Voles, Volas, Values

Matthias R. Fengler
Wolfgang K. Härdle

Institut für Statistik und Ökonometrie
Humboldt-Universität zu Berlin
Center for Applied Statistics and Economics
(CASE)

Microtus Ochrogaster

Prairie Vole

Order Rodentia (Nager): Family Muridae (echte Mäuse) : *Microtus Ochrogaster*

Figure 1: *Microtus Ochrogaster*

Description & habits:

- dark brownish or blackish mouse; total length 146 mm, tail 34 mm on average
- inhabits Hardin County in southeastern Texas, and in the extreme northern Panhandle.
- lives in tall-grass prairies in colonies, utilizing underground burrows and surface runways under lodged vegetation for concealment
- food almost entirely vegetable including green parts of plants, seeds, bulbs, and bark, much of which they store for winter use

Microtus Californicus

California Vole

Order Rodentia (Nager): Family Muridae (echte Mäuse): *Microtus Californicus*

Figure 2: *Microtus Californicus*

Description & habits:

- grizzled brownish mouse, gray below; total length, 157-214 mm; tail, 39-68 mm
- known in Southwestern Oregon through much of California
- inhabits grassy meadows from sea level to mountains
- is a burrower, but it also forms surface runways
- food is almost entirely vegetable including green parts of plants, seeds, bulbs, and bark

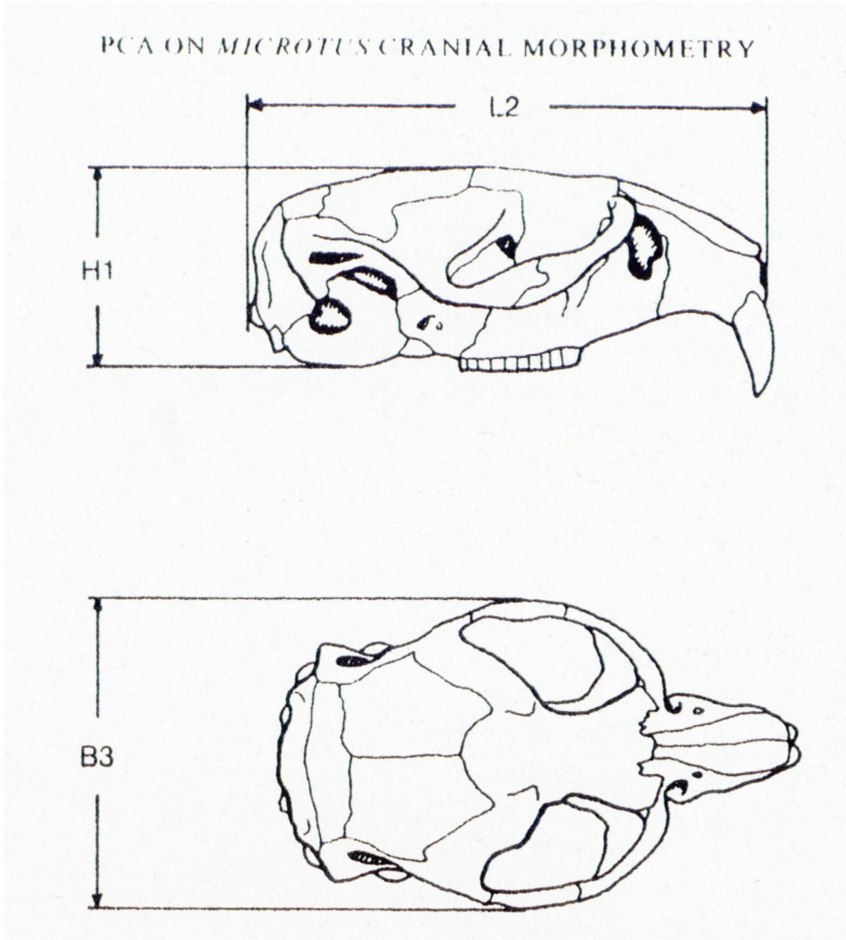


Figure 3: *Diagram of cranial measurements; L2 condylo-incisive length, B3: zygomatic width, H1: skull height. Taken from Airoldi and Flury (1988).*

Common Principle Components

Common principle components has been used for morphometric purposes to estimate a joint eigenstructure for the cranial measurements of voles, Airoldi and Flury (1988).

This data contains cranial measurements for four natural groups of the animals: two sexes in two species. The measurements include the condylo-incisive length (L2),the zygomatic width (B3) and the skull height (H1) (Figure 3).

Key Hypothesis of CPC

Impose

- a joint eigenstructure Γ on population covariance matrices S_i,
- while in-group variances ($=$ eigenvalues λ_i) are not restricted.

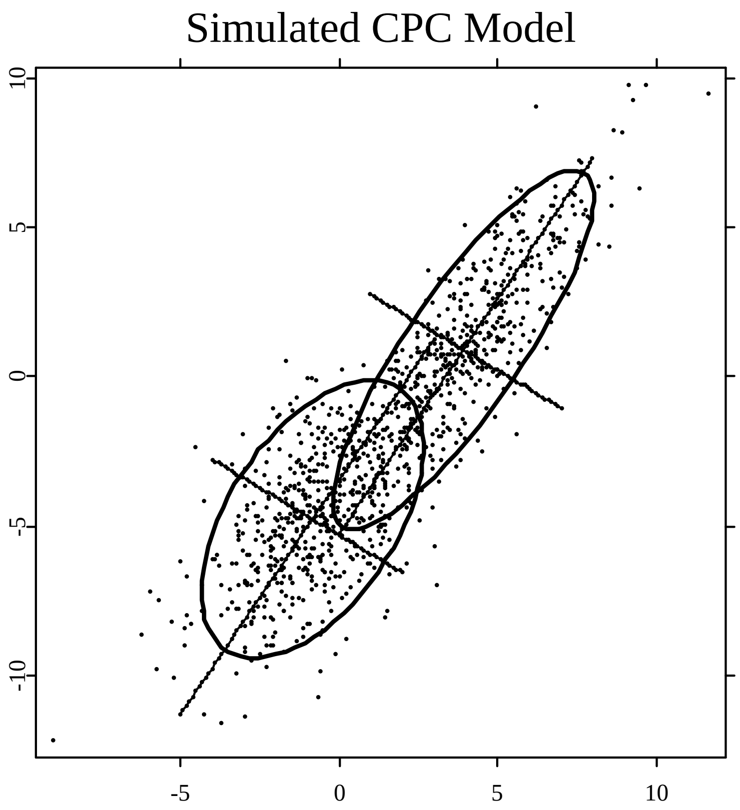


Figure 4: *Simulated CPC model as observable in vole data; compare Airoldi and Flury (1988)*

Voles: What did we learn?

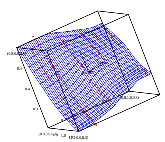
CPC

- allows for estimating a common eigenstructure in the presence of different group variances.
- helps identify morphometric structures across different species and sexes.

Using a simple PCA instead in grouped data may lead to biased estimates.

Overview

1. Voles: Zoological Motivation ✓
2. Volas: Implied Volatility Surface Dynamics
3. Principal Components Analysis
4. *Common* Principal Components Analysis
5. Estimation, Selection, Prediction
6. Values: Trading Strategies, Risk Management



Implied Volatility Surface Dynamics

The Black-Scholes Model, Implied Volatilities and the Smile

Based on the assumption that asset prices follow a geometric Brownian motion, the Black and Scholes (BS) formula values European options:

$$C_t^{BS} = S_t \Phi(d_1) - K e^{-r\tau} \Phi(d_2)$$

BS Formula

$$C_t^{BS} = S_t \Phi(d_1) - K e^{-r\tau} \Phi(d_2)$$

$$d_1 = \frac{\ln(S_t/K) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} \qquad d_2 = d_1 - \sigma\sqrt{\tau}$$

Φ	CDF of the standard normal distribution	r	Interest rate
S_t	Asset price	$\tau = T - t$	Time to maturity
K	Strike price	σ	Constant volatility parameter

BS Formula

Suppose $S_t = 230$, $K = 210$, $r = 5\%$, $\tau = 0.5$, and $\sigma = 25\%$.
Then the call price is given by $C_t^{BS} = 30.98$ and the put price $P_t^{BS} = 5.92$.
You can derive the P_t^{BS} also by the put-call-parity:

$$\begin{aligned} C_t - P_t &= S_t - Ke^{-r\tau} \\ 30.98 - 5.92 &= 230 - 210e^{(-0.05\cdot 0.5)} \end{aligned}$$

Implied Volatilities

However, σ is unknown! Hence define the volatility $\hat{\sigma}$ *implied* by observed market prices $\tilde{C}_t$ as

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

This solution may be found by using a Newton-Raphson or a bisection algorithm. It is unique as the BS formula is globally concave in σ.

Implied Volatilities

Empirical Findings

- Implied volatility is not constant across time t.
- Implied volatility is not flat across strikes.
- Implied volatility is not flat across time to maturity.
- Implied volatility became asymmetric since the 1987 stock market crash.

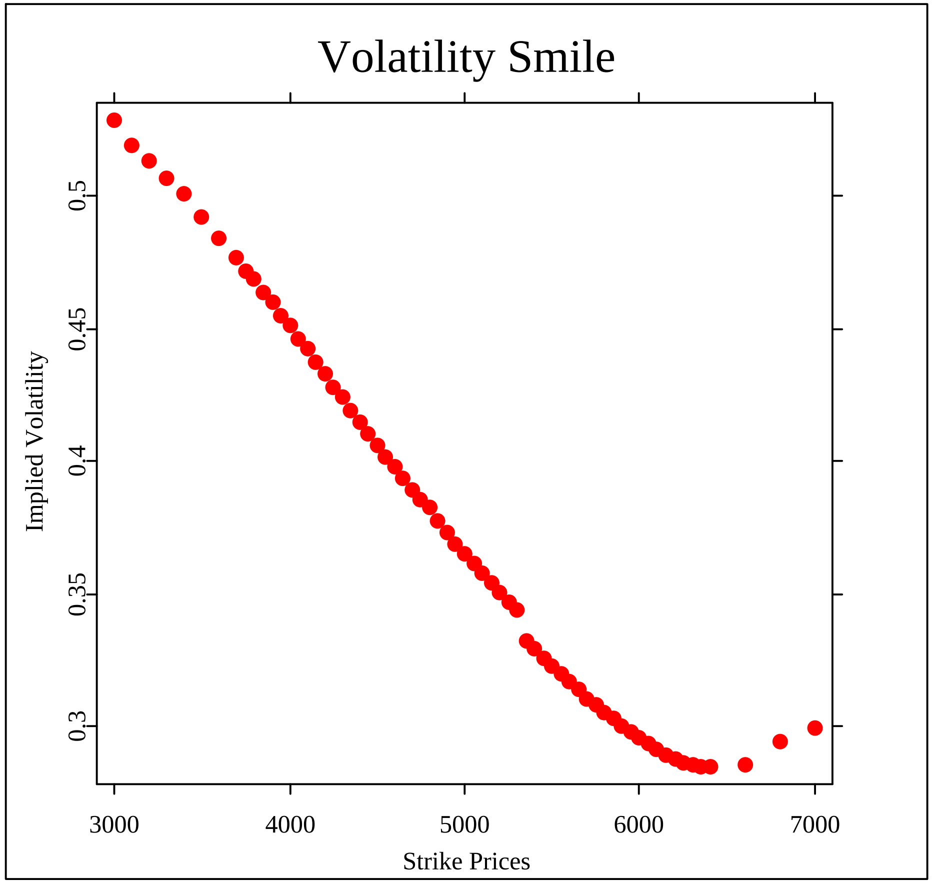


Figure 5: *Vola smile/smirk: 3 months to expiry,* $t = 990104$*, ODAX*

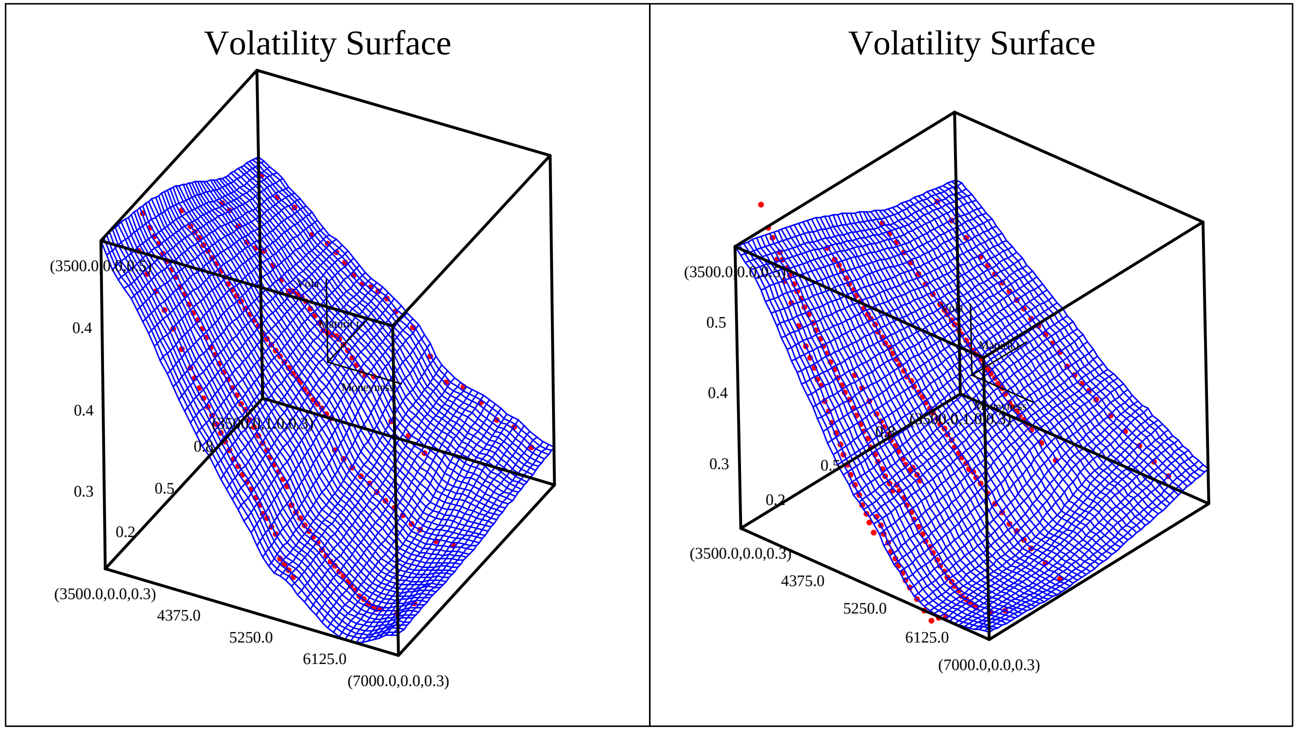


Figure 6: *Implied Volatility Surfaces:* $t_1 = 990104$ *and* $t_2 = 990201$, *ODAX*

CPCdoubleSurf.xpl

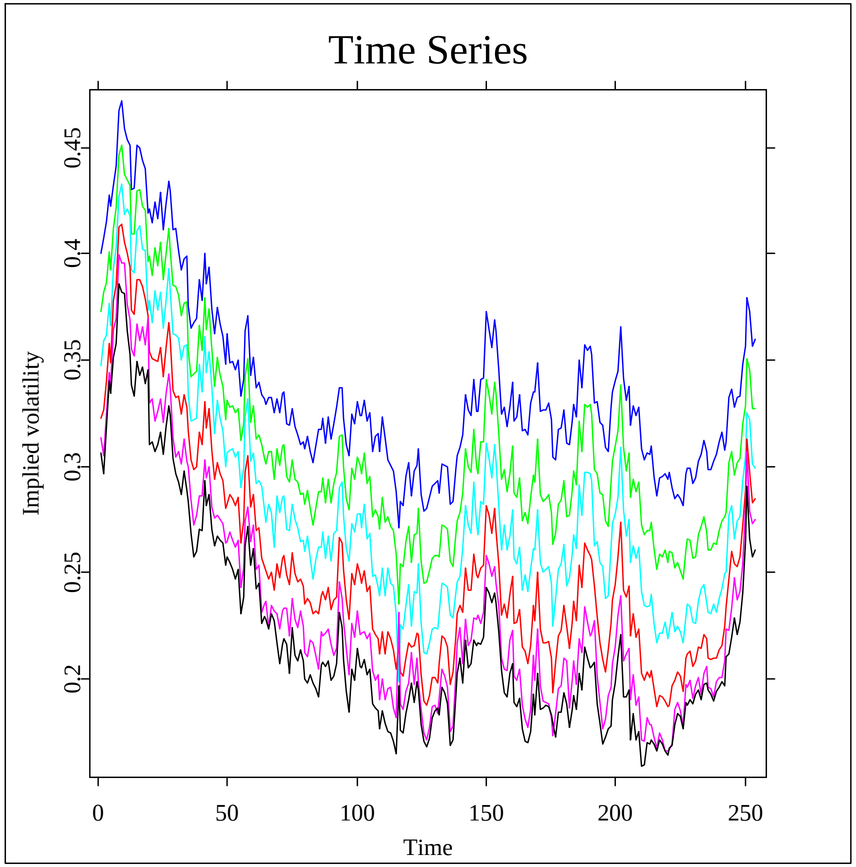


Figure 7: *Time series 1999 of implied volatilities across the smile: 3 months maturity – $\kappa = 1.10$ up to $\kappa = 0.85$, ODAX*

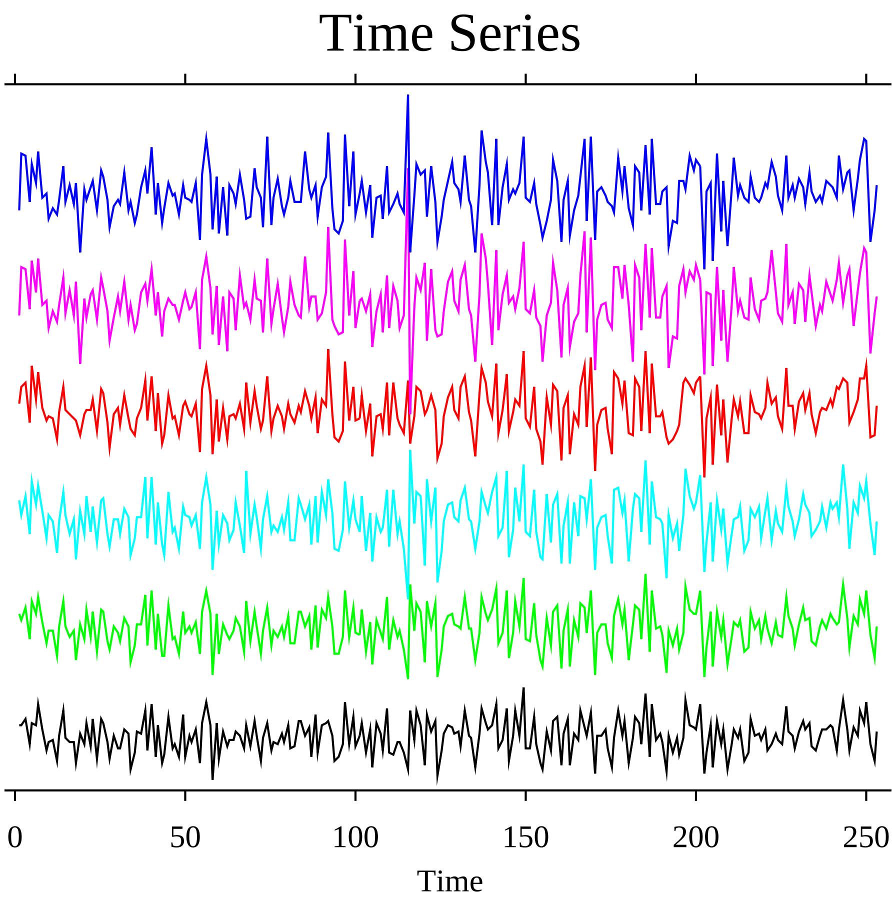
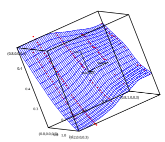


Figure 8: *Time series 1999 of log-returns of implied volatilities across the smile: 3 months maturity –* $\kappa = 1.10$ *up to* $\kappa = 0.85$*, ODAX*



Importance of Implied Volatilities

Practitioners' point of view

- BS-formula serves as a convenient mapping from the spaces of prices, maturities, interest rate, strikes to the real line
- Trading rules can be based on implied volatilities
- Volatility contracts (e.g. VDAX) are based on implied volatilities

Theoretical point of view

- Pricing of illiquid or exotic options by directly modeling implied volatilities: Market Models of Volatility (Dupire, 1994, Derman and Kani, 1989, Schönbucher, 1999)

Purpose of the study

Understand the dynamics of implied volatilities:

- identify the number and shape of shocks driving the surface
- reduce the dimension of the surface vector time series

Plan

- Estimate nonparametrically the implied volatility surface $\hat{\sigma}_t(\kappa, \tau)$ on a fixed grid of moneyness $\kappa_i = \frac{K}{F_{\tau t}}$ and maturity τ_j ($F_{\tau t} = S_t e^{r\tau}$ is the implied future price).
- Apply (Common) Principle Components Analysis to $\Delta \ln \hat{\sigma}_t$
- Study common factors and their dynamics

Nonparametric Smoothing

For a partition of explanatory variables $(x_1, x_2)^\top = (\kappa, \tau)^\top$, i.e. of moneyness and maturities, the two-dimensional Nadaraya-Watson kernel estimator is given by

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

where $\hat{\sigma}_{ti}$ is the volatility implied by the observed option prices $\tilde{C}_{ti}(\kappa, \tau)$ or $\tilde{P}_{ti}(\kappa, \tau)$ respectively, K_1 and K_2 are univariate kernel functions, and h_1 and h_2 are bandwidths.

Nonparametric Smoothing

Kernel choice

An order 2 quartic kernel:

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

Bandwidth choice

A penalizing function technique yields asymptotically optimal bandwidths h_1 and h_2 as a starting point.

Principal Components Analysis

For illustration we pick two time series of implied volatility returns from different parts of the smile ($\kappa = 0.90$ and $\kappa = 1.10$) at a fixed one-month maturity:

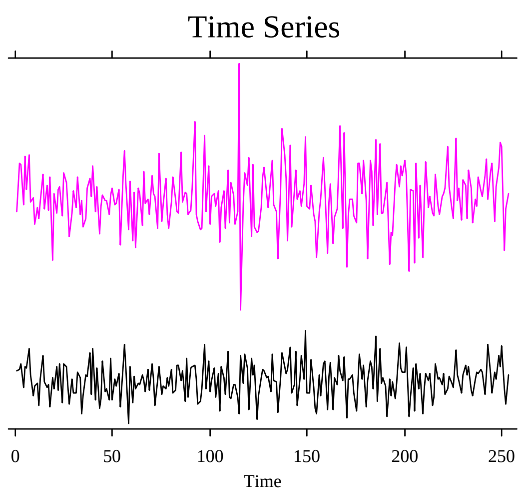


Figure 9: *1 months maturity - moneyness is* $\kappa = 0.90$ *against* $\kappa = 1.10$, *ODAX*

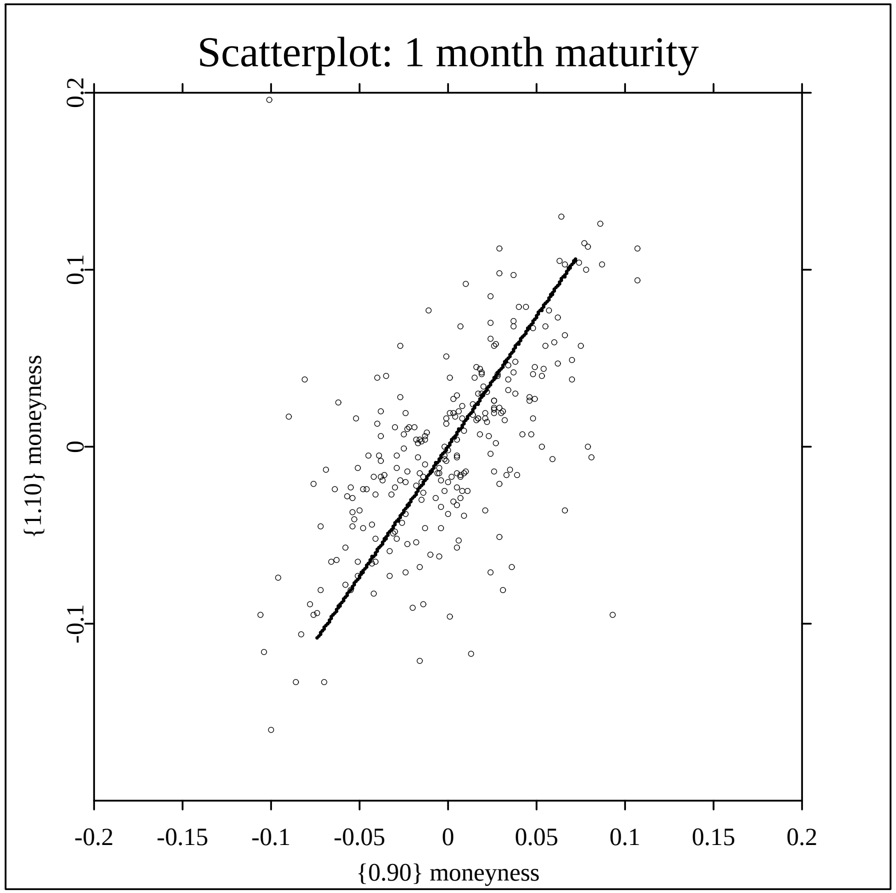


Figure 10: *1 months maturity - moneyness is* $\kappa = 0.90$ *against* $\kappa = 1.10$*, ODAX*

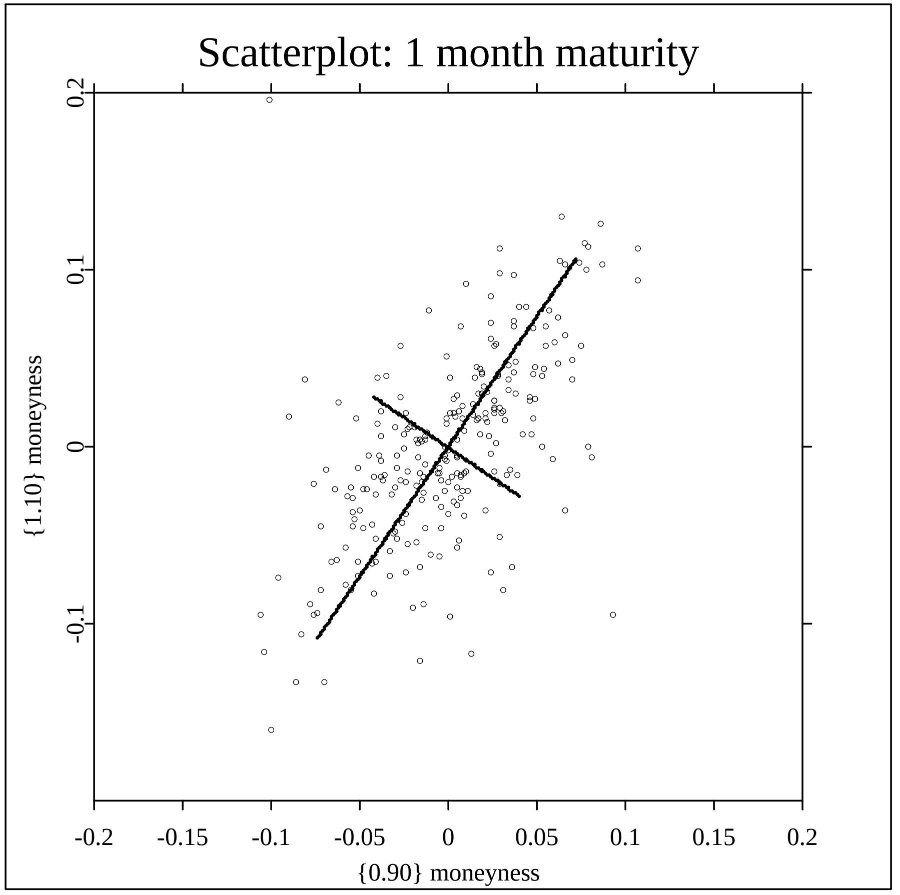
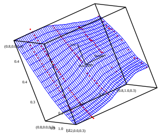


Figure 11: *1 months maturity - moneyness is* $\kappa = 0.90$ *against* $\kappa = 1.10$*, ODAX*



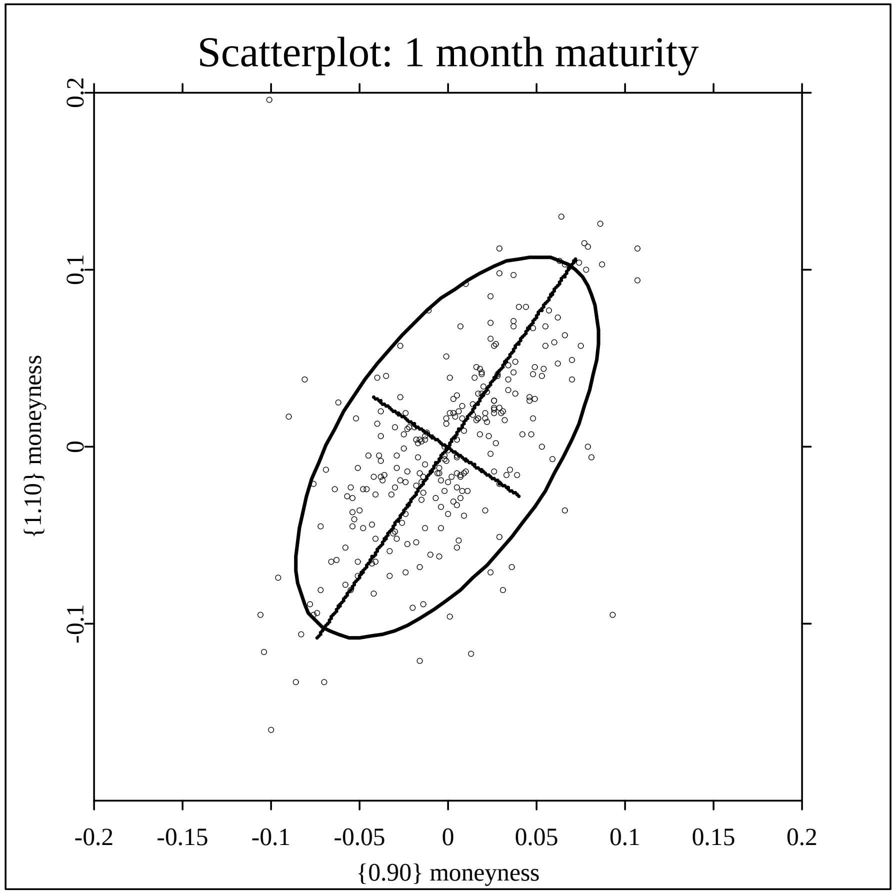
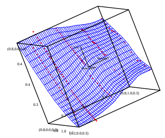


Figure 12: *1 months maturity - moneyness is* $\kappa = 0.90$ *against* $\kappa = 1.10$*, ODAX*



Solution of this dimension reduction problem:

The spectral decomposition of the covariance matrix Ψ, i.e.

$$\Psi = \Gamma \Lambda \Gamma^\top$$

- $\Gamma = (\gamma_1 \vdots \gamma_2 \vdots \cdots \vdots \gamma_p)$ the matrix of eigenvectors. Eigenvectors are principle axes of the hyper-ellipsoid.
- $\Lambda =$ diag$(\lambda_1, \lambda_2, ..., \lambda_p)$ are the eigenvalues. Eigenvalues are the variances of principal components.
- $Y = \Gamma^\top X$ are the principal components.

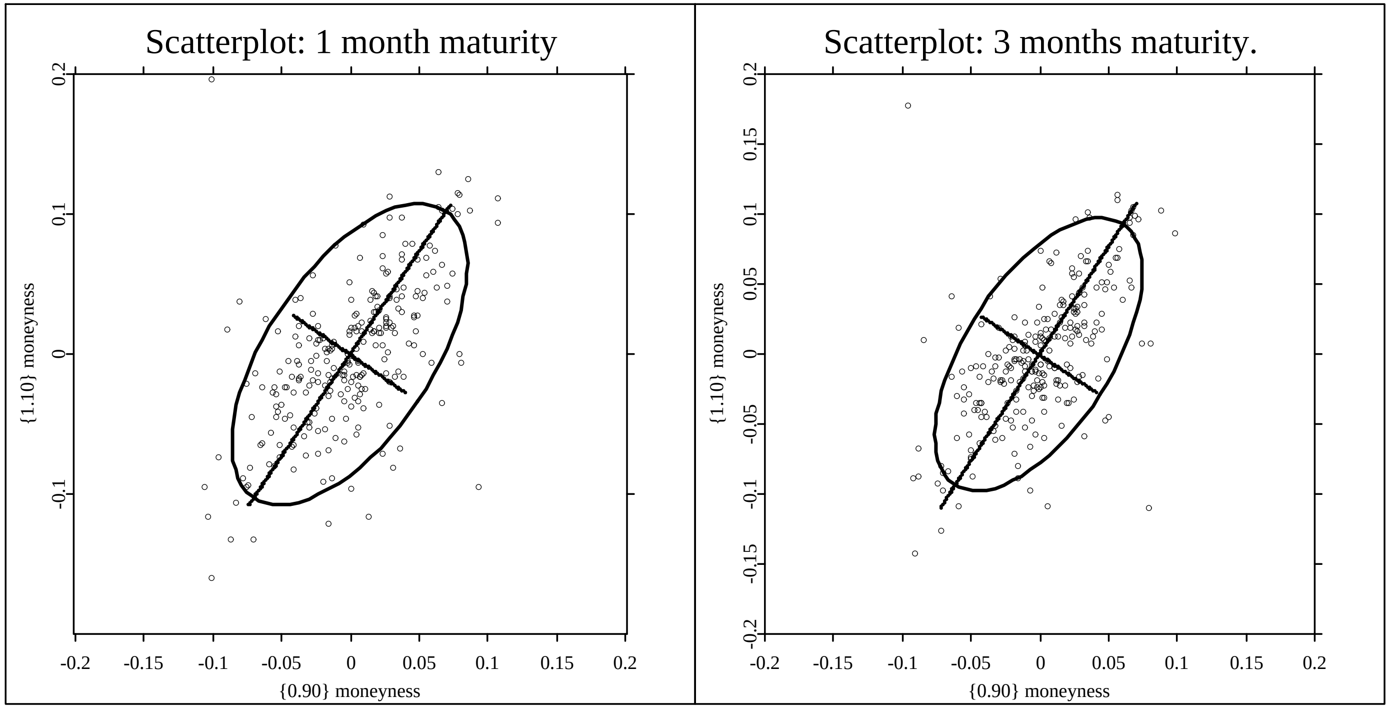


Figure 13: *1 and 3 months maturity - moneyness is $\kappa = 0.90$ against $\kappa = 1.10$, separate PCA; ODAX*

CPCpca.xpl

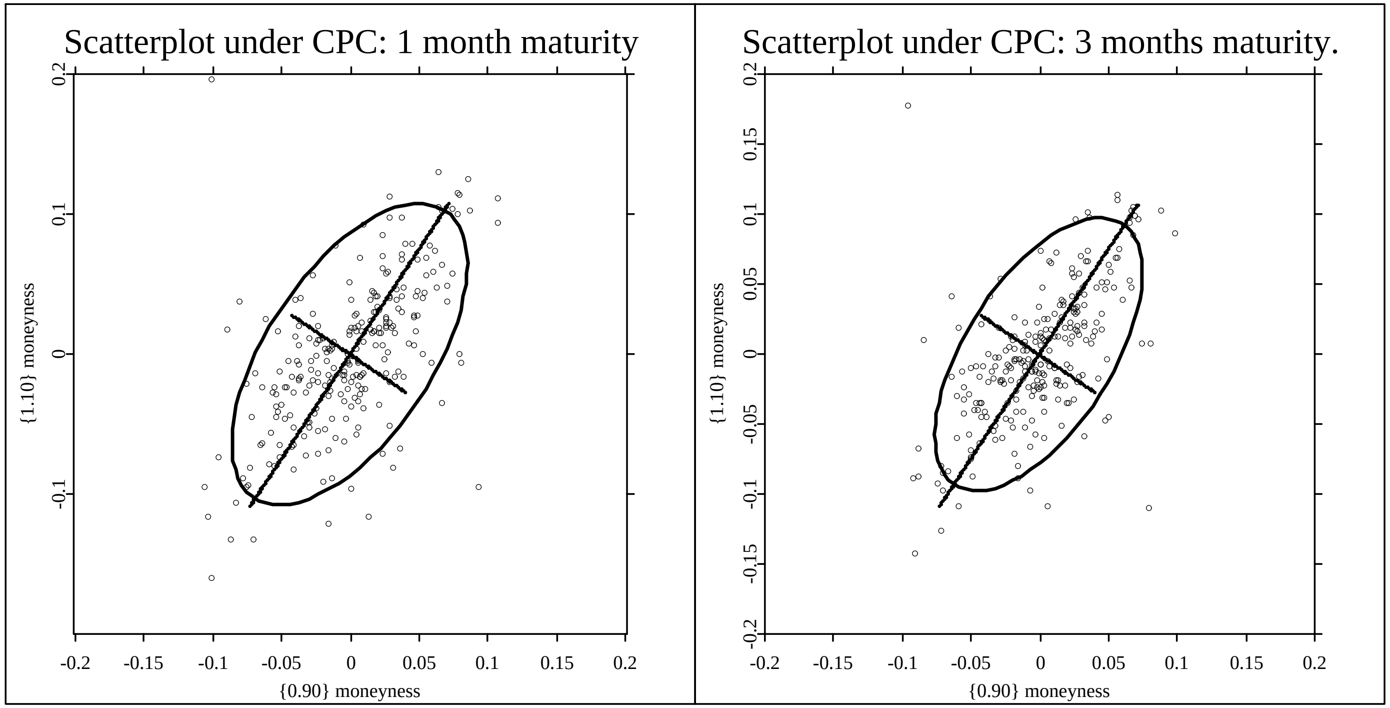


Figure 14: *1 and 3 months maturity - moneyness is* $\kappa = 0.90$ *against* $\kappa = 1.10$, *common PCA, ODAX*

CPCcpc.xpl

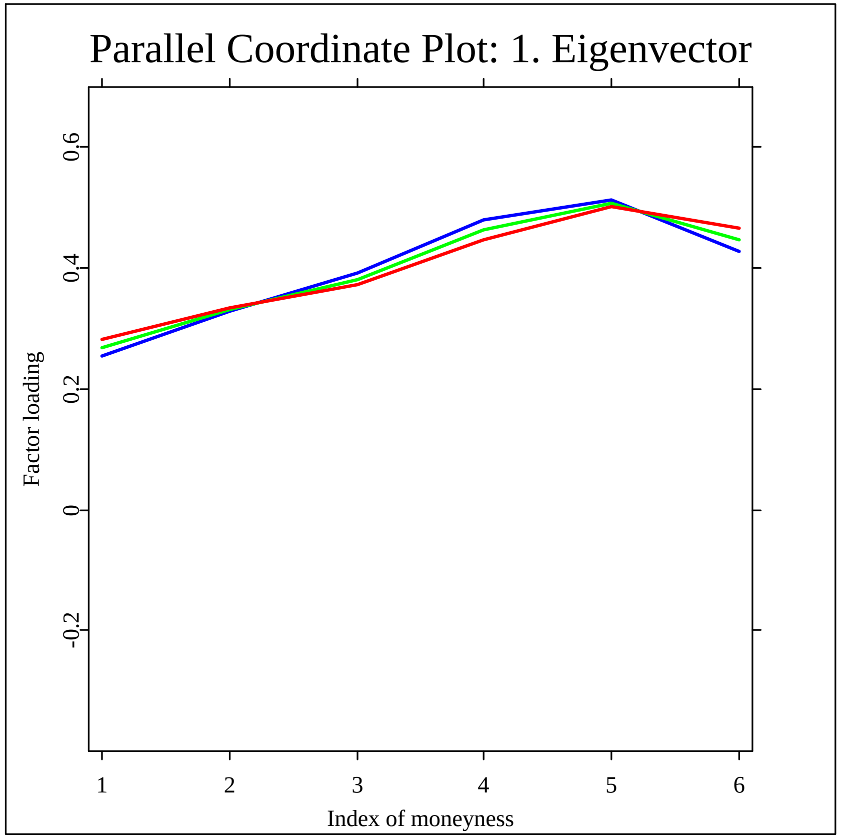


Figure 15: 1^{st} *eigenvectors (sep. PCA) for 1, 2 and 3 months maturity – index 1 to 6 is* $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$

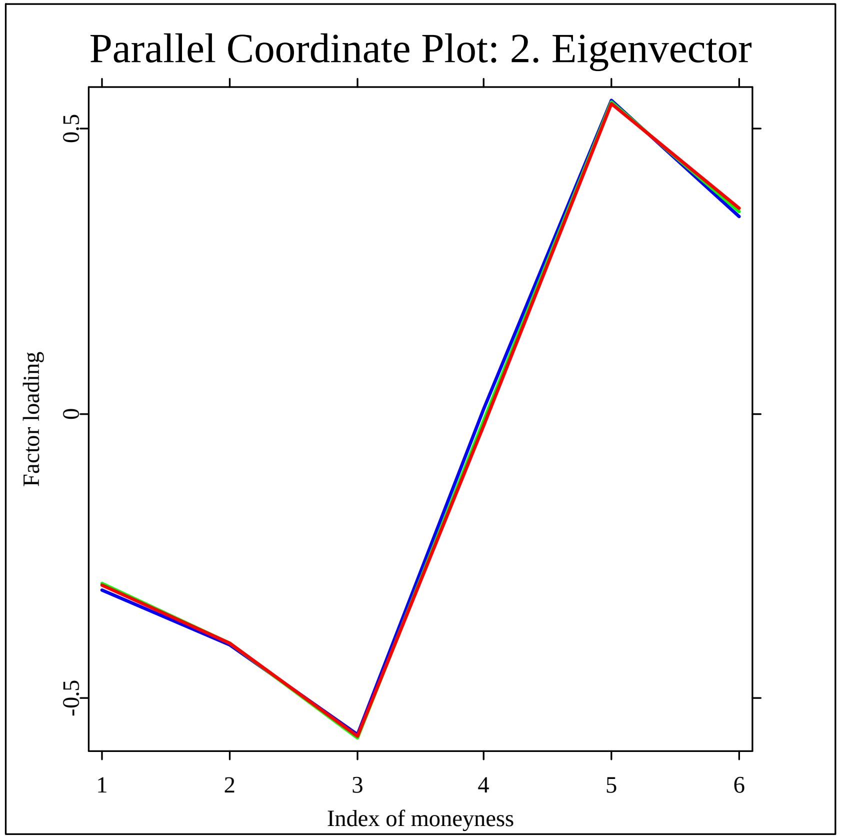


Figure 16: 2^{nd} *eigenvectors (sep. PCA) for 1, 2 and 3 months maturity – index 1 to 6 is* $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$

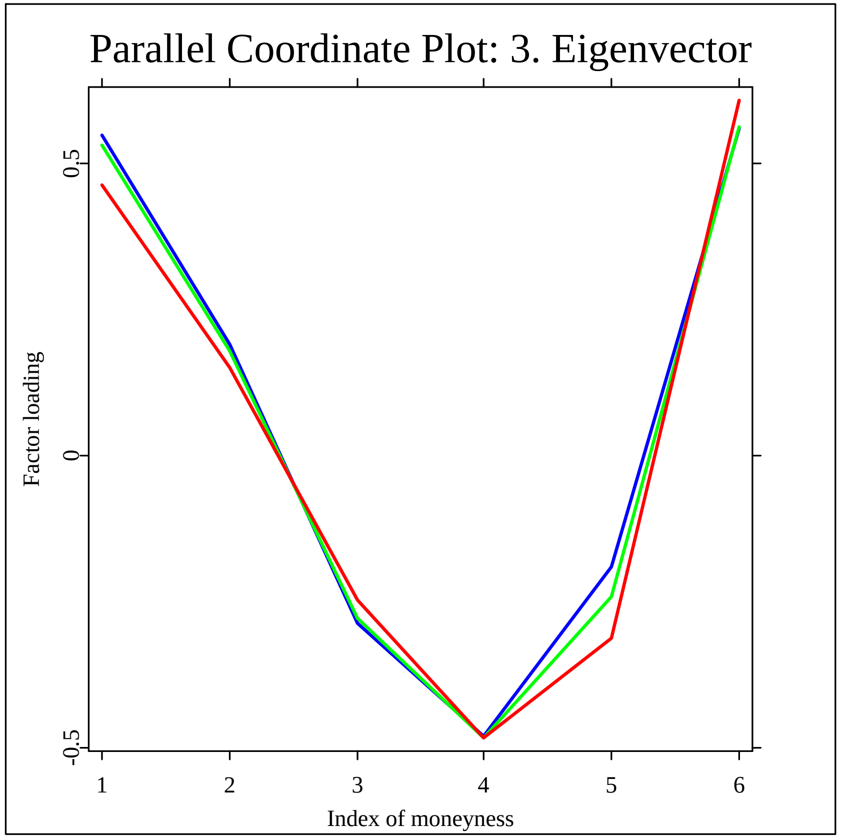


Figure 17: 3^{rd} *eigenvectors (sep. PCA) for 1, 2 and 3 months maturity – index 1 to 6 is* $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$

Common Principle Components

Essential Idea

- As eigenvectors are quite similar across maturity groups, restrict them to be equal.
- As eigenvalues differ between groups, allow them to vary.
- Therefore, estimate principal axes common to all maturity groups, but allow for different variability of principal components.

The CPC-Hypothesis

$$H_{\mathsf{CPC}} : \Psi_i = \Gamma \Lambda_i \Gamma^\top, \qquad i = 1, ..., k. \tag{1}$$

Ψ_i are positive definite $p \times p$ population covariance matrices, Γ is an orthogonal $p \times p$ matrix and $\Lambda_i = \text{diag}(\lambda_{i1}, ..., \lambda_{ip})$.

Let S_i be the (unbiased) sample covariance matrix of implied volatility returns, which are assumed to stem from an underlying p-variate normal distribution $N_p(\mu, \Psi_i)$. Sample size is $n_i (> p)$. Then the distribution of S_i is a generalization of the chi-squared variate, the Wishart distribution (Muirhead, 1982, p.86) with $n_i - 1$ degrees of freedom, denoted by

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

For the k Wishart matrices S_i the likelihood function is

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

where C is a constant not depending on the parameters. The likelihood function has to be maximized under the orthogonality conditions

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

Maximizing the likelihood is equivalent to minimizing the function

$$\begin{aligned} g(\Psi_1, ..., \Psi_k) &= -2\log L + 2\log C \\ &= \sum_{i=1}^{k} (n_i - 1)\Big\{\ln|\Psi_i| + \mathrm{tr}(\Psi_i^{-1} S_i)\Big\}. \end{aligned} \tag{3}$$

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

$$g(\Gamma, \Lambda_1, ..., \Lambda_k) = \sum_{i=1}^{k} (n_i - 1) \sum_{j=1}^{p} \left(\ln \lambda_{ij} + \frac{\gamma_j^\top S_i \gamma_j}{\lambda_{ij}} \right). \tag{4}$$

We impose the orthogonality constraints by the Lagrange method, where μ_j denotes the Lagrange multiplyer of the p constraints $\gamma_j^\top \gamma_j = 1$, and μ_{hj} the Lagrange multiplyer for the $p(p-1)/2$ constraints $\gamma_h^\top \gamma_j = 0 \quad (h \neq j)$. It follows that the function to be minimized is given by

$$g^*(\Gamma, \Lambda_1, ..., \Lambda_k) = g(\cdot) - \sum_{j=1}^{p} \mu_j(\gamma_j^\top \gamma_j - 1) - 2 \sum_{h<j}^{p} \mu_{hj} \gamma_h^\top \gamma_j. \tag{5}$$

ebook W. Härdle, L. Simar(2003): Applied Multivariate Statistical Analysis

Solution

By taking partial derivatives w.r.t. all λ_{im} and γ_m and some manipulations, the solution of the CPC model can be written as the generalized system of characteristic equations

$$\gamma_m^\top \left(\sum_{i=1}^{k} (n_i - 1) \frac{\lambda_{im} - \lambda_{ij}}{\lambda_{im}\lambda_{ij}} S_i \right) \gamma_j = 0, \qquad m, j = 1, ..., p, \quad m \neq j, \tag{6}$$

which needs to be solved using

$$\lambda_{im} = \gamma_m^\top S_i \gamma_m, \qquad i = 1, ..., k, \quad m = 1, ..., p$$

and the constraints

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

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

Partial Common Principle Components

The partial CPC-Hypothesis

For a partial CPC (pCPC) model of order q, the hypothesis is given by

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

where the Ψ_i are positive definite population covariance matrices, and $\Lambda_i = \text{diag}(\lambda_{i1}, ..., \lambda_{ip})$. $\Gamma^{(i)} = (\Gamma_c, \Gamma_s^{(i)})$ are orthogonal $p \times p$ matrices, where Γ_c is $p \times q, \quad q \leq p - 2$ and denotes the matrix of eigenvectors common to all groups, and $\Gamma_s^{(i)}$ the $p \times (p - q)$ matrix of eigenvectors that are specific.

A Hierarchy of Covariance Matrix Structures

Level 1:	Equality	$\Psi_i = \Psi$
Level 2:	Proportionality	$\Psi_i = \rho_i \Psi_1$
Level 3:	CPC	$\Psi_i = \Gamma \Lambda_i \Gamma^\top$
Level 4:	partial CPC(q)	$\Psi_i = \Gamma^{(i)} \Lambda_i \Gamma^{(i)\top}$
Level 5:	Unrelatedness	

Table 1: Possible hypotheses for all $i = 1, ..., k$ groups

Estimation, Selection, Prediction

Our Database

- German DAX Options 1999, daily settlement prices, European style
- Calculate implied volatilities by solving the Black Scholes formula for $\hat{\sigma}$ with observed market prices
- Replace all in-the-money call (put) options by their implicit out-of-the-money put (call)
- Omit prices less than $1/10$ Euro, and maturities less then 10 days
- Smooth the implied volatility surface 1999 nonparametrically (stored in MD*base database containing the volatility surface from 1995-2001, http://www.mdtech.de)

Akaike and Schwarz Information Criteria (AIC, SIC)

The AIC is defined by

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

Assume there are I hierarchically ordered models, with $r_1 < r_i < ... < r_I \quad (i = 1, ..., I)$ parameters in model i.

Define a modified AIC (Flury, 1988) as

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

where L_i is the maximum of the log-likelihood function of model i.

We have

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

such that

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

Define a modified SIC as

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

where $N = \sum_{i=1}^{k} n_i$ (sum of all observations across k groups).

Results: 1, 2, and 3 months maturity

Model						
higher	lower	χ^2	df	p-val	AIC	SIC
Equality	Proport.	237.0	2	0.00	352.0	352.0
Proport	CPC	82.7	10	0.00	118.0	127.7
CPC	pCPC(4)	7.1	2	0.03	55.7	111.3*
pCPC(4)	pCPC(3)	0.2	4	1.00*	52.6*	117.4
pCPC(3)	pCPC(2)	8.1	6	0.23	60.4	143.8
pCPC(2)	pCPC(1)	4.5	8	0.81	64.4	175.2
pCPC(1)	Unrelated	11.9	10	0.29	75.9	223.4
Unrelated					84.0	278.5

CPCFluryShort.xpl

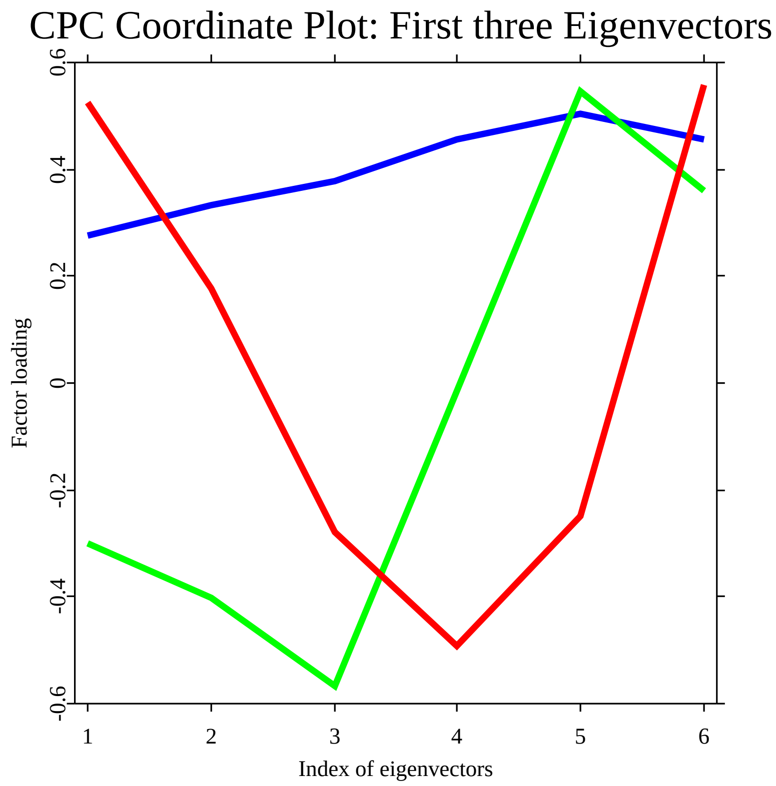


Figure 18: *First three eigenvectors under CPC for 1, 2 and 3 months maturity – index 1 to 6 is* $\kappa \in \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10\}$

`CPCpcpCPC.xpl`

Interpretation of Factor Loadings

- Factor loadings of the first eigenvector have the same sign across moneyness and have almost similar size for each moneyness.

$\Rightarrow$ Linear combination of volatility returns have almost equal weights across moneyness. Hence, the biggest source of shocks are up and down shocks of volatility returns
(Shift-Interpretation).

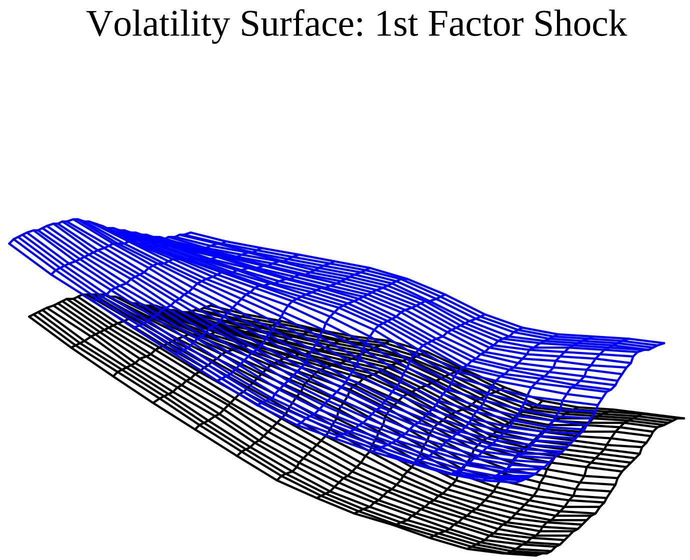


Figure 19: *Simulated Shift Shock: black original, blue after shift shock*

Interpretation of Factor Loadings

- Factor loadings of the second eigenvector have the opposite sign across moneyness, while the weight of ATM options is near to zero.

⇒ Volatility returns enter linear combinations with opposite weights at each end of the smile. Therefore, the second biggest source of shocks affects the slope of volatility returns ((Z-)Slope-Interpretation).

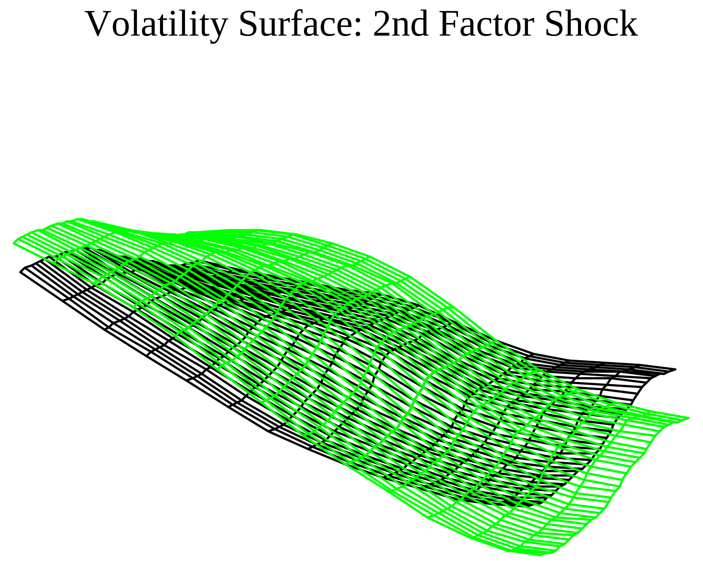


Figure 20: *Simulated Slope Shock: black original, green after slope shock*

Interpretation of Factor Loadings

- Factor loadings of the third eigenvector have the same sign at both ends of the smile and an opposite sign for ATM options.

$\Rightarrow$ Volatility returns enter linear combinations with almost the same weights at each end of the smile, and a large opposite one for ATM options. Hence, the third biggest source of shocks affects the curvature of volatility returns
(Twist-Interpretation).

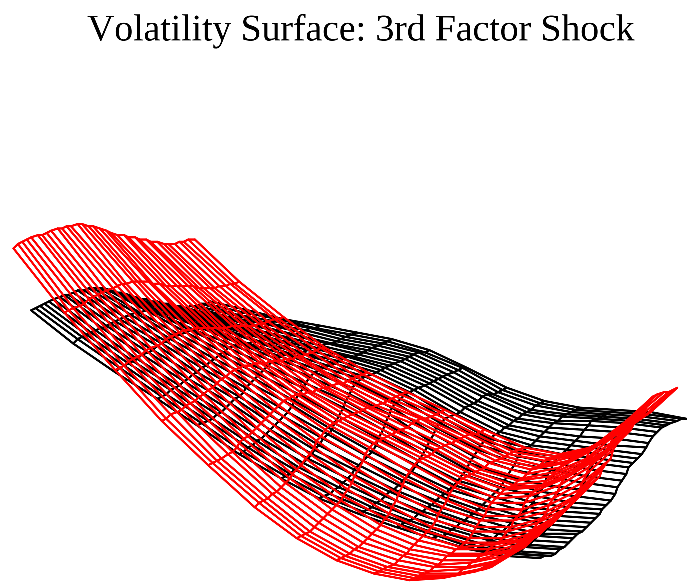


Figure 21: *Simulated Twist Shock: black original, red after twist shock*

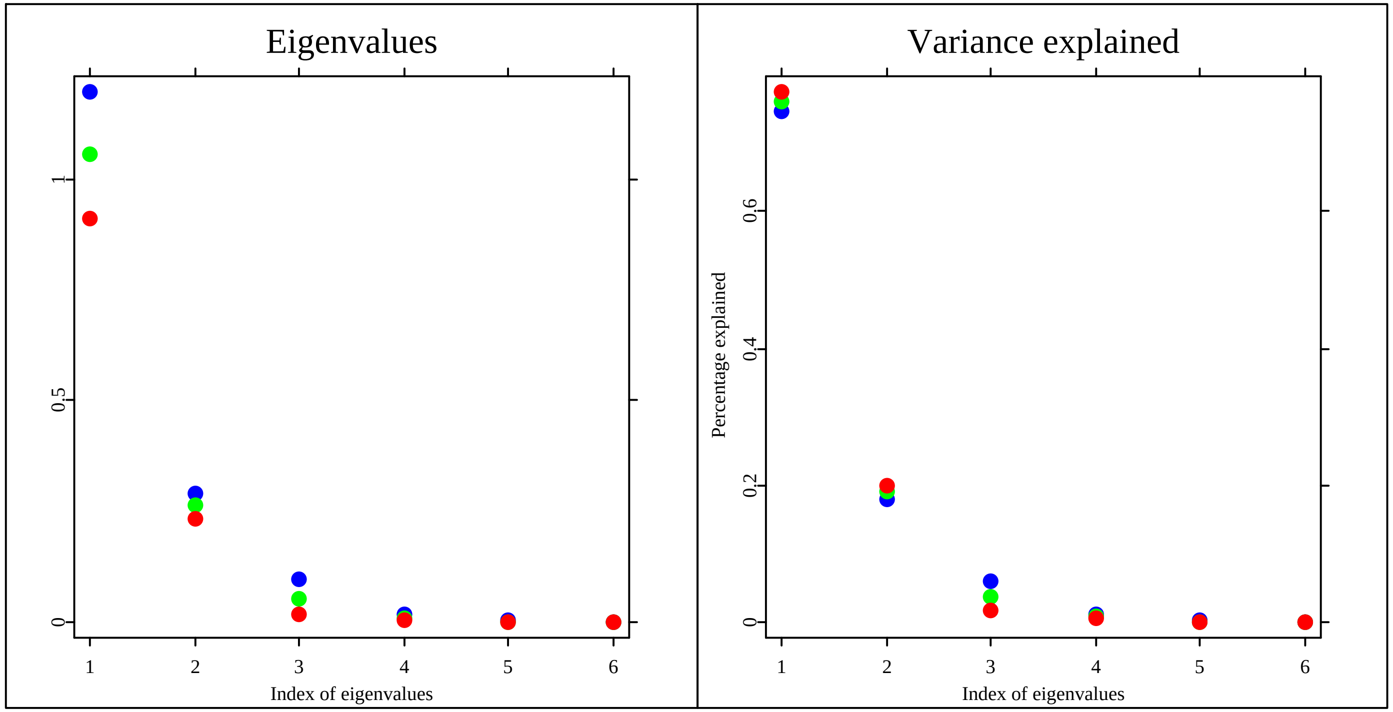


Figure 22: *Eigenvalues and the variance explained as obtained in the CPC model, 1, 2 and 3 months maturity*

CPCpcpCPC.xpl

Results: 6, 9, 12 months maturity

Model						
higher	lower	χ^2	df	p-val	AIC	SIC
Equality	Proport.	250.8	2	0.00	486.0	486.0
Proport	CPC	81.0	10	0.00	239.0	248.5
CPC	pCPC(4)	5.3	2	0.07	178.0	233.8
pCPC(4)	pCPC(3)	4.0	4	0.40	177.0	241.8
pCPC(3)	pCPC(2)	109.5	6	0.00	182.0	264.3
pCPC(2)	pCPC(1)	19.2	8	0.01	89.4	194.6*
pCPC(1)	Unrelated	16.2	10	0.09	83.6*	228.4
Unrelated					84.0	278.5

Interpretation of Factor Loadings

- Eigenvectors exhibit similar patterns as seen for short maturities, hence interpretation stays the same. Only shift factor is common across groups, while factor loadings for the other shocks may differ.
- Between the same principle components of different groups a scaling property is visible.
- The expiry behavior is mostly captured by the third component: observe the regular spikes in the black line of Figure 58.

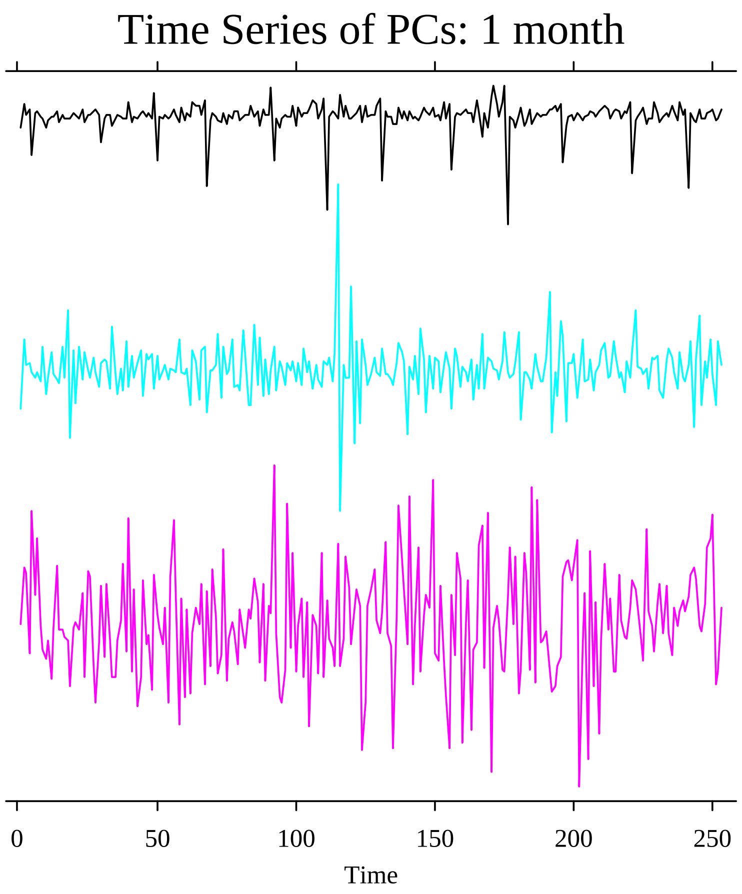
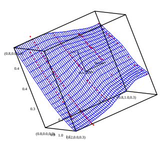


Figure 23: *1st, 2nd and 3rd principal component of the 1 months maturity*



General Statistics of PCs (3 months)

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

Table 2: *Descriptive statistics of principal components (daily); ODAX.*

Summary: General Statistics of PCs

- skewness is close to zero for the three PCs
- evidence for excess kurtosis especially in the second and third PC
- evidence for 'leverage effect': correlation with the returns of underlying is around -0.5 for the first component. When there is a negative shock in the market value of the firm, (implied) volatility rises, since the shock results into an increase of the debt-equity ratio
- negligible correlation with underlying in the second the third component

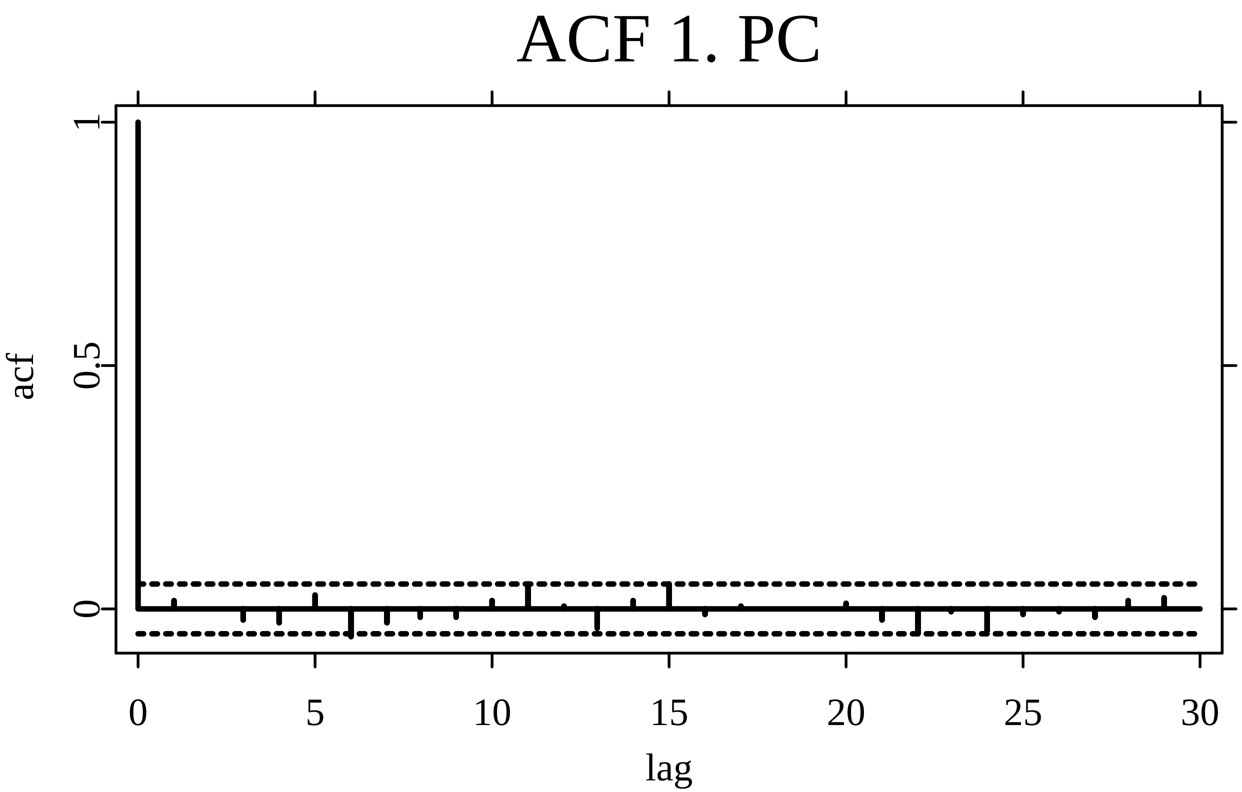


Figure 24: *Autocorrelation function of the 1. PC.*

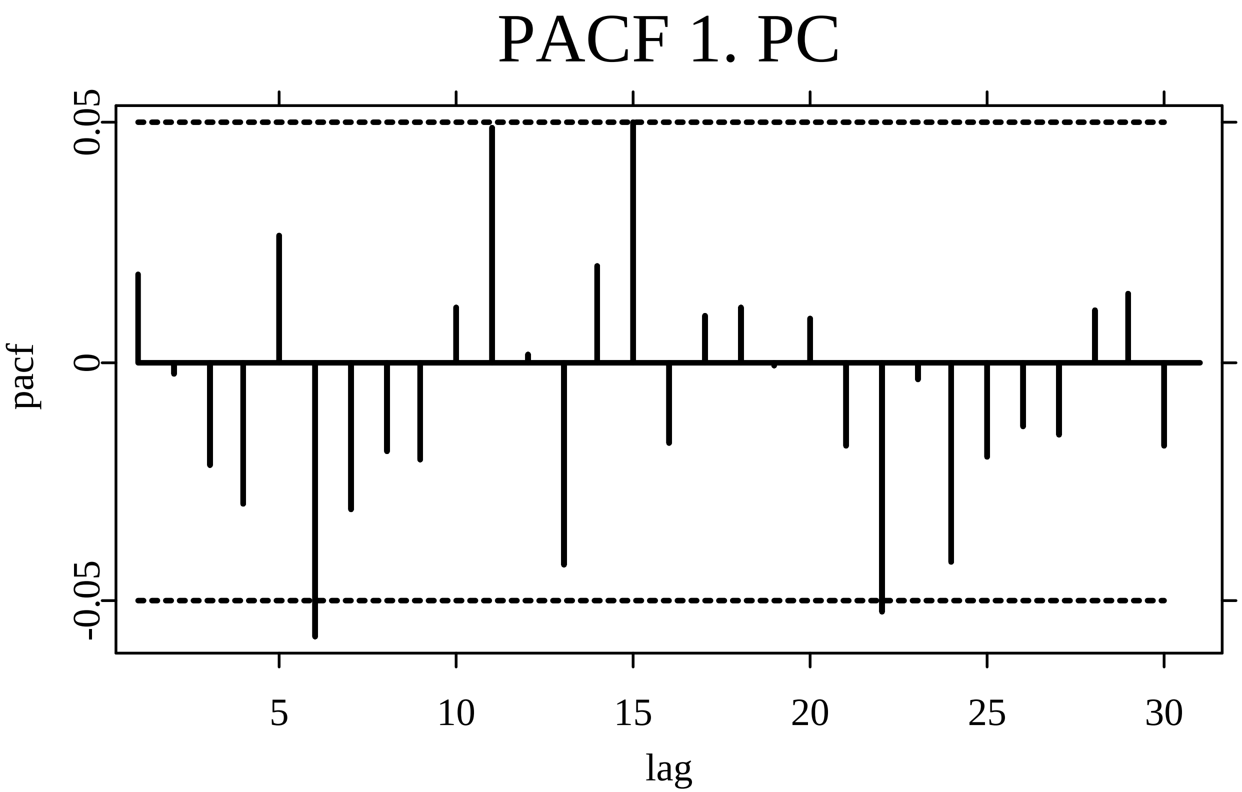


Figure 25: *Partial autocorrelation function of the 1. PC.*

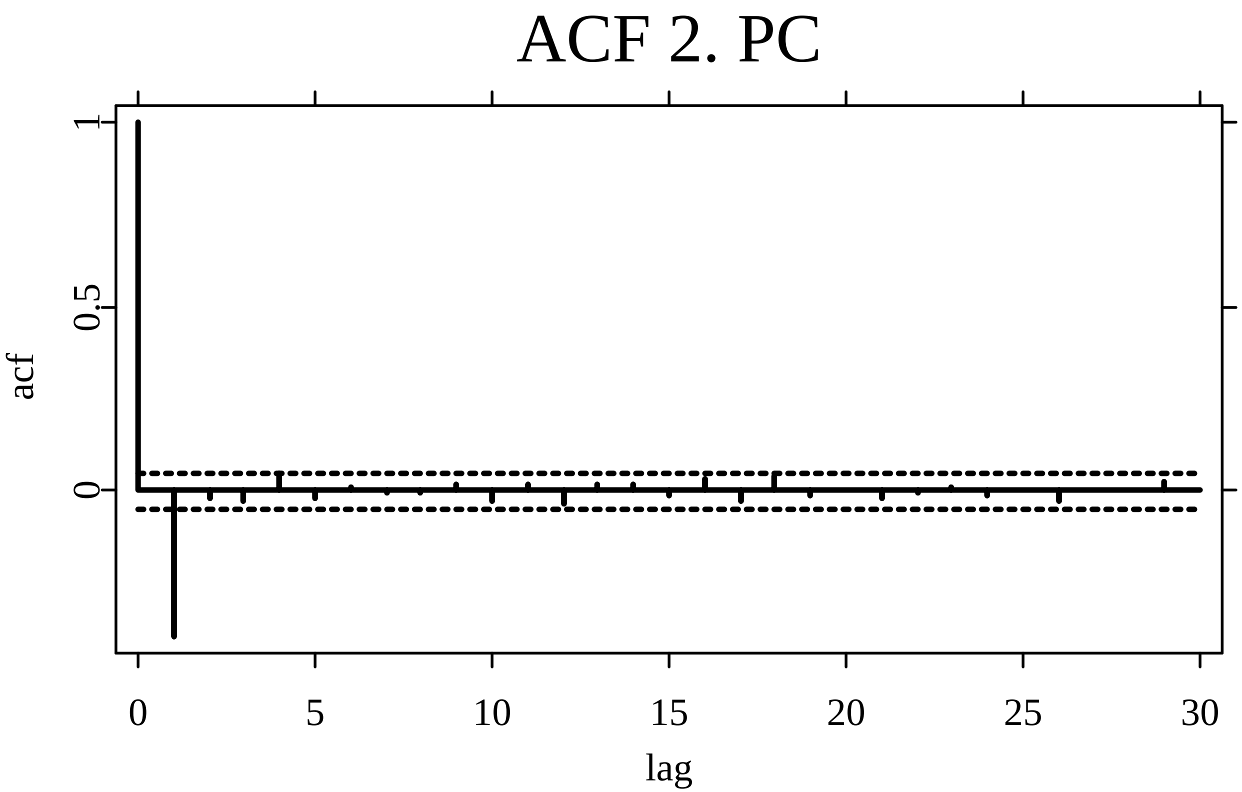


Figure 26: *Autocorrelation function of the 2. PC.*

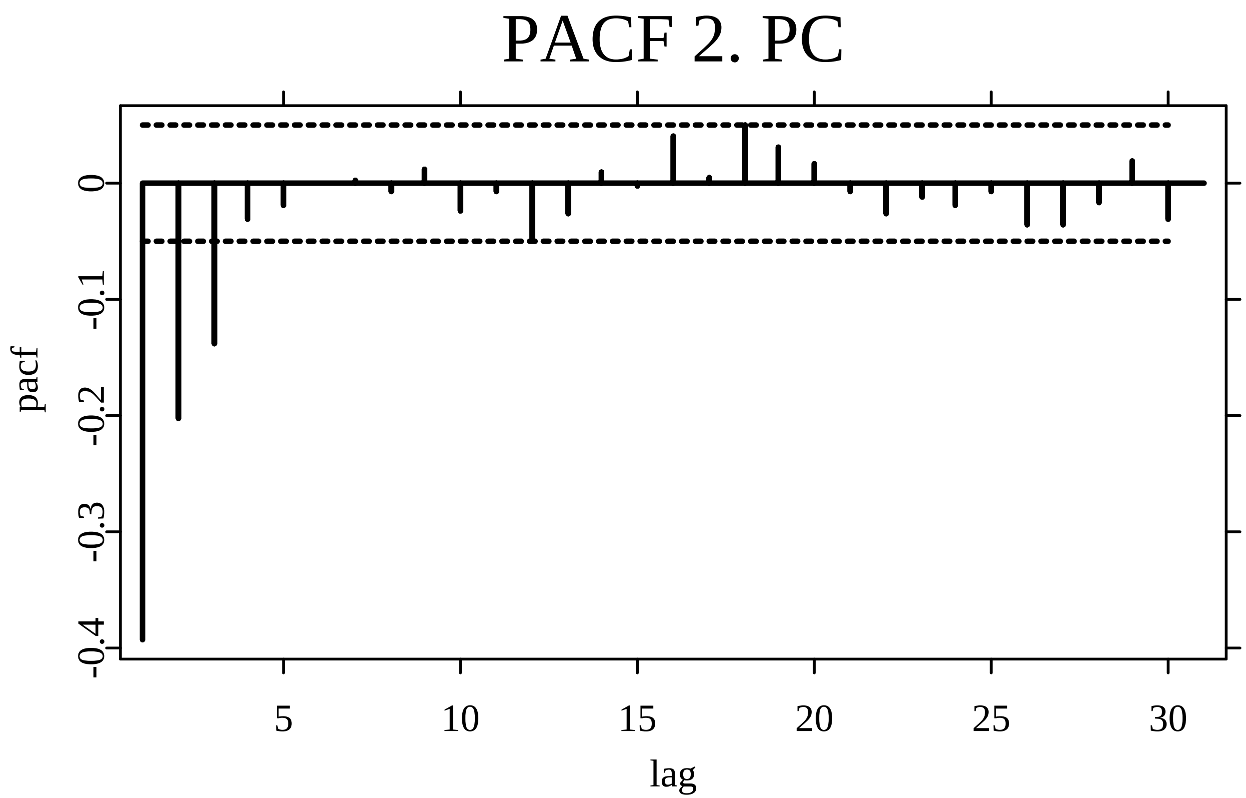


Figure 27: *Partial autocorrelation function of the 2. PC.*

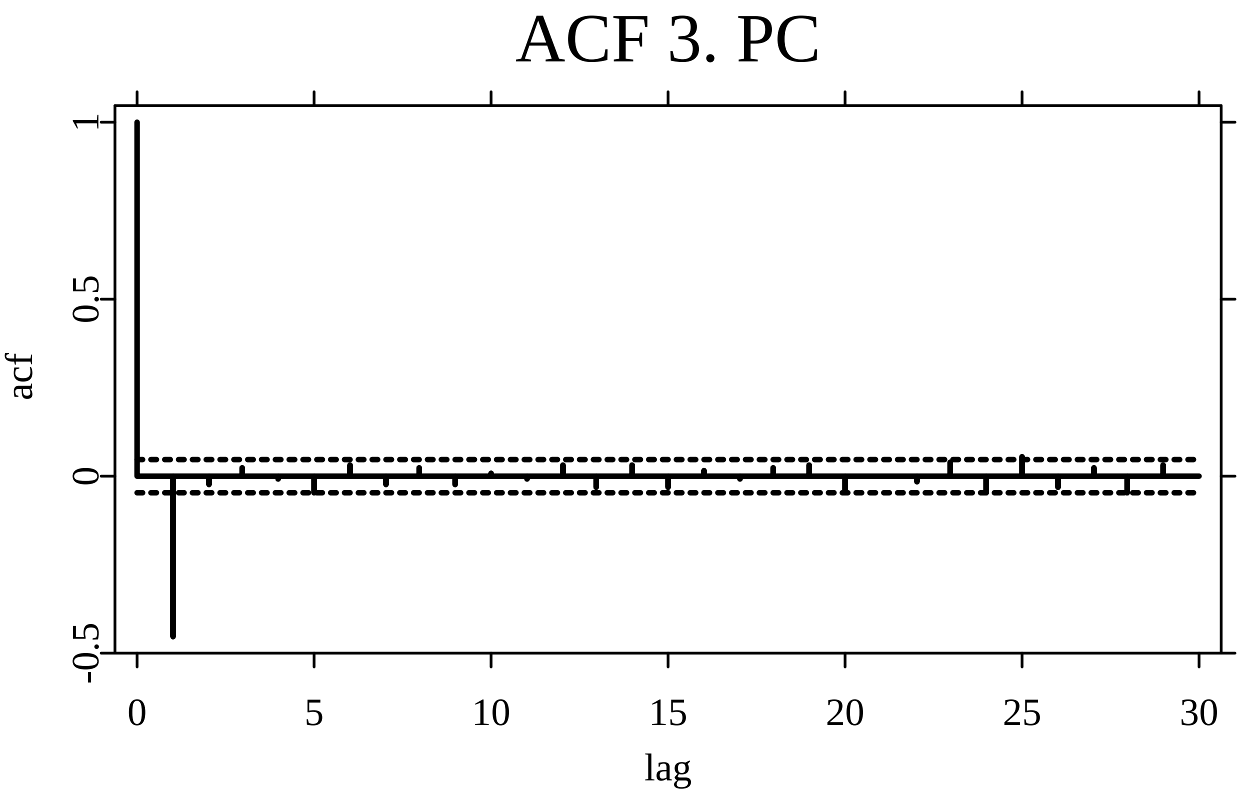


Figure 28: *Autocorrelation function of the 3. PC.*

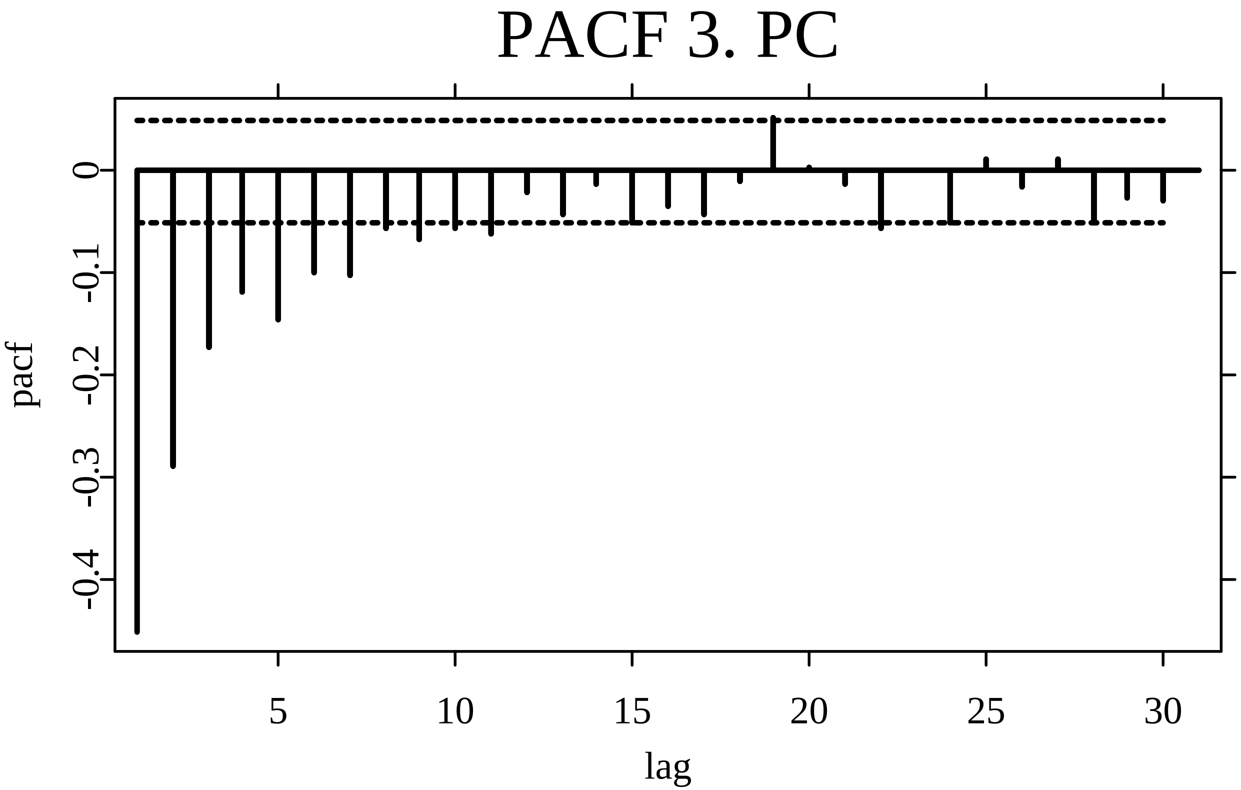
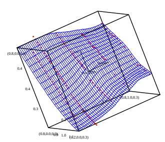


Figure 29: *Partial autocorrelation function of the 3. PC.*



From the autocorrelation and partial autocorrelation function we propose MA(q)-GARCH(r, s) models:

$q = 0 \quad r = 1, 2 \quad s = 1, 2$ for y_{1t} (1. PC), and

$q = 1 \quad r = 1, 2 \quad s = 1, 2$ for y_{2t} (2. PC) and y_{3t} (3. PC)

$$
\begin{aligned}
y_{it} &= c + a_1 z_t + \varepsilon_{it} + b_1 \varepsilon_{i,t-1}, \qquad (7) \\
&\quad \varepsilon_{it} \sim \mathcal{N}(0, \sigma_{it}^2), \\
\sigma_{it}^2 &= \omega + \sum_{j=1}^{k} \alpha_j \sigma_{i,t-j} + \sum_{j=1}^{s} \beta_j \varepsilon_{i,t-j}^2 + \gamma z_t^2, \qquad (8)
\end{aligned}
$$

where z_t denotes log returns in the DAX index.

We conduct AIC-SIC searches over a large variety of models:

- For y_{1t} both AIC and SIC suggest an GARCH(1,2) specification.
- For y_{2t} and y_{3t}, a MA(1)-GARCH(1,1) is preferred.

cond. mean	Factor		
	y_{1t}	y_{2t}	y_{3t}
c	0.001	$1.9E^{-4}$	-$3.8E^{-05}$
	(0.407)	(1.170)	(-0.592)
a_1	-2.920	0.086	0.005
	(-24.46)	(4.860)	(0.457)
b_1		-0.733	-0.733
		(-35.50)	(-35.50)

Table 3: *Mean equation: estimation results of GARCH models for the three principal components, t-statistics in parenthesis.*

cond. var.	Factor		
	y_{1t}	y_{2t}	y_{3t}
ω	$1.4E^{-4}$	$6.7E^{-5}$	$1.7E^{-05}$
	(3.945)	(7.515)	(8.687)
α_1	0.803	0.425	0.686
	(32.09)	(6.774)	(24.41)
β_1	0.246	0.200	0.147
	(7.112)	(6.840)	(8.027)
β_2	-0.130		
	(-4.110)		
γ	1.480		
	(4.991)		
$\bar{R}^2$	0.23	0.22	0.33

Table 4: *Variance equation: estimation results of GARCH models for the three principal components, t-statistics in parenthesis.*

Summary: Model estimates of 1. PC

- mean equation: index returns have a highly significant impact on 1. PC
- sign of a_1 in line with the 'leverage effect' hypothesis
- variance equation: $\beta_2 < 0$ may be interpreted as an 'over-reaction correction' in terms of variance: High two-period lagged returns have a dampening impact on variance
- volatility increases also when volatility in the underlying is high ($\gamma > 0$)
- adjusted $\bar{R}^2$ around 23% – however: this is due to index returns: leaving z_t out of the mean equations reduces $\bar{R}^2$ to around 0.2%, only

Summary: Model estimates of 2. and 3. PC

- mean equations of y_{2t} and y_{3t}: MA(1) components are negative and significant
- index returns are only significant for y_{2t} and positively influence the slope structure in the surface.
- positive shocks in the underlying reduce implied volatility levels, while at the same time the slope of the surface is intensified

Checking for model robustness

Model robustness is essential for trading strategies or risk computations.

Two directions of robustness analysis:

1. Choice of data: Settlement prices may be artificially quoted by the exchange. Do we only recover the model of the EUREX?
2. Choice of time period: Is CPC a particular feature of the year 1999?

Perform a CPC analysis for data from 1995 to May 2001 separately in each year, using tick data (= contract data) of puts, calls, and futures observed on the EUREX.

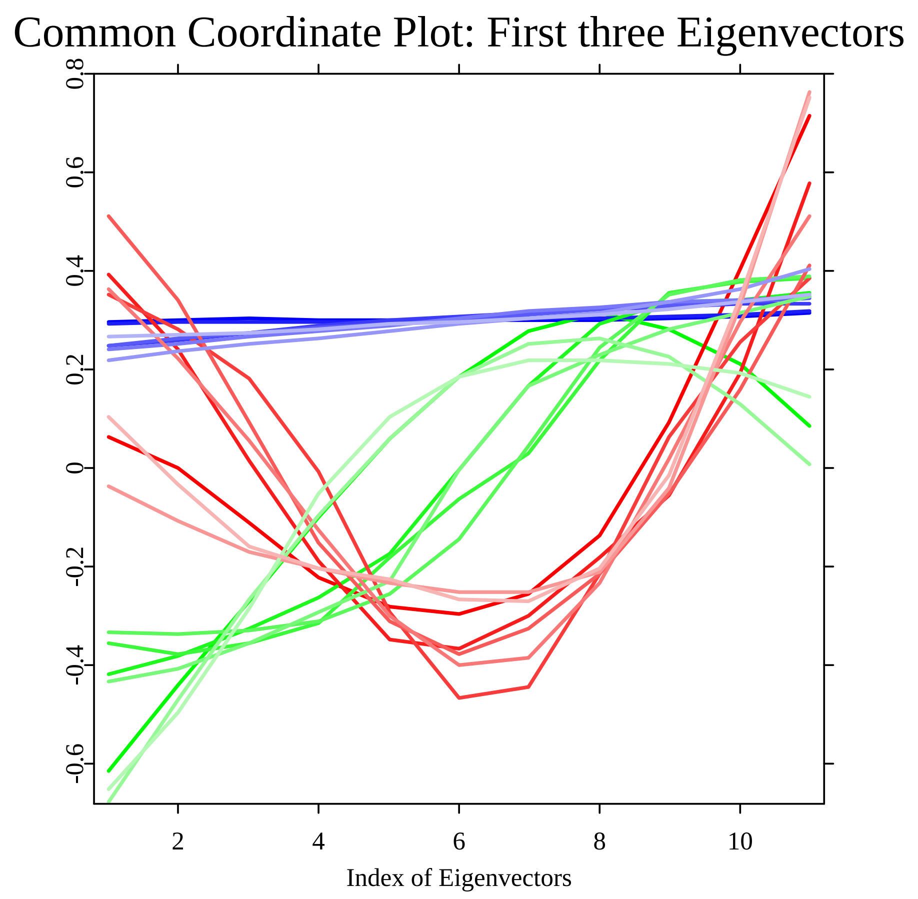


Figure 30: *First, second, and third CPC eigenvectors through the years 1995 to May 2001; increasing color intensity with more recent data, ODAX, EUREX.*

Results on robustness:

- CPC holds in each year from 1995 to 2001,
- settlement data inherits tick data characteristics,
- shift, slope and twist interpretations remain valid,
- shift component is subject to little time variability,
- slope and twist factors changes slowly over time, and not completely in a non-systematic manner,
- time to maturity component is still captured in the third component.

Tests of time homogeneity of eigenvectors across different sub-samples indicate that it can be necessary to re-estimate the model regularly.

Volas: What did we learn?

- CPC is the preferred modeling strategy for implied volatility returns
- Factor loadings have a natural interpretation (shift, slope, twist)
- CPC yields the desired dimension reduction of the implied volatility surface

Trading Strategies, Risk Management

Values: CPC and State Price Density Dynamics

To find the price H_t of an option take the discounted expected value of the pay-off function with respect to a risk-neutral pricing measure $f^*(S_T, S_t, \tau)$:

$$H_t = e^{-r\tau}\mathbb{E}[\psi(S_T, K, \tau)|\mathcal{F}_t] = e^{-r\tau}\int_0^\infty \psi(S_T, K, \tau) f^* dS_T \ ,$$

where ψ is the payoff function, e.g. $\psi = \max(S_T - K, 0)$ in case of the call.

$f^*(S_T, S_t, \tau)$ is also called (implied) State Price Density (SPD).

$f^*(S_T, S_t, \tau)$ can be obtained by taking the second derivative of the option price function $H(S_t, K, r, \tau)$ w.r.t. K:

$$f^*(S_T, S_t, \tau) = e^{r\tau} \frac{\partial^2 H_t}{\partial K^2}_{|K=S_T},$$

when time to maturity τ, the current underlying asset price $S_t = S$ are fixed, Breeden and Litzenberger (1987).

This derivative can be expressed in terms of moneyness $M = S/K$ and first and second derivative of the implied volatility surface $\sigma(M)$, $\sigma^{'}(M)$, $\sigma^{''}(M)$ only, Rookley (1997).

Adopt the following procedure:

1. Set up a $q < p$ factor model for the smile at maturity τ_i

$$\hat{\sigma}_t(\kappa, \tau_i) = \hat{\sigma}_0(\kappa, \tau_i) + \sum_{j=1}^{q} y_{it} \gamma_j^\top ,$$

   where PCs are modeled as a function in lagged values and exogeneous variables Z as $y_{it} = F(y_{t-1}, y_{t-2}, \ldots ; Z)$, e.g.

$$\Delta y_{it} = \beta(\bar{y}_i - y_{it-1}) + \varepsilon_t ,$$

   where $\bar{y}$ is a long run mean.

2. Other maturity groups are obtained by an appropriate scaling factor $c(\tau_i)$
3. From the smile estimate $\sigma^{'}(M)$, $\sigma^{''}(M)$ by a local polynomial method
4. Obtain $f^*(S_T, S_t, \tau) = e^{r\tau} \frac{\partial^2 H_t}{\partial K^2}_{|K=S_T}$
5. Generate trading signals.

An Example of SPD Estimates

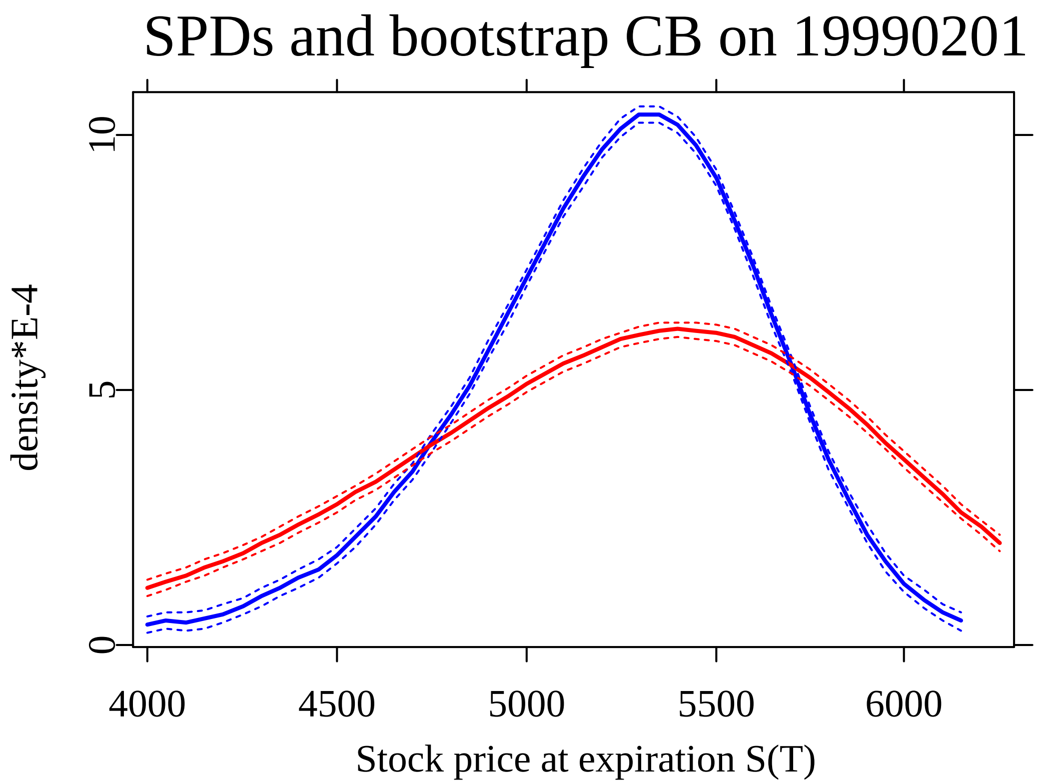


Figure 31: *SPD estimates from ODAX 1999 data by Rookley's method:* $\tau = 1$ *month and* $\tau = 2$ *months; solid line: density, dashed lines: bootstrap confidence intervals*

Based on this procedure certain trading strategies in options are possible, e.g. skewness and kurtosis trades Aït-Sahalia et al. (2001), Blaskowitz (2001), Härdle and Zheng (2001). They are based on the following idea:

From option data we can extract an **implied** SPD f^*, based on a cross section of options. However, there is also the **historical** SPD g^* given by underlying asset's time series data.

Estimation of historical SPD

Suppose S_t follows the diffusion process

$$dS_t \quad = \quad \mu(S_t)dt + \sigma(S_t)dW_t$$

Consider now the conditional density g^* generated by the dynamics

$$dS_t^* \quad = \quad (r_{t,\tau} - \delta_{t,\tau})S_t^* dt + \sigma(S_t^*)dW_t^* \ .$$

The transformation from W_t to W_t^*, and S_t to S_t^* is an application of Girsanov's Theorem. W^* is a Brownian Motion under the risk neutral measure, r the interest rate and δ the dividend yield.

Idea: Compare g^* and f^*.

Estimation of the diffusion function

Florens–Zmirou's (1993) nonparametric estimator for σ (time scale is [0,1] for expository convenience):

$$\hat{\sigma}_{FZ}(S) \quad = \quad \frac{\sum_{i=1}^{N^*-1} K_{FZ}\left(\frac{S_i - S}{h_{FZ}}\right) N^* \{S_{(i+1)/N^*} - S_{i/N^*}\}^2}{\sum_{i=1}^{N^*} K_{FZ}\left(\frac{S_i - S}{h_{FZ}}\right)},$$

where K_{FZ} is a kernel function, h_{FZ} a bandwidth parameter, and N^* the number of observed index values.

$\hat{\sigma}_{FZ}$ is an unbiased estimator of σ and does not impose any restrictions on the drift.

Computation of historical SPD g^*

We use a Monte–Carlo simulation with a Milstein scheme given by

$$\begin{aligned} S_i &= S_{i-1} + rS_{i-1}\Delta t + \sigma(S_{i-1})\Delta W_i + \\ &\quad \frac{1}{2}\sigma(S_{i-1})\frac{\partial \sigma}{\partial S}(S_{i-1})\{(\Delta W_{i-1})^2 - \Delta t\}, \end{aligned}$$

where ΔW_i is the increment of a Wiener Process, Δt time between two grid points. The drift is set to r and $\frac{\partial \sigma}{\partial S}$ is approximated by $\frac{\Delta \sigma}{\Delta S}$.

SPD g^* may be now obtained by means of a nonparametric kernel density estimation

$$g^*(S) \quad = \quad \frac{\hat{p_t}^*\{\log(S/S_t)\}}{S}$$

where

$$\hat{p}_t^*(u) = \frac{1}{Mh} \sum_{m=1}^{M} K\Big(\frac{u_m - u}{h}\Big),$$

$u = \log(S/S_t)$ returns and M is the number of simulated Monte Carlo paths.

g^* is $\sqrt{N^*}$– consistent.

Suppose one knew f^* and g^*. Are there profitable trading strategies to exploit differences in f^* and g^*? Consider the following situation:

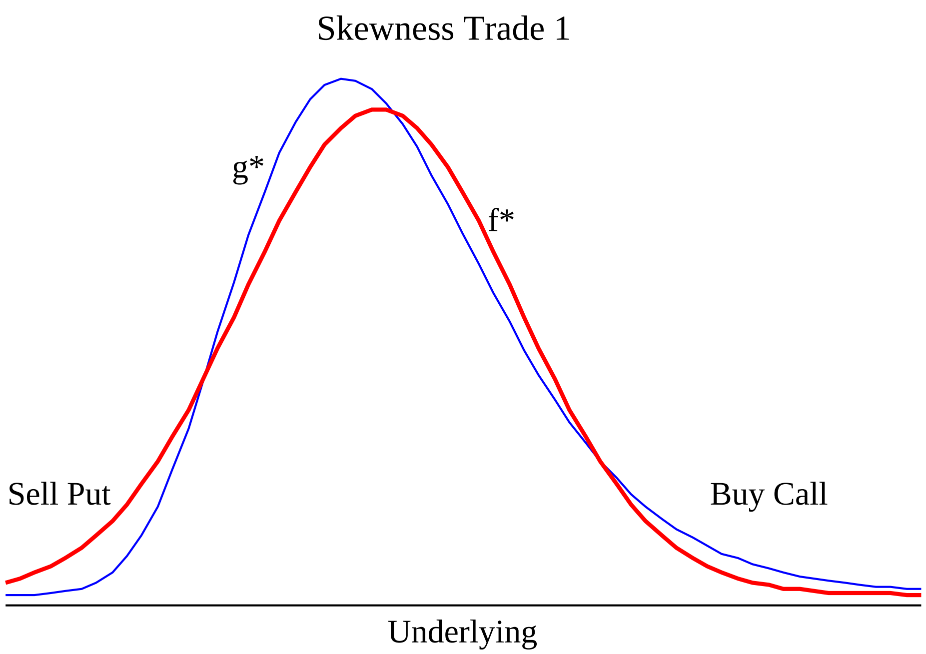


Figure 32: *Skewness trade*

This may be exploited by the following strategies:

Skewness Trade 1	Skewness Trade 2
skew(f^*) $<$ skew(g^*)	skew(f^*) $>$ skew(g^*)
Sell OTM Puts	Buy OTM Puts
Buy OTM Calls	Sell OTM Calls

Similarly, kurtosis trades depending on the discrepancies between the two densities f^* and g^* can be developed.

Historical simulations show that positive net cash flows may be generated by these kinds of strategies, Aït-Sahalia et al. (2001), Blaskowitz (2001).

However, risk adjusted performance measurement needs to be done and a fine tuning of trading signals remains to be developed.

Values: CPC and Maximum Loss Analysis

The parsimony of the CPC model may also be exploited in the context of Maximum Loss analysis of vega-sensitive, delta-gamma-neutral portfolios (e.g. Fengler, Härdle, Schmidt, 2002).

Consider a Taylor series expansion of a portfolio P_t built out of N options:

$$\begin{aligned}\Delta P_t \quad &\approx \quad \sum_{i=1}^{N} \left(\frac{\partial H_{it}}{\partial \sigma_{it}} \Delta \sigma_{it}(\kappa, \tau) \right. \\ &+ \quad \left. \frac{\partial H_{it}}{\partial t} \Delta t + \frac{\partial H_{it}}{\partial r_t} \Delta r_t + \frac{\partial H_{it}}{\partial S_t} \Delta S_t + \frac{1}{2} \frac{\partial^2 H_{it}}{\partial S_t^2} (\Delta S_t)^2 \right)\end{aligned}$$

If the portfolio is delta-gamma neutral and if rho and theta-risks can be neglected due to their negligible size, the expression reduces to

$$\Delta P_t \approx \sum_{i=1}^{N} \frac{\partial H_{it}}{\partial \sigma_{it}} \Delta \sigma_{it}(\kappa, \tau)$$

The CPC model allows us to write the returns of the implied volatilities $\hat{\sigma}_t(\kappa, \tau)$ as a linear combination of PCs. Thus, taking the respective nearby fixed grid point of the volatility surface $\hat{\sigma}_t(\kappa_i, \tau_j)$ as a proxy for $\hat{\sigma}_{it}(\kappa, \tau)$, one gets:

$$\Delta P_t \approx \sum_{i=1}^{N} \frac{\partial H_{it}}{\partial \sigma_{it}} \left(\sum_k \gamma_{jk} y_{kt} \right) \hat{\sigma}_{i,t-1}(\kappa, \tau)$$

Definition of Maximum Loss

Maximum loss (ML) is defined as the maximum possible loss

- over a given risk factor space $A_{\tilde{\tau}}$, where $A_{\tilde{\tau}}$ will be assumed a closed set with confidence level $P(y|y \in A_{\tilde{\tau}}) = \alpha$
- for some holding period $\tilde{\tau}$.

In contrast to Value at Risk which requires the profit and loss distribution to be known, ML is directly defined in the risk factor space, Studer (1995).

Constructing $A_{\tilde{\mathcal{T}}}$

Assuming multi-normally distributed PCs, i.e. the y obey the joint density function

$$\varphi(y) = \frac{1}{\sqrt{2\pi|\Lambda_i|}} \exp\left(-\frac{1}{2} y^\top \Lambda_i^{-1} y\right),$$

where Λ_i is diagonal matrix of eigenvalues of group i, construction of the trust region $A_{\tilde{\mathcal{T}}}$ is straightforward:

$y^\top \Lambda_i^{-1} y$ is chi-square distributed with q degrees of freedom, where q is the number of factors retained for modeling.

Trust region $A_{\tilde{\tau}}$ is the ellipse given by

$$A_{\tilde{\tau}} = (y | y^\top \Lambda_i^{-1} y \leq c_\alpha),$$

where c_α denotes the α-quantile of a chi-squared distribution with p degrees of freedom.

Fengler, Härdle and Schmidt (2002) consider a simple straddle portfolio over a horizon of one day, where an ATM straddle of short maturities is sold and an ATM straddle of long maturities is bought.

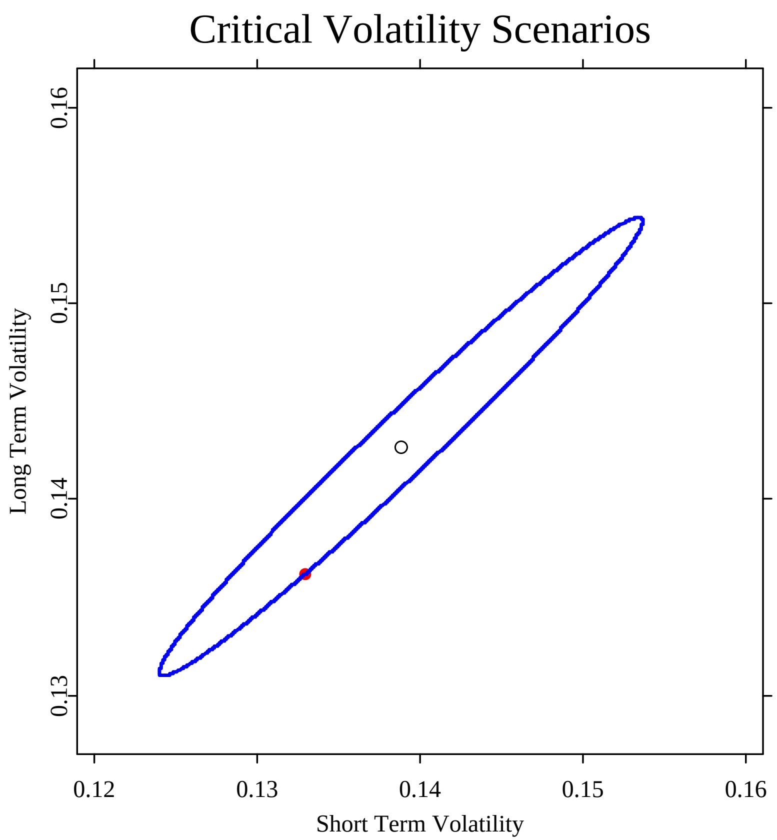


Figure 33: *Critical volatility scenarios for an straddle portfolio on 29/03/96; black circle current level, red circle ML scenario; two factors modeled at $\alpha = 99\%$.*

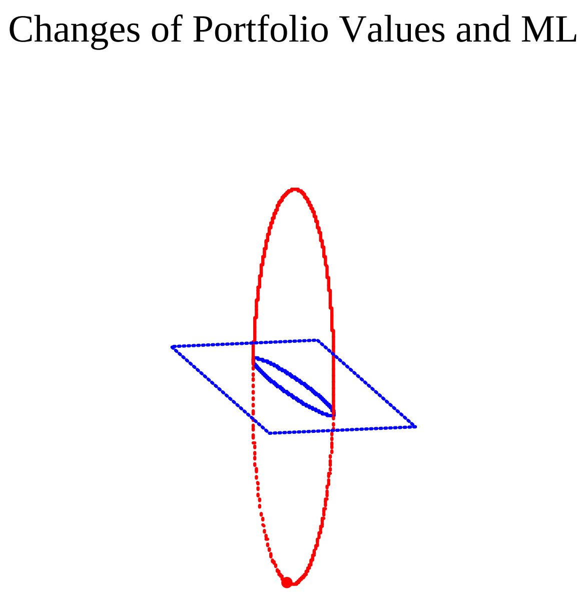


Figure 34: *Critical volatility scenarios for a straddle portfolio on 29/03/96 (blue) portfolio changes (red, gains solid, losses dashed) and ML (red ball); two factors modeled at* $\alpha = 99\%$.

Fengler, Härdle and Schmidt (2002) argue that

- the procedure can be a convenient guideline tool for daily risk management analysis at trading desks
- the procedure is capable to identify critical volatility scenarios for the portfolio under consideration, even during the Asian crisis 1997
- although the true confidence level of the modelling approach remains unknown, the procedure performs empirically better than is suggested by the number of retained factors
- adaptive methods, notably in the context of *Common Principle Components Analysis* need to be developed to enhance predictability of the the model.

Volas: What did we learn?

CPC

- faciliates a high dimensional modeling task by working in a low dimensional manifold,
- factor loadings and common PC factors have natural interpretations in Finance,
- due its generality it is widely applicable in other contexts.

Voles, Volas, Values: What did we learn?

Biology and Finance are cross-pollinated by Statistics!

References

[1] Airoldi, J.P. and B. Flury (1988) "An application of common principal component analysis to cranial morphometry of Microtus californicus and M. ochrogaster (Mammalia, Rodentia)." *J. Zool., London* 216, 21-36.

[2] Aït-Sahalia Y. and A. Lo (1998) "Nonparametric estimation of state-price densities implicit in financial asset prices." *Journal of Finance* 53, 499-548.

[3] Aït-Sahalia Y. and A. Lo (2000) "Nonparametric risk management and implied risk-aversion." *Journal of Econometrics* 94, 9-51.

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

[5] Barle, S. and Cakici, N. (1998), "How to Grow a Smiling Tree" *The Journal of Financial Engineering*, 7, 127–146.

[6] Black F. (1976), "Studies of stock price volatility changes." *Proceedings of the American Statistical Association*, 177-181.

[7] Blaskowitz, O. (2001), "Trading on Deviations of Implied from Empirical Distributions", Diplomarbeit, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin.

[8] Breeden and Litzenberger (1987), "Prices of state-contingent claims implicit in option prices." *Journal of Business*, 51, 621–651.

[9] Derman E. and I. Kani (1997), "Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility." Quantitative Strategies Technical Notes, Goldman Sachs.

[10] Dupire, B. (1994), "Pricing with a Smile", *Risk* 7(1), pp. 18-20.

[11] Fengler, M., Härdle, W. and C. Villa (2001), "The Dynamics of Implied Volatilities: A Common Principle Component Approach", Discussion Paper 38/2001, SfB 373, Humboldt-Universität zu Berlin.

[12] Fengler, M., Härdle, W. and P. Schmidt (2002), "Common Factors Governing VDAX Movements and the Maximum Loss", forthcoming in: Journal of Financial Markets and Portfolio Management.

[13] Florens-Zmirou, D. (1993), "On Estimating the Diffusion Coefficient from Discrete Observations", *Journal of Applied Probability* 30, pp. 790–804.

[14] Flury, B. (1988), "Common Principal Components and Related Multivariate Models". Wiley, New York.

[15] Flury, B. and Gautschi, W. (1986), "An Algorithm for Simultaneous Orthogonal Transformations of Several Positive Definite Matrices to

nearly Diagonal Form", *SIAM Journal on Scientific and Statistical Computing*, 7, 169-184.

[16] Härdle, W. and J. Zheng (2001), "How Precise are Price Distributions Predicted by Implied Binomial Trees?", mimeo.

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

[18] Jarrow R.A. and O'Hara (1989), "Primes and scores: an essay on market imperfections". *Journal of Finance* 44, 1265-1287.

[19] Rookley, C. (1997), "Fully Exploiting the Information Content of Intra Day Option Quotes: Applications in Option Pricing and Risk Management". University of Arizona, mimeo.

[20] Rubinstein M. (1985), "Nonparametric tests of alternative option-pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through

August 31, 1978." *Journal of Finance* 40, 455-480.

[21] Schönbucher P.J.(1999), " A Market Model for Stochastic Implied Volatility ", *Working papers*, Department of Statistics, Bonn University.

[22] Studer, G. (1995), "Value at Risk and Maximum Loss Optimization". RiskLab, Technical Report, Zürich.

Pictures of voles have been taken from
http://www.nsrl.ttu.edu/tmot1/microchr.htm
and http://www.enature.com/