095)	(0.0108)	(0.0076)	(0.0052)	(0.0032)

$\hat{\lambda}_{i1}$	$\hat{\lambda}_{i2}$	$\hat{\lambda}_{i3}$	$\hat{\lambda}_{i4}$	$\hat{\lambda}_{i5}$	$\hat{\lambda}_{i6}$
16.39	0.90	0.55	0.11	0.04	0.01
(0.578)	(0.032)	(0.019)	(0.004)	(0.001)	(0.0003)
10.14	0.41	0.16	0.07	0.03	0.01
(0.357)	(0.014)	(0.006)	(0.002)	(0.001)	(0.0004)
7.20	0.33	0.14	0.07	0.04	0.02
(0.254)	(0.012)	(0.005)	(0.002)	(0.001)	(0.001)
6.01	0.40	0.23	0.09	0.06	0.02
(0.211)	(0.014)	(0.008)	(0.003)	(0.002)	(0.001)

Table 5.2: *Eigenvectors*  $\hat{\Gamma} = (\hat{\gamma}_1, \dots, \hat{\gamma}_6)$  and *eigenvalues*  $\hat{\lambda}_{ij} \times 10^3$  (from top to bottom increasing maturity  $\tau_i \in \{0.06205, 0.12500, 0.18750, 0.25000\}$ ) of the common principle components model, standard errors in parenthesis; sample period 1995 to 05/2001; ODAX.

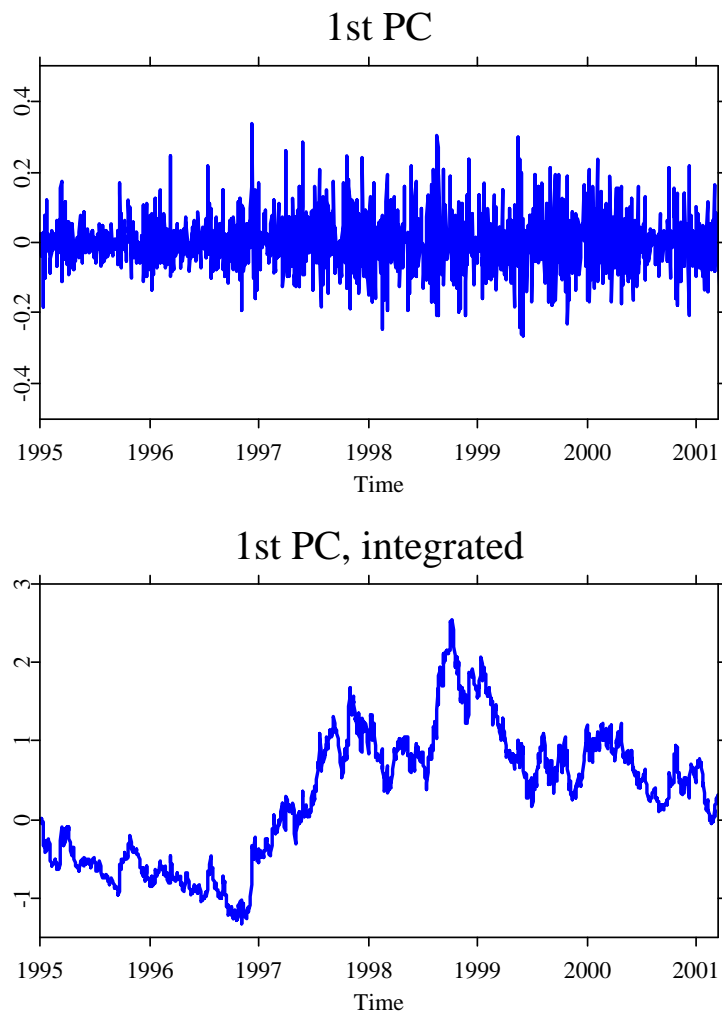


Figure 5.2: *Projection of the longest maturity group on the first eigenvector; upper panel shows returns, lower panel the integrated series.*

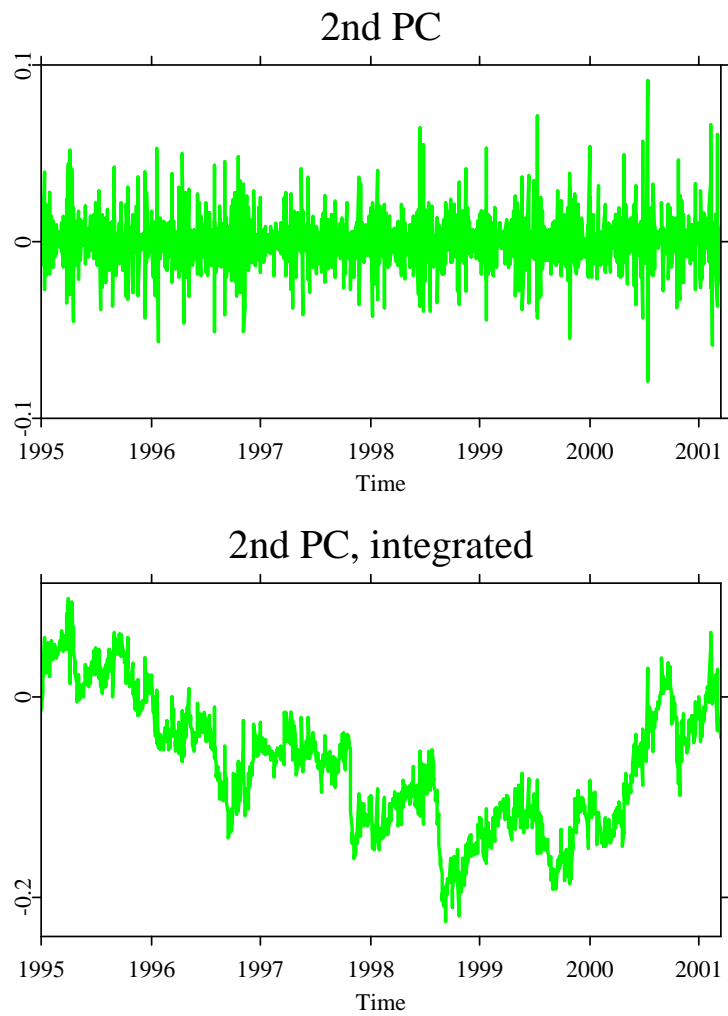


Figure 5.3: *Projection of the longest maturity group on the second eigenvector; upper panel shows returns, lower panel the integrated series.*

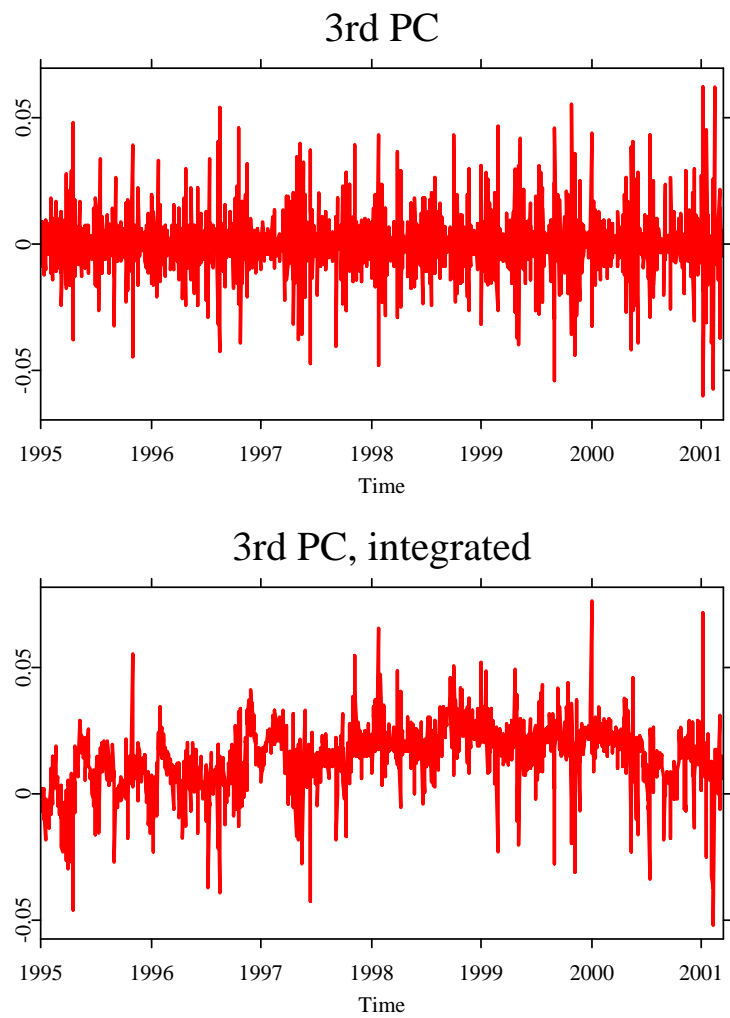


Figure 5.4: *Projection of the longest maturity group on the third eigenvector; upper panel shows returns, lower panel the integrated series.*

## 5.2 Model selection

After presenting the empirical results of the CPC model for the implied volatility data, a discussion of model selection seems to be appropriate. As explained in Section (4) due to their structure CPC models are nested in a hierarchy of models with increasing flexibility of the eigenstructure, Table (5.3). This allows for sequentially testing each model against the next more flexible model via likelihood ratio tests. For example, the log-likelihood ratio statistic for testing the  $H_{\text{CPC}}$  against the unrestricted model (unrelatedness between covariance matrices) is given by

$$T = -2 \ln \frac{L(\hat{\Psi}_1, \dots, \hat{\Psi}_k)}{L(\mathbf{S}_1, \dots, \mathbf{S}_k)} = \sum_{i=1}^k (n_i - 1) \ln \frac{|\hat{\Psi}_i|}{|\mathbf{S}_i|}, \quad (5.1)$$

where  $L(\mathbf{S}_1, \dots, \mathbf{S}_k)$  denotes the unrestricted maximum of the log-likelihood. The number of parameters estimated in the CPC model is  $p(p-1)/2$  for the orthogonal matrix  $\mathbf{\Gamma}$  plus  $kp$  for the eigenvalues  $\mathbf{\Lambda}_i$ , and the number of parameters in the unrelated case is given by  $kp(p-1)/2 + kp$ . Hence the test is asymptotically  $\chi^2$  with  $(k-1)p(p-1)/2$  degrees of freedom as  $\min n_i \rightarrow \infty$ , Rao (1973).

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 5.3: *Testing sequentially against the next lower model in the model hierarchy and corresponding degrees of freedom, Flury (1988).*

Alternatively, one can use Akaike and Schwarz information criteria (AIC, SIC) for model selection, Akaike (1973); Schwarz (1978). Our AIC is slightly modified: Given  $U$  hierarchically ordered models to compare, with numbers of parameters  $a_u$  in model  $u$ , we define the AIC as

$$AIC(u) = -2(L_u - L_U) + 2(a_u - a_1), \quad (5.2)$$

where  $L_u$  is the maximum of the log-likelihood function of model  $u$ .

The SIC takes the finite dimensionality better into account and gives a more severe complexity penalty than the AIC. Our modified SIC is defined as

$$SIC(u) = -2(L_u - L_U) + 2(a_u - a_1) \ln(N), \quad (5.3)$$

where  $N = \sum_{i=1}^k n_i$  denotes the overall sum of observations across the  $k$  groups. For both criteria the model with the lowest AIC or SIC is the best fitting one. A more detailed discussion on choice and definition of these criteria is found in [Fengler et al. \(2001\)](#).

Model		Chi. Sqr	df	$p$ -val	AIC	SIC
higher	lower					
Equality	Proportionality	1174.9	3	0.00	3529.3	3529.3
Proportionality	CPC	1488.8	15	0.00	2360.3	2407.0
CPC	pCPC(4)	122.6	3	0.00	901.5	1181.2
pCPC(4)	pCPC(3)	210.6	6	0.00	784.9	1111.2
pCPC(3)	pCPC(2)	115.5	9	0.00	586.2	1005.8
pCPC(2)	pCPC(1)	398.9	12	0.00	488.7	1048.2
pCPC(1)	Unrelated	17.6	15	0.28	113.7	859.7
Unrelated					126.0	1105.1

Table 5.4: *Step-up  $\mathcal{E}$  model building approach of CPC models*

The results of these test procedures are displayed in Table (5.4). According to the sequential chi-square tests the model to be preferred is a partial CPC(1) model, since this test is the first not to be rejected against the next more flexible model. Also when testing directly against the unrelated model (these tests are obtained by adding up the test statistics and the corresponding degrees from freedom between the model of interest and the unrelated model in Table (5.4)) it is the partial CPC(1) model which is not rejected. AIC and SIC both recommend the partial CPC(1). Note also that according to the SIC all CPC( $q$ ) models with  $q \leq 3$  are superior to the unrelated model. For the remaining CPC models SIC is slightly higher than for the unrelated case, whereas for the proportional and the equality models the information criteria increase tremendously. Since from Section (5) it was obvious that for an approximation up to 88% one component is sufficient, while the second and third only add 6% and 3% of explained variance, we believe also for computational and practical simplicity that a CPC model can be chosen as a valid description of IVS dynamics.

### 5.3 Stability analysis

For any application in trading or risk management model stability is a decisive model characteristic, since otherwise ‘model risk’ becomes unreckonable. In terms of the CPC models under consideration there are two types of stability of interest: the more important one refers to the stability of the transformation matrices  $\mathbf{\Gamma}$ . Stability of  $\mathbf{\Gamma}$  implies that the model can be estimated in a given (historical) sample period and contemporaneous PCs can be obtained by daily updating the data base of implied volatilities and by projecting them into the same space without explicitly estimating  $\mathbf{\Gamma}$  again. The second type of stability refers to the variances of the components collected in  $\mathbf{\Lambda}_i$ . Instability of  $\mathbf{\Lambda}_i$  does not imply the need to often re-estimate the model, since the CPC model places no restrictions on  $\mathbf{\Lambda}_i$ :  $\mathbf{\Psi}_i = \mathbf{\Gamma}\mathbf{\Lambda}_i\mathbf{\Gamma}^\top$  may very well hold across time in the sense of time-dependent variances  $\mathbf{\Psi}_{i,t}$  and  $\mathbf{\Lambda}_{i,t}$ . However, instability of  $\mathbf{\Lambda}_i$  implies the need of time series models capturing heteroscedasticity of the PCs and thus has significant impact on model choice in Section (6).

To assess stability we split the entire sample into  $R = 7$  annual, non-overlapping sub-samples with around 250 observations in each sub-sample, except of the last one with 105 observations. In each of the annual samples we estimate the CPC model separately.

#### 5.3.1 Stability of eigenvectors

In Figure (5.5) we display the estimation results, where again blue refers to the first, green to the second and red to the third eigenvector. To highlight time dependence, colors move gradually from light to intensive tones the more recent the sub-sample.

As is immediately seen, the general structure of the eigenvectors is not altered: shift, slope and twist interpretation are visible for each sample. However, estimates display variability of different degrees. The first eigenvectors (blue) changes to some degree and seems to slightly rotate around ATM implied volatility with giving more equal weights to implied volatilities across different moneyness for the most recent samples. The second and third eigenvector have greater variance through time. For the years 1995, 1999, 2000, 2001 the second eigenvector appears concave, for 1996, 1997, 1998 convex. From the color intensity it is also seen that the third eigenvectors appear hardly altered for the most recent samples,

Common Coordinate Plot: First three Eigenvectors

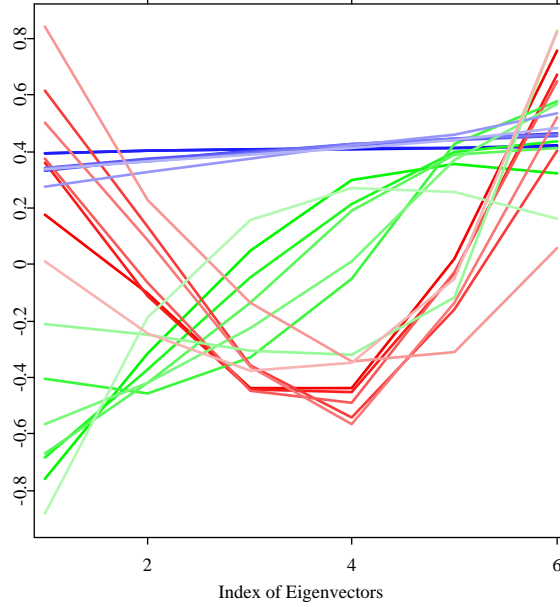


Figure 5.5: *CPC model estimated separately in each annual sample 1995 - May 2001; ODAX.*

whereas only those belonging to the samples 1995 and 1996 are largely different.

In Table (5.5) we display the test results for stability of the eigenvectors. All test statistics are rejected. Even the slight variation of the first eigenvectors has to be rejected based on the asymptotic distribution. Thus, for our time series models, we will use projections of the time series based on yearly estimates of  $\hat{\Gamma}$ .

$\gamma_1$	$\gamma_2$	$\gamma_3$
433.3	7318.3	4089.3

Table 5.5: *Test statistics for stability of eigenvectors  $\gamma_j^{(r)}$  across  $R = 7$  sub-samples. The critical value at the 1% level is  $\chi^2(0.99; 30) = 50.9$ .*

### 5.3.2 Stability of eigenvalues

The prominent virtue of the CPC model is that it allows for different variances in each group, while eigenvectors are restricted to be constant through groups. As will presently be seen, this is an adequate model feature as variances of long run and short run implied volatilities are indeed different, and their relative size is subject to changes through time. In Figure (5.6), we display estimated eigenvalues for the first, i.e. shortest maturity group (filled circles), and the fourth group, i.e. the longest maturity group (plain circles). Figure (5.7) reiterates Figure (5.6) for  $\lambda \leq 0.002$ .

As reported above, the first component (blue circles) largely outweighs the second and third ones (green and red circles) in terms of variance, but the prominence of the first component (filled circles) decays with increasing time to maturity: Whereas for the first group eigenvalues are roughly around 0.02, they drop to around 0.005 in group four (plain circles). This reflects the empirical fact that long run implied volatilities behave less volatile than those calculated from front contracts. Furthermore, we find that the relative importance of slope and twist shocks increases slightly for long maturities: The ratio of the first eigenvalue to the overall sum of eigenvalues decreases from around 0.95% to around 85% from group one to four.

group $i$	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
1	115.9	184.4	241.3
2	132.3	215.3	207.2
3	161.9	104.0	148.0
4	170.0	96.1	153.4

Table 5.6: *Test statistics for stability of eigenvalues  $\lambda_{ij}^{(r)}$  for eigenvalues  $j = 1, 2, 3$ , in  $k = 4$  groups across  $R = 7$  sub-samples. The critical value at the 1% level is  $\chi^2(0.99; 6) = 16.81$ .*

Another obvious empirical fact is a strong variability of the first eigenvalues (blue circles) over the seven sample periods: From 1995 to its 1998 volatility of the first component is increasing, while from 1998 it gradually declined. Thus, in line with increased volatility levels during the Asian market turmoil 1997 and its aftermath, also volatility of implied volatility increased. This effect is only reflected in the first component, whereas second

(green circles) and third eigenvalues (red circles) stay roughly constant during from 1995 to 1999, and increase sharply in the years 2000 and 2001. Generally, this implies that relative importance of slope and twist shocks grows during less volatile markets.

In Table (5.6), we present our test statistics of the tests for stability of the first three eigenvalues in each maturity group across the seven sub-samples. As can readily be seen, one can reject the hypothesis of homogeneity of variances in all tests. This finding implies that modeling conditional heteroscedasticity will be the appropriate model choice in the time series analysis in Section (6).

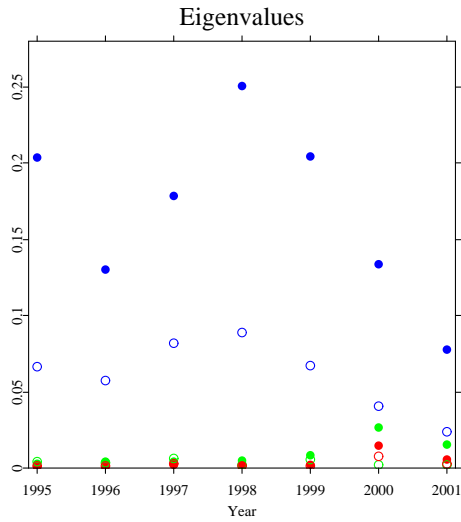


Figure 5.6: *Eigenvalues  $\lambda \times 10$  of the CPC model for each sample from 1995 - May 2001, ODAX. Group 1, filled circle, group 4, plain circle.*

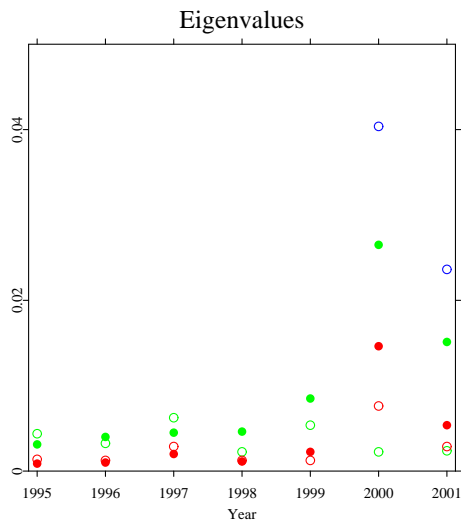


Figure 5.7: *Eigenvalues  $\lambda \times 10$  of the CPC model for each sample from 1995 - May 2001, ODAX. Group 1, filled circle, group 4, plain circle. Same as Figure (5.6) for  $\lambda \leq 0.002$ , only.*

## 6 Time series models

The time series of the first three principal components are obtained from the projection  $\mathbf{Y}_k = \hat{\mathbf{\Gamma}}^\top \mathbf{X}_k$ , i.e. from the longest times to maturity group, only. Since stability of the eigenvectors was rejected, we reestimate the model in each sub-samples, project according to the new matrices  $\hat{\mathbf{\Gamma}}^{(r)}$  to obtain the time series of principal components. To understand the dynamics of the components, our first approach to modeling aims at the principal components itself, a second one to the integrated processes. Based on autocorrelation and partial autocorrelation plots, we propose adequate models for each series. By a AIC and SIC searches we will identify a best fitting model, and present estimation results in more detail.

From Figures (5.2) to (5.4) it is seen that the first three principal components  $y_{1t}$ ,  $y_{2t}$ , and  $y_{3t}$  display a very white noise like behavior. This impression is reinforced when inspecting the autocorrelation and partial autocorrelation functions as displayed in Figures (6.1)–(6.12). From Figure (6.1) it is seen that the first component  $y_{1t}$  exhibits no autocorrelation: it immediately dies off. Also the partial autocorrelation function in Figure (6.7) shows no significant structure. Thus  $y_{1t}$ , which explains up to 88% of the variance, is likely to be considered as pure noise.

For the second and third components  $y_{2t}$  and  $y_{3t}$  a different picture arises: From Figures (6.3) and (6.5) a negative first order correlation is visible hinting towards an MA(1) model. Also the partial autocorrelation functions in Figures (6.9) and (6.11) display the typical patterns of an MA process.

With this preliminary analysis at hand, we perform AIC and SIC searches over MA( $q$ )-GARCH( $r, s$ ) models, where  $q = 0, 1, 2$ ,  $r = 1, 2$ ,  $s = 1, 2$  for  $y_{1t}$ , and  $q = 1, 2$ ,  $r = 1, 2$ ,  $s = 1, 2$  for  $y_{2t}$  and  $y_{3t}$ . We also estimate different types of GARCH specifications such as TARARCH specifications in order to investigate asymmetries in shocks. Since Table (5.1) suggests substantial correlation with contemporaneous index returns we additionally include index returns into mean equations of all processes, and additionally into the variance of  $y_{1t}$ . Thus

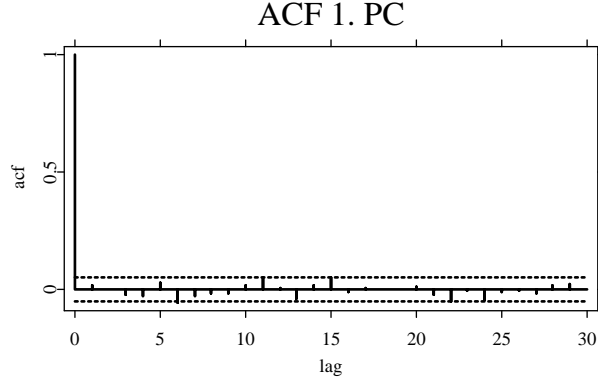


Figure 6.1: *Autocorrelation function of the 1. PC.*

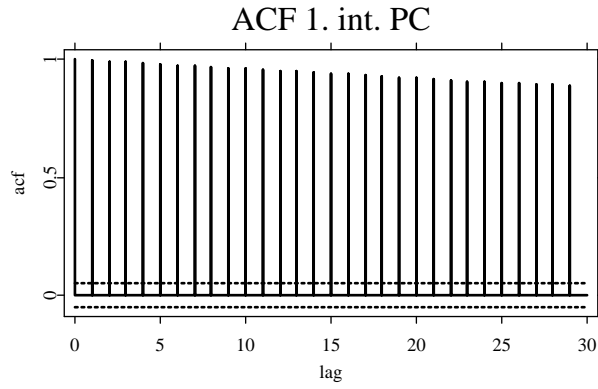


Figure 6.2: *Autocorrelation function of the 1. integrated PC.*

the MA-GARCH models are given by

$$y_{it} = c + a_1 z_t + \varepsilon_{it} + b_1 \varepsilon_{i,t-1}, \quad (6.1)$$

$$\varepsilon_{it} \sim \mathcal{N}(0, \sigma_{it}^2),$$

$$\sigma_{it}^2 = \omega + \sum_{j=1}^k \alpha_j \sigma_{i,t-j} + \sum_{j=1}^s \beta_j \varepsilon_{i,t-j}^2 + \gamma z_t^2, \quad (6.2)$$

where  $z_t$  denotes log returns in the DAX index.

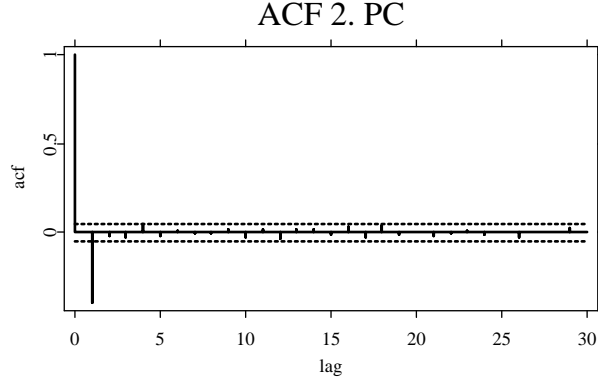


Figure 6.3: *Autocorrelation function of the 2. PC.*

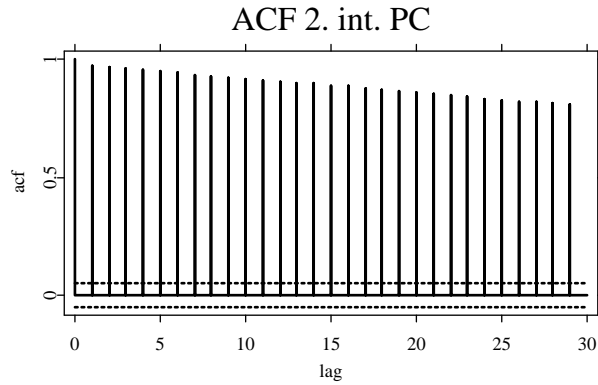


Figure 6.4: *Autocorrelation function of the 2. integrated PC.*

Table (6.1) displays model selection criteria over the models under consideration. For  $y_{1t}$  both AIC and SIC suggest an GARCH(1,2) specification. For  $y_{2t}$  and  $y_{3t}$ , the results are not as clear cut. However, since differences of model selection criteria are very much the same, we decided for the most parsimonious models, i.e. MA(1)-GARCH(1,1) for both. Given these results, one may like to alter the variance equation to allow for asymmetries in shocks: Under the TARARCH model (Glosten et al., 1993; Zakoian, 1994) the variance equation (6.2) becomes

$$\sigma_{it}^2 = \omega + \sum_{j=1}^r \alpha_j \sigma_{i,t-j} + \sum_{j=1}^s \beta_j \varepsilon_{i,t-j}^2 + \beta_1^- \varepsilon_{i,t-1}^2 I(\varepsilon_{i,t-1} < 0) + z_t^2. \quad (6.3)$$

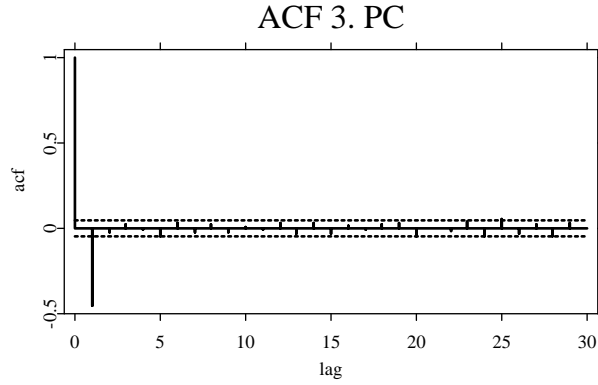


Figure 6.5: *Autocorrelation function of the 3. PC.*

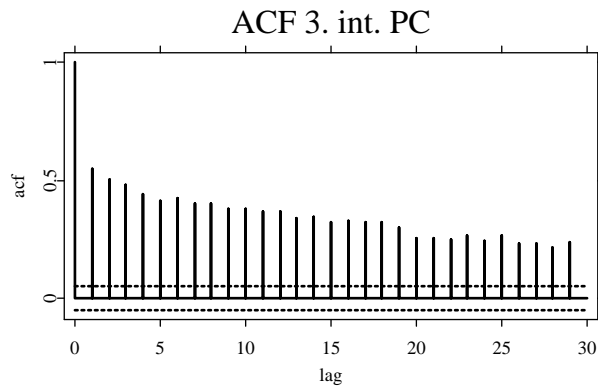


Figure 6.6: *Autocorrelation function of the 3. integrated PC.*

In this model, good news,  $\varepsilon_t > 0$ , and bad news,  $\varepsilon_t < 0$ , have differential effects on the conditional variance – good news have an impact of  $\sum_{j=1}^s \beta_j$ , while bad news have an impact of  $\sum_{j=1}^s \beta_j + \beta_1^-$ . If  $\beta_1^- > 0$  a leverage effect exists, and the news impact is asymmetric if  $\beta_1^- \neq 0$ . We also estimated EGARCH models (Nelson, 1991), however since they did not bring any substantial gain compared to the other models, we do not report estimation results here.

In Table (6.2) estimation results are displayed in more detail. From the mean equation for  $y_{1t}$  it is clear that index returns have a highly significant impact on the first principal

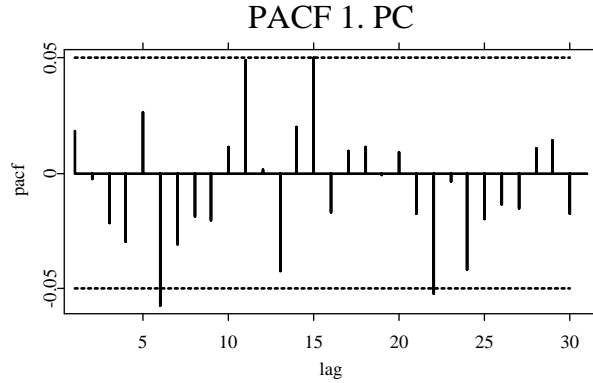


Figure 6.7: *Partial autocorrelation function of the 1. PC.*

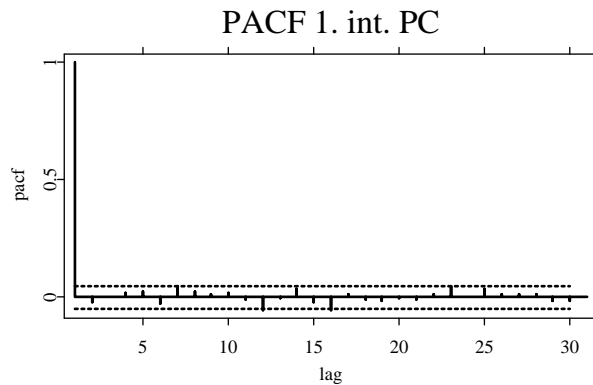


Figure 6.8: *Partial autocorrelation function of the 1. integrated PC.*

component. The sign – as explained in Section (5) – is in line with the ‘leverage effect’ hypothesis: negative shocks in the underlying increase volatility. In the variance equation all parameters are significant.  $\beta_2 < 0$  may be interpreted as an ‘over-reaction correction’ in terms of variance: High two-period lagged returns have a dampening impact on variance. As is to be expected volatility increases also when volatility in the underlying is high ( $\gamma > 0$ ). From the TGARCH model, no evidence for the ‘GARCH type’ leverage effect is found, since  $\beta_1^- < 0$ . The other parameter estimates for the TGARCH are of same size and significance level. Adjusted  $\bar{R}^2$  is around 23%. This seems to be quite high, however it must be noted that this is entirely due to the index returns. Leaving  $z_t$  out of the mean equations reduces

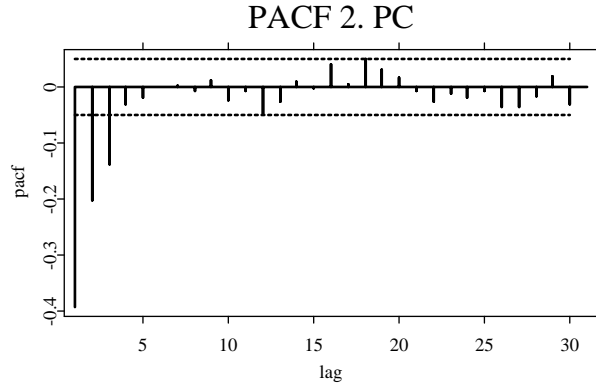


Figure 6.9: *Partial autocorrelation function of the 2. PC.*

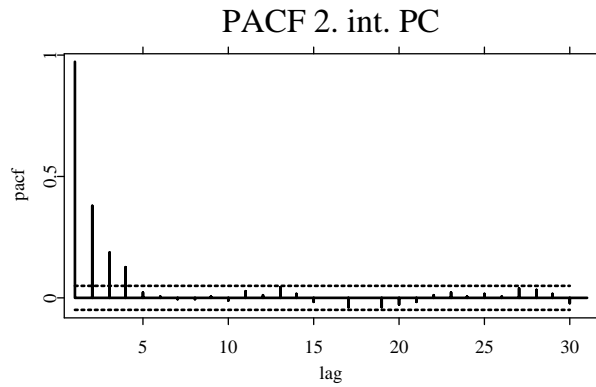


Figure 6.10: *Partial autocorrelation function of the 2. integrated PC.*

$\bar{R}^2$  to around 0.2%, only.

In the mean equations of  $y_{2t}$  and  $y_{3t}$  the MA(1) components are negative and significant. Index returns are only significant for  $y_{2t}$  and positively influence the slope structure in the surface. Thus, together with the result from  $y_{1t}$ , the picture emerges that positive shocks in the underlying reduce implied volatility levels, while at the same time the slope of the surface is intensified.

The variance equations do not exhibit special features, however it is interesting to remark that a ‘GARCH type’ leverage effect is present, as  $\beta_1^- > 0$ : lagged negative shocks increase

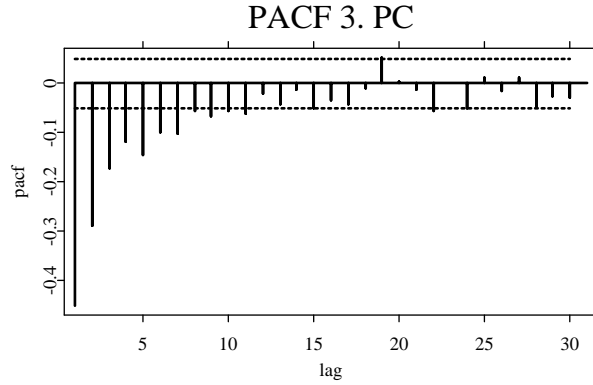


Figure 6.11: *Partial autocorrelation function of the 3. PC.*

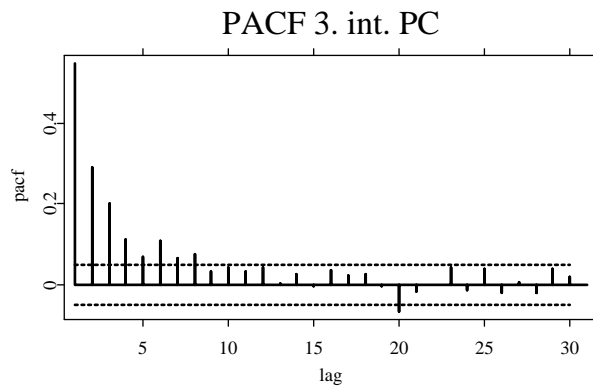


Figure 6.12: *Partial autocorrelation function of the 3. integrated PC.*

the variance of both processes.

We now turn to the analysis of the integrated series. From the (partial) autocorrelation plots, it is seen that they exhibit distinct autoregressive patterns. Especially for the first component, Figures (6.2) and (6.8) indicate even the presence of a unit root. Indeed, the augmented Dickey-Fuller test on the integrated series of  $y_{1t}$  with four lags yields -1.84, which means that the hypothesis of a unit root cannot be rejected at any sensible level of significance. The same conclusion is drawn from a Phillips-Perron test. The trouble with such a result is that it is economically not plausible: It cannot be believed that the first PC,

Model	AIC			SIC		
	$y_{1t}$	$y_{2t}$	$y_{3t}$	$y_{1t}$	$y_{2t}$	$y_{3t}$
GARCH(1,1)	-2.674			-2.654		
GARCH(1,2)	<b>-2.681</b>			<b>-2.657</b>		
GARCH(2,1)	-2.677			-2.654		
GARCH(2,2)	-2.681			-2.654		
MA(1)-GARCH(1,1)		<b>-5.872</b>	-6.460		<b>-5.849</b>	-6.436
MA(1)-GARCH(1,2)		-5.871	-6.460		-5.844	<b>-6.443</b>
MA(1)-GARCH(2,1)		-5.871	<b>-6.461</b>		-5.844	-6.434
MA(1)-GARCH(2,2)		-5.870	-6.461		-5.840	-6.431

Table 6.1: *Univariate model selection: Akaike and Schwarz Information Criteria (AIC, SIC) over a variety of MA(q)-GARCH(r, s) models of  $y_{it}$ .*

i.e. a linear combination of implied volatilities, can reach any level in finite time. However, results indicate that the process behaves locally as a unit root.

Augmented Dickey-Fuller tests on the integrated series of  $y_{2t}$  and  $y_{3t}$  deliver -2.37 and -8.58. For  $y_{2t}$  the hypothesis can nearly be rejected at the 10% level, while for  $y_{3t}$  it is rejected at the 1% level. Phillips-Perron tests support the hypothesis of stationarity in both series.

For our time series model of the integrated series we use the same set-up as before, except that now AR( $p$ )-GARCH( $r, s$ ),  $p = 1$  for  $y_{1t}$ ,  $p = 1, \dots, 4$  for  $y_{2t}$  and  $y_{3t}$ , and  $r, s = 1, 2$ , models seem to be appropriate model choices. Furthermore, there does not seem to be an economic hypothesis how levels of implied volatilities and returns or levels in the underlying are related. Indeed, as was seen in Section (3.3), there are various market regimes in which different relations between these variables are plausible. Hence, in our time series models we do not include DAX returns or DAX index values in the mean equation:

$$y_{it} = c + \sum_{j=1}^p \tilde{\alpha}_j y_{i,t-j} + \varepsilon_{it}, \quad (6.4)$$

$$\varepsilon_{it} \sim \mathcal{N}(0, \sigma_{it}^2),$$

$$\sigma_{it}^2 = \omega + \sum_{j=1}^r \alpha_j \sigma_{i,t-j} + \sum_{j=1}^s \beta_j \varepsilon_{i,t-j}^2 + \gamma z_t^2, \quad (6.5)$$

where  $z_t$  again denotes log returns in the DAX index.

As is seen from Table (6.3) the preferred models are in all cases GARCH(1,1) processes, with an AR(1) for  $y_{1t}$  and AR(4) for  $y_{2t}$  and  $y_{3t}$ . In Table (6.4) estimation results are displayed.  $y_{1t}$  shows the typical characteristics when a unit root is present:  $\tilde{a}_1 \approx 1$  and the  $t$ -statistic and  $\bar{R}^2$  are extremely high. For the variance, we find significant correlation with the volatility of the underlying, and the TGARCH specification leads to the same conclusions as for the simple process: no GARCH leverage effect.

The estimations results for  $y_{2t}$  and  $y_{3t}$  show clear examples of AR(4) processes the impact of which mostly decays the higher the lag. The variance equation is in line with the findings of the simple processes. For these two series, we find evidence for a leverage effect in the TGARCH model.

Summarizing the time series models, we have to conclude two main results: The first, and admittedly disappointing result is the presence of the unit root in the first component, which captures most of the variance in IVS's. Since the best predictor of a random walk, is today's value, there is little that can be done in terms of direct prediction and forecasting. In analysis of volatility functions [Dumas et al. \(1998\)](#) already documented that the 'naïve' volatility predictor, i.e. today's implied volatility, excels all other specifications. Our results confirm this evidence. The second result concerning the second the third component is more positive, since stationary AR(4) processes have some predictability. However, since variance from these sources is small, it must remain a question to further research two which extend this can be exploited in term of trading strategies.

	Factor					
	$y_{1t}$		$y_{2t}$		$y_{3t}$	
cond. mean						
$c$	0.001 (0.407)	0.001 (1.048)	$1.9E^{-4}$ (1.170)	$1.0E^{-4}$ (0.566)	$-3.8E^{-05}$ (-0.592)	$-5.8E^{-05}$ (-0.907)
$a_1$	-2.920 (-24.46)	-2.930 (-24.21)	0.086 (4.860)	0.079 (4.564)	0.005 (0.457)	0.004 (0.351)
$b_1$			-0.733 (-35.50)	-0.501 (-21.78)	-0.733 (-35.50)	-0.729 (-34.81)
cond. var.						
$\omega$	$1.4E^{-4}$ (3.945)	$1.6E^{-4}$ 4.141	$6.7E^{-5}$ (7.515)	$6.4E^{-5}$ (7.353)	$1.7E^{-05}$ (8.687)	$2.2E^{-05}$ (8.681)
$\alpha_1$	0.803 (32.09)	0.797 (29.07)	0.425 (6.774)	0.462 (7.791)	0.686 (24.41)	0.631 (17.11)
$\beta_1$	0.246 (7.112)	0.284 (7.598)	0.200 (6.840)	0.115 (3.505)	0.147 (8.027)	0.082 (3.206)
$\beta_2$	-0.130 (-4.110)	-0.124 (-3.611)				
$\beta_1^-$		-0.950 (-3.706)		0.150 (3.239)		0.142 (3.916)
$\gamma$	1.480 (4.991)	1.580 (4.909)				
$\bar{R}^2$	0.23	0.23	0.22	0.21	0.33	0.33

Table 6.2: *Estimation results of GARCH models for the three principal components, t-statistics in parenthesis.*

Model	AIC			SIC		
	$y_{1t}$	$y_{2t}$	$y_{3t}$	$y_{1t}$	$y_{2t}$	$y_{3t}$
AR(1)-GARCH(1,1)	<b>-2.439</b>	-5.735	-6.233	<b>-2.422</b>	-5.718	-6.217
AR(1)-GARCH(1,2)	-2.438	-5.732	-6.240	-2.418	-5.175	-6.220
AR(1)-GARCH(2,1)	-2.432	-5.734	-6.234	-2.421	-5.714	-6.214
AR(1)-GARCH(2,2)	-2.430	-5.734	-6.240	-2.417	-5.711	-6.217
AR(2)-GARCH(1,1)		-5.822	-6.359		-5.802	-6.339
AR(2)-GARCH(1,2)		-5.821	-6.358		-5.800	-6.334
AR(2)-GARCH(2,1)		-5.821	-6.332		-5.801	-6.321
AR(2)-GARCH(2,2)		-5.822	-6.361		-5.795	-6.335
AR(3)-GARCH(1,1)		-5.850	-6.415		-5.827	-6.391
AR(3)-GARCH(1,2)		-5.849	-6.414		-5.820	-6.387
AR(3)-GARCH(2,1)		-5.849	-6.415		-5.822	-6.389
AR(3)-GARCH(2,2)		-5.850	-6.415		-5.819	-6.385
AR(4)-GARCH(1,1)		<b>-5.866</b>	-6.442	<b>-5.840</b>	<b>-6.415</b>	
AR(4)-GARCH(1,2)		-5.865	-6.443	-5.835	-6.413	
AR(4)-GARCH(2,1)		-5.864	-6.444	-5.822	-6.412	
AR(4)-GARCH(2,2)		-5.864	<b>-6.446</b>	-5.831	-6.412	

Table 6.3: *Univariate model selection: Akaike and Schwarz Information Criteria (AIC, SIC) over a variety of MA(q)-GARCH(r, s) models of integrated  $y_{it}$ .*

	Factor					
	$y_{1t}$		$y_{2t}$		$y_{3t}$	
cond. mean						
$c$	0.003 (-1.611)	-0.002 (-1.029)	$-6.74E^{-5}$ -0.146	$-3.4E^{-4}$ (-0.741)	0.002 (4.811)	0.002 (4.257)
$\tilde{a}_1$	0.996 (472.87)	0.996 (482.4)	0.548 (19.00)	0.525 (18.94)	0.339 (11.82)	0.340 (11.50)
$\tilde{a}_2$			0.196 (6.122)	0.202 (6.519)	0.211 (6.819)	0.210 (6.783)
$\tilde{a}_3$			0.114 (4.019)	0.122 (4.211)	0.191 (6.575)	0.187 (6.423)
$\tilde{a}_4$			0.132 (5.416)	0.140 (5.656)	0.107 (3.769)	0.108 (3.865)
cond. var.						
$\omega$	0.001 (7.895)	0.001 (8.384)	$6.8E^{-5}$ (7.375)	$6.2E^{-5}$ (7.269)	$1.8E^{-05}$ (9.368)	$2.2E^{-05}$ (9.178)
$\alpha_1$	0.344 (7.897)	0.342 (8.957)	0.415 (6.184)	0.460 (7.417)	0.680 (24.14)	0.622 (17.29)
$\beta_1$	0.150 (5.266)	0.244 (5.179)	0.196 (6.511)	0.106 (3.242)	0.147 (8.022)	0.076 (3.186)
$\beta_1^-$		-0.168 (-3.293)		0.047 (3.394)		0.151 (0.037)
$\gamma$	10.15 (9.319)	9.796 (9.221)				
$\bar{R}^2$	0.99	0.99	0.95	0.96	0.39	0.38

Table 6.4: *Estimation results of AR-GARCH models for the three integrated principal components, t-statistics in parenthesis.*

## 7 Conclusions

In this study we describe the phenomenology of German DAX index implied volatility surfaces (IVS) from 1995 to May 2001. A particular phenomenon discovered by [Fengler et al. \(2001\)](#) is the common eigenstructure inherent to implied volatility maturity groups. In this study we extend this analysis to a larger data set. Special concern is put on model stability through the sample period under consideration and time series modeling.

We identify a common eigenstructure and can reduce model complexity to three factors, obtained by projecting implied volatilities into the space spanned by a common orthogonal transformation matrix. Variability of the estimated eigenvectors is little for the first and most important eigenvector, and larger for the second and third ones. However, based on the asymptotic distribution we reject the hypothesis of stability across the seven years in all cases. This implies that factor models in implied volatility modeling need to be regularly updated. As is to be expected, also variance stationarity of the factors is rejected in the sample period.

In our time series analysis, we discovered MA models for the obtained factor dynamics. Accordingly, integrated factors can be modeled by AR processes. For the first component there is strong indication of a unit root, while the second and the third factors can be modeled by AR(4) models. In all cases we used a heteroscedastic error specification, motivated by the variance instationarity of the factors. Strong GARCH and TGARCH effects seem to be present in the IVS data.

Several implications can be drawn from these results: Firstly, a common eigenstructure is the preferred model strategy for IVS dynamics. Secondly, from a theoretical point of view, only multi-factor models of stochastic volatility can be sufficient for modeling adequately volatility dynamics. Thirdly, vega risk shows more complex dynamics than simple models assume, which has significant impact on hedging volatility risk, [Engle and Rosenberg \(2000\)](#). Exploiting the CPC models in terms trading strategies based on the factors will remain for further research.

## References

- Akaike, H.**, “Information Theory and an Extension of the Maximum Likelihood Principle,” in “2nd International Symposium on Information Theory,” Akademiai Kiado, Budapest, 1973.
- Alexander, C.**, “Principles of the Skew,” *Risk - Equity Risk Special Report*, 2001, *January*, s29–s32.
- Anderson, T. W.**, “Asymptotic theory for principal component analysis,” *Annals of Mathematical Statistics*, 1963, *34*, 122–148.
- Avellaneda, M. and Y. Zhu**, “An E-ARCH Model for the Term-Structure of Implied Volatility of FX Options,” *Applied Mathematical Finance*, 1997, *4*, 81–100.
- Bakshi, G., C. Cao, and Z. Chen**, “Do call and underlying prices always move in the same direction?,” *Review of Financial Studies*, 2000, *13* (3), 549–584.
- Black, F. and M. Scholes**, “The pricing of options and corporate liabilities.,” *Journal of Political Economy*, 1973, *81*, 637–654.
- Cont, R. and J. da Fonseca**, “The Dynamics of Implied Volatility Surfaces,” *Quantitative Finance*, 2002, *2* (1), 45–602.
- Derman, E.**, “Regimes of Volatility,” Quantitative Strategies Research Notes, Goldman Sachs 1999.
- **and I. Kani**, “Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility,” *International Journal of Theoretical and Applied Finance*, 1998, *1* (1), 61–110.
- Deutsche Börse**, *Leitfaden zu den Aktienindizes der Deutschen Börse*, 4.3 ed., Deutsche Börse AG, 2002.
- Dumas, B., J. Fleming, and R. E. Whaley**, “Implied Volatility Functions: Empirical Tests,” *Journal of Finance*, 1998, *LIII* (6), 2059–2106.
- Engle, R. and J. Rosenberg**, “Testing the volatility term structure using option hedging criteria,” *Journal of Derivatives*, 2000, *8*, 10–28.

- Fengler, M. R., W. Härdle, and C. Villa**, “The dynamics of implied volatilities: A common principle components analysis,” Discussion Paper 38, Sfb 373, Humboldt-University, Berlin 2001.
- , – , and **E. Mammen**, “Implied Volatility String Dynamics,” Technical Report, Humboldt-University, Berlin 2002.
- , – , and **P. Schmidt**, “The Analysis of Implied Volatilities,” in W. Härdle, T. Kleinow, and G. Stahl, eds., *Applied Quantitative Finance*, Heidelberg: Springer, 2002, chapter 6.
- , – , and – , “Common Factors Governing VDAX Movements and the Maximum Loss,” *Journal of Financial Markets and Portfolio Management*, 2002, 16 (1), 16–29.
- Flury, B.**, *Common Principal Components and Related Multivariate Models* Wiley Series in Probability and Mathematical Statistics, New York: John Wiley & Son, 1988.
- and **W. Gautschi**, “An Algorithm for Simultaneous Orthogonal Transformations of Several Positive Definite Matrices to nearly Diagonal Form,” *Journal on Scientific and Statistical Computing*, 1986, 7, 169–184.
- Glosten, L., R. Jagannathan, and D. Runkle**, “Relationship between the expected value and the volatility of the nominal excess return on stocks,” *Journal of Finance*, 1993, 48, 1779–1801.
- Hafner, R. and M. Wallmeier**, “The Dynamics of DAX Implied Volatilities,” *International Quarterly Journal of Finance*, 2001, 1 (1), 1–27.
- Härdle, W.**, *Applied Nonparametric Regression*, Cambridge, UK: Cambridge University Press, 1990.
- and **L. Simar**, *Applied Multivariate Statistical Analysis*, Springer Verlag, 2003. Forthcoming.
- , **M. Müller, S. Sperlich, and A. Werwatz**, *Non - and Semiparametric Modelling*, <http://www.i-xplore.de/>: Springer Verlag, 2002.
- Johnson, R. A. and D. W. Wichern**, *Applied Multivariate Statistical Analysis*, 4 ed., Englewood Cliffs, N.J.: Prentice-Hall, 1998.

- Ledoit, O. and P. Santa-Clara**, “Relative Option Pricing With Stochastic Volatility,” Working paper, UCLA, Los Angeles, USA 1998.
- Lintner, J.**, “The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *Review of Economics and Statistics*, 1965, 47, 13–37.
- Markowitz, H.**, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley, New York, 1959.
- Nelson, Daniel B.**, “Conditional heteroskedasticity in asset returns: A new approach,” *Econometrica*, 1991, 59, 347–370.
- Rao, C. R.**, *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York, 1973.
- Rebonato, R.**, *Interest-Rate Option Models: Understanding, Analyzing and Using Models for Exotic Interest-Rate Options* Wiley Series in Financial Engineering, 2nd ed., John Wiley & Son Ltd., 1998.
- Rosenberg, J.**, “Implied volatility functions: A reprise,” *Journal of Derivatives*, 2000, 7, 51–64.
- Schönbucher, P. J.**, “A Market Model for Stochastic Implied Volatility,” Working paper, Department of Statistics, Bonn University 1999.
- Schwarz, G.**, “Estimating the Dimension of a Model,” *Annals of Statistics*, 1978, 6, 461–464.
- Sharpe, W.**, “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, 1964, 19, 425–442.
- Skiadopoulos, G., S. Hodges, and L. Clewlow**, “The Dynamics of the S&P 500 Implied Volatility Surface,” *Review of Derivatives Research*, 1999, 3, 263–282.
- Tompkins, R.**, “Stock Index Futures Markets: Stochastic Volatility Models and Smiles,” *The Journal of Futures Markets*, 2001, 21 (1), 43–78.
- Zakoian, J.M.**, “Threshold Heteroskedastic Functions,” *Journal of Economic Dynamics and Control*, 1994, 18, 931–955.

## Appendix

In this appendix we summarize a number of plots of the general data description in Section (3.3) and two important results on the distribution theory of the CPC estimates which are integral to the tests derived in Section (4).

### A Additional implied volatility surface plots

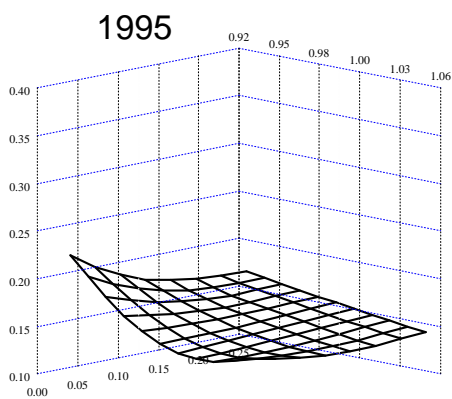


Figure A.1: *Mean implied volatility surface 1995, ODAX.*

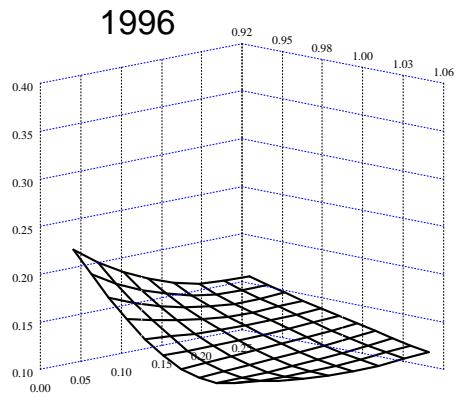


Figure A.2: *Mean implied volatility surface 1996, ODAX.*

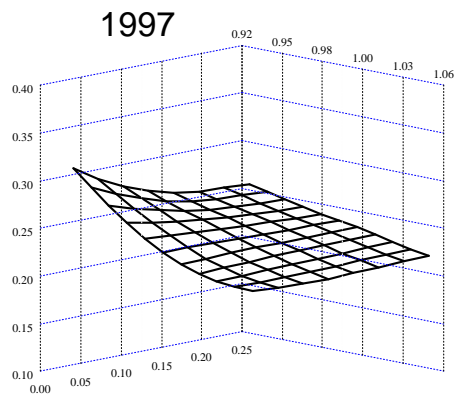


Figure A.3: *Mean implied volatility surface 1997, ODAX.*

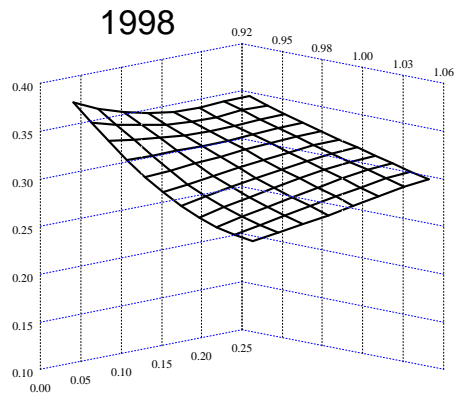


Figure A.4: *Mean implied volatility surface 1998, ODAX.*

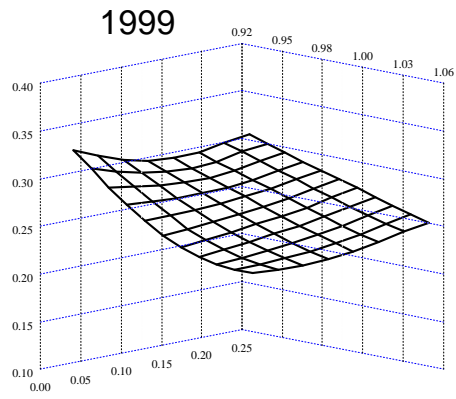


Figure A.5: *Mean implied volatility surface 1999, ODAX.*

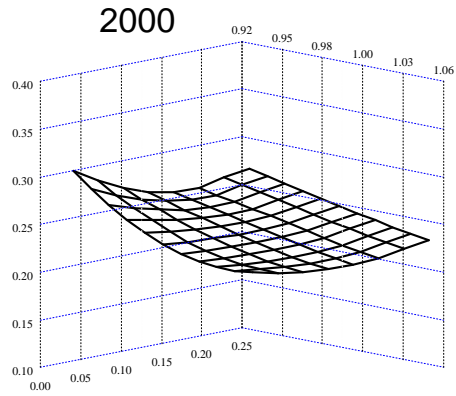


Figure A.6: *Mean implied volatility surface 2000, ODAX.*

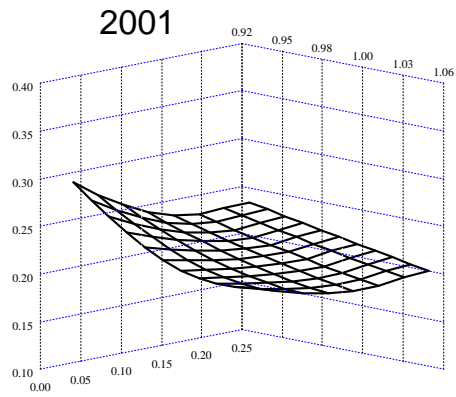


Figure A.7: *Mean implied volatility surface from January to May 2001, ODAX.*

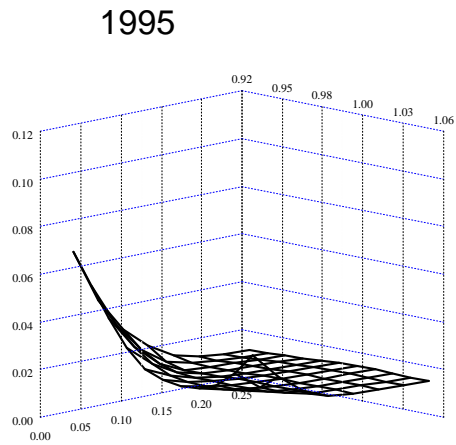


Figure A.8: *Standard deviation of implied volatility surfaces 1995, ODAX.*

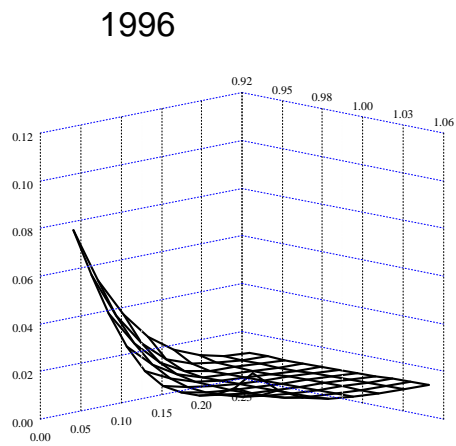


Figure A.9: *Standard deviation of implied volatility surfaces 1996, ODAX.*

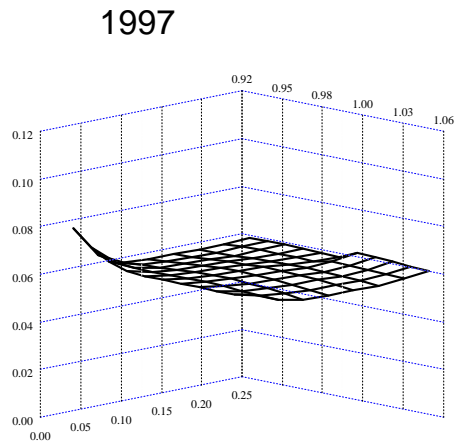


Figure A.10: *Standard deviation of implied volatility surfaces 1997, ODAX.*

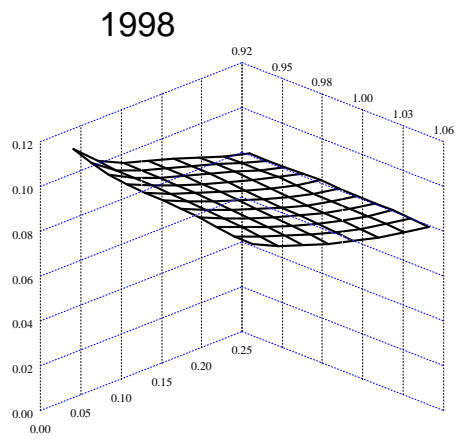


Figure A.11: *Standard deviation of implied volatility surfaces 1998, ODAX.*

1999

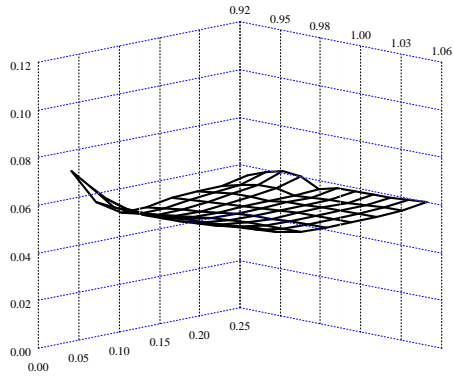


Figure A.12: *Standard deviation of implied volatility surfaces 1999, ODAX.*

2000

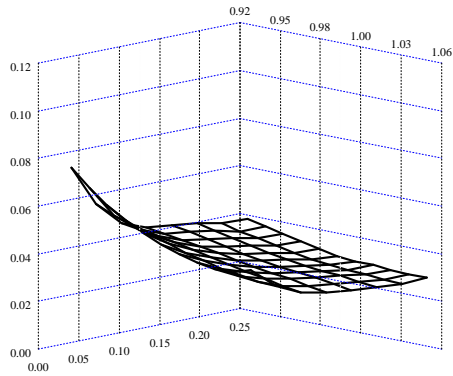


Figure A.13: *Standard deviation of implied volatility surfaces 2000, ODAX.*

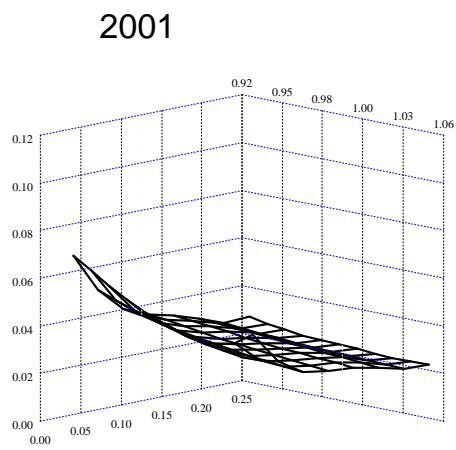


Figure A.14: *Standard deviation of implied volatility surfaces from January to May 2001, ODAX.*

## B Asymptotic theory

The estimates  $\hat{\Lambda}_i$  and  $\hat{\Gamma}$  obtained by optimizing the likelihood function (4.4) have the following properties:

**Proposition B.1.** *The estimated eigenvalues  $\hat{\lambda}_{ij}$ ,  $i = 1, 2, \dots, k$ ,  $j = 1, 2, \dots, p$  behave asymptotically as*

$$\sqrt{n_i - 1}(\hat{\lambda}_{ij} - \lambda_{ij}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2\lambda_{ij}^2)$$

as  $\min n_i \rightarrow \infty$ , and are independent of each other and independent of  $\hat{\Gamma}$ .

**Proof B.1.** *Flury (1988).*

**Proposition B.2.** *Denote  $N = \sum_{i=1}^k n_i$  the overall sum of observations. The asymptotic distribution of the  $p$  eigenvectors is given by*

$$\sqrt{N - k} \text{vec}(\hat{\Gamma} - \Gamma) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{V}_{\hat{\Gamma}}),$$

where  $\mathbf{V}_{\hat{\Gamma}}$  is the  $p^2 \times p^2$  matrix

$$\mathbf{V}_{\hat{\Gamma}} = \begin{pmatrix} \sum_{\substack{j=1 \\ j \neq 1}}^p \theta_{1j} \gamma_j \gamma_j^\top & -\theta_{12} \gamma_2 \gamma_1^\top & \cdots & -\theta_{1p} \gamma_p \gamma_1^\top \\ -\theta_{21} \gamma_1 \gamma_2^\top & \sum_{\substack{j=1 \\ j \neq 2}}^p \theta_{1j} \gamma_j \gamma_j^\top & \cdots & -\theta_{2p} \gamma_p \gamma_1^\top \\ \vdots & \vdots & \ddots & \vdots \\ -\theta_{p1} \gamma_1 \gamma_p^\top & -\theta_{p2} \gamma_2 \gamma_p^\top & \cdots & \sum_{\substack{j=1 \\ j \neq p}}^p \theta_{pj} \gamma_j \gamma_j^\top \end{pmatrix} \quad (\text{B.1})$$

and  $\theta_{jm} = \left\{ \sum_{i=1}^k \left( \frac{N-k}{n_i-1} \frac{\lambda_{ij} \lambda_{im}}{(\lambda_{ij} - \lambda_{im})^2} \right)^{-1} \right\}^{-1}$  with  $m \neq j$ .

**Proof B.2.** *Flury (1988).*