# Implied Basket Correlation Dynamics

Wolfgang Karl Härdle *
Elena Silyakova *

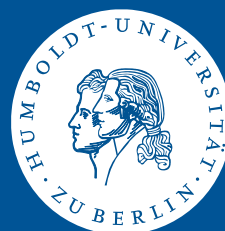* Humboldt-Universität zu Berlin, Germany

BERLIN

ECONOMIC RISK

SFB 649

# Implied Basket Correlation Dynamics *

Wolfgang Karl Härdle[†]and Elena Silyakova[‡]

November 14, 2012

**Abstract**

Equity basket correlation is an important risk factor. It characterizes the strength of linear dependence between assets and thus measures the degree of portfolio diversification. It can be estimated both under the physical measure from return series, and under the risk neutral measure from option prices. The difference between the two estimates motivates a so called "dispersion strategy". We study the performance of this strategy on the German market over the recent 2 years and propose several hedging schemes based on implied correlation (IC) forecasts. Modeling IC is a challenging task both in terms of computational burden and estimation error. First the number of correlation coefficients to be estimated would grow with the size of the basket. Second, since the IC is implied from option prices it is not constant over maturities and strikes. Finally, the IC changes over time. The dimensionality of the problem is reduced by an assumption that the correlation between all pairs of equities is constant (equicorrelation). The IC surface (ICS) is then approximated from implied volatilities of stocks and implied volatility of the basket. To analyze this structure and the dynamics of the ICS we employ a dynamic semiparametric factor model (DSFM).

*JEL classification codes:* C14, C32, G12, G13, G15, G17

*Keywords:* correlation risk, dimension reduction, dispersion strategy, dynamic factor models, implied correlation

[†]Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. Email: haerdle@wiwi.hu-berlin.de.

[‡]Corresponding author. Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. Email: silyakoe@cms.hu-berlin.de.

# 1 Introduction

In a basket of $N$ assets correlation $\rho_{i,j}$, $i, j \in \{1, \ldots, N\}$, is an important (linear) measure of co-movements between two return series. It is an input for many pricing models, plays a key role in portfolio optimization and risk management. The concept of a time-varying correlation is frequently used in studies that describe the joint dynamics of assets, Bollerslev et al. (1988), Engle (2002). However the idea to consider the correlation an asset on its own is relatively new and has recently gained its popularity together with emerging such derivative instruments as variance, volatility, correlation swaps and trading strategies with them. In this context being able to predict correlation patterns might help to reveal profitable trading opportunities. One of the most common ways to obtain the desired correlation exposure is to replicate it with volatility derivatives such as variance swaps. Here we study the behavior of a particular correlation trading strategy known as "dispersion strategy", in which one sells the volatility of the index and buys volatilities of index constituents. We propose several ways of improving the profitability of the strategy by extracting information from the dynamic model of the implied correlation.

Unlike asset prices, correlation is not directly observed in the market and needs to be estimated in the context of a particular model. Obtaining a well-conditioned and invertible estimate of an empirical correlation matrix is often a complicated task, in particular when dimensionality of basket elements $N$ is higher than the time series length $T$. Here some work has been done in the field of random matrix theory (RMT), where the case "large $N$, small $T$" is studied in an asymptotic setting, Bai (1999), Laloux et al. (1999), Plerou et al. (2002). A further segment of research has moved in the direction of developing various regularization methods for sample covariance and correlation matrices, such as shrinkage technique proposed in Ledoit and Wolf (2003), regularization via thresholding in Bickel and Levina (2008a), bending in Bickel and Levina (2008b), factor models in Fan et al. (2008) and many others. There are some studies proposing a dynamic model for returns' correlation such as DCC model by Engle (2002), and in high-dimensional setting, Engle et al. (2008). The common feature of all these studies is that the empirical correlation matrix is estimated under the physical measure from the time series of asset returns. Alternatively, instead of relying on historical data, one can infer correlation from the current snapshot of the option market. Option prices reflect expectations of market participants about the future price (volatility) and disclose their perception of market risk, Bakshi et al. (2000), Britten-Jones and Neuberger (2000). Some recent studies have shown that the implied volatility (IV), that equates the model option price and the one taken from the market, contains incremental information beyond the historical estimate and outperforms it in forecasting future volatility, Christensen and Prabhala (1998), Fleming (1998), Blair et al. (2001). Yet only few papers have studied the predic-

tive content of the correlation, implied by option prices. Some work has been done for foreign exchange (FX) options, Campa and Chang (1998), Lopez and Walter (2000), who showed that correlation implied from FX options is useful for forecasting future currency correlations. Skintzi and Refenes (2005) investigated average correlation implied by equity options and introduced the Implied Correlation index (ICX). They show that ICX, computed from current option prices, is a useful proxy for the future realized correlation. Driessen et al. (2009) investigate the power of options implied correlation to explain the future realized correlation and conclude that its predictive power is quite high.

Here we model the implied correlation (IC), which is an object of a very high dimensionality. Similar to the IV, every day one recovers a IC surface. We model the IC with a dynamic semiparametric factor model (DSFM), Fengler et al. (2007), Park et al. (2009) and Song et al. (2010). It yields a low dimensional representation as a linear combination of a small number of time-invariant basis functions (surfaces), whose time evolution is driven by series of coefficients. We produce an IC forecast and use it in several hedging schemes for dispersion strategy. For empirical analysis we chose the German market represented by DAX portfolio over the 2-years sample period from 20100802 until 20120801 (dates are written as YYYYMMDD). Backtesting shows that the hedging allows to reduce potential losses and increase the average profitability of the strategy.

The paper is structured as follows. In Section 2 we introduce the notions of realized, model-implied and model-free implied volatility and correlation and describe the basic setup of a dispersion strategy with variance swaps. The DSFM model for IC is introduced in Section 3 starting with general description in Section 3.1 followed by the description of the functional principal component analysis (FPCA) approach to finding the basis functions in Section 3.2 and the estimation procedure for both factors and factor loadings in Section 3.3. Section 4 presents the dataset taken for the empirical study, followed by description of the estimation results in Section 5. Here first we interpret obtained factors and factor loadings and propose a time series model for low-dimensional factors in 5.1. Finally in Section 5.2 we propose and compare alternative dispersion strategy setups: no hedge, naïve approach and advanced hedge. Section 6 concludes.

# 2  Correlation trading

## 2.1  Average basket correlation

Standard statistical analysis yields that the basket variance $\sigma_B^2$ can be decomposed as:

$$\sigma_B^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{ij}, \tag{1}$$

where $\sigma_i^2$ denotes the variance of the $i$-th asset and $w_i$ its weight in the basket. Now, assuming that $\rho_{ij}$ is constant for every pair $(i, j)$, one can imply the equicorrelation $\rho$ from (1):

$$\rho = \frac{\sigma_B^2 - \sum_i w_i^2 \sigma_i^2}{\sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j}. \tag{2}$$

Later we call $\rho$ a basket correlation or simply correlation. The corresponding correlation matrix has all off-diagonal elements equal to $\rho$ and offers thus several advantages. First, plugging $\rho_{i,j} = \rho$ into (1) reproduces the basket variance $\sigma_B^2$. Second, if $-\frac{1}{N-1} < \rho < 1$ then the correlation matrix is positive semidefinite, Härdle and Simar (2012). This property becomes particularly important if $N$ is large. A closer look also reveals that (2) is in fact a nonlinear weighted average over all $\rho_{i,j}$ in the basket:

$$\rho = \sum_i \sum_{j \neq i} c_{i,j} \rho_{i,j} \tag{3}$$

with weights $c_{i,j}$ defined by:

$$c_{i,j} = \frac{w_i w_j \sigma_i \sigma_j}{\sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j}. \tag{4}$$

Bourgoin (2001) showed that if the correlation matrix is positive semidefinite, for sufficiently large baskets it holds that $0 \leq \rho \leq 1$. Using this property maximum and minimum variances of a basket, $\sigma_{B,min}^2$ and $\sigma_{B,max}^2$ respectively, are defined as follows:

$$\sigma_{B,min}^2 = \sum_i w_i^2 \sigma_i^2, \tag{5}$$

$$\sigma_{B,max}^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j. \tag{6}$$

$\sigma_{B,min}^2$ is achieved when $\rho = 0$ that is when the assets in the basket are fully diversified. In case of no diversification one observes the maximal possible basket variance $\sigma_{B,max}^2$ corresponding to $\rho = 1$.

Further we can rewrite $\rho$ by substituting (5) and (6) to (2):

$$\rho = \frac{\sigma_B^2 - \sigma_{B,min}^2}{\sigma_{B,max}^2 - \sigma_{B,min}^2} \tag{7}$$

and obtain an additional interpretation as a measure for degree of diversification, Skintzi and Refenes (2005). In fact (7) shows how far is $\sigma_B^2$ from its minimal value $\sigma_{B,min}^2$ relative to the possible value range $\sigma_{B,max}^2 - \sigma_{B,min}^2$, or in other words, how far is the basket from the perfect diversification. High $\rho$ is the sign of a poorly diversified portfolio, which is typical for the market downturn, when asset prices simultaneously drop driving $\sigma_B^2$ up. It means diversification benefits disappear in times when they are at most needed. To hedge against correlation risk investors look for derivative securities that offer higher payoffs (premia) when the correlation decreases.

If a basket is constructed from constituents of an equity index with weights equal to index weights, then the corresponding basket correlation would serve as a benchmark for a sector, an industry or a whole market average correlation. Figure 1 shows an example of the DAX correlation together with the volatility of DAX and some of its components. First, we clearly see that correlation and volatility vary over time. Second, the volatility of the basket (DAX) is smaller than almost any individual volatility of its constituents, which illustrates the impact of the diversification effect on the portfolio risk. Finally, there is a clear linear dependence of the correlation of the basket and its volatility. However the strength of this dependence changes when the volatility exceeds a certain threshold. We investigate this phenomenon and propose a dataset correction scheme in Section 4.

## 2.2 Implied versus realized correlation

Based on (2) we conclude that the exposure to the basket correlation $\rho$ can be achieved by exposures to the variances of a basket $\sigma_B^2$ and its constituents, $\sigma_i^2$. Such trades can be realized via a combination of variance swaps, an over-the-counter forward contract opened at $t$, which at $t + \tau$ pays the difference between the variance cumulated over the
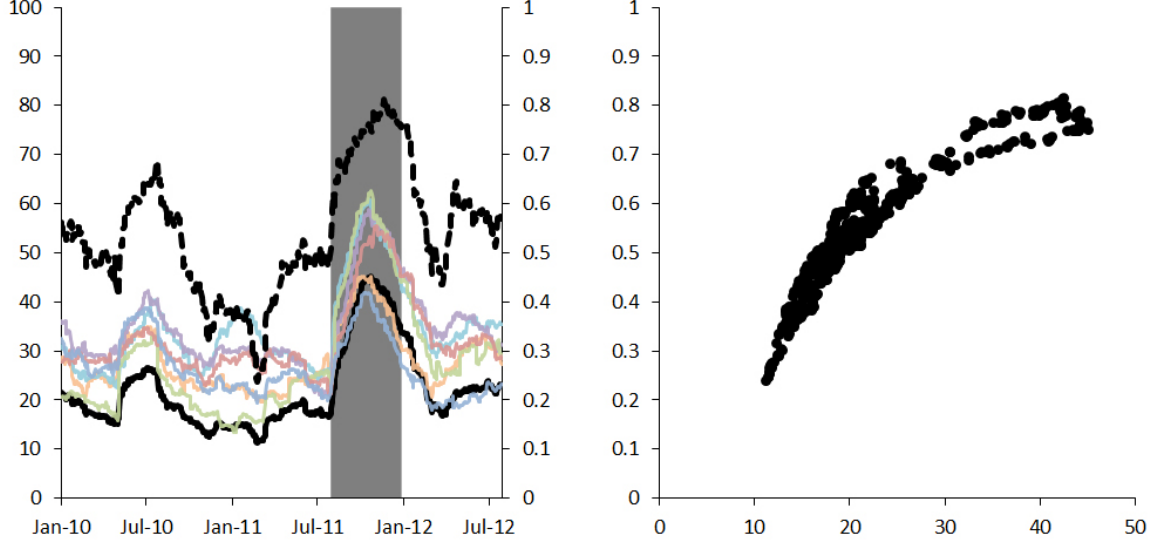
Figure 1: Left panel: series of Adidas, BMW, Siemens, Daimler, E.ON, Lufthansa volatilities (11) - color, DAX volatility (11) - solid black, equicorrelation (2) - dashed black, the stock market fall 2011 - shaded area. Right panel: scatter plot DAX volatility vs. equicorrelation. Estimation period - from 20100104 to 20120801, estimation window - 3 month.

life time of the swap $\sigma_{t+\tau}^2$ and the strike $\tilde{\sigma}_t^2(\tau)$. Thus the payoff is defined by

$$\left\{\sigma_{t+\tau}^2 - \tilde{\sigma}_t^2(\tau)\right\} N_{var}, \tag{8}$$

where $N_{var}$ is the notional amount. Here and later $t$ and $\tau$ are given in fractions of a year.

The strike of the variance swap, the model-free implied variance (MFIV), is the risk-neutral expectation at $t$ of the integrated variance from $t$ to $t+\tau$. "Model-free" indicates that the expectation does not depend on the specification of the underlying price process, Britten-Jones and Neuberger (2000). MFIV can be approximated by a function of current option prices, Breeden and Litzenberger (1978), Carr and Madan (1998), Britten-Jones and Neuberger (2000), which has the following form

$$\tilde{\sigma}_t^2(\tau) = \mathrm{E}_t^Q\left[\int_t^{t+\tau} \sigma^2(s)ds\right] =$$

$$\frac{2e^{r\tau}}{\tau}\left\{\int_0^{S_t} \frac{P_t(K,\tau)dK}{K^2} + \int_{S_t}^{\infty} \frac{C_t(K,\tau)dK}{K^2}\right\}, \tag{9}$$

6

where $\mathrm{E}_t^Q$ expected value at $t$ under the risk-neutral measure $Q$, $P_t(K, \tau)$ $\{C_t(K, \tau)\}$ price of put (call) with exercise price $K$ and time to maturity $\tau$ traded at $t$, $S_t$ price of the asset in $t$, $r$ the annualized continuously compounded risk-free interest rate.

MFIV can be opposed to the implied variance $\widehat{\sigma}_t^2(\kappa, \tau)$, the implied volatility (IV) squared, which is obtained by solving

$$V_t(\widehat{\sigma}, \kappa, \tau) - \widecheck{V}_t(\kappa, \tau) = 0, \tag{10}$$

where $V_t$ is the theoretical (model) option price, $\widecheck{V}_t$ option price taken from the market, $\kappa = \dfrac{K}{S_t e^{r\tau}}$ moneyness of the option. IV, in comparison to MFIV, is a function of both $\kappa$ and $\tau$, meaning that every $t$ one recovers a cloud of points, which can be approximated by a surface, Cont and Da Fonseca (2002), Fengler et al. (2007).

The floating leg of the variance swap, the realized variance (RV) of an asset from $t$ to $t+\tau$, can be computed from the time series of daily asset returns in different ways, depending on the contract specification. Here we use the most common following form

$$\sigma_{t+\tau}^2 = \tau^{-1} \sum_{i=252t}^{252(t+\tau)} \left( \log \frac{S_i}{S_{i-1}} \right)^2. \tag{11}$$

In Carr and Wu (2009) $\sigma_{t+\tau}^2 - \tilde{\sigma}_t^2(\tau)$ is referred to as the variance risk premium (VRP), which is shown to be strongly negative for major US stock indexes over the sample period from January 1996 until December 2003. The negative sign indicates that investors are willing to pay extra to hedge themselves against possible future market turmoils. Bakshi et al. (2003), who investigated the S&P100 index and its largest constituents from 1991 until 1995, also found significant negative difference between realized and option implied volatilities for the average of 25 stocks and stressed that this difference is less pronounced than for the index. Driessen et al. (2009) study each S&P100 constituent individually. Their $t$-test for $H_0$ that on average RV=MFIV was not rejected for the majority of stocks in the sample from January 1996 until December 2003.

We check the same hypothesis on the German market for DAX and its 23 selected constituents over the most recent sample period from 20100802 until 20120801. Appendix A summarizes the results of a $t$-test for the null hypothesis that RV and MFIV are on average equal against the alternative RV<MFIV. $H_0$ is strongly rejected for DAX index, however for individual stocks the results differ: for 5 out of 23 stocks we cannot reject the $H_0$ at 5% significance level, for others the hypothesis over the studied 2-years interval was strongly rejected in favor of the alternative. Appendix A also reports sample averages of
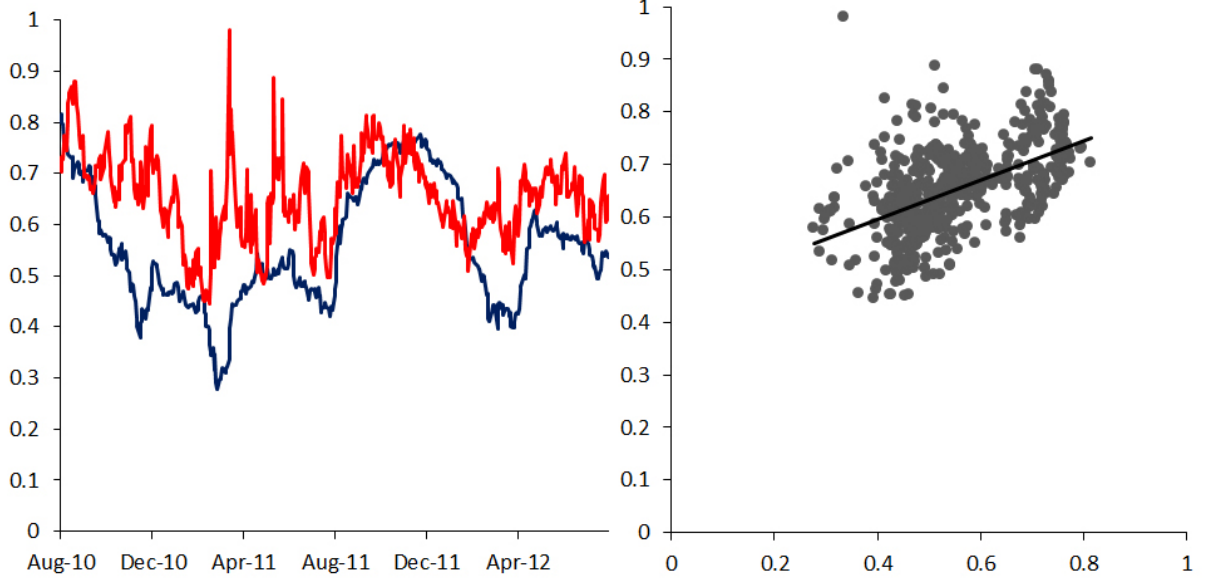
Figure 2: Left panel: $\rho_{t+\tau}$ (blue) $\tilde{\rho}_t(\tau)$ (red), right panel: scatter plot $\rho_{t+\tau}$ vs $\tilde{\rho}_t(\tau)$, for $t + 0.25$ from 20100802 till 20120801

RV and MFIV and their differences, which are found to be negative for all stocks and the DAX index.

Driessen et al. (2009) interpret their $t$-test results as indirect evidence that there exists a negative correlation risk premium (CRP). To see this in the DAX dataset we compute the model free implied correlation (MFIC) $\tilde{\rho}_t(\tau)$ from the MFIVs of DAX and its constituents and the realized correlation (RC) $\rho_{t+\tau}$ from the corresponding RV by applying(2):

$$\tilde{\rho}_t(\tau) = \frac{\tilde{\sigma}_{t,DAX}^2(\tau) - \sum_i w_i^2 \tilde{\sigma}_{t,i}^2(\tau)}{\sum_i \sum_{j \neq i} w_i w_j \tilde{\sigma}_{t,i}(\tau) \tilde{\sigma}_{t,j}(\tau)}, \tag{12}$$

$$\rho_{t+\tau} = \frac{\sigma_{t+\tau,DAX}^2 - \sum_i w_i^2 \sigma_{t+\tau,i}^2}{\sum_i \sum_{j \neq i} w_i w_j \sigma_{t+\tau,i} \sigma_{t+\tau,j}}. \tag{13}$$

Figure 2 plots the of MFIC and RC of DAX computed over the 3-month window and with 3 month maturity respectively ($\tau = 0.25$). The $H_0$ : RC=MFIC of the $t$-test is strongly rejected (Appendix B). Using this finding and taking into account results in the literature we would expect the CRP $\rho_{t+\tau} - \tilde{\rho}_t(\tau)$ to be negative most of the times. One of the ways of exploiting this observation is making a bet on the market correlation by entering a dispersion strategy.

## 2.3 Dispersion strategy with variance swaps

We study one of the variations of the dispersion strategy, which consists in selling the variance of the basket (DAX) and buying variances of basket constituents. We implement the strategy by taking a short position in the variance swap (8) on an index and long positions in variance swaps on its constituents with notional amounts proportional to index weights. The payoff a dispersion strategy at $t + \tau$ is then defined by

$$D_{t+\tau} = -\left\{\sigma_{t+\tau,B}^2 - \tilde{\sigma}_{t,B}^2(\tau)\right\} + \sum_{i=1}^{N} w_i \left\{\sigma_{t+\tau,i}^2 - \tilde{\sigma}_{t,i}^2(\tau)\right\}. \tag{14}$$

Then we apply (2) and rewrite (14) in the following form

$$D_{t+\tau} = \tilde{\rho}_t(\tau) \sum_i \sum_{j \neq i} w_i w_j \tilde{\sigma}_{t,i}(\tau)\tilde{\sigma}_{t,j}(\tau) - \rho_{t+\tau} \sum_i \sum_{j \neq i} w_i w_j \sigma_{t+\tau,i}\sigma_{t+\tau,j}. \tag{15}$$

Based on empirical findings described in Section 2.2 we assume $\tilde{\sigma}_{t,i}(\tau) \approx \sigma_{t+\tau,i}$ for each constituent stock and simplify the payoff (15), as follows

$$D_{t+\tau} \approx \sum_i \sum_{j \neq i} w_i w_j \tilde{\sigma}_{t,i}(\tau)\tilde{\sigma}_{t,j}(\tau) \left\{\tilde{\rho}_t(\tau) - \rho_{t+\tau}\right\}, \tag{16}$$

which illustrates that by entering the dispersion strategy one obtains exposure to $\rho_{t+\tau} - \tilde{\rho}_t(\tau)$, where the floating leg $\rho_{t+\tau}$ is computed with (11) and (2) at expiry, and the fixed leg $\tilde{\rho}_t(\tau)$ is a function of variance swap strikes (9). Test results described in Section 2.2 suggest that we should on average expect $\rho_{t+\tau} - \tilde{\rho}_t(\tau) < 0$. It also means the dispersion strategy with payoff $D_{t+\tau}$ on average would have a profit. However, as one can see in Figure 2, there might be days when $\rho_{t+\tau} - \tilde{\rho}_t(\tau) \geq 0$. In order to hedge against these potential losses one needs a forecast of the floating leg of the dispersion strategy.

# 3 Modeling and forecasting correlation dynamics

To determine the amount of hedge for $D_{t+\tau}$ we model the implied correlation (IC) and use the forecast to approximate the floating leg of the dispersion strategy $\rho_{t+\tau}$. By applying (2) to IV of a basket $\hat{\sigma}_{t,B}(\kappa, \tau)$ and its $N$ constituents $\hat{\sigma}_{t,i}(\kappa, \tau)$, $i \in \{1, \dots, N\}$, every $t$
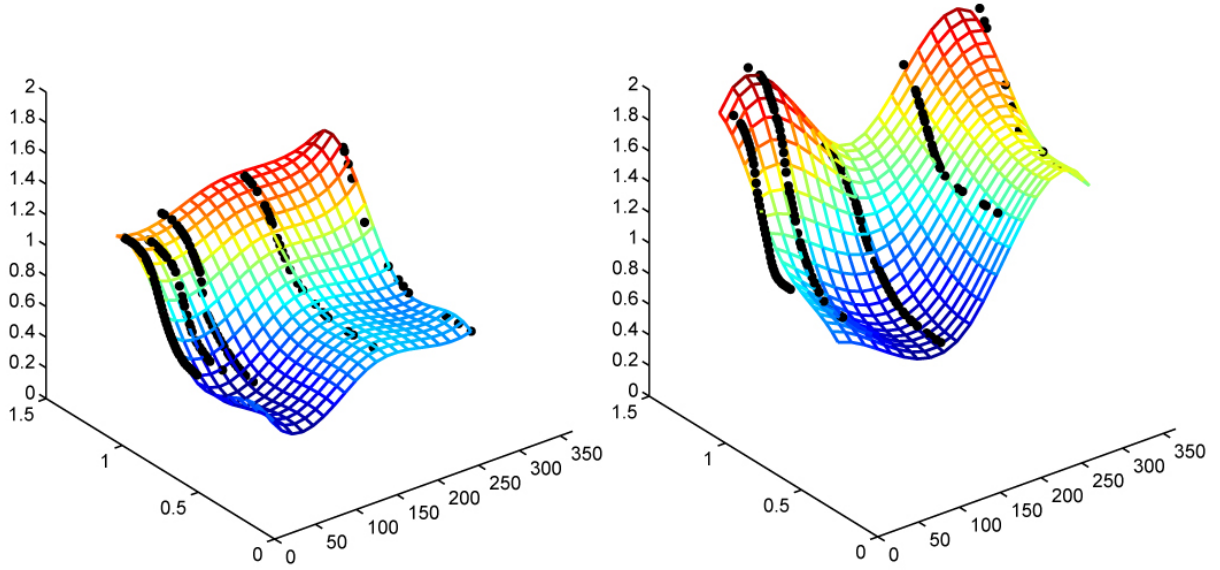
Figure 3: ICS implied by prices of DAX options traded on the 20111209, 20120710, surfaces recovered by the Nadaraya-Watson smoothing

we obtain the IC surface (ICS):

$$\widehat{\rho}_t(\kappa, \tau) = \frac{\widehat{\sigma}^2_{t,B}(\kappa, \tau) - \sum_i w_i^2 \widehat{\sigma}^2_{t,i}(\kappa, \tau)}{\sum_i \sum_{j \neq i} w_i w_j \widehat{\sigma}_{t,i}(\kappa, \tau) \widehat{\sigma}_{t,j}(\kappa, \tau)}. \tag{17}$$

Figure 3 displays $\widehat{\rho}_t(\kappa, \tau)$ in different trading days: 20111209, 20120710. Due to the specific option data structure, every day one observes a "cloud of strings" that visually resembles a surface and can be recovered by applying nonparametric smoothing. One can clearly see that surfaces have shape similarities and vary in levels, slopes and curvatures, and thus may be treated as daily realizations of a random function. In addition one can observe that the strings do not have fixed spacial locations. In order to model the dynamics of such a complicated multidimensional object we apply a dynamic semiparametric factor model that reduces the dimensionality of the problem and allows to study the ICS in a conventional time-series context.

## 3.1    Model Characterization

At every day $t$ one observes ICs $\widehat{\rho}(\kappa_{t,j}, \tau_{t,j})$, $t = 1, \ldots, T$, $j = 1, \ldots, J_t$ (index of observations at day $t$). Prior to introducing the model we exclude the case of a fully undiversified basket, with $\widehat{\rho} = 1$, from the analysis and apply a variance stabilizing transformation.

Fisher's Z-transformation (Härdle and Simar (2012)) gives:

$$T(u) \stackrel{\text{def}}{=} \frac{1}{2} \log \frac{1+u}{1-u} \tag{18}$$

with $Y_{t,j} \stackrel{\text{def}}{=} T\{\widehat{\rho}(\kappa_{t,j}, \tau_{t,j})\}$.

Our aim is to model the dynamics of $\{(Y_{t,j}, X_{t,j}), 1 \leq t \leq T, 1 \leq j \leq J_t\}$, where $X_{t,j} = (\kappa_{t,j}, \tau_{t,j})$. The technique we are employing allows to reduce the dimensionality and to simultaneously study the dynamics of $Y_t$ by approximation through an $L$-dimensional object with $L << J$. The DSFM, first introduced by Fengler et al. (2007) in application to IV surfaces dynamics, and then extended by Park et al. (2009) and Song et al. (2010) has these desired properties.

The basic idea is to approximate $\mathsf{E}(Y_t|X_t)$ by the sum of $L+1$ smooth basis functions $m \stackrel{\text{def}}{=} \{m_0, \ldots, m_L\}^\top$ (factor loadings) weighted by time dependent coefficients $Z_t \stackrel{\text{def}}{=} (1, Z_{t,1}, \ldots, Z_{t,L})^\top$ (factors):

$$Y_{t,j} = m_0(X_{t,j}) + \sum_{l=1}^{L} Z_{t,l} m_l(X_{t,j}) + \varepsilon_{t,j}. \tag{19}$$

In representation (19) $m$ are chosen data driven and do not have a particular (parametric) form.

Here two important remarks are appropriate. First, the unknown basis functions $m$ have to be estimated. Fengler et al. (2007) estimate both $m$ and $Z_t$ iteratively using kernel smoothing techniques, Park et al. (2009) approximate $m$ by tensor B-splines basis functions weighted by a coefficients matrix. Here we employ functional principal component analysis (FPCA) approach that will be described in Section 3.2. Nonparametric estimation procedure we use is introduced in Section 3.3. The basics for this technique is in Song et al. (2010).

The second issue is estimation of the latent factors $Z_t$. Having the data-driven basis $\widehat{m}_l$ in hand we can estimate daily factors by the ordinary least squares (OLS) method. Afterwards one fits the econometric model to $\widehat{Z}_t$, as it has been done by Cont and Da Fonseca (2002) and Hafner (2004), who fitted AR(1) to every $Z_{t,l}$, $l \in \{1, \ldots, L\}$, or in Fengler et al. (2007) who considered a multivariate VAR(2) process.

## 3.2 Correlation surface with FPCA

We are approximating the ICS by the sum of orthogonal functions. By doing so we involve the FPCA theory looking at the ICS as a stationary random function $f : \mathbb{R}^2 \to \mathbb{R}$.

Let $\mathscr{J} = [\kappa_{min}, \kappa_{max}] \times [\tau_{min}, \tau_{max}]$ the range of possible values of $\kappa_{t,j}$ and $\tau_{t,j}$. We introduce $(\rho_t)$, $t \in \{1, \ldots, T\}$, the sample of i.i.d. smooth random functions (surfaces). Every $\rho_t$ is a smooth map $\rho_t : \mathscr{J} \to \mathbb{R}$ and satisfies $\int_{\mathscr{J}} \mathsf{E}(\rho_t^2) < \infty$. Also for every $\rho_t$ we assume a well-defined mean function $\mu(u) = \mathsf{E}\{\rho_t(u)\}$ and existence of covariance function $\psi(u,v) = \mathsf{E}\left[\{\rho_t(u) - \mu(u)\}\{\rho_t(v) - \mu(v)\}\right]$. With $\phi(u,v) = \mathsf{E}\{\rho_t(u)\rho_t(v)\}$ the covariance function can be expressed as

$$\psi(u,v) = \phi(u,v) - \mu(u)\mu(v), \tag{20}$$

which can be also interpreted as a covariance coefficient of two points on the surface with coordinates $u$ and $v \in \mathscr{J}$. Since (20) is a symmetric positive definite function we can use it as a nucleus of the integral transform, performed by the linear operator. Define the covariance operator $\Gamma$:

$$(\Gamma f)(u) = \int_{\mathscr{J}} \psi(u,v)f(v)dv \tag{21}$$

that transforms $f$ into $(\Gamma f)$. $\Gamma$ is a symmetric positive operator with orthonormal eigenfunctions $\{\gamma_j\}_{j=1}^{\infty}$, $\gamma_j : \mathscr{J} \to \mathbb{R}$, and associated eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$. Now we can express (20) in terms of eigenfunctions and eigenvalues of the covariance operator $\Gamma$ by applying Mercer's theorem, e.g. Indritz (1963):

$$\psi(u,v) = \sum_{j=1}^{\infty} \lambda_j \gamma_j(u)\gamma_j(v). \tag{22}$$

Taking eigenfunctions $\{\gamma_j\}_{j=1}^{\infty}$ as basis, we represent $\rho_t(u) - \mu(u)$ as a generalized Fourier series with coefficients given by $\zeta_{tj} = \int_{\mathscr{J}} \{\rho_t(u) - \mu(u)\} \gamma_j(u)du$ called the $j$-th principal component score with $\mathsf{E}(\zeta_{tj}) = 0$, $\mathsf{E}(\zeta_{tj}^2) = \lambda_j$ and $\mathsf{E}(\zeta_{tj}\zeta_{ik}) = 0$ for $j \neq k$, Ramsay and Silverman (2010). Thus one may rewrite $\rho_t(u) - \mu(u)$ in the Karhunen-Loève form:

$$\rho_t(u) - \mu(u) = \sum_{j=1}^{\infty} \zeta_{tj}\gamma_j(u). \tag{23}$$

Here $\zeta_{tj}$ indicate how strong is the influence of the $j$-th basis function on the shape of

the $t$-th surface. The higher the score, the closer will the shape of $\rho_t$ resemble the shape of the $j$-th eigenfunction.

In practice one needs to take $L$ eigenfunctions to replace the infinite sum in (23) by the finite sum of $L$ basis functions, corresponding to the highest eigenvalues. One calls $\{\gamma_j\}_{j=1}^{L}$ the empirical orthonormal basis, Ramsay and Silverman (2010). In the next Section we discuss the estimation procedure for $\{\gamma_j\}_{j=1}^{L}$ as well as criteria for the $L$ selection.

## 3.3 Estimation Algorithm

In model (19) both $Z_t$ and $m$ have to be estimated. We do that in two steps.

At the **first step** we estimate the covariance operator introduced in Section (3.2) and take $\widehat{\mu}$ as $\widehat{m}_0$ and $\widehat{\gamma}_l$ as $\widehat{m}_l$, $l \in \{1, \ldots, L\}$ .

The covariance function (20) is estimated as described in Yao et al. (2005) and Hall et al. (2006). The procedure consists in least-squares fitting of two local linear models, for $\widehat{\mu}$ and $\widehat{\psi}$.

Given $u \in \mathscr{J}$ we choose $(\widehat{a}_\mu, \widehat{b}_\mu) = (a_\mu, b_\mu)$ to minimize

$$\sum_{t=1}^{T} \sum_{j=1}^{J_t} \{Y_{t,j} - a_\mu - b_\mu(u - X_{t,j})\}^2 \mathcal{K}_{h_\mu}(X_{t,j} - u), \tag{24}$$

and take $\widehat{\mu}(u) = \widehat{a}_\mu$. Then, given $u, v \in \mathscr{J}$ we choose $(\widehat{a}_\phi, \widehat{b}_{\phi,1}, \widehat{b}_{\phi,2}) = (a_\phi, b_{\phi,1}, b_{\phi,2})$ to minimize

$$\sum_{t=1}^{T} \sum_{j,k:1 \leq j \neq k \leq J_t} \{Y_{t,j}Y_{t,k} - a_\phi - b_{\phi,1}(u - X_{t,j}) - b_{\phi,2}(v - X_{t,k})\}^2 \tag{25}$$

$$\times \mathcal{K}_{h_\phi}(X_{t,j} - u) \mathcal{K}_{h_\phi}(X_{t,k} - v),$$

and take $\widehat{\phi}(u,v) = \widehat{a}_\phi$.

Here $\mathcal{K}_h$ denotes the two-dimensional product kernel, $\mathcal{K}_h(\bar{q}) = k_{h_1}(\bar{q}_1) \times k_{h_2}(\bar{q}_2)$, $h = (h_1, h_2)^\top$, based on one-dimensional $k_h(\bar{q}) = h^{-1}k(h^{-1}\bar{q})$. For our application we selected the quartic kernel, where $k(\bar{q}) = 15/16(1-\bar{q}^2)^2$ for $|\bar{q}| < 1$ and 0 otherwise. For both (24) and (25) kernel bandwidths $h_\mu = (h_{\mu,1}, h_{\mu,2})^\top$ and $h_\phi = (h_{\phi,1}, h_{\phi,2})^\top$ are to be selected. The procedure is described in Appendix C. Figure 4 shows an example of $\widehat{\mu}(u)$ estimated using the dataset described in Section 4 for DAX ICS from August 2010 until July 2011.
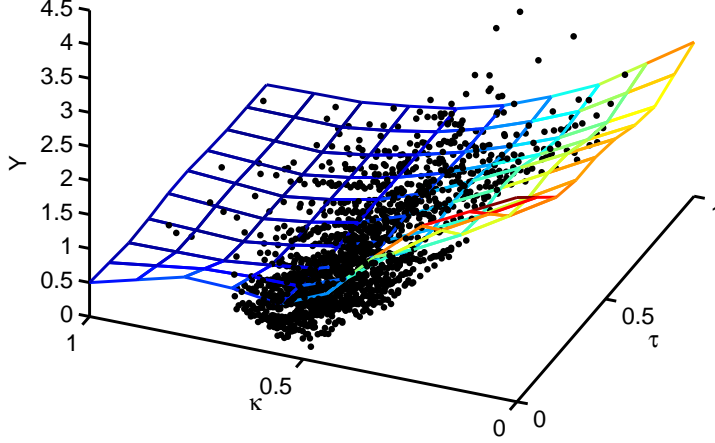
Figure 4: $\widehat{\mu}(u)$ of the DAX ICS with corresponding data points, estimated from 20100802 until 20110801 with $h_\mu = (h_{\mu,1}, h_{\mu,2})^\top = (0.12, 0.17)^\top$

Finally, having estimates $\widehat{\mu}(u)$ and $\widehat{\phi}(u,v)$, we compute $\widehat{\psi}(u,v)$ using (20) and take its $L$ eigenfunctions corresponding to the largest eigenvalues as $\widehat{m}_l$, $l \in \{1, \ldots, L\}$. Parameter $L$ is chosen in such a way that the selected eigenfunctions explain a big share of variability in original data. It is also necessary to mention that $\widehat{\psi}(u,v)$ is a matrix of a very large dimensionality. To obtain its consistent estimator, suitable for further spectral decomposition, various matrix regularization techniques can be used., e.g. banding as in Bickel and Levina (2008b), thresholding in Bickel and Levina (2008a), eigenvalues cleaning as in Laloux et al. (1999) and factor models described in Fan et al. (2008). We use the later in this step.

In the **second step** using $\widehat{m}$ we obtain the estimates $\widehat{Z}_t = (1, \widehat{Z}_{t,1}, \ldots, \widehat{Z}_{t,L})^\top$ as minimizers of the following least squares criterion:

$$\widehat{Z}_t = \arg \min_{Z_t} \sum_{t=1}^{T} \sum_{j=1}^{J_t} \left\{ Y_{t,j} - Z_t^\top \widehat{m}(X_{t,j}) \right\}^2. \tag{26}$$

## 4  Data

We study the dispersion strategy over the two years sample period from 20100802 to 20120801 on the German market represented by the DAX basket. The basket is composed of 23 stocks, constituents of DAX, with the most liquidly traded options and weights proportional to the current market capitalization. To model the dynamics of IC and construct the dispersion trade we operate with tree main variables representing different correlation estimates. MFIC, RC, and IC. The datasets are described in Table 1.

14

|        |                        | Min.   | Max.   | Mean   | Median | Stdd.  | Skewn.  | Kurt    |
|--------|------------------------|--------|--------|--------|--------|--------|---------|---------|
| IC     | $\kappa$               | 0.8000 | 1.2000 | 0.9825 | 0.9825 | 0.0986 | 0.0690  | 2.0661  |
|        | $\tau$                 | 0.0274 | 0.9671 | 0.2442 | 0.1753 | 0.1979 | 1.3717  | 4.3941  |
|        | $\widehat{\rho}_t(\kappa,\tau)$ | 0.0587 | 0.9998 | 0.6150 | 0.6290 | 0.1566 | -0.2739 | 2.6115  |
| MFIC   | $\tilde{\rho}_t(0.083)$ | 0.3895 | 0.4860 | 0.6061 | 0.6193 | 0.0834 | 0.0696  | 0.1957  |
|        | $\tilde{\rho}_t(0.25)$  | 0.4446 | 0.9795 | 0.6549 | 0.6573 | 0.0850 | 0.0613  | 0.1631  |
|        | $\tilde{\rho}_t(0.5)$   | 0.4997 | 1.4730 | 0.7037 | 0.6953 | 0.0866 | 1.8188  | 0.1305  |
|        | $\tilde{\rho}_t(1)$     | 0.5611 | 1.0851 | 0.7496 | 0.7422 | 0.0905 | 0.7764  | 0.6788  |
| RC     | $\rho_{t+0.083}$        | 0.1754 | 0.8955 | 0.5373 | 0.5013 | 0.1331 | 0.5221  | -0.2154 |
|        | $\rho_{t+0.25}$         | 0.2774 | 0.8149 | 0.5566 | 0.5363 | 0.1192 | 0.2489  | -0.8083 |
|        | $\rho_{t+0.5}$          | 0.3794 | 0.7343 | 0.5759 | 0.5713 | 0.1053 | -0.0243 | -1.4012 |
|        | $\rho_{t+1}$            | 0.4312 | 0.6581 | 0.5924 | 0.6050 | 0.0522 | -1.2443 | 0.9875  |

Table 1: Summary statistics: IC data computed from DAX index and constituents options over the period from 20090803 to 20120801 including the 1 year estimation period (3 years, 770 trading days, 135 obs./day). MFIC computed from daily variance swaps rates. RC computed from daily stock returns from 20100802 to 20120801 (2 years, 515 trading days). The figures are given after filtering and data preparation.

The *MFIC dataset* contains daily series of MFICs with maturities 0.083, 0.25, 0.5 and 1 years computed via (12) from variance swap rates given by Bloomberg as discrete approximation of (9).

The *RC dataset* contains daily series of RCs computed with (11) and (13) from the Bloomberg end-of-day stock prices over estimation windows 0.083, 0.25, 0.5 and 1 years.

The *IC dataset* is constructed using out-of-the-money (OTM) DAX and single stock options from the EUREX database. To estimate the DSFM model and produce forecasts for the sample period the dataset covers one additional year from 20090803 to 20100730. The dataset is transaction-based, meaning every trade is registered with the date it occurred, expiry date, underlying ticker, exercise price (strike) and settlement price . To obtain IV from option prices via (10) we distinguish between index and single stock options. For index options, which have the European type of option payoff, Black-Sholes (BS) model is used. To account for dividends and early execution in options on single stocks (American payoff) we use binomial trees, Cox et al. (1979), and bisection algorithm. Another necessary model parameters, such as stock prices, index levels, dividend amounts for constituent stocks, interest rates and stock market capitalization are taken from the Bloomberg database. As a risk free rate proxy we take daily values of EURIBOR (Euro Interbank Offered Rate) with 1 week up to 1 year maturities and use linear interpolation

to obtain values for required option $\tau$. We use the most liquid segment of data with $\kappa$ ranging from 0.8 to 1.2 and $\tau$ from 10 days to 1 year. Options from original EUREX dataset are not given on a regular $(\kappa, \tau)$-grid, required in (17). In $\tau$-dimension maturities are standardized by market regulation, so every $t$ one can find several $\tau_t$, similar for the index and all constituents. However, in $\kappa$-dimension one needs to interpolate. At every $t$ we use the original $(\kappa_t, \tau_t)$ grid of the index and linearly interpolate IVs of all constituents to obtain values corresponding to this grid. To avoid computational problems with highly skewed empirical distribution of $(\kappa_t, \tau_t)$, we transform the initial space $[0.8, 1.2] \times [0.03, 1]$ to $[0, 1]^2$ using empirical distribution function. Also we remove options with extremely high IVs (bigger than 50%) considering them the misprints in trade registration. After this we use (17) to obtain IC, which produces on average 135 observations per day.

Figure 1 shows there is a linear dependence between basket correlation and volatility. We check this finding in the RC dataset for different estimation windows and in IC dataset for different maturities. The RC data allows to identify a breakpoint, a threshold, after which the strength of the dependence changes, Appendix G. This phenomenon is persistent over different estimation windows. The IC dataset does not show any clear change in correlation/volatility dependence, Appendix H. Since the IC is used to obtain a forecast of a floating leg of the dispersion strategy, that is RC, we propose to make a regime dependent correction of the IC dataset as described in Appendix I.

# 5 Empirical results

## 5.1 Estimation Results and Factor Modeling

Using the IC dataset described in Section 4 we estimate the DSFM model for three non-overlapping subsamples 20090803 - 20100730, 20100802 - 20110729, 20110802 - 20120801, and for the entire sample 20090803 - 20120801. All subsamples include particularly volatile periods caused by the stock market falls in May 2010, "Flash Crash 2010", and more pronounced drop in August 2011.

After obtaining the common factor loadings $\widehat{m}_0, \widehat{m}_1, \widehat{m}_2, \widehat{m}_3$, Figure 5, and the daily time series of factors $\widehat{Z}_{t,1}, \widehat{Z}_{t,2}, \widehat{Z}_{t,3}$, Figure 6, the modeling task is simplified to the low-dimensional analysis of factor series. We fit the VAR model of order $p$ for $\widehat{Z}_{t,1}, \widehat{Z}_{t,2}, \widehat{Z}_{t,3}$. Before proposing a proper VAR specification, we check if $\widehat{Z}_t$ have characteristics that violate assumptions for linear multiple time series models. We perform the augmented Dickey-Fuller (ADF) test to check each $\widehat{Z}_{t,1}, \widehat{Z}_{t,2}, \widehat{Z}_{t,3}$ for stationarity, Appendix D. For $\widehat{Z}_{t,2}$ in subsample 20100802 - 20110729 we cannot reject the hypothesis of a unit root and use
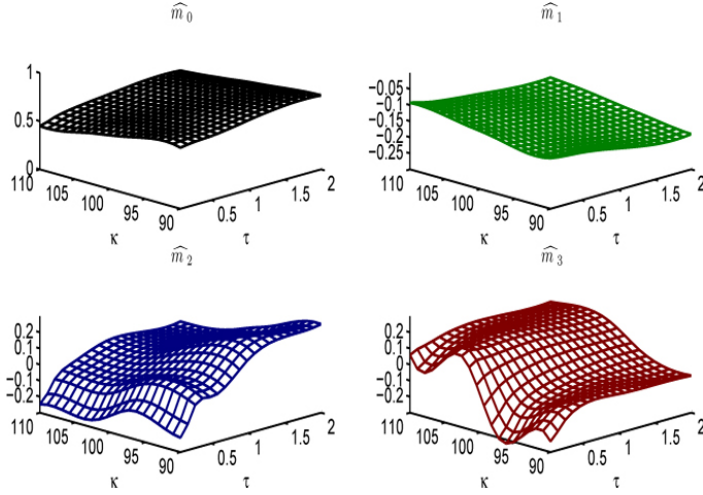
Figure 5: Factor loadings $\widehat{m}_0, \widehat{m}_1, \widehat{m}_2, \widehat{m}_3$ estimated from 20090803 to 20100730
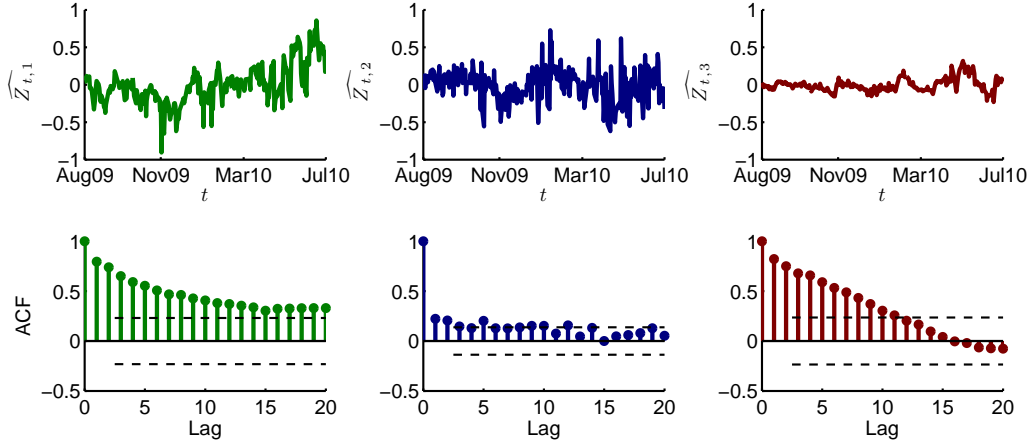


Figure 6: Driving factors of the DAX ICS $\widehat{Z}_{t,1}, \widehat{Z}_{t,2}, \widehat{Z}_{t,3}$ and ACF up to the 20th lag from 20090803 to 20100730

its first differences instead. Then we define the appropriate number of lags, or order $p$, by computing Akaike's information criterion (AIC), Schwarz's Bayesian information criterion (SBIC), and the Hannan and Quinn information criterion (HQIC) values, Appendix E. An * appearing next to the test statistics indicates the optimal lag. Except the subsample 20110802 - 20120801, the test statistics suggest $p = 2$, so we make a choice in favor of this specification. The estimation results are summarized in Appendix F. We also conducted a portmanteau (Q) test the null hypothesis that a series of residuals exhibits no autocorrelation. The test does not indicate the presence of serial correlation in residuals in subsample regressions.

We can clearly distinguish the influence of each factor on the time evolution of the ICS.

The first factor can be interpreted as level, the second as moneyness and the third as maturity effect. The relative size of the largest eigenvalues of the estimated (20) suggest that $\widehat{m}_1$ is capable to capture the biggest share of the surface variability. The variation captured by the second $\widehat{m}_2$ has a smaller influence, since it is only responsible for the surface shape transformation in the $\tau$ dimension. Finally, since the variation of the ICS in the $\kappa$ dimension is relatively small, the $\widehat{m}_3$ has a smaller impact, which is also reflected in the $\widehat{Z}_{t,3}$ series.

The forecast of $\widehat{Z}_{t,1}, \widehat{Z}_{t,2}, \widehat{Z}_{t,3}$ modeled with VAR(2) together with estimated fixed $\widehat{m}_0, \widehat{m}_1, \widehat{m}_2, \widehat{m}_3$ give a forecast of the ICS.

## 5.2   Backtesting the dispersion strategy

Here we show that using the correlation forecast one can improve the original dispersion strategy (14) and test it empirically over the 2-years sample period 20100801 - 20120802. We compare the payoff of the strategy without hedging with the *naïve hedging strategy* and propose its improvement, the *advanced strategy.*

To obtain the value of the naive hedge position to be held over $\Delta t$ days from $t + \tau - \Delta t$ till $t + \tau$ we make a $\Delta t$-days ahead DSFM forecast $\widehat{\rho}_{t+\tau}(1, t+\tau)$ and use it as $\rho_{t+\tau}$ in (14). Thus the size of the position is defined by

$$D_{t+\tau}^h = \sum_i \sum_{j \neq i} w_i w_j \tilde{\sigma}_{t,i}(\tau) \tilde{\sigma}_{t,j}(\tau) \left\{ \tilde{\rho}_t(\tau) - \widehat{\rho}_{t+\tau}(1, t+\tau) \right\}. \tag{27}$$

The corresponding relative hedging error is given by

$$\varepsilon_{t+\tau}^h = \frac{D_{t+\tau}^h - D_{t+\tau}}{D_{t+\tau}} = -\frac{\widehat{\rho}_{t+\tau}(1, t+\tau) - \rho_{t+\tau}}{\tilde{\rho}_t(\tau) - \rho_{t+\tau}}, \tag{28}$$

where $\varepsilon_{t+\tau}^h < 0 (> 0)$ means that the hedge (27) under-(over-)estimates the actual position (14). Table 2 gives summary statistics for the (28) over the studied sample period for 3 trades with four different maturities: 0.083, 0.25, 0.5 and 1 years. The statistic includes 515 trades originated every day and expired over the given 2 years sample period.

The improved version of the strategy uses the DSFM forecast $\widehat{\rho}_{t+\tau}(1, t+\tau)$ as a trigger which defines whether one should hedge or not. If $\widehat{\rho}_{t+\tau}(1, t+\tau) \geq \tilde{\rho}_t(\tau)$ (DSFM predicts loss in dispersion strategy), take an offsetting (with negative sign) position in (27); if $\widehat{\rho}_{t+\tau}(1, t+\tau) < \tilde{\rho}_t(\tau)$ (DSFM predicts gain in dispersion strategy), don't hedge. Thus we

| $\tau$ | Min. | Max. | Mean. | Median | Stdd. | Skew. | Kurt. |
|---|---|---|---|---|---|---|---|
| 0.083 | -108.04 | 72.30 | -1.14 | -0.71 | 8.00 | -6.61 | 100.49 |
| 0.25 | -255.48 | 49.53 | -1.20 | -0.41 | 11.49 | -17.58 | 372.33 |
| 0.5 | -216.04 | 32.78 | -0.74 | -0.30 | 9.37 | -18.66 | 425.86 |
| 1 | -64.84 | 76.59 | -0.01 | -0.38 | 7.47 | 2.74 | 46.85 |

Table 2: Performance of naïve hedge, summary statistics for $\varepsilon^h_{t+\tau}$ from 20100101 until 20120801

can write the payoff of the advanced strategy at $t + \tau$ as follows:

$$D^{adv}_{t+\tau} = \begin{cases} D_{t+\tau} - D^h_{t+\tau} & \text{,if} \quad \widehat{\rho}_{t+\tau}(1, t+\tau) \geq \tilde{\rho}_t(\tau) \\ D_{t+\tau} & \text{,if} \quad \widehat{\rho}_{t+\tau}(1, t+\tau) < \tilde{\rho}_t(\tau). \end{cases} \tag{29}$$

Since variance swap contract costs nothing to initiate (we ignore transactions costs), the presented series of daily payoffs correspond to daily P&L of the hypothetical trade where swaps expire daily over the whole period from 20100801 till 20120802. We compare the cash flows from three strategies. As one can see in Table 3, the advanced strategy outperforms the other two by having the smallest maximal losses, highest maximal gains ($\tau = 0.25, 0.5$) and the highest (second highest for $\tau = 1$) average payoff over the studied sample period.

# 6    Conclusions

In this study we investigated the implied correlation (IC) of the DAX index basket and introduced a hedging approach for the dispersion trading strategy using the IC forecast. We apply the dynamic semiparametric factor model (DSFM) to the IC dataset from January 2010 to August 2012, recover four basis functions and three time series of factors and use them to forecast the IC. The advanced dispersion strategy we employ using the IC forecast shows the smallest maximal losses, the highest maximal gains and the highest average payoff over the studied sample period and in comparison to the alternative strategies. So we conclude that our modeling approach can be of potential use in equity dispersion trading.

The choice of DSFM as a model for the IC surface (ICS) dynamics is motivated by the degenerated dataset design, which has to be modeled nonparametrically. On the other hand we were driven by the necessity to reduce the dimensionality of the problem

| Strategy | $\tau$ | Min. | Max. | Mean. | Stdd. |
|---|---|---|---|---|---|
| $D_{t+\tau}$ | 0.083 | -1502.58 | *1080.23* | 87.09 | 356.94 |
| (no hedge) | 0.25 | -1531.94 | *1282.31* | 101.92 | 440.54 |
| | 0.5 | -1270.90 | *1301.28* | 136.91 | 456.75 |
| | 1 | -872.76 | *760.92* | *134.26* | 299.01 |
| $D_{t+\tau} - D_{t+\tau}^h$ | 0.083 | -3237.72 | 617.40 | 15.35 | *203.09* |
| (naïve hedge) | 0.25 | -1726.53 | 413.28 | 35.90 | *110.14* |
| | 0.5 | -1301.47 | 344.91 | 41.13 | *91.91* |
| | 1 | -914.27 | 327.03 | 79.62 | *93.14* |
| $D_{t+\tau}^{adv}$ | 0.083 | *-1375.99* | 1011.38 | *100.93* | 256.50 |
| (advanced hedge) | 0.25 | *-1137.79* | 1282.31 | *195.09* | 248.41 |
| | 0.5 | *-760.85* | *1301.28* | *231.35* | 281.66 |
| | 1 | *-367.89* | 623.38 | 123.04 | 190.80 |

Table 3: Summary statistics for $D_{t+\tau}$ (no hedge), $D_{t+\tau} - D_{t+\tau}^h$ (naïve hedge), $D_{t+\tau}^{adv}$ (advanced hedge) from 20100101 until 20120801, best results (highest min, max,mean and smallest stdd.) are given in italic

and facilitate the forecasting. DSFM satisfies both requirements. It captures well the form of the ICS by its nonparametric part and allows using simple parametric model for dynamics. At the later modeling stage we fit the three-dimensional VAR(2) model, which as is a good choice to carry out the forecasting exercise. In addition we found that it is possible to separate the influence of each recovered basis function on the ICS shape. The functions allow their interpretation as level, moneyness and maturity effects. The strength of these effects is defined by the time series of corresponding factors, which can be characterized as drivers of the correlation risk. An interesting task would be to study the presence, size and magnitude of the correlation risk premia, captured by these factors. We consider this findings to be important topics for further research.

# References

Bai, Z. D. (1999). Methodologies in spectral analysis of large dimensional random matrices, *Statistica Sinica* **9**: 611–677.

Bakshi, G., Cao, C. and Chen, Z. (2000). Do call and underlying prices always move in the same direction?, *Review of Financial Studies* **13**(3): 549–584.

Bakshi, G., Kapadia, N. and Madan, D. (2003). Stock return characteristics, skew laws, and differential pricing of individual equity options, *Review of Financial Studies* **16**: 101–143.

Bickel, P. J. and Levina, E. (2008a). Covariance regularization by thresholding, *Ann. Statist.* **36**(6): 2577–2604.

Bickel, P. J. and Levina, E. (2008b). Regularized estimation of large covariance matrices, *Ann. Statist.* **36**(1): 199–227.

Blair, B. J., Poon, S.-H. and Taylor, S. J. (2001). Forecasting S&P100 volatility: the incremental information content of implied volatilities and high-frequency index returns, *Journal of Econometrics* **105**(1): 5 – 26.

Bollerslev, T., Engle, R. F. and Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances, *Journal of Political Economy* **96**(1): 116–31.

Bourgoin, F. (2001). Stress-testing correlations: An application to portfolio risk management, *in* C. Dunis, A. Timmermann and J. Moody (eds), *Developments in forecast combination and portfolio choice*, New York: Wiley.

Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices, *Journal of Business* **51**(4): 621–51.

Britten-Jones, M. and Neuberger, A. J. (2000). Option prices, implied price processes, and stochastic volatility, *Journal of Finance* **55**(2): 839–866.

Campa, J. M. and Chang, P. (1998). The forecasting ability of correlations implied in foreign exchange options, *Journal of International Money and Finance* **17**(6): 855 – 880.

Carr, P. and Madan, D. (1998). Towards a theory of volatility trading, *Reprinted in Option Pricing, Interest Rates, and Risk Management, Musiella, Jouini, Cvitanic*, University Press, pp. 417–427.

Carr, P. and Wu, L. (2009). Variance risk premiums, *Review of Financial Studies* **22(3)**(3): 1311–1341.

Christensen, B. and Prabhala, N. (1998). The relation between implied and realized volatility, *Journal of Financial Economics* **50**(2): 125 – 150.

Cont, R. and Da Fonseca, J. (2002). Dynamics of implied volatility surfaces, *Quantitative Finance* **2**: 45–60.

Cox, J. C., Ross, S. A. and Rubinstein, M. (1979). Option pricing: A simplified approach, *Journal of Financial Economics* **7**: 229–263.

Driessen, J., Maenhout, P. J. and Vilkov, G. (2009). The price of correlation risk: Evidence from equity options, *Journal of Finance* **9**: 1377–1406.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business & Economic Statistics* **20**(3): 339–50.

Engle, R. F., Shephard, N. and Sheppard, K. (2008). Fitting vast dimensional time-varying covariance models, *Economics Series Working Papers 403*, University of Oxford, Department of Economics.
**URL:** *http://ideas.repec.org/p/oxf/wpaper/403.html*

Fan, J., Fan, Y. and Lv, J. (2008). High dimensional covariance matrix estimation using a factor model, *Journal of Econometrics* **147**: 186–197.

Fengler, M. R., Härdle, W. K. and Mammen, E. (2007). A semiparametric factor model for implied volatility surface dynamics, *Journal of Financial Econometrics* **5**(2): 189–218.

Fleming, J. (1998). The quality of market volatility forecasts implied by S&P100 index option prices, *Journal of Empirical Finance* **5**(4): 317 – 345.

Hafner, R. (2004). *Stochastic Implied Volatility*, Springer Verlag, Heidelberg.

Hall, P., Müller, H.-G. and Wang, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis, *Ann. Statist.* **34**(3): 1493–1517.

Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*, Springer Verlag, Heidelberg.

Härdle, W. and Simar, L. (2012). *Applied Multivariate Statistical Analysis*, 3rd edn, Springer Verlag, Heidelberg.

Indritz, J. (1963). *Methods in analysis*, Macmillan New York.

Laloux, L., Cizeau, P., Bouchaud, J.-P. and Potters, M. (1999). Noise dressing of financial correlation matrices, *Physical Review Letters* **83**: 1467–1470.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance* **10**(5): 603–621.

Lopez, J. and Walter, C. (2000). Is implied correlation worth calculating? Evidence from foreign exchange options and historical data, *Journal of Derivatives* **7**(3): 65 – 81.

Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points, *Statistics in Medicine* **22**(19): 3055–3071.
**URL:** *http://dx.doi.org/10.1002/sim.1545*

Park, B., Mammen, E., Härdle, W. and Borak, S. (2009). Dynamic semiparametric factor models, *J. Am. Statist. Assoc.* **104(485)**: 284–298.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T. and Stanley, H. E. (2002). Random matrix approach to cross correlations in financial data, *Phys. Rev. E* **65**: 066126.
**URL:** *http://link.aps.org/doi/10.1103/PhysRevE.65.066126*

Ramsay, J. and Silverman, B. W. (2010). *Functional Data Analysis (Springer Series in Statistics)*, 2nd edn, Springer Verlag, Heidelberg.

Skintzi, V. and Refenes, A. (2005). Implied correlation index: A new measure of diversification, *Journal of Futures Markets* **25(2)**: 171–197.

Song, S., Härdle, W. K. and Ritov, Y. (2010). High dimensional nonstationary time series modelling with generalized dynamic semiparametric factor model, *SFB 649 Discussion Paper 2010-039*, Sonderforschungsbereich 649, Humboldt Universität zu Berlin, Germany. available at http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2010-039.pdf, resubmitted to Econometrics Journal.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data, *J. Am. Statist. Assoc.* **100**: 577–590.

# A  Variance risk premia in DAX and its constituents

|  | $\tau = 0.25$ | | | $\tau = 0.5$ | | | $\tau = 1$ | | |
|---|---|---|---|---|---|---|---|---|---|
|  | $\sigma$ | $\tilde{\sigma}$ | $\sigma - \tilde{\sigma}$ | $\sigma$ | $\tilde{\sigma}$ | $\sigma - \tilde{\sigma}$ | $\sigma$ | $\tilde{\sigma}$ | $\sigma - \tilde{\sigma}$ |
| Adidas AG | 27.62 | 31.26 | -3.64 | 28.33 | 31.88 | -3.55 | 28.60 | 31.96 | -3.37 |
| Allianz SE | 31.21 | 34.25 | -3.04 | 32.16 | 34.93 | -2.77 | 32.61 | 34.98 | -2.37 |
| BASF SE | 29.74 | 32.67 | -2.92 | 30.43 | 32.89 | -2.45 | 30.79 | 32.70 | -1.92 |
| Bayer AG | 28.13 | 31.92 | -3.78 | 28.28 | 32.01 | -3.73 | 28.19 | 31.77 | -3.58 |
| BMW AG | 35.74 | 39.01 | -3.27 | 36.22 | 39.84 | -3.63 | 35.92 | 39.61 | -3.69 |
| Commerzbank AG | 49.04 | 55.85 | -6.80 | 48.97 | 54.61 | -5.64 | 48.53 | 53.40 | -4.87 |
| Daimler AG | 34.86 | 39.13 | -4.27 | 35.68 | 39.83 | -4.15 | 35.82 | 39.88 | -4.05 |
| Deutsche Bank AG | 39.19 | 45.08 | -5.89 | 40.08 | 45.12 | -5.04 | 40.89 | 44.23 | -3.35 |
| Lufthansa AG | 32.15 | 35.57 | -3.42 | 32.96 | 35.80 | -2.84 | 33.05 | 36.01 | -2.96 |
| Deutsche Post AG | 24.98 | 28.25 | -3.27 | 25.71 | 28.64 | -2.93 | 26.23 | 29.22 | -2.98 |
| DeutscheTelekom AG | 24.02 | 26.93 | -2.91 | 24.57 | 27.07 | -2.50 | 24.78 | 27.94 | -3.16 |
| E.ON AG | 29.31 | 31.07 | -1.75 | 30.00 | 31.17 | -1.17 | 30.13 | 31.50 | -1.37 |
| Fresenius | 19.11 | 23.41 | -4.30 | 19.47 | 23.79 | -4.31 | 19.79 | 24.02 | -4.23 |
| Henkel AG | 23.48 | 27.23 | -3.74 | 23.97 | 27.46 | -3.49 | 24.25 | 27.61 | -3.36 |
| Infineon AG | 39.50 | 43.01 | -3.51 | 40.17 | 43.75 | -3.57 | 41.06 | 44.37 | -3.30 |
| Linde AG | 22.83 | 26.60 | -3.77 | 23.08 | 27.51 | -4.43 | 23.20 | 27.73 | -4.53 |
| Metro AG | 32.23 | 33.30 | -1.07 | 32.69 | 33.61 | -0.92 | 32.16 | 33.48 | -1.32 |
| Munich Re | 25.01 | 28.16 | -3.15 | 25.47 | 28.84 | -3.37 | 24.94 | 29.43 | -4.48 |
| RWE AG | 30.34 | 31.39 | -1.04 | 30.65 | 31.64 | -0.99 | 30.29 | 31.84 | -1.55 |
| SAP AG | 21.67 | 24.32 | -2.66 | 21.51 | 25.51 | -3.99 | 21.71 | 26.93 | -5.23 |
| Siemens AG | 25.23 | 29.79 | -4.56 | 26.22 | 30.77 | -4.55 | 27.64 | 30.92 | -3.28 |
| ThyssenKrupp AG | 37.68 | 39.23 | -1.55 | 37.97 | 39.57 | -1.60 | 37.78 | 39.57 | -1.79 |
| Volkswagen AG | 33.89 | 37.55 | -3.67 | 34.38 | 37.97 | -3.59 | 35.35 | 37.93 | -2.57 |
| DAX Index | 22.51 | 26.45 | -3.94 | 23.10 | 27.54 | -4.45 | 23.44 | 28.47 | -5.02 |

Table 4: Mean of the $\sqrt{RV}$ ($\sigma$) and $\sqrt{MFIV}$ ($\tilde{\sigma}$) and their difference $\sqrt{RV} - \sqrt{MFIV}$ ($\sigma - \tilde{\sigma}$), for DAX index and 23 selected constituent stocks computed over the time period 20100802 - 20120801 for 3 different maturities/estimation windows: $\tau = 0.25, 0.5, 1$)

|  | $\tau = 0.25$ | $\tau = 0.5$ | $\tau = 1$ |
|---|---|---|---|
| Allianz SE | 0.0563 | 0.0526 | 0.0225 |
| E.ON AG | 0.2519 | 0.3176 | 0.0814 |
| Metro AG | 0.2931 | 0.1884 | 0.0196 |
| RWE AG | 0.6322 | 0.5655 | 0.0707 |
| ThyssenKrupp AG | 0.1964 | 0.0700 | 0.0100 |

Table 5: The results of $t$-test for $H_0$ that on average $RV = MFIV$ against the alternative $RV < MFIV$ of stocks for which the the $H_0$ is not rejected at 5% significance level. Results are presented for DAX index and 23 selected constituent stocks computed over the time period 20100802 - 20120801 for 3 different maturities/estimation windows: $\tau = 0.25, 0.5, 1$)

# B  Realized versus model free implied correlation

|  | $\tau = 0.25$ | $\tau = 0.5$ | $\tau = 1$ |
|---|---|---|---|
| $\rho$ | 0.5566 | 0.5759 | 0.5924 |
| $\tilde{\rho}$ | 0.6549 | 0.7037 | 0.7496 |
| $\rho - \tilde{\rho}$ | -0.0983 | -0.1278 | -0.1572 |
| p-value | 0.0000 | 0.0000 | 0.0000 |

Table 6: The results of $t$-test for $H_0$ that on average $RC = MFIC$ against the alternative $RC < MFIC$, which is rejected for all maturities/estimation windows: $\tau = 0.25, 0.5, 1$ at 5% significance level. RC and MFIC are computed using (2), (11) and (9) correspondingly over the time period 20100802 - 20120801

# C    Smoothing Parameters Selection

For both (24) and (25) kernel bandwidths $h_\mu = (h_{\mu,1}, h_{\mu,2})^\top$ and $h_\phi = (h_{\phi,1}, h_{\phi,2})^\top$ are to be selected. As suggested in Härdle et al. (2004) we use the penalizing function approach to select optimal $h_\mu^{opt}$, minimizing mean integrated squared error (MISE):

$$\frac{1}{T} \sum_{t=1}^{T} \frac{1}{J_t} \sum_{j=1}^{J_t} \left\{ Y_{t,j} - \sum_{l=1}^{L} \widehat{Z}_{t,l} \widehat{m}_l(X_{t,j}) \right\}^2 w_{h^*,t}(X_{t,j}) \Xi_{AIC} \left\{ \frac{W_{h^*,t,j}(X_{t,j})}{T J_t} \right\}, \qquad (30)$$

with the Akaike (1970) Information Criterion (AIC) as penalizing function $\Xi_{AIC}(q) = \exp(2q)$ and $W_{h^*,t,j}(X_{t,j})$ defined by

$$W_{h^*,t,j}(X_{t,j}) = \frac{\mathcal{K}_h(0)}{J_t^{-1} \sum_{k=1}^{J_t} \mathcal{K}_h(X_{t,k} - X_{t,j})}, \qquad (31)$$

for every $X_{t,j}$, $1 \leq t \leq T$, $1 \leq j \leq J_t$.

Since the distribution of the observations is very uneven, we are using the weighted version of the criterion with weights $w_{h^*,t}(\bar{u}) \stackrel{\text{def}}{=} p_{h^*,t}^{-1}(\bar{u})$, where $p_{h^*,t}(\bar{u})$ is the average design density. For every $X_{t,j}$, $1 \leq t \leq T$, $1 \leq j \leq J_t$ it is defined by:

$$p_{h^*,t}(X_{t,j}) = J_t^{-1} \sum_{k=1}^{J_t} \mathcal{K}_h(X_{t,k} - X_{t,j}), \qquad (32)$$

The bandwidth $h_{\mu AIC}^{opt} = (h_{\mu 1}, h_{\mu 2})^\top$ corresponding to the minimal criterion 30 is taken as optimal. The bandwidth $h^*$ of the weighting function is constant and does not depend of choice of $h_\mu$.

# D    Testing Factor Time Series for Stationarity

|  | $\widehat{Z}_{t,1}$ | $\widehat{Z}_{t,2}$ | $\widehat{Z}_{t,3}$ |
|---|---|---|---|
| 20090803 - 20100730 (1st year) | -2.991 (1) | -6.982 (1) | -5.710 (3) |
| 20100802 - 20110729 (2nd year) | -1.666* (3) | -3.090 (2) | -4.480 (1) |
| 20110802 - 20120801 (3rd year) | -3.511 (2) | -3.796 (3) | -3.480 (2) |
| 20090803 - 20120801 (entire sample) | -4.025 (1) | -6.912 (3) | -8.979 (1) |

Table 7: Augmented Dickey-Fuller (ADF) test on $\widehat{Z}_{t,1}, \widehat{Z}_{t,2}, \widehat{Z}_{t,3}$. Number of lags included in the ADF regression (in brackets) is chosen by starting with 3 lags and subsequently deleting lag terms, until the last one is significant at 5% level. Test statistics that does not reject the hypothesis of a unit root at 5% level are denoted by *.

# E    Determining Number of Lags for VAR Model

|  |  | AIC | HQIC | SBIC |
|---|---|---|---|---|
| 20090803 - 20100730 (1st year) | 1 | 1.923 | 2.061 | 2.162 |
|  | 2 | 1.839* | 1.975* | 2.152* |
|  | 3 | 1.856 | 2.052 | 2.304 |
|  | 4 | 1.882 | 2.060 | 2.389 |
| 20100802 - 20110729 (2nd year) | 1 | -2.868 | -2.800 | -2.699 |
|  | 2 | -3.075* | -2.932* | -2.755* |
|  | 3 | -3.068 | -2.898 | -2.645 |
|  | 4 | -3.051 | -2.854 | -2.525 |
| 20110802 - 20120801 (3rd year) | 1 | -0.118 | -0.051 | 0.048 |
|  | 2 | -0.355 | -0.238* | -0.064* |
|  | 3 | 0.361* | -0.193 | 0.055 |
|  | 4 | -0.360 | -0.144 | 0.179 |
| 20090803 - 20120801 (entire sample) | 1 | 0.745 | 0.773 | 0.818 |
|  | 2 | 0.384* | 0.461* | 0.539* |
|  | 3 | 0.397 | 0.467 | 0.579 |
|  | 4 | 0.412 | 0.475 | 0.621 |

Table 8: Akaike's information criterion (AIC), Schwarz's Bayesian information criterion (SBIC), and the Hannan and Quinn information criterion (HQIC) for defining the optimal lag order $p$ of a VAR model for DAX and S&P100 ICS factors $\widehat{Z}_{t,1}, \widehat{Z}_{t,2}, \widehat{Z}_{t,3}$. * appearing next to the test statistics indicates the optimal lag at 5% significance level.

# F   Parameters of VAR model for DAX ICS factors

| | 20090803 - 20120801 (entire sample) | | | 20090803 - 20100730 (1st year) | | |
|---|---|---|---|---|---|---|
| | $Z_{1,t}$ | $Z_{2,t}$ | $Z_{3,t}$ | $Z_{1,t}$ | $Z_{2,t}$ | $Z_{3,t}$ |
| $Z_{1,t-1}$ | 0.645* | -0.012 | -0.019 | 0.630* | -0.032 | 0.029 |
| $Z_{1,t-2}$ | 0.310* | 0.008 | 0.029 | 0.276* | 0.013 | -0.060* |
| $Z_{2,t-1}$ | -0.104* | 0.259* | 0.156 | -0.036 | 0.047 | 0.036 |
| $Z_{2,t-2}$ | 0.057* | 0.406* | -0.014 | -0.039 | 0.339 | -0.104* |
| $Z_{3,t-1}$ | -0.07 | 0.140* | 0.471* | -0.091 | -0.494* | 0.525* |
| $Z_{3,t-2}$ | 0.149* | 0.118* | 0.251* | 0.046 | 0.181 | -0.208* |
| $c$ | 0.004 | 0.006 | -0.003 | -0.004 | -0.001 | 0.001 |

| | 20100802 - 20110729 (2nd year) | | | 20110802 - 20120801 (3rd year) | | |
|---|---|---|---|---|---|---|
| | $Z_{1,t}$ | $Z_{2,t}$ | $Z_{3,t}$ | $Z_{1,t}$ | $Z_{2,t}$ | $Z_{3,t}$ |
| $Z_{1,t-1}$ | 0.809* | 0.048 | -0.202* | 0.339* | 0.191* | -0.001 |
| $Z_{1,t-2}$ | 0.254* | -0.029 | 0.112* | 0.351* | 0.041 | -0.036 |
| $Z_{2,t-1}$ | 0.188* | 0.687* | -0.223* | 0.355* | 0.264* | 0.132* |
| $Z_{2,t-2}$ | 0.091 | 0.262* | 0.018 | 0.084 | 0.302* | -0.045 |
| $Z_{3,t-1}$ | 0.453* | 0.051 | 0.162* | -0.197 | -0.008 | 0.623* |
| $Z_{3,t-2}$ | 0.118 | 0.118 | 0.275* | 0.044 | 0.240* | 0.204* |
| $c$ | -0.004 | 0.001 | -0.002 | 0.003 | -0.001 | -0.002 |

Table 9: The estimated parameters for VAR(2) model for DAX ICS factors. * marks estimates which are not significant at 5% level.

# G  Piecewise linear dependence of $\rho_{t+\tau}$ and $\sigma_{B,t+\tau}$
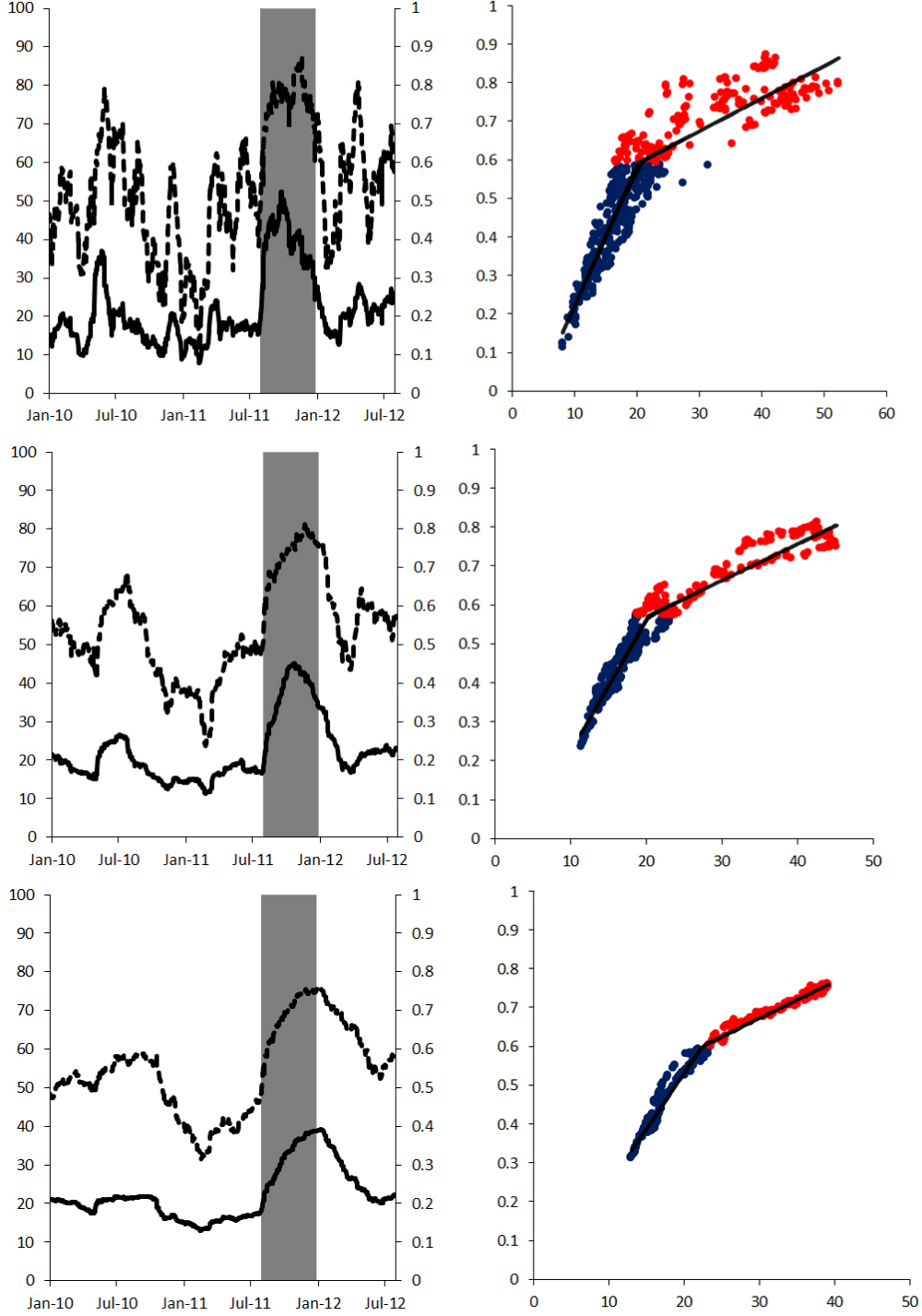


Figure 7: DAX $\sigma_{B,t+\tau}$ (solid line) vs $\rho_{t+\tau}$ (dashed line), scatter plot $\sigma_{B,t+\tau}$ vs $\rho_{t+\tau}$, for $t+\tau$ from 20100104 till 20120801, estimated with (11) and (2) respectively over 1 month ($\tau = 0.083$), 3 months ($\tau = 0.25$) and 6 months ($\tau = 0.5$) window. Shaded area: Aug 2011 market fall. Switch point for two regression line is defined as described in Appendix I

# H Linear dependence of $\widehat{\sigma}_{t,B}(\kappa, \tau)$ and $\widehat{\rho}_t(\kappa, \tau)$
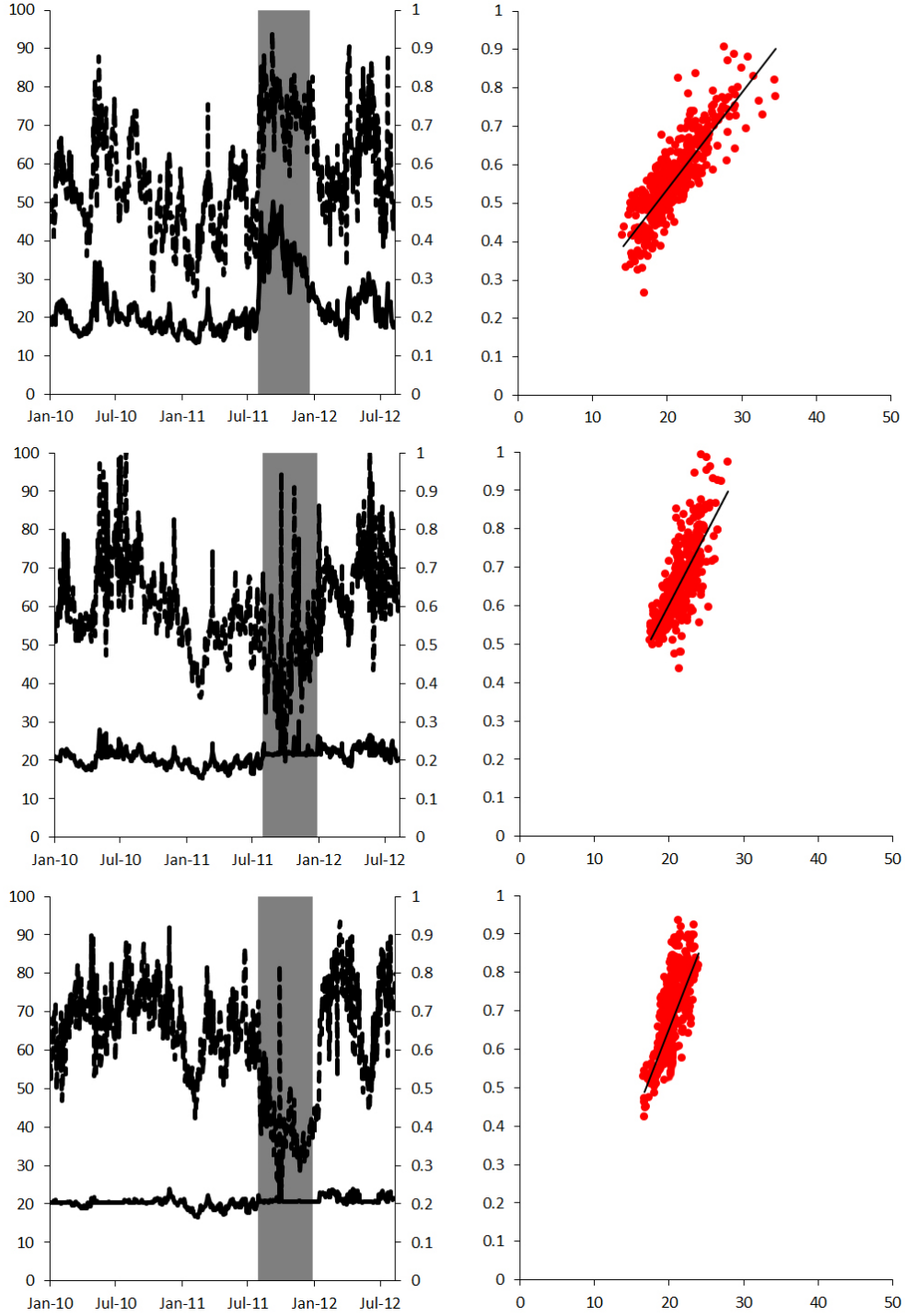


Figure 8: DAX $\widehat{\sigma}_{t,B}(1, \tau)$ (solid line) vs $\widehat{\rho}_t(1, \tau)$ (dashed line), scatter plot $\widehat{\sigma}_{t,B}(1, \tau)$ and $\widehat{\rho}_t(1, \tau)$, for $t + \tau$ from 20100104 till 20120801, estimated from IVs with (17) for option with 1 month ($\tau = 0.083$), 3 months ($\tau = 0.25$) and 6 months ($\tau = 0.5$) maturity. Shaded area: Aug 2011 market fall.

# I   Switch point selection for correlation regimes

| $\tau$ | $\sigma_{B,t+\tau}$ | $\rho_{t+\tau}$ | Slope 1 | Slope 2 |
|---|---|---|---|---|
| 0.083 | 20.24 | 0.5917 | 0.0361 | 0.0085 |
| 0.25 | 20.34 | 0.5728 | 0.0336 | 0.0093 |
| 0.5 | 22.42 | 0.6008 | 0.0286 | 0.0094 |
| Average | 21.00 | 0.5884 | 0.0328 | 0.0091 |

Table 10: Segmented linear regression of $\rho_{t+\tau}$ on $\sigma_{B,t+\tau}$ with one break point, $\tau = 0083, 0.25, 0.5$ for $t + \tau$, from 20100104 till 20120801

The dependence of $\rho$ and $\sigma_B$ observed in RV and RC (Appendix G) is not pronounced in case of ATM IV and IC (Appendix H). Therefore we propose a market regime correction scheme for the IC dataset. The idea is to find a breakpoint between two segments of a piecewise linear regression of $\rho_{t+\tau}$ on $\sigma_{B,t+\tau}$. Using the procedure described in Muggeo (2003) we fit a segmented linear regression with one break point.

Based on results summarized in Table 10 we make a following state dependent correction: if $\widehat{\sigma}_{B,t}(1,\tau) > 21$ (high volatility regime), then $\widehat{\rho}_t(\kappa,\tau) = 0.0091\widehat{\sigma}_{B,t}(\kappa,\tau)$
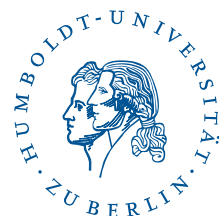
# SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649,
please visit http://sfb649.wiwi.hu-berlin.de.

001   "HMM in dynamic HAC models" by Wolfgang Karl Härdle, Ostap Okhrin and Weining Wang, January 2012.
002   "Dynamic Activity Analysis Model Based Win-Win Development Forecasting Under the Environmental Regulation in China" by Shiyi Chen and Wolfgang Karl Härdle, January 2012.
003   "A Donsker Theorem for Lévy Measures" by Richard Nickl and Markus Reiß, January 2012.
004   "Computational Statistics (Journal)" by Wolfgang Karl Härdle, Yuichi Mori and Jürgen Symanzik, January 2012.
005   "Implementing quotas in university admissions: An experimental analysis" by Sebastian Braun, Nadja Dwenger, Dorothea Kübler and Alexander Westkamp, January 2012.
006   "Quantile Regression in Risk Calibration" by Shih-Kang Chao, Wolfgang Karl Härdle and Weining Wang, January 2012.
007   "Total Work and Gender: Facts and Possible Explanations" by Michael Burda, Daniel S. Hamermesh and Philippe Weil, February 2012.
008   "Does Basel II Pillar 3 Risk Exposure Data help to Identify Risky Banks?" by Ralf Sabiwalsky, February 2012.
009   "Comparability Effects of Mandatory IFRS Adoption" by Stefano Cascino and Joachim Gassen, February 2012.
010   "Fair Value Reclassifications of Financial Assets during the Financial Crisis" by Jannis Bischof, Ulf Brüggemann and Holger Daske, February 2012.
011   "Intended and unintended consequences of mandatory IFRS adoption: A review of extant evidence and suggestions for future research" by Ulf Brüggemann, Jörg-Markus Hitz and Thorsten Sellhorn, February 2012.
012   "Confidence sets in nonparametric calibration of exponential Lévy models" by Jakob Söhl, February 2012.
013   "The Polarization of Employment in German Local Labor Markets" by Charlotte Senftleben and Hanna Wielandt, February 2012.
014   "On the Dark Side of the Market: Identifying and Analyzing Hidden Order Placements" by Nikolaus Hautsch and Ruihong Huang, February 2012.
015   "Existence and Uniqueness of Perturbation Solutions to DSGE Models" by Hong Lan and Alexander Meyer-Gohde, February 2012.
016   "Nonparametric adaptive estimation of linear functionals for low frequency observed Lévy processes" by Johanna Kappus, February 2012.
017   "Option calibration of exponential Lévy models: Implementation and empirical results" by Jakob Söhl und Mathias Trabs, February 2012.
018   "Managerial Overconfidence and Corporate Risk Management" by Tim R. Adam, Chitru S. Fernando and Evgenia Golubeva, February 2012.
019   "Why Do Firms Engage in Selective Hedging?" by Tim R. Adam, Chitru S. Fernando and Jesus M. Salas, February 2012.
020   "A Slab in the Face: Building Quality and Neighborhood Effects" by Rainer Schulz and Martin Wersing, February 2012.
021   "A Strategy Perspective on the Performance Relevance of the CFO" by Andreas Venus and Andreas Engelen, February 2012.
022   "Assessing the Anchoring of Inflation Expectations" by Till Strohsal and Lars Winkelmann, February 2012.

# SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649,
please visit http://sfb649.wiwi.hu-berlin.de.

# SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649,
please visit http://sfb649.wiwi.hu-berlin.de.

045    "Additive Models: Extensions and Related Models." by Enno Mammen, Byeong U. Park and Melanie Schienle, July 2012.

046    "A uniform central limit theorem and efficiency for deconvolution estimators" by Jakob Söhl and Mathias Trabs, July 2012

047    "Nonparametric Kernel Density Estimation Near the Boundary" by Peter Malec and Melanie Schienle, August 2012

048    "Yield Curve Modeling and Forecasting using Semiparametric Factor Dynamics" by Wolfgang Karl Härdle and Piotr Majer, August 2012

049    "Simultaneous test procedures in terms of p-value copulae" by Thorsten Dickhaus and Jakob Gierl, August 2012

050    "Do Natural Resource Sectors Rely Less on External Finance than Manufacturing Sectors? " by Christian Hattendorff, August 2012

051    "Using transfer entropy to measure information flows between financial markets" by Thomas Dimpfl and Franziska J. Peter, August 2012

052    "Rethinking stock market integration: Globalization, valuation and convergence" by Pui Sun Tam and Pui I Tam, August 2012

053    "Financial Network Systemic Risk Contributions" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, August 2012

054    "Modeling Time-Varying Dependencies between Positive-Valued High-Frequency Time Series" by Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig, September 2012

055    "Consumer Standards as a Strategic Device to Mitigate Ratchet Effects in Dynamic Regulation" by Raffaele Fiocco and Roland Strausz, September 2012

056    "Strategic Delegation Improves Cartel Stability" by Martijn A. Han, October 2012

057    "Short-Term Managerial Contracts and Cartels" by Martijn A. Han, October 2012

058    "Private and Public Control of Management" by Charles Angelucci and Martijn A. Han, October 2012

059    "Cartelization Through Buyer Groups" by Chris Doyle and Martijn A. Han, October 2012

060    "Modelling general dependence between commodity forward curves" by Mikhail Zolotko and Ostap Okhrin, October 2012

061    "Variable selection in Cox regression models with varying coefficients" by Toshio Honda and Wolfgang Karl Härdle, October 2012

062    "Brand equity – how is it affected by critical incidents and what moderates the effect" by Sven Tischer and Lutz Hildebrandt, October 2012

063    "Common factors in credit defaults swaps markets" by Yi-Hsuan Chen and Wolfgang Karl Härdle, October 2012

064    "Measuring the impact of critical incidents on brand personality" by Sven Tischer, October 2012

065    "Covered bonds, core markets, and financial stability" by Kartik Anand, James Chapman and Prasanna Gai, October 2012

066    "Implied Basket Correlation Dynamics" by Wolfgang Karl Härdle and Elena Silyakova, November 2012