Implied Volatility Surface Modeling for KOSPI 200 option and ODAX with DSFM

A Master Thesis Presented

by

Ji Cao

(507100)

to

Prof. Dr. Wolfgang Härdle

CASE - Center of Applied Statistics and Economics

Humboldt-Universität zu Berlin



in partial fulfillment of the requirements

for the degree of

Master of Science

Berlin, April 13, 2008

Declaration of Authorship

I hereby confirm that I have authored this master thesis independently and without use of others than the indicated resources. All passages, which are literally or in general matter taken out of publications or other resources, are marked as such.

Ji Cao

Berlin, April 13, 2008

Acknowledgement

I would like to thank Professor Dr. Wolfgang Härdle for giving me the opportunity and motivation to write this thesis.

I'm especially indebted to Szymon Borak for his excellent guidance all the time.

Furthermore I'm also grateful to my parents, without their support it would be impossible to finish this work.

Abstract

Implied volatility is one of the important topics in financial markets. Due to option data's characteristics, estimating implied volatility is a challenging task for both academia and industry. Dynamic Semiparametric Factor Model (DSFM) is method to model high-dimensional data with dynamic context. It employs semiparametric factor functions and time-varying loadings. One of its application is implied volatility surface (IVS) modeling. This master thesis applys DSFM to estimate IVS of Korean Stock Index (KOSPI 200) options and ODAX. Estimation result is discussed and a comparison between two markets from view of DSFM is studied.

Keywords: implied volatility surface, dynamic semiparametric factor model, option pricing

Contents

| 1 | Intro | oduction | 9 |
|---|-------|--|----|
| 2 | Blac | k-Scholes formula and volatility | 12 |
| | 2.1 | Black-Scholes formula | 12 |
| | 2.2 | Volatility | 13 |
| 3 | Dyn | amic semiparametric factor model | 16 |
| | 3.1 | Estimation | 16 |
| | 3.2 | Orthogonalization | 18 |
| | 3.3 | Model selection | 20 |
| 4 | IVS | modeling for KOSPI 200 option and ODAX | 21 |
| | 4.1 | Data description | 21 |
| | 4.2 | Close implied volatility and interpolation | 24 |
| | 4.3 | Model fit | 29 |
| | 4.4 | Time series of factors loadings | 33 |
| 5 | Con | clusion | 37 |

List of Figures

| 1.1 | Upper left: implied volatility of ODAX on May 02, 2003. (moneyness lower left axis, time to maturity lower right axis). Upper right: Data design of ODAX on May 02, 2003. Lower left: implied volatility of ODAX on June 02, 2003. Lower right: Data design of ODAX on June | |
|-----|--|----|
| | 02, 2003 | 10 |
| 4.1 | Kernel density of KOSPI 200 option (left) and ODAX (right) on Septem- | |
| | ber 1, 2003 (moneyness lower left axis, time to maturity lower right axis), bandwidths for moneyness (in year) and time to maturity are | |
| | both 0.05 | 23 |
| 4.2 | Implied volatility of last traded ODAX with 27 days to maturity on | |
| | September 22, 2003. Left panel: original data, before interpolation. | |
| | Call options are crosses, put options are circles. Right panel: with | |
| | linear interpolation and taking average. | 28 |
| 4.3 | Left panel: DSFM for KOSPI 200 option on September 9, 2003 with | |
| | bandwidths (0.02, 0.2). Right panel: DSFM for ODAX on September | |
| | 9, 2003 with bandwidths (0.02, 0.04). (moneyness lower left axis, time | |
| | to maturity lower right axis) | 30 |
| 4.4 | Factor \hat{m}_1 for KOSPI 200 option (left) and for ODAX (right). (monev- | |
| | ness lower left axis, time to maturity lower right axis) | 31 |
| 4.5 | Factor \hat{m}_2 for KOSPI 200 option (left) and for ODAX (right). (monev- | |
| - | ness lower left axis, time to maturity lower right axis) | 31 |
| 4.6 | Factor \hat{m}_3 for KOSPI 200 option (left) and for ODAX (right). (money- | - |
| | ness lower left axis, time to maturity lower right axis) | 32 |
| 4.7 | Time series of factor loadings β_1 , β_2 , β_2 for KOSPI 200 option (blue) | - |
| 1.1 | and ODAX (red) in year 2003 | 33 |
| | | 00 |

| List | of | Figures |
|------|----|---------|
|------|----|---------|

| 4.8 | $\hat{\beta}_1$ (blue) and underlying price (red, logarithmized and centralized) of | |
|-----|---|----|
| | KOSPI 200 option (left) and ODAX (right) | 34 |
| 4.9 | $\hat{\beta}_1$ (blue) and negative underlying price (red, logarithmized and central- | |
| | ized) of KOSPI 200 option (left) and ODAX (right). | 35 |

List of Tables

| 4.1 | Descriptive statistics on KOSPI 200 option and ODAX in September | |
|-----|---|----|
| | 2003, time to maturity in days. \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots | 22 |
| 4.2 | Descriptive statistics on last traded KOSPI 200 option and ODAX in | |
| | September 2003, time to maturity in days | 25 |
| 4.3 | DSFM modeling result for both markets | 29 |
| 4.4 | Explained variance for the model size. | |
| | *Bandwidth (0.02, 0.22). | |
| | **Bandwidth (0.02, 0.28) | 29 |
| 4.5 | Descriptive statistics for factors loading $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ | 33 |
| 4.6 | ADF test statistics for factor loading series. Tests include constant. | |
| | Critical values are -3.46 (1%), -2.87 (5%) and -2.57 (10%). | 36 |
| 4.7 | Two cointegration tests for KOSPI $\hat{\beta}_1$ and ODAX $\hat{\beta}_1$ | 36 |

1 Introduction

As one of the most important parameters in financial markets, implied volatility has attracted a lot of attentions. In derivatives pricing and hedging, volatility is a key issue. For example, implied volatility is usually derived from quoted prices for certain derivatives on an underlying and then used to price other derivatives on the same underlying - perhaps options that are not actively traded or for which prices are otherwise not readily available. In delta-hedging, implied volatility can be used to calculate different deltas, when underlying price changes.

A feature of implied volatility is that it is not constant and can be affected by underlying price S_t or by option's time to maturity τ . Implied volatility is also dynamic - it changes with time. Figure 1.1 displays implied volatility of ODAX on May 2 and June 2, 2003. Implied volatility changes with moneyness (we define moneyness as $\kappa \stackrel{\text{def}}{=} \frac{K}{F_t}$, where K is strike, F_t is future price of underlying.) and time to maturity. When time to maturity is small, implied volatilities display curvature like smile, while when time to maturity increases, implied volatilities become flat as straight line. Moreover, implied volatility is about 1.12, largest implied volatility on June 2 is about 2. June 2 tends to have more observations. There are three different expiration times on May 2 but eight on June 2. In order to price and hedge options with certain strike and expiration which doesn't appear in the observed implied volatility strings, a entire implied volatility surface should be estimated. Furthermore, due to its change with time, the estimation of implied volatility surface has to be dynamic.

Dynamic semiparametric factor model is designed to capture this feature of IVS with low-dimensional factor functions and time-varying factor loadings. IVS is approximated by unknown factor functions moving in a finite dimensional function space. The dynamics can be understood by using vector autoregression (VAR) techniques on the time-varying loadings. The finite dimensional fits are obtained in the local





Figure 1.1: Upper left: implied volatility of ODAX on May 02, 2003. (moneyness lower left axis, time to maturity lower right axis). Upper right: Data design of ODAX on May 02, 2003. Lower left: implied volatility of ODAX on June 02, 2003. Lower right: Data design of ODAX on June 02, 2003.

neighborhood of strikes and maturities, for which implied volatilities are recorded on the specific day. Surface estimation and dimension reduction is achieved in one single

step. This technology can be seen as a combination of functional principal component analysis, nonparametric curve estimation and backfitting for additive models.

$$Y_{i,j} = m_0(X_{i,j}) + \sum_{l=1}^{L} \beta_{i,l} m_l(X_{i,j})$$
(1.1)

In DSFM (1.1), *i* is index of day (i = 1, ..., I). *j* is index of options on day *i* (j = 1, ..., J). $Y_{i,j}$ is log-implied volatility for option *j* on day *i*. m_l is time-invariant factor function. $X_{i,j}$ is independent variables, a 2-dimensional vector $X_{i,j} = (\kappa_{i,j}, \tau_{i,j})$ with $\kappa_{i,j}$ the moneyness and $\tau_{i,j}$ time to maturity. $\beta_{i,l}$ is loading for factor m_l . It changes with day index *i*. After estimating unknown factor function m_l and corresponding loading $\beta_{i,l}$, implied volatility can be calculated given desired point $(\kappa_{i,j}, \tau_{i,j})$.

This master thesis is organized as follows: Section 2 reviews Black-Scholes formula and other concepts related with implied volatility. Section 3 introduces DSFM in detail. An empirical study with DSFM on German and Korean option markets is applied in Section 4. Section 5 is conclusion.

2 Black-Scholes formula and volatility

2.1 Black-Scholes formula

The publication of option pricing formula by Fischer Black and Myron Scholes in 1973 was a great step in finance. Since then option pricing model has developed into a standard tool for pricing and hedging derivatives. This theory requires an perfect market assumption:

- 1. Risk-free interest rates for borrowing and lending cash are the same and constant.
- 2. There is no transaction cost.
- 3. There is no taxes.
- 4. Short selling is allowed.
- 5. All securities are perfectly divisible.
- 6. There exists no arbitrage opportunity.

It is also assumed that the underlying price S_t follows a geometric Brownian motion and the trading is continuous:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{2.1}$$

where W_t is standard wiener process.

There are six parameters in Black-Scholes formula (2.2) for pricing vanilla European call options in an perfect market:

- 1. Spot price S_t
- 2. Strike price K
- 3. Time to maturity τ

- 4. Risk-free interest rate \boldsymbol{r}
- 5. Dividend rate d
- 6. Volatility σ .

The first three parameters are given when an option is quoted in the market. The risk-free interest rate and dividend rate are always known or can be easily estimated. Only the volatility is unknown and it is the most important parameter in Black-Scholes formula.

$$C_t^{BS}(S_t, K, \tau, r, d, \sigma) = e^{d\tau} S_t \Phi(y + \sigma \sqrt{\tau}) - e^{-r\tau} K \Phi(y), \qquad (2.2)$$

with

$$y = \frac{\ln \frac{S}{K} + (b - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}.$$

 Φ is the standard normal cumulative distribution function.

2.2 Volatility

Volatility measures uncertainty, or dispersion about the return provided by the underlying - it is simply a measure of the degree of price movement in a stock, futures contract or any other market. If underlying follows Brownian motion, volatility increases with square root of times as the time increase.

There are different calculations of volatility:

Historical Volatility Historical volatility estimate σ based on the variability of the underlying in the past. It takes the standard deviation of the underlying's log-returns and times the time length.

$$R_t = \ln \frac{S_t}{S_{t-1}},$$

is the log-return of underlying asset. Its sample variance

$$\hat{v} = \frac{1}{n-1} \sum_{t=1}^{n} (R_t - \bar{R}_n)^2,$$

with $\bar{R}_n = \frac{1}{n} \sum_{t=1}^n R_t$ being the sample mean.

 $\sigma_H = \sqrt{T \cdot \hat{v}}$ is then the annualized volatility.

When daily data is used, time length T is always chosen to be 252. When it is monthly data, T = 12.

Implied volatility Implied volatility of an option is calculated from its market price observed on an exchange and not from the prices of the underlying as it is the case for the historical volatility. Consider a European call on a stock, which has a quoted market price of C_t , then its implied volatility σ_I is given by solving

$$C_t^{BS}(S_t, K, \tau, r, d, \sigma_I) - C_t = 0, (2.3)$$

 σ_I is the volatility which if substituted into the Black-Scholes formula (2.2) would give a price equal to the observed market C_t . σ_I is implicitly defined as a solution of the above equation, and has to be computed numerically due to the fact that the Black-Scholes formula cannot be inverted.

Although historical volatility and implied volatilities both estimate the volatility of the underlying asset over the life of the option, the two estimates differ from that they use different data and different models. Implied methods use current data on market prices of options, so the implied volatility contains all of the expectations of investors about possible future price path of the underlying. Moreover, implied volatility assumes that the underlying's price path is continuous. Historical volatility use past data of the underlying returns in a discrete time.

Other methods to calculate volatility are like *stochastic volatility*, which assumes the volatility follows some stochastic process.

Look at Figure 1.1 again. The left panels are typical structure of implied volatility. One can find curvatures across the strike dimension when time to maturity is small though not very clearly. In the maturity dimension, data are sparse. Consequently, implied volatilities appear like several strings. Moreover, the curvatures across strike dimension don't stay unchanged when the strings move through the maturity axis. They change both in levels and shapes. Also in the moneyness axis, observations have their ranges. Thus, even when the data sets are huge, for a large number of cases implied volatility observations are missing for certain sub-regions of the desired estimation grid. This is particularly virulent when transaction based data are used.

However, despite their appearance as strings, implied volatilities are thought to have a structure of smooth surface. This is because in practice one needs to price and hedge OTC options whose expiry dates do not coincide with the expiry dates of the options that are traded at the futures exchange.

3 Dynamic semiparametric factor model

In the dynamic semiparametric factor model (3.1), the index *i* is the number of the day, while the total number of days is denoted by I (i = 1, ..., I). The index *j* represents an intra-day trade on day *i* and the number of trades on that day is J_i ($j = 1, ..., J_i$). Let $X_{i,j}$ be a two-dimensional variable containing moneyness $\kappa_{i,j}$ and maturity $\tau_{i,j}$. Among many moneyness settings we define it as $\kappa_{i,j} \stackrel{\text{def}}{=} \frac{K_{i,j}}{F_{t_i}}$, where $K_{i,j}$ is strike, F_{t_i} is future price of underlying at time t_i . $Y_{i,j} \stackrel{\text{def}}{=} \log\{\sigma_I(\kappa, \tau)\}$ is regressed on $X_{i,j} = (\kappa_{i,j}, \tau_{i,j})$ with nonparametric methods. We take log-implied volatility data, since the data appear less skewed and potential outliers are scaled down after taking logs.

$$Y_{i,j} = m_0(X_{i,j}) + \sum_{l=1}^{L} \beta_{i,l} m_l(X_{i,j}), \qquad (3.1)$$

where m_0 is an invariant basis function, m_l (l = 1, ..., L) are the dynamic basis functions and $\beta_{i,l}$ are the factor weights depending on time *i*.

3.1 Estimation

The estimates \hat{m}_l (l = 0, ..., L) and $\hat{\beta}_{i,l}$ (i = 1, ..., I; l = 1, ..., L) are defined as minimizers of the following least squares criterion $(\hat{\beta}_{i,0} \stackrel{\text{def}}{=} 1)$:

$$\sum_{i=1}^{I} \sum_{j=1}^{J_i} \int \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(u) \right\}^2 K_h(u - X_{i,j}) du.$$
(3.2)

Here, K_h denotes a two-dimensional product kernel, $K_h(u) = k_{h_1}(u_1) \times k_{h_2}(u_2)$, $h = (h_1, h_2)$, based on a one-dimensional kernel $k_h(v) \stackrel{\text{def}}{=} h^{-1}k(h^{-1}v)$.

In (3.2) the minimization runs over all functions $\hat{m}_l \colon \mathbb{R}^2 \to \mathbb{R}$ and all values $\hat{\beta}_{i,l} \in \mathbb{R}$. When L = 0, implied volatility $Y_{i,j}$ are approximated by a surface \hat{m}_0 which does not depend on time *i*. It is the Nadaraya-Watson estimate. Moreover, using (3.2), the estimates \hat{m}_l are not uniquely defined: they can be replaced by functions that give the same affine space. Thus, \hat{m}_l are selected such that they are orthogonal.

Replacing in (3.2) \hat{m}_l by $\hat{m}_l + \delta g$ with arbitrary functions g and taking derivatives with respect to δ yields, for $0 \leq l' \leq L$:

$$\sum_{i=1}^{I} \sum_{j=1}^{J_i} \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(u) \right\} \hat{\beta}_{i,l'} K_h(u - X_{i,j}) = 0.$$
(3.3)

By replacing $\hat{\beta}_{i,l}$ by $\hat{\beta}_{i,l} + \delta$ in (3.2) and again taking derivatives with respect to δ , we get for $0 \leq l' \leq L$ and $1 \leq i \leq I$:

$$\sum_{j=1}^{J_i} \int \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(u) \right\} \hat{m}_{l'}(u) K_h(u - X_{i,j}) du = 0.$$
(3.4)

Using the following notation, for $1 \leq i \leq I$

$$\hat{p}_i(u) = \frac{1}{J_i} \sum_{j=1}^{J_i} K_h(u - X_{i,j}), \qquad (3.5)$$

$$\hat{q}_i(u) = \frac{1}{J_i} \sum_{j=1}^{J_i} K_h(u - X_{i,j}) Y_{i,j}, \qquad (3.6)$$

from (3.3) - (3.4) we get, for $0 \le l' \le L$ and $1 \le i \le I$:

$$\sum_{i=1}^{I} J_i \hat{\beta}_{i,l'} \hat{q}_i(u) = \sum_{i=1}^{I} J_i \sum_{l=0}^{L} \hat{\beta}_{i,l'} \hat{\beta}_{i,l} \hat{p}_i(u) \hat{m}_l(u)$$
(3.7)

$$\int \hat{q}_i(u)\hat{m}_{l'}(u)du = \sum_{l=0}^L \hat{\beta}_{i,l} \int \hat{p}_i(u)\hat{m}_{l'}(u)\hat{m}_l(u)du.$$
(3.8)

We calculate the estimates by iterative use of (3.7) and (3.8). We start with initial values $\hat{\beta}_{i,l}^{(0)}$ for $\hat{\beta}_{i,l}$. Define the matrix $B^{(r)}(u)$ by its elements:

$$(B^{(r)}(u))_{l,l'} \stackrel{\text{def}}{=} \sum_{i=1}^{I} J_i \hat{\beta}_{i,l'}^{(r-1)} \hat{\beta}_{i,l}^{(r-1)} \hat{p}_i(u), \quad 0 \le l, l' \le L,$$
(3.9)

and introduce a vector $Q^{(r)}(u)$ with elements

$$Q^{r}(u)_{l} \stackrel{\text{def}}{=} \sum_{i=1}^{I} J_{i} \hat{\beta}_{i,l}^{(r-1)} \hat{q}_{i}(u), \quad 0 \le l \le L.$$
(3.10)

In the r-th iteration the estimate $\hat{m} = (\hat{m}_0, \dots, \hat{m}_L)^\top$ is given by:

$$\hat{m}^{(r)}(u) = B^{(r)}(u)^{-1}Q^{(r)}(u).$$
 (3.11)

This update step is motivated by (3.7). The value of $\hat{\beta}$ are updated in the r-th cycle as follows: define the matrix $M^{(r)}(i)$

$$(M^{(r)}(i))_{l,l'} \stackrel{\text{def}}{=} \int \hat{p}_i(u)\hat{m}_{l'}^{(r)}(u)du, \quad 0 \le l, l' \le L,$$
(3.12)

and define a vector $S^{(r)}(i)$

$$S^{(r)}(i)_{l} \stackrel{\text{def}}{=} \int \hat{q}_{i}(u)\hat{m}_{l}(u)du - \int \hat{p}_{i}(u)\hat{m}_{0}^{(r)}(u)\hat{m}_{l}^{(r)}(u)du, \quad 0 \le l, l' \le L.$$
(3.13)

Motivated by (3.8), put

$$(\hat{\beta}_{i,1}^{(r)}, \dots, \hat{\beta}_{i,L}^{(r)})^{\top} = M^{(r)}(i)^{-1} S^{(r)}(i).$$
(3.14)

The algorithm is run until only minor changes occur. In the implementation, we choose a grid of points and calculate \hat{m}_l at these points.

3.2 Orthogonalization

As discussed above, \hat{m}_l and $\hat{\beta}_{i,l}$ are not uniquely defined. Therefore, we orthogonalize $\hat{m}_0, \ldots, \hat{m}_L$ in $L^2(\hat{p})$, where $\hat{p}(u) = I^{-1} \sum_{i=1}^{I} \hat{p}_i(u)$, such that $\sum_{i=1}^{I} \hat{\beta}_{i,1}^2$ is maximal, and given $\hat{\beta}_{i,1}, \hat{m}_0, \hat{m}_1, \sum_{i=1}^{I} \hat{\beta}_{i,2}^2$ is maximal, and so forth. These aims can be achieved by the following two steps: first replace

$$\hat{m}_{0} \quad \text{by} \quad \hat{m}_{0}^{new} = \hat{m}_{0} - \gamma^{\top} \Gamma^{-1} \hat{m},$$

$$\hat{m} \quad \text{by} \quad \hat{m}^{new} = \Gamma^{-1/2} \hat{m},$$

$$\begin{pmatrix} \hat{\beta}_{i,1} \\ \vdots \\ \hat{\beta}_{i,L} \end{pmatrix} \quad \text{by} \quad \begin{pmatrix} \hat{\beta}_{i,1}^{new} \\ \vdots \\ \hat{\beta}_{i,L}^{new} \end{pmatrix} = \Gamma^{-1/2} \left\{ \begin{pmatrix} \hat{\beta}_{i,1} \\ \vdots \\ \hat{\beta}_{i,L} \end{pmatrix} + \Gamma^{-1} \gamma \right\},$$
(3.15)

where $\hat{m} = (\hat{m}_1, \dots, \hat{m}_L)^\top$ and the $(L \times L)$ matrix $\Gamma = \int \hat{m}(u)\hat{m}(u)^\top \hat{p}(u)du$, $\gamma = (\gamma_{l,l'})$, with $\gamma_{l,l'} = \int \hat{m}_l(u)\hat{m}_{l'}(u)\hat{p}(u)du$. Finally, we have $\gamma = (\gamma_l)$, with $\gamma_l = \int \hat{m}_0(u)\hat{m}_l(u)\hat{p}(u)du$.

By applying (3.15), \hat{m}_0 is replaced by a function that minimizes $\int \hat{m}_0^2(u)\hat{p}(u)du$. This is evident because \hat{m}_0 is orthogonal to the linear space spanned by $\hat{m}_1, \ldots, \hat{m}_L$. By the second equation of (3.15), $\hat{m}_1, \ldots, \hat{m}_L$ are replaced by orthogonal functions in $L^2(\hat{p})$.

Second step, we proceed as in PCA and define a matrix B with $B_{l,l'} = \sum_{i=1}^{I} \hat{\beta}_{i,l} \hat{\beta}_{i,l'}$ and calculate the eigenvalues of $B, \lambda_1 > \ldots > \lambda_L$, and the corresponding eigenvectors z_1, \ldots, z_L . Put $Z = (z_1, \ldots, z_L)$. Replace

$$\hat{m}$$
 by $\hat{m}^{new} = Z^{\top} \hat{m},$ (3.16)

(i.e. $\hat{m}^{new} = z_l^{\top} \hat{m}$), and

$$\begin{pmatrix} \hat{\beta}_{i,1} \\ \vdots \\ \hat{\beta}_{i,L} \end{pmatrix} \quad \text{by} \quad \begin{pmatrix} \hat{\beta}_{i,1}^{new} \\ \vdots \\ \hat{\beta}_{i,L}^{new} \end{pmatrix} Z^{\mathrm{T}} \begin{pmatrix} \hat{\beta}_{i,1} \\ \vdots \\ \hat{\beta}_{i,L} \end{pmatrix}.$$
(3.17)

After application of (3.16) and (3.17) the orthogonal basis $\hat{m}_1, \ldots, \hat{m}_L$ is chosen such that $\sum_{i=1}^{I} \hat{\beta}_{i,1}^2$ is maximal, and - given $\hat{\beta}_{i,1}, \hat{m}_0, \hat{m}_1 - \sum_{i=1}^{I} \hat{\beta}_{i,2}^2$ is maximal, ..., i.e. \hat{m}_1 is chosen such that as much as possible is explained by $\hat{\beta}_{i,1}\hat{m}_1$. Next \hat{m}_2 is chosen to achieve maximum explanation by $\hat{\beta}_{i,1}\hat{m}_1 + \hat{\beta}_{i,2}\hat{m}_2$, and so forth.

The functions \hat{m}_l are not eigenfunctions of an operator as in usual functional PCA. This is because different norm is used, namely $\int f^2(u)\hat{p}_i(u)du$, for each day. Through the norming procedure the functions are chosen as eigenfunctions in an *L*-dimensional approximating linear space. The *L*-dimensional approximating spaces are not necessarily nested for increasing *L*. For this reason the estimates cannot be calculated by an iterative procedure that starts by fitting a model with one component, and that uses the old L - 1 components in the iteration step from L - 1 to *L* to fit the next component. The calculation of $\hat{m}_0, \ldots \hat{m}_L$ has to be fully redone for different choices of *L*.

3.3 Model selection

For the choice of the model size the residual sum of squares is calculated:

$$RV(L) = \frac{\sum_{i} \sum_{j} \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_{l}(X_{i,j}) \right\}^{2}}{\sum_{i} \sum_{j} \left\{ Y_{i,j} - \bar{Y} \right\}^{2}},$$
(3.18)

where \bar{Y} is the overall mean of the observation. One may increase the parameter L until the explained variance 1 - RV(L) is sufficiently high. However if the model was fitted for L dynamic functions, the new fit for the size L + 1 requires repeating of almost entire procedure.

For the data-driven choice of bandwidths we take like Härdle et al. (2005) a weighted AIC. For the weight function w one needs to minimize:

$$\frac{1}{N} \sum_{i,j} \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(X_{i,j}) \right\}^2 w(X_{i,j})$$
(3.19)

with respect to bandwidths. This is equivalent to minimizing:

$$\Xi_{AIC_1} = \sum_{i,j} \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(X_{i,j}) \right\}^2 w(X_{i,j}) \exp\left\{ \frac{2L}{N} K_h(0) \int w(u) du \right\}$$
(3.20)

or computationally more easy criterion:

$$\Xi_{AIC_2} = \sum_{i,j} \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(X_{i,j}) \right\}^2 \exp\left\{ \frac{2L}{N} K_h(0) \frac{\int w(u) du}{\int w(u) \hat{p}(u) du} \right\}.$$
 (3.21)

Since the distribution of the data is very unequal the weight function w should give greater weight for the regions where data is sparse. One possible selection of w is $w(u) = \frac{1}{\hat{p}(u)}$. Then the two criteria are:

$$\Xi_{AIC_1} = \sum_{i,j} \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(X_{i,j}) \right\}^2 \hat{p}(X_{i,j}) \exp\left\{ \frac{2L}{N} K_h(0) \int \frac{1}{\hat{p}(u) du} \right\}$$
(3.22)

and

$$\Xi_{AIC_2} = \sum_{i,j} \left\{ Y_{i,j} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(X_{i,j}) \right\}^2 \hat{p}(X_{i,j}) \exp\left\{ \frac{2L}{N} K_h(0) \mu^{-1} \int \frac{1}{\hat{p}(u) du} \right\}$$
(3.23)

where μ is the measure of the design set.

4 IVS modeling for KOSPI 200 option and ODAX

4.1 Data description

The data are KOSPI 200 option from Korea Exchange and ODAX from Eurex in whole year of 2003. KOSPI 200 (Korea Composite Stock Price Index) is the index of 200 big companies' stocks traded on the Stock Market Division of the Korea Exchange. It is one of the most actively traded index in the world. The Korea Exchange (KRX) is by transactional volume the largest derivatives exchange in the world.¹ Deutscher Aktien IndeX 30 (German stock index) is a Blue Chip stock market index consisting of the 30 major German companies traded on the Frankfurt Stock Exchange. Its option ODAX is traded on Eurex (Germany and Switzerland) which is by far the world's largest international market organizer for the trading and settlement of futures and options on shares and share indices, as well as of interest rate derivatives.²

Both KOSPI 200 option data and ODAX data contain tick statistics of option contracts traded in year 2003. They are contract based data, i.e. each contract is recorded together with its price, maturity, time of settlement, strike and so on.

Let's take one month from the entire data for example. Table 4.1 are descriptive statistics for data of two markets in September 2003. KOSPI 200 option has much more contracts than ODAX. For KOSPI 200 option, there are more calls than puts, while for ODAX calls are less than puts, although these differences are slight. For KOSPI 200 option, time to maturity has a median of 15 days and ranges from 3 to 102 days. Time to maturity for ODAX has a median of 25 days and ranges from 2 to 475 days. Moneyness of KOSPI 200 option has a median of 1.01 and ranges from 0.56

¹Source: Wikipedia, KRX

²Source: Wikipedia, Eurex

| | | Sum of Trades | Mean | Median | Min. | Max. | Stdd. | Kurt. | Skew. |
|--------------|---|---|---|---|--|---|--|--|--|
| Time to mat. | KOSPI | 2707016 | 15.84 | 15.00 | 3 | 102 | 10.30 | 12.25 | 2.61 |
| | ODAX | 37578 | 50.92 | 25.00 | 2 | 475 | 75.71 | 13.01 | 3.36 |
| Moneyness | KOSPI | 2707016 | 1.03 | 1.02 | 0.56 | 1.24 | 0.04 | 5.23 | 0.62 |
| | ODAX | 37578 | 1.03 | 1.02 | 0.27 | 2.21 | 0.11 | 12.35 | -0.27 |
| Time to mat. | KOSPI | 2682454 | 16.00 | 15.00 | 3 | 102 | 10.02 | 14.92 | 2.87 |
| | ODAX | 41580 | 53.95 | 26.00 | 2 | 475 | 76.92 | 13.11 | 3.37 |
| Moneyness | KOSPI | 2682454 | 0.98 | 0.98 | 0.56 | 1.24 | 0.04 | 2.67 | -0.65 |
| | ODAX | 41580 | 0.96 | 0.97 | 0.28 | 2.24 | 0.11 | 9.81 | 0.12 |
| Time to mat. | KOSPI | 5389470 | 15.92 | 15.00 | 3 | 102 | 10.16 | 13.52 | 2.73 |
| | ODAX | 79158 | 52.51 | 25.00 | 2 | 475 | 76.37 | 13.06 | 3.37 |
| Moneyness | KOSPI | 5389470 | 1.00 | 1.01 | 0.56 | 1.24 | 0.05 | 2.40 | -0.07 |
| | ODAX | 79158 | 0.99 | 1.00 | 0.27 | 2.24 | 0.12 | 8.46 | -0.08 |
| | Time to mat. Moneyness Time to mat. Moneyness Time to mat. Moneyness | Time to mat. KOSPI ODAX Moneyness KOSPI ODAX Time to mat. KOSPI ODAX Moneyness KOSPI ODAX Time to mat. KOSPI ODAX Moneyness KOSPI ODAX | Sum of Trades Time to mat. KOSPI 2707016 ODAX 37578 Moneyness KOSPI 2707016 ODAX 37578 Time to mat. KOSPI 2682454 ODAX 41580 Moneyness KOSPI 2682454 ODAX 41580 Moneyness KOSPI 2682454 ODAX 41580 Time to mat. KOSPI 5389470 ODAX 79158 Moneyness KOSPI 5389470 ODAX 79158 | Sum of Trades Mean Time to mat. KOSPI 2707016 15.84 ODAX 37578 50.92 Moneyness KOSPI 2707016 1.03 ODAX 37578 1.03 ODAX 37578 1.03 ODAX 37578 1.03 Time to mat. KOSPI 2682454 16.00 ODAX 41580 53.95 Moneyness KOSPI 2682454 0.98 ODAX 41580 0.96 Time to mat. KOSPI 2682454 0.98 ODAX 41580 0.96 Time to mat. KOSPI 5389470 15.92 ODAX 79158 52.51 Moneyness KOSPI 5389470 1.00 ODAX 79158 0.99 | Sum of Trades Mean Median Time to mat. KOSPI 2707016 15.84 15.00 ODAX 37578 50.92 25.00 Moneyness KOSPI 2707016 1.03 1.02 ODAX 37578 1.03 1.02 ODAX 37578 1.03 1.02 ODAX 37578 1.03 1.02 Time to mat. KOSPI 2682454 16.00 15.00 ODAX 41580 53.95 26.00 Moneyness KOSPI 2682454 0.98 0.98 ODAX 41580 0.96 0.97 Time to mat. KOSPI 5389470 15.92 15.00 ODAX 79158 52.51 25.00 Moneyness KOSPI 5389470 1.00 1.01 ODAX 79158 0.99 1.00 | Sum of Trades Mean Median Min. Time to mat. KOSPI 2707016 15.84 15.00 3 ODAX 37578 50.92 25.00 2 Moneyness KOSPI 2707016 1.03 1.02 0.56 ODAX 37578 1.03 1.02 0.27 Time to mat. KOSPI 2682454 16.00 15.00 3 ODAX 41580 53.95 26.00 2 Moneyness KOSPI 2682454 0.98 0.98 0.56 ODAX 41580 53.95 26.00 2 Moneyness KOSPI 2682454 0.98 0.98 0.56 ODAX 41580 0.96 0.97 0.28 Time to mat. KOSPI 5389470 15.92 15.00 3 ODAX 79158 52.51 25.00 2 Moneyness KOSPI 5389470 1.00 1.01 0.56 ODA | Sum of Trades Mean Median Min. Max. Time to mat. KOSPI 2707016 15.84 15.00 3 102 ODAX 37578 50.92 25.00 2 475 Moneyness KOSPI 2707016 1.03 1.02 0.56 1.24 ODAX 37578 1.03 1.02 0.27 2.21 Time to mat. KOSPI 2682454 16.00 15.00 3 102 ODAX 41580 53.95 26.00 2 475 Moneyness KOSPI 2682454 0.98 0.98 0.56 1.24 ODAX 41580 0.96 0.97 0.28 2.24 Moneyness KOSPI 2682454 0.98 0.98 0.56 1.24 ODAX 41580 0.96 0.97 0.28 2.24 Time to mat. KOSPI 5389470 15.92 15.00 3 102 ODAX 79158 | $\begin{array}{c c c c c c c c c c c c c c c c c c c $ | $\begin{array}{c c c c c c c c c c c c c c c c c c c $ |

4 IVS modeling for KOSPI 200 option and ODAX

Table 4.1: Descriptive statistics on KOSPI 200 option and ODAX in September 2003, time to maturity in days.

to 1.24. For ODAX, moneyness has a median of 1.00 and ranges from 0.27 to 2.24. For both time to maturity and moneyness, ODAX has larger ranges than KOSPI 200 option.

Figure 4.1 are kernel densities on September 1, 2003 of two markets. ODAX has a flatter distribution than KOSPI 200 option. While for moneyness, distributions are symmetric for both markets, distributions of time to maturity are right-skewed contracts tend to have short expiration time in both markets.



Figure 4.1: Kernel density of KOSPI 200 option (left) and ODAX (right) on September 1, 2003 (moneyness lower left axis, time to maturity lower right axis), bandwidths for moneyness (in year) and time to maturity are both 0.05.

4.2 Close implied volatility and interpolation

One of the problems we face is intraday underlying price. Using intraday underlying price is a convention for IVS estimation. However, neither KOSPI data nor ODAX data in our case contains underlying price. In Härdle et al. (2005) intraday future price within minute interval is used to derive the underlying price. While for ODAX data corresponding future price can be found in our case, we haven't got any information of intraday future price for KOSPI 200.

Another approach to get intraday unerlying price is to use put-call parity:

$$C_t = P_t + S_t - D_t - K e^{-r\tau}, (4.1)$$

where C_t and P_t are prices of European call and put option with the same strike price K, same maturity T and on the same underlying. D_t is the discounted value of dividend payed by underlying during the time to maturity $\tau = T - t$. r and S_t are interest rate and underlying price. Based on corresponding put and call option prices, underlying price S_t can be derived.

Unfortunately, there are merely few cases that pair of corresponding call and put contracts can be found in our data. Thus, intraday implied volatility cannot be calculated in this circumstance.

While it is difficult to find intraday underlying price for both markets, acquiring close price is relatively convenient. Instead of using intraday underlying price, close price of underlying is used in our case. Together with the last traded contract for each series of options - options with the same type, same maturity, same strike and traded on the same day, a *close implied volatility* is calculated. Last traded contracts are extracted from the whole data set according to their settlement time. Most of these last traded contracts are settled between 14:00 and 15:15 on Korea Exchange and between 17:00 and 19:00 on Eurex - both near the close time of their underlying markets. At the same time, it is however also possible that for one series of same contracts, the last traded one is settled as early as 9:15 - near market opens, so a implied volatility based on this contract, together with close underlying price doesn't necessarily lead to a close implied volatility. To improve accuracy, some adjustment or smooth techniques should be used in further study.

| | | | Sum of Trades | Mean | Median | Min. | Max. | Stdd. | Kurt. | Skew. |
|------|--------------|-------|---------------|-------|--------|------|------|--------|-------|-------|
| Call | Time to mat. | KOSPI | 533 | 52.96 | 53.00 | 10 | 102 | 27.10 | -1.21 | -0.04 |
| | | ODAX | 639 | 86.28 | 34.00 | 10 | 475 | 105.98 | 4.36 | 2.17 |
| | Moneyness | KOSPI | 533 | 1.04 | 1.04 | 0.76 | 1.24 | 0.09 | -0.35 | -0.04 |
| | | ODAX | 639 | 0.95 | 0.95 | 0.31 | 2.23 | 0.23 | 2.98 | 0.85 |
| | Implied V. | KOSPI | 533 | 0.22 | 0.20 | 0.12 | 0.75 | 0.05 | 25.38 | 3.89 |
| | | ODAX | 639 | 0.31 | 0.31 | 0.00 | 1.78 | 0.15 | 17.63 | 1.66 |
| Put | Time to mat. | KOSPI | 680 | 53.36 | 53.00 | 10 | 102 | 28.04 | -1.28 | -0.03 |
| | | ODAX | 567 | 85.67 | 33.00 | 10 | 475 | 115.71 | 4.20 | 2.24 |
| | Moneyness | KOSPI | 680 | 0.98 | 0.98 | 0.56 | 1.24 | 0.10 | -0.10 | -0.13 |
| | | ODAX | 567 | 1.01 | 1.01 | 0.27 | 1.79 | 0.23 | 2.47 | -0.43 |
| | Implied V. | KOSPI | 680 | 0.27 | 0.25 | 0.15 | 1.24 | 0.08 | 57.27 | 6.08 |
| | | ODAX | 567 | 0.28 | 0.24 | 0.00 | 3.56 | 0.25 | 73.15 | 7.00 |
| All | Time to mat. | KOSPI | 1213 | 53.18 | 53.00 | 10 | 102 | 27.62 | -1.25 | -0.03 |
| | | ODAX | 1206 | 85.99 | 34.00 | 10 | 475 | 110.61 | 4.30 | 2.21 |
| | Moneyness | KOSPI | 1213 | 1.01 | 1.01 | 0.56 | 1.24 | 0.10 | -0.03 | -0.22 |
| | | ODAX | 1206 | 0.98 | 0.98 | 0.27 | 2.23 | 0.23 | 2.27 | 0.25 |
| | Implied V. | KOSPI | 1213 | 0.25 | 0.23 | 0.12 | 1.24 | 0.08 | 53.13 | 5.38 |
| | | ODAX | 1206 | 0.29 | 0.27 | 0.00 | 3.56 | 0.20 | 79.89 | 6.31 |

4 IVS modeling for KOSPI 200 option and ODAX

Table 4.2: Descriptive statistics on last traded KOSPI 200 option and ODAX in September 2003, time to maturity in days.

For calculating implied volatility, contracts with time to maturity less than 10 days are removed since their behaviors in this range are irregular due to expiry effect. Interest rate for KOSPI 200 option is daily Korean treasury bill rate, for ODAX is daily EURIBOR rate in the same period. Table 4.2 shows the summaries of these contracts in September 2003. Both markets turn to have similar amount of contracts. ODAX still has wider range of moneyness than KOSPI 200 option. ODAX also has a larger range of implied volatility.

After extracting the last traded contracts from the whole data set, our data size reduce largely - only about 40 observations left each day in average. An linear interpolation is used in order to represent the data on a regular grid. Options on the same day and with the same maturity time are grouped together and interpolated with respect to the moneyness grid $\kappa \in [0.8, 1.2]$ with 1000 points. Furthermore, call option and put option are interpolated separately.

A new problem arises here: implied volatility calculated from put option and call option are different, even if they have the same strike, same maturity and same spot.

In Härdle et al. (2005) a correction algorithm is used to obtain an adjusted underlying price, based on which the put and call implied volatilities are the same. This algorithm (4.2) is based on future price formula and put-call parity.

$$\tilde{S}_t = e^{-r_F(T_F - t)} F_t + \Delta D_{t, T_H, T_F},$$
(4.2)

where \tilde{S}_t is the adjusted underlying price, r_F is interest rate used in deriving underlying price from future price, T_F is future's maturity time, F_t is future price, T_H is option's maturity time, $\Delta D_{t,T_H,T_F}$ is dividend difference calculated from put-call parity (4.1) and future price formula $F_t = e^{r_F(T_F-t)}S_t - \Delta D_{t,T_F}$, $\Delta D_{t,T_F}$ is dividend used for calculating future price.

In our case, it is a different situation. Firstly, our underlying price is not obtained form future price, but directly from close price in the underlying markets. Secondly, as stated before, there are only few corresponding put and call options in our case, thus put-call parity cannot be used. Consequently, we don't apply algorithm (4.2) in our case. Instead, after interpolating put and call option separately we take the average of these two implied volatility series. Our algorithm of interpolation and average taking is described in detail as follows:

- 1. For each day i, options (implied volatilities) with the same maturity are grouped together as $IV_{i,j}^C(\kappa_m)$ and $IV_{i,j}^P(\kappa_n)$. *i* is day index. *j* is group index on day i index for different maturity. *C* is call and *P* is put. κ_m and κ_l are moneyness for calls and puts respectively.
- 2. For each group j of these options $IV_{i,j}^C(\kappa_m)$ and $IV_{i,j}^P(\kappa_n)$,
 - a) if both the number of calls $\sum m \geq 3$ and the number of puts $\sum n \geq 3$, call implied volatility and put implied volatility are linearly interpolated separately with respect to moneyness κ on grid $\kappa \in [0.8, 1.2]$ with 1000 points.

$$IV_{i,j}^C(\kappa_l) = \frac{\kappa_l - \kappa_m}{\kappa_{m+1} - \kappa_m} IV_{i,j}^C(\kappa_{m+1}) + \frac{\kappa_{m+1} - \kappa_l}{\kappa_{m+1} - \kappa_m} IV_{i,j}^C(\kappa_m), \qquad (4.3)$$

and

$$IV_{i,j}^P(\kappa_l) = \frac{\kappa_l - \kappa_n}{\kappa_{n+1} - \kappa_n} IV_{i,j}^P(\kappa_{n+1}) + \frac{\kappa_{n+1} - \kappa_l}{\kappa_{n+1} - \kappa_n} IV_{i,j}^P(\kappa_m), \qquad (4.4)$$

where κ_l is desired moneyness point, $l \in [1, 1000]$, $\kappa_l \in [0.8, 1.2]$. κ_m and

 κ_{m+1} are two closest observations to κ_l , with $\kappa_m < \kappa_l < \kappa_{m+1}$. If $\kappa_{m+1} < \kappa_l$ or $\kappa_l < \kappa_m$, a constant extrapolation is used. For κ_n and κ_{n+1} is the same.

Then, an average of $IV_{i,j}^C(\kappa_l)$ and $IV_{i,j}^P(\kappa_l)$ is taken as the final implied volatility,

$$IV_{i,j}(\kappa_l) = \frac{IV_{i,j}^C(\kappa_l) + IV_{i,j}^P(\kappa_l)}{2}$$
(4.5)

b) if only the number of calls $\sum m \ge 3$, while the number of puts $\sum n < 3$, call implied volatilities are interpolated with (4.3). Put implied volatilities are dropped. Final implied volatilities are these interpolated call implied volatilities,

$$IV_{i,j}(\kappa_l) = IV_{i,j}^C(\kappa_l) \tag{4.6}$$

c) similarly, if only the number of puts $\sum n \geq 3$, but the number of calls $\sum m < 3$, put implied volatilities are interpolated with (4.4). Call implied volatilities are dropped. Final implied volatilities are these interpolated put implied volatilities,

$$IV_{i,j}(\kappa_l) = IV_{i,j}^P(\kappa_l) \tag{4.7}$$

d) if both the number of calls $\sum m < 3$ and the number of puts $\sum n < 3$, final implied volatilities are the original observations,

$$IV_{i,j}(\kappa_m) = IV_{i,j}^C(\kappa_m)$$

$$IV_{i,j}(\kappa_n) = IV_{i,j}^P(\kappa_n)$$
(4.8)

Figure 4.2 is an example of one group ODAX data on September 22, 2003. These contracts have 27 days to maturity. Left panel is original data - lasted traded contracts taken from the entire data set. Right panel is based on the left panel, with interpolating call and put options separately and then taking average.



Figure 4.2: Implied volatility of last traded ODAX with 27 days to maturity on September 22, 2003. Left panel: original data, before interpolation. Call options are crosses, put options are circles. Right panel: with linear interpolation and taking average.

| | Bandwidths | 1 - RV(3) | Ξ_{AIC_1} | Ξ_{AIC_2} |
|------------------|--------------|-----------|---------------|---------------|
| KOSPI 200 option | (0.02, 0.2) | 0.853 | 0.00131 | 0.00663 |
| ODAX | (0.02, 0.05) | 0.853 | 0.00360 | 0.01418 |

4 IVS modeling for KOSPI 200 option and ODAX

Table 4.3: DSFM modeling result for both markets.

| L | | 1 - RV(L) | $\triangle RV$ |
|---|---------|-----------|----------------|
| 2 | KOSPI | 0.836 | |
| | DAX | 0.805 | |
| 3 | KOSPI | 0.853 | 0.017 |
| | DAX | 0.853 | 0.048 |
| 4 | KOSPI* | 0.870 | 0.017 |
| | DAX | 0.888 | 0.035 |
| 5 | KOSPI** | 0.865 | -0.005 |
| | DAX | 0.910 | 0.022 |

Table 4.4: Explained variance for the model size. *Bandwidth (0.02, 0.22). **Bandwidth (0.02, 0.28).

4.3 Model fit

Before applying DSFM, observations with implied volatility smaller than 0.04 or larger than 0.8 are removed. Consequently, there are 653655 observations in KOSPI 200 option data - about 2880 per day and 931662 observations in ODAX data, about 3818 per day. Log-implied volatilities are modeled on moneyness and time to maturity ($\kappa_{i,j}$, $\tau_{i,j}$). The grid covers in moneyness $\kappa \in [0.8, 1.2]$ and time to maturity $\tau \in [0.02, 0.5]$ measured in year. L=3 factor functions are employed for both markets. According to variance explained, Ξ_{AIC_1} and Ξ_{AIC_2} criterions, bandwidths (0.02, 0.2) for KOSPI 200 option and (0.02, 0.05) for ODAX are chosen, see Table 4.3. We have recalculated the model with the same bandwidths with 2, 3, 4 and 5 dynamic factor functions, see Table 4.4. A choice of 3 factor functions is based on a balance between model economy and efficiency - capturing as much as possible explained variances while keeping small number of factors.



Figure 4.3: Left panel: DSFM for KOSPI 200 option on September 9, 2003 with bandwidths (0.02, 0.2). Right panel: DSFM for ODAX on September 9, 2003 with bandwidths (0.02, 0.04). (moneyness lower left axis, time to maturity lower right axis)

Figure 4.3 display Models on September 9, 2003 for both markets. IVS of KOSPI 200 option is flatter than ODAX on this day. Figure 4.4, 4.5 and 4.6 are factor functions \hat{m}_i , i = 1, 2, 3. \hat{m}_1 are positive in both markets. They both are relatively flat. Factor \hat{m}_1 can be interpreted as the time dependent mean of the (log-)IVS, a *shift effect*. \hat{m}_2 for both markets have visible upward trends in moneyness axis, from negative to positive. The surface is close to 0 when moneyness near 1. This trend is strong when time to maturity is small while it becomes weak when time to maturity increases. For \hat{m}_3 of KOSPI 200 option, a downward trend exists in maturity direction when moneyness is in the neighborhood of 0.8 (this could be more clear if the plot is seen in another direction). The surface slope down when time to maturity increases. This trend becomes weak when moneyness is increases. \hat{m}_3 of ODAX appears like wave.



Figure 4.4: Factor \hat{m}_1 for KOSPI 200 option (left) and for ODAX (right). (moneyness lower left axis, time to maturity lower right axis)



Figure 4.5: Factor \hat{m}_2 for KOSPI 200 option (left) and for ODAX (right). (moneyness lower left axis, time to maturity lower right axis)



Figure 4.6: Factor \hat{m}_3 for KOSPI 200 option (left) and for ODAX (right). (moneyness lower left axis, time to maturity lower right axis)

| 1 | 4 | IVS | modeling | for | KOSPI | 200 | option | and | ODA |
|---|---|-----|----------|-----|-------|-----|--------|-----|-----|
|---|---|-----|----------|-----|-------|-----|--------|-----|-----|

| | | Mean | Median | Min | Max | Stdd | Kurt | Skew |
|-------|---------------|-------|----------|---------|-------|-------|--------|-------|
| | ^ | wican | Wieulali | 101111. | max. | biuu. | Ituro. | DRCW. |
| KOSPI | β_1 | -1.13 | -1.09 | -1.45 | -0.84 | 0.16 | 1.68 | -0.16 |
| | \hat{eta}_2 | 0.00 | 0.03 | -0.35 | 0.13 | 0.08 | 5.78 | -1.50 |
| | \hat{eta}_3 | 0.00 | 0.01 | -0.22 | 0.13 | 0.04 | 7.65 | -1.21 |
| ODAX | \hat{eta}_1 | -1.03 | -1.09 | -1.36 | -0.56 | 0.20 | 2.01 | 0.49 |
| | \hat{eta}_2 | 0.00 | -0.01 | -0.27 | 0.24 | 0.11 | 2.18 | -0.08 |
| | \hat{eta}_3 | 0.00 | 0.00 | -0.27 | 0.21 | 0.07 | 4.32 | -0.14 |

Table 4.5: Descriptive statistics for factors loading $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$.



Figure 4.7: Time series of factor loadings β_1 , β_2 , β_3 for KOSPI 200 option (blue) and ODAX (red) in year 2003.

4.4 Time series of factors loadings

Table 4.5 are descriptive statistics of factor loadings $\hat{\beta}_i$, i = 1, 2, 3. Figure 4.7 plots $\hat{\beta}_i$ of both markets together. $\hat{\beta}_1$ have larger scales than $\hat{\beta}_2$, $\hat{\beta}_3$ and display trends in both markets. Moreover, $\hat{\beta}_1$ of both markets seem to have common trend. $\hat{\beta}_2$ and $\hat{\beta}_3$ of ODAX have stronger fluctuation than that of KOSPI 200 option.

ADF tests in Table 4.6 suggest $\hat{\beta}_1$ of both markets have unit roots. $\hat{\beta}_2$ of KOSPI 200 option has a unit root. Other $\hat{\beta}_i$ are stationary.

In Härdle et al. (2005) VAR(2) model is used to model $\hat{\beta}_i$ within one markets. What interests us more here is the relation between $\hat{\beta}_i$ of different markets. Since Figure 4.7 suggests trend of $\hat{\beta}_1$ in both markets, we firstly compare $\hat{\beta}_1$ with underlying price of the





Figure 4.8: $\hat{\beta}_1$ (blue) and underlying price (red, logarithmized and centralized) of KOSPI 200 option (left) and ODAX (right).

their own market. Since underlying price and $\hat{\beta}_1$ have different scales, underlying prices are logarithmized and centralized (minus their mean). Figure 4.8 shows these two series on two markets. Two series on both markets seem to move in opposite direction. We then plot $\hat{\beta}_1$ and negative underlying price series (logarithmized and centralized) together. Results are showed in Figure 4.9. On KOSPI 200 market, although about from day 175, $\hat{\beta}_1$ and underlying price begin to move in opposite direction, they move in same trend before day 175. On ODAX market, $\hat{\beta}_1$ and underlying price always move in same trend. Correlation coefficients between $\hat{\beta}_1$ and underlying price (logarithmized and centralized) are -0.82877 for Korean market and -0.89434 for German market.

Moreover, the two trends in left panel of Figure 4.7 seem to move together, so cointegration tests are implemented next in Table 4.7 for $\hat{\beta}_1$ of two markets. Both trace test and maximum eigenvalue test cannot reject a cointegration between $\hat{\beta}_1$ of ODAX $(\hat{\beta}_1^k)$ and $\hat{\beta}_1$ of KOSPI 200 option $(\hat{\beta}_1^d)$. Cointegration matrix $\beta = (-9.80, 11.63)$ for $(\hat{\beta}_1^d, \hat{\beta}_1^k)$, loading matrix $\alpha = (-0.01, 0.01)^{\top}$. This is, a linear combination z of $\hat{\beta}_1^d$ and $\hat{\beta}_1^k$, $z = -9.8\hat{\beta}_1^d + 11.63\hat{\beta}_1^k$ is a stationary series. The existence of cointegration means that there is long-term relationship between two variables and that they are influenced by the same stochastic trend. Since $\hat{\beta}_1$ is the loading for \hat{m}_1 , which is thought





Figure 4.9: $\hat{\beta}_1$ (blue) and negative underlying price (red, logarithmized and centralized) of KOSPI 200 option (left) and ODAX (right).

to be the time dependent mean of the (log-)IVS, a long-term relationship between two markets' $\hat{\beta}_1$ could probably indicate long-term relationship between two markets' IVS. From theoretical point of view, this result is not surprising. According to no-arbitrage relation, for global exchanges like KRX and Deutsche Börse/Eurex, there shouldn't exist any arbitrage opportunity between them. Two exchanges are both influenced by the global markets and macro environment. Relations or equilibriums should exist between them. As we can see from Figure 4.8, underlying price series in left panel and right panel have similar shapes and trends. Implied volatility as a state indicator for market situation, should also reflect this no-arbitrage relation. Consequently, when global markets have large fluctuation, so do KRX and Deutsche Börse/Eurex, implied volatilities of these two markets will become large. If global markets have small fluctuation, implied volatilities of these two markets will become small.

| Number of lags | | 1 | 2 | 3 | 4 |
|----------------|-----------------|--------|-------|-------|-------|
| KOSPI | \hat{eta}_1 | -1.75 | -1.54 | -1.47 | -1.27 |
| | $\hat{\beta}_2$ | -4.26 | -2.87 | -2.71 | -2.55 |
| | \hat{eta}_3 | -6.23 | -5.49 | -5.09 | -4.69 |
| ODAX | $\hat{\beta}_1$ | -1.90 | -1.63 | -1.35 | -1.24 |
| | $\hat{\beta}_2$ | -11.20 | -8.45 | -7.39 | -6.01 |
| | \hat{eta}_3 | -10.65 | -9.24 | -8.55 | -7.57 |

Table 4.6: ADF test statistics for factor loading series. Tests include constant. Critical values are -3.46 (1%), -2.87 (5%) and -2.57 (10%).

| Trace test | | | | | Maximum eigenvalue test | | | | |
|------------|----------------|-------|-----------|-------|-------------------------|----------------|----------------|-------|-------|
| | | Cr | itical va | lue | | | Critical value | | |
| Rank | Test statistic | 10% | 5% | 1% | Rank | Test statistic | 10% | 5% | 1% |
| 1 | 2.28 | 6.50 | 8.18 | 11.65 | 1 | 2.28 | 6.50 | 8.18 | 11.65 |
| 0 | 21.90 | 15.66 | 17.95 | 23.52 | 0 | 19.62 | 12.91 | 14.90 | 19.19 |

Table 4.7: Two cointegration tests for KOSPI $\hat{\beta}_1$ and ODAX $\hat{\beta}_1$.

5 Conclusion

In this master thesis, we apply dynamic semiparametric factor model (DSFM) to estimate implied volatility surfaces (IVS) of KOSPI 200 option and ODAX. Due to lack of information for intraday underlying price, daily close underlying price is used. Together with the last traded contract, daily close implied volatility is calculated. For data preparation, a smooth procedure is employed. Close implied volatility is linearly interpolated on regular grid with respect to moneyness. Due to the difference between call and put implied volatility, an average of them is taken.

We choose L = 3 factor functions for estimation. 85.3% of the variances are explained for both markets. The first factor function are similar for two markets and can be interpreted as the time dependent mean of the (log-)IVS. We then study the factor loadings. Loadings for first factor have unit roots for both markets. They are highly correlated with their own markets' underlying prices. Cointegration tests indicate the existence of cointegration between two markets' first factor loadings. This result furthermore indicates the long-term relationship between two markets' IVS, which confirms the no-arbitrage theory.

Bibliography

- Ahoniemi, K. and Lanne, M. (2007). Joint Modeling of Call and Put Implied Volatility, *Discussion Paper 198*, Helsinki Center of Economic Research, University of Helsinki, Finland.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, Journal of Political Economy 81: 637–654.
- Borak, S. (2005). Implementation of dynamic semiparametric factor model for implied volatilities, *Master thesis*, Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin, Germany.
- Borak, S., Härdle, W. and Fengler, M. R. (2005). DSFM fitting of Implied Volatility Surfaces, *Sfb 649 discussion paper*, Sonderforschungsbereich 649, Humboldt-Universität zu Berlin, Germany.
- Borak, S., Härdle, W., Mammen, E. and Park, B. U. (2007). Time Series Modelling with Semiparametric Factor Dynamics, *Sfb 649 discussion paper*, Sonderforschungsbereich 649, Humboldt-Universität zu Berlin, Germany.
- Franke, J., Härdle, W. and Hafner, C. (2004). Statistics of Financial Markets: An introduction, Springer, Heidelberg.
- Franses, P. H. (1998). Time Series Models for Business and Economic Forecasting, Cambridge University Press.
- Härdle, W. K., Fengler, M. R. and Mammen, E. (2005). A Dynamic Semiparametric Factor Model for Implied Volatility String Dynamics, *Sfb 649 discussion paper*, Sonderforschungsbereich 649, Humboldt-Universität zu Berlin, Germany.
- Härdle, W., Klinke, S. and Müller, M. (2000). *Xplore Learning Guide*, Springer, Heidelberg.

- Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2003). Non- and Semiparametric Models, Springer, Heidelberg.
- Härdle, W. and Simar, L. (2003). Applied Multivariate Statistical Analysis, Springer, Heidelberg.
- Hull, J. C. (2000). Options, Futures and other Derivatives, Prentice Hall.
- Kühl, M. (2007). Cointegration in the Foreign Exchange Market and Market Efficiency since the Introduction of the Euro: Evidence based on bivariate Cointegration Analyses, *Discussion Paper 68*, Center for European, Governance and Economic Development Research, Georg-August-Universität Göttingen, Germany.

Tsay, R. S. (2002). Analysis of Financial Time Series, Wiley.