

# CDO Surfaces Dynamics

Barbara Choroś-Tomczyk\*  
Wolfgang Karl Härdle\*  
Ostap Okhrin\*



\* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>  
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin  
Spandauer Straße 1, D-10178 Berlin



# CDO Surfaces Dynamics\*

Barbara Choroś-Tomczyk<sup>†‡</sup>, Wolfgang Karl Härdle<sup>†§</sup>, Ostap Okhrin<sup>†</sup>

July 2, 2013

## Abstract

Modelling the dynamics of credit derivatives is a challenging task in finance and economics. The recent crisis has shown that the standard market models fail to measure and forecast financial risks and their characteristics. This work studies risk of collateralized debt obligations (CDOs) by investigating the evolution of tranche spread surfaces and base correlation surfaces using a dynamic semiparametric factor model (DSFM). The DSFM offers a combination of flexible functional data analysis and dimension reduction methods, where the change in time is linear but the shape is nonparametric. The study provides an empirical analysis based on iTraxx Europe tranches and proposes an application to curve trading strategies. The DSFM allows us to describe the dynamics of all the tranches for all available maturities and series simultaneously which yields better understanding of the risk associated with trading CDOs and other structured products.

*Keywords:* base correlation, collateralized debt obligation, curve trade, dynamic factor model, semiparametric model.

*JEL classification:* C14, C51, G11, G17

## 1 Introduction

The recent financial crisis began in 2007 with the subprime mortgage crisis in U.S. Then it spread globally and gathered intensity in 2008. The financial system weakened and remained frozen. In 2009 the global economy stabilized but has not returned to its pre-crisis levels.

Collateralized debt obligations (CDOs) played a significant role in the global financial crisis. A CDO is a credit derivative used by financial institutions to repackage individual assets into a product that can be sold to investors on the secondary market. The

---

\*The financial support from the German Research Foundation, Project HA2229/7-2 and via CRC 649 *Economic Risk* at Humboldt-Universität zu Berlin is gratefully acknowledged.

<sup>†</sup>Ladislav 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.

<sup>‡</sup>Corresponding author. Email: barbara.choros@wiwi.hu-berlin.de

<sup>§</sup>School of Business, Singapore Management University, 50 Stamford Road, Singapore 178899

assets may be mortgages, auto loans, credit card debt, corporate debt or credit default swaps (CDS). CDOs were initially constructed for securitization of big portfolios. The entire portfolio risk is sliced into tranches and then transferred to investors. Prior to the credit crisis, CDOs provided outstanding investment opportunities to market participants. Banks used CDOs to reduce the amount of debt on their balance sheets. Tranching made it possible to create new securities of different risk classes that met the needs of a wide range of clients. The market observed an excess demand for senior CDO tranches because they were considered as safe and offered unusual high returns. As we know now, the rating agencies underestimated default risk of CDOs. Consequently, investors were exposed to more risk than the ratings of these CDOs implied. When the market collapsed, CDO investors faced enormous losses that led some of them to bankruptcy (Lehman Brothers) or a takeover (Bear Stearns, Merrill Lynch) by another institutions. The CDO market has significantly shrunk the beginning of the financial crisis. However, the methodology proposed in our study can be used in modelling and trading other financial instruments, especially non-standardised and bespoke structured products.

Developments in the CDO research mainly concern finding an accurate and flexible pricing model, see a comparison of popular models in Bluhm & Overbeck (2006) and Burtshell, Gregory & Laurent (2009). Papers that investigate implied correlations concentrate on the replication of the shape of the implied curve, e.g. Ağca, Agrawal & Islam (2008). Primarily because of the high dimensionality of the CDO problem the vast majority of papers consider only CDOs of one particular maturity, see e.g. Hamerle, Igl & Plank (2012). Up to our knowledge, the available literature do not look at the CDO market as a whole. Since CDOs are quoted for different maturities, we should consider the effect of the CDO term structure.

The empirical research of this study was performed using the iTraxx indices and their tranches of Series 2 to 10. The iTraxx Europe is the most widely traded credit index in Europe. Its reference portfolio consists of 125 equally weighted, most liquid credit default swaps (CDS) on European companies. Twice a year, every March and September, a new series of iTraxx is issued and the current constituents are reconsidered. The iTraxx Europe trades at maturities of 3, 5, 7 and 10 years. The tranches with 5 years maturity are the most liquid, unlike those with 3 years maturity that are rarely quoted. Because of the regular index roll, each day one observes on the market tranches with various times to expiration. By plotting prices of all available tranches at one day as a function of the time to maturity and the tranche seniority, one gets a surface that represents entire market information about spreads, see Figure 1. Similarly, an implied base correlation surface can be constructed. When we record these surfaces every day, we can follow how they change their shape and level. The dynamics over time of such surfaces is the main goal of this paper. From an investor's point of view, it is desirable to have an insight into the behaviour of spreads and base correlations in the future. The forecasting has useful applications in hedging and trading CDOs, like computation of Greeks, risk measures, or construction of investment strategies. One of the simplest solutions would be to consider the classic time series analysis for each tranche of each series for every maturity. Due to illiquidity and due to the short history of particular series, this methodology is not applicable. Thus, the natural choice for CDO data is to first interpolate observations

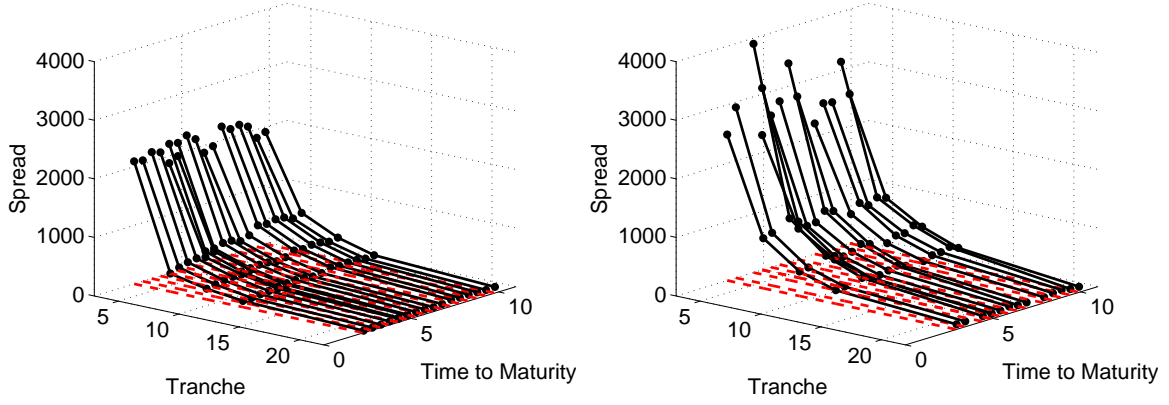


Figure 1: Spreads of all tranches of all series observed on 20080909 (left) and 20090119 (right).

for each time point by creating a smooth surface and afterwards to forecast this new high-dimensional object.

High-dimensional data sets naturally appear in many fields of science ranging from finance to genetics. Worth mentioning statistical approaches for handling complex high-dimensional problems are a structural analysis of curves by Kneip & Gasser (1992), a stochastic warping model by Liu & Müller (2004), penalized splines by Kauermann, Krivobokova & Fahrmeir (2009), and a functional principal components approach by Gromenko, Kokoszka, Zhu, & Sojka (2012). One of the most popular methods are factor type models as they effectively reduce the dimensionality. Factor models assume that the comovements of big number of variables are generated by a small set of latent factors. When data disclose a dynamic structure then one needs a technique that is able to correctly detect and describe the observed behaviour, e.g. Gouriéroux & Jasiak (2001), Hallin & Liška (2007), Koopman, Lucas & Schwaab (2012). In this study we employ a dynamic semiparametric factor model (DSFM).

In the DSFM the observed variables are expressed as linear combinations of the factors. The factors and the factor loadings are estimated from the data. The first ones represent the spatial, time-invariant component. The later ones form multidimensional time series that reflect the dynamics. The inference on the original variables reduces to the inference on the factors and the factor loadings.

The DSFM was introduced by Fengler, Härdle & Mammen (2007) for modelling the dynamics of implied volatility surfaces. Further, Giacomini, Härdle & Krättschmer (2009) apply the DSFM to analyse risk neutral densities implied by prices of DAX options, Härdle, Hautsch & Mihoci (2012) use it for limit order book analysis, Detlefsen & Härdle (2013) for variance swaps, and Myšičková, Song, Majer, Mohr, Heekeren & Härdle (2013) for fMRI images. In this work we study the dynamics of CDO surfaces with the DSFM and propose an application to curve trading strategies.

The paper is structured as follows. Section 2 discusses the CDO valuation. Section 3 describes the DSFM. Section 4 shows results of the empirical modelling. Section 5 presents applications in CDO trading. Section 6 concludes.

## 2 Collateralized Debt Obligations

Consider a CDO with a maturity of  $T$  years,  $J$  tranches and a pool of  $d$  entities at the valuation day  $t_0$ . A tranche  $j = 1, \dots, J$  absorbs losses between  $l_j$  percent and  $u_j$  percent of the total portfolio loss.  $l_j$  and  $u_j$  are called an attachment and a detachment point respectively and  $l_j < u_j$ . For the iTraxx Europe, successive tranches have the following attachment points: 0%, 3%, 6%, 9%, 12%, 22%. The corresponding detachment points are 3%, 6%, 9%, 12%, 22%, 100%.

The protection buyer pays periodically to the protection seller a predetermined premium, called a spread on the outstanding tranche notional and is compensated for losses that occur within the range of the tranche. Each default in the portfolio reduces the notional on which the payments are based. This leads to a decline in the value of the periodic fee. The cash exchange takes place until the maturity of the CDO or until the portfolio losses exceed the detachment point.

This section briefly discusses key points of the CDO pricing. First, we describe the cash-flow structure. Then we specify the distribution of the portfolio losses. At the end we present a concept of a base correlation.

### 2.1 Valuation

We assume that there exists a risk-neutral measure  $P$  under which the discounted asset prices are martingales. The expectations in the formulas below are taken with respect to this measure.

The loss of the portfolio of  $d$  assets at time  $t$  is defined as

$$L(t) = \frac{\text{LGD}}{d} \sum_{i=1}^d \Gamma_i(t), \quad t \in [t_0, T], \quad (1)$$

where LGD is a common loss given default and  $\Gamma_i(t) = \mathbf{1}(\tau_i \leq t)$ ,  $i = 1, \dots, d$ , is a default indicator showing that the credit  $i$  defaults at time  $t$  within the period  $[t_0, T]$  if the time of default random variable  $\tau_i \leq t$ . The loss of a tranche  $j = 1, \dots, J$  at time  $t$  is expressed as

$$L_j(t) = L^u(t, u_j) - L^u(t, l_j),$$

with  $L^u(t, x) = \min\{L(t), x\}$ ,  $x \in [0, 1]$ . The outstanding notional of the tranche  $j$  is given by

$$F_j(t) = F^u(t, u_j) - F^u(t, l_j),$$

with  $F^u(t, x) = x - L^u(t, x)$ ,  $x \in [0, 1]$ . At the predefined dates  $t = t_1, \dots, T$ ,  $t_1 > t_0$ , the protection seller and the protection buyer exchange the payments. The protection leg  $\text{DL}_j$  is defined as the present value of all expected payments made upon defaults

$$\text{DL}_j(t_0) = \sum_{t=t_1}^T \beta(t_0, t) \mathbf{E}\{L_j(t) - L_j(t - \Delta t)\}, \quad j = 1, \dots, J, \quad (2)$$

where  $\beta$  is a discount factor and  $\Delta t$  is a time between  $t$  and the previous payment day. The premium leg  $PL_j$  is expressed as the present value of all expected premium payments

$$PL_j(t_0) = \sum_{t=t_1}^T \beta(t_0, t) s_j(t_0) \Delta t \mathbf{E}\{F_j(t)\}, \quad j = 2, \dots, J, \quad (3)$$

where  $s_j$  denotes the spread of tranche  $j$ . The first tranche, called the equity is traded with an upfront payment  $\alpha$  and a fixed spread of 500 bp. Its premium leg (3) turns into

$$PL_1(t_0) = \alpha(t_0)(u_1 - l_1) + \sum_{t=t_1}^T \beta(t_0, t) \cdot 500 \cdot \Delta t \mathbf{E}\{F_1(t)\}.$$

A spread  $s_j$  is calculated once, at  $t_0$  so that the marked-to-market value of the tranche is zero, i.e. the value of the premium leg equals the value of the protection leg

$$s_j(t_0) = \frac{\sum_{t=t_1}^T \beta(t_0, t) \mathbf{E}\{L_j(t) - L_j(t - \Delta t)\}}{\sum_{t=t_1}^T \beta(t_0, t) \Delta t \mathbf{E}\{F_j(t)\}}, \quad \text{for } j = 2, \dots, J. \quad (4)$$

The upfront payment of the equity tranche is computed as

$$\alpha(t_0) = \frac{100}{u_1 - l_1} \sum_{t=t_0}^T (\beta(t, t_0) [\mathbf{E}\{L_1(t) - L_1(t - \Delta t)\} - 0.05 \Delta t \mathbf{E}\{F_1(t)\}]).$$

For more details we refer to Bluhm & Overbeck (2006) and Kakodkar, Galiani, Jónsson & Gallo (2006).

The iTraxx data used in this study cover years 2005–2009 when the tranches were priced in the way presented above. However, since 2009 the quoting convention of the iTraxx Europe tranches has changed. Now all the tranches (0-3%, 3-6%, 6-9%, 9-12%, 12-22%, 22-100%) have a structure of the equity tranche and trade with an upfront fee and a fixed running spread of 500 bp, 500 bp, 300 bp, 100 bp, 100 bp, and 25 bp respectively.

## 2.2 Credit risk models

The main challenge in calculating the fair tranche spread (4) is the correct calculation of the expected losses. This task requires the analysis of how the portfolio entities are likely to default together. At the core of the CDO pricing lies a dependency model for portfolio credit risk. There are two main types of credit risk models: structural and reduced form models. The structural model is motivated from a Merton style approach where a default occurs when the value of an asset drops below a certain level. In the reduced form approach a default is modeled with an intensity process. A third class is based on copula theory and is connected with the first two approaches. For a comprehensive overview we refer to Bielecki & Rutkowski (2004).

There has been a multitude of CDO risk models proposed that apply different dependency concepts. The market standard for pricing CDOs is the large pool Gaussian copula model

that has been introduced to the valuation of multi-name credit derivatives by Li (2000). The large pool concept is also the basis of the Gaussian one-factor model proposed by Vasicek (1987).

In the Vasicek model, an obligor  $i = 1, \dots, d$  defaults before time  $T$  if the value of a random variable  $\mathcal{X}_i$  drops below a threshold  $C_i$

$$\mathbf{P}(\tau_i \leq T) = \mathbf{P}(\mathcal{X}_i < C_i).$$

The variable  $\mathcal{X}_i$  is defined as a linear combination of the systematic risk factor  $\mathcal{Y}$  and the idiosyncratic risk factor  $\mathcal{Z}_i$

$$\mathcal{X}_i = \sqrt{\rho}\mathcal{Y} + \sqrt{1-\rho}\mathcal{Z}_i, \quad (5)$$

In this setting  $\mathcal{Y}$  and  $\{\mathcal{Z}_i\}_{i=1}^d$  are i.i.d.  $N(0, 1)$  variables. Thus,  $\{\mathcal{X}_i\}_{i=1}^d$  follow a multivariate Gaussian distribution with an equal pairwise correlation  $\rho$ .

The model assumes that the portfolio is large and homogeneous, i.e. it possesses an infinite number of assets that have the same exposure, default probability, loss given default, correlation and that these values are constant over all time horizons. The individual default probability determines the default threshold  $C_i = C = \Phi^{-1}(p)$  for all  $i = 1, \dots, d$ , where  $\Phi$  denotes the cdf of the standard normal distribution.

The portfolio loss distribution is approximated by the conditional probability given the common factor  $\mathcal{Y}$ . When  $\mathcal{Y}$  is fixed, the conditional default probability of any obligor is given by

$$p(y) = \mathbf{P}(\mathcal{X}_i < C | \mathcal{Y} = y) = \Phi\left(\frac{C - \sqrt{\rho}y}{\sqrt{1-\rho}}\right). \quad (6)$$

Conditional on the realization of the common factor, the variables (5) are independent. The portfolio loss conditional on  $\mathcal{Y}$  converges, by the law of large numbers, to its expectation  $p(y)$ . The cdf of the loss of a very large portfolio is in the limit equal

$$\mathbf{P}(L \leq x) = \mathbf{P}\{p(\mathcal{Y}) \leq x\} = \Phi\left\{\frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(p)}{\sqrt{\rho}}\right\}. \quad (7)$$

The expected tranche loss in (2) and (3) is calculated as an integral with respect to the distribution (7) and because of absence of the explicit solution it has to be evaluated numerically.

The main drawback of the Gaussian copula is that it exhibits no tail dependence and in consequence it cannot model the extreme events accurately. However, due to its analytical tractability and numerical simplicity, the Gaussian copula model still remains the benchmark on the market.

## 2.3 Base correlation

In the Gaussian copula model the main driver of the tranche price is the correlation coefficient. The correlations can be computed from market data by inverting the pricing

formula (4). If we keep the value of other parameters fixed, then the correlation parameter that matches the quoted tranche spread is called an implied compound correlation. It is observed that implied compound correlations are not constant across the tranches. This phenomenon is called an implied correlation smile. Still, the main disadvantage of the compound coefficient is that the mezzanine tranches are not monotonic in correlation and two parameters might result in the same spread value. The second problem that we might encounter is a nonexistence of the implied correlation. These disadvantages caused the enhanced popularity of base correlations proposed by McGinty & Ahluwalia (2004).

The main idea behind the concept of the base correlation is that each tranche  $[l_j, u_j]$  can be represented as a difference of two, equity type tranches that have the lower attachment point zero:  $[0, u_j]$  and  $[0, l_j]$ . Here we use a property that the equity tranche is monotone in correlation. The base correlations can be implied from the market spreads using standard bootstrapping techniques. One needs the spread value of the tranche  $[u_{j-1}, u_j]$  and the base correlation of the tranche  $[0, u_{j-1}]$  in order to imply the base correlation  $[0, u_j]$ . In this approach, (4) is calculated as

$$s_j(t_0) = \frac{\sum_{t=t_1}^T \beta(t_0, t) [\mathbb{E}_{\rho(0, u_j)} \{L_j^u(t, u_j) - L_j^u(t - \Delta t, u_j)\} - \mathbb{E}_{\rho(0, l_j)} \{L_j^u(t, l_j) - L_j^u(t - \Delta t, l_j)\}]}{\sum_{t=t_1}^T \beta(t_0, t) \Delta t [\mathbb{E}_{\rho(0, u_j)} \{F_j^u(t, u_j)\} - \mathbb{E}_{\rho(0, l_j)} \{F_j^u(t, l_j)\}]} \quad (8)$$

for  $j = 2, \dots, J$ , where the expected value  $\mathbb{E}_{\rho(0, u_j)}$  is calculated with respect to the distribution (7) determined by the base correlation  $\rho(0, u_j)$  of the tranche  $[0, u_j]$ . In the Gaussian copula model the base correlations are nondecreasing with respect to the seniority of tranches and the implied correlation smile turns into a correlation skew.

### 3 Dynamic Semiparametric Factor Model

Let  $Y_{t,k}$  be a data point, a tranche spread or a base correlation, observed on a day  $t$ ,  $t = 1, \dots, T$ . The index  $k$  represents an intra-day numbering of observations on that day,  $k = 1, \dots, K_t$ . The observations  $Y_{t,k}$  are regressed on two-dimensional covariates  $X_{t,k}$  that contain the tranche seniority and the remaining time to maturity

$$Y_{t,k} = m_0(X_{t,k}) + \sum_{l=1}^L Z_{t,l} m_l(X_{t,k}) + \varepsilon_{t,k}, \quad (9)$$

where  $m_l : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $l = 0, \dots, L$ , are factor loading functions,  $Z_{t,l} \in \mathbb{R}$  are factors, and  $\varepsilon_{t,k}$  are error terms with zero means and finite variances.

The additive structure of (9) is a typical approach in regression models. Here, the functions  $m$  are estimated nonparametricly and represent the time-invariant, spatial component. The factors  $Z_t$  drive the dynamics of  $Y_t$ . The number of factors  $L$  is fixed and should be small relative to the number of observed data points so that we achieve a significant reduction in the dimension. The investigation of the dynamics of the entire system boils down to the analysis of the factors' variability. These arguments justify calling (9) a dynamic semiparametric factor model.



Fengler et al. (2007) estimate  $m$  and  $Z_t$  iteratively using kernel smoothing methods, Song, Härdle & Ritov (2013) apply functional principal component analysis, Park, Mammen, Härdle & Borak (2009) estimate  $m$  with a series based estimator. For numerical convenience we follow the last paper and define functions  $\psi_b : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $b = 1, \dots, B$ ,  $B \geq 1$ , such that  $\int_{\mathbb{R}^2} \psi_b^2 dx = 1$ . Then, a tuple of functions  $(m_0, \dots, m_L)^\top$  may be approximated by  $A\psi$ , where  $A$  is a  $(L+1 \times K)$  matrix of coefficients  $\{\{a_{l,b}\}_{l=1}^{L+1}\}_{b=1}^B$  and  $\psi = (\psi_1, \dots, \psi_B)^\top$ . We take  $\{\psi_b\}_{b=1}^B$  to be a tensor B-spline basis. For a survey over the mathematical foundations of splines we refer to de Boor (2001). With this parametrization (9) turns into

$$Y_{t,k} = Z_t^\top m(X_{t,k}) + \varepsilon_{t,k} = Z_t^\top A\psi(X_{t,k}) + \varepsilon_{t,k},$$

where  $Z_t = (Z_{t,0}, \dots, Z_{t,L})^\top$  with  $Z_{t,0} = 1$  and  $m = (m_0, \dots, m_L)^\top$ .

The estimates  $\hat{Z}_t = (\hat{Z}_{t,0}, \dots, \hat{Z}_{t,L})^\top$  and  $\hat{A}$  are obtained by

$$(\hat{Z}_t, \hat{A}) = \arg \min_{Z_t, A} \sum_{t=1}^T \sum_{k=1}^{K_t} \{Y_{t,k} - Z_t^\top A\psi(X_{t,k})\}^2, \quad (10)$$

yielding estimated basis functions  $\hat{m} = \hat{A}\psi$ . The minimization is carried out using an iterative algorithm. However, the estimates of  $m$  and  $Z_t$  are not uniquely defined. Therefore, the final estimates of  $m$  are orthonormalized and  $Z_t$  are centered. Park et al. (2009) also prove that the difference of the inference based on the estimated  $\hat{Z}_{t,l}$  and the true, unobserved  $Z_{t,l}$  is asymptotically negligible. This result justifies fitting an econometric model, like a vector autoregressive to the estimated factors for further analysis of the data.

The number of factors  $L$  as well as the numbers of spline knots in both maturity and tranche directions  $R_1, R_2$ , and the orders of splines  $r_1, r_2$  have to be chosen in advance. A common approach is to maximize a proportion of the variation explained by the model among the total variation. We propose a following criterion

$$\text{EV}(L, R_1, r_1, R_2, r_2) = 1 - \frac{\sum_{t=1}^T \sum_{k=1}^{K_t} \left\{ Y_{t,k} - \sum_{l=1}^L Z_{t,l} m_l(X_{t,k}) \right\}^2}{\sum_{t=1}^T \sum_{k=1}^{K_t} \{Y_{t,k} - \tilde{m}_0(X_{t,k})\}^2}, \quad (11)$$

where

$$\tilde{m}_0(X_\ell) = \frac{1}{T} \sum_{t=1}^T \frac{\sum_{k=1}^{K_t} Y_{t,k} \mathbf{1}\{X_\ell = X_{t,k}\}}{\sum_{k=1}^{K_t} \mathbf{1}\{X_\ell = X_{t,k}\}}, \quad \ell = 1, \dots, K_{\max}, \quad (12)$$

is an empirical mean surface and  $K_{\max}$  is the number of all different  $X_{t,k}$  observed during  $T$  days. The criterion (11) is a modified version of this considered in Fengler et al. (2007) and other literature on the DSFM, where instead of the empirical mean surface, the overall mean of the observations is used. The mean surface (12) makes more sense, since our data reflect monotonous behaviour w.r.t. the time to maturity.

The  $\tilde{m}_0$  factor in (9) is usually interpreted as a mean function of the data. We propose to first subtract the estimate (12) from the data and then fit the DSFM. The extraction of the empirical mean  $\tilde{m}_0$  leads to the following model

$$Y_{t,k} = \tilde{m}_0(X_{t,k}) + \sum_{l=1}^L Z_{t,l} m_l(X_{t,k}) + \varepsilon_{t,k} = \tilde{m}_0(X_{t,k}) + Z_t^\top A\psi(X_{t,k}) + \varepsilon_{t,k}, \quad (13)$$

where  $m_l$  are factor functions,  $l = 1, \dots, L$ ,  $Z_{t,l}$  are factor loadings, and  $A$  is a  $(L \times B)$  coefficient matrix. The representation (13) reduces the number of the factor functions estimated in the iterative algorithm (10). The model (9) is the classic DSFM and we will refer to it as the DSFM. The model (13) is hereafter called the DSFM without the mean factor.

## 4 Modelling the Dynamics of CDO Surfaces

### 4.1 Data Description

The data set analysed in this study contains daily spreads of iTraxx tranches of Series 2 to 10 between 30 March 2005 (hereafter denoted 20050330) and 2 February 2009 (denoted 20090202) obtained from **Bloomberg**. We have in total  $\sum_{t=1}^T K_t = 49502$  data points over  $T = 1004$  days. As far as we know, this is the first study on CDOs that consider such an extensive data set.

Each index has 3, 5, 7, or 10 years maturity. Figure 2 shows the market spreads of Series 6 and the corresponding implied Gaussian base correlations for the maturity of 5 and 10 years. We see similarities in the general evolution between these two maturities. However, since their exposure in terms of duration differs, they represent different levels.

Due to the issuing scheme one observes every day a bunch of indices from various series and different maturities. Here we analyse tranche spreads and also base correlations, both denoted  $Y_t$ , as a function of the tranche seniority  $\xi_t$  and the remaining time to maturity  $\tau_t$ . The seniority of a tranche  $\xi_t$  is represented by its corresponding detachment point. The remaining time to maturity of an index is an actual time left till its expiration and takes values between zero and 10.25. For every day a separate surface representing the market information is available. The number of observed every day indices is low (minimum 4, maximum 17, median 12, see Figure 3). This results in a string structure in the data. Each string corresponds to one  $\tau_t \in [0, 10.25]$  and is composed out of at most five points. The market quotes five out of six tranches as the most senior tranche is usually not traded. Figure 1 and Figure 4 present the curves of market spreads and corresponding implied base correlations on 20080909 and 20090119. As time passes, the curves move through the space towards expiry and simultaneously change their skewness and level. As previously mentioned, the main aim of this research is to model the evolution of the iTraxx spreads and base correlations simultaneously in space and time dimensions.

Since the shortest maturity is 3 years and every half a year new four indices are issued, the number of indices present on the market grows in time. Table 1 shows the number of observed market values of iTraxx tranches for every possible maturity during the entire period considered and during the annual subperiods. Table 2 outlines a percentage of missing values for every maturity and for every tranche. We see that the CDO market was booming in 2006 and 2007. However, since the beginning of the financial crisis in 2008 the demand for credit derivatives had been shrinking meaningfully. In the first quarter of

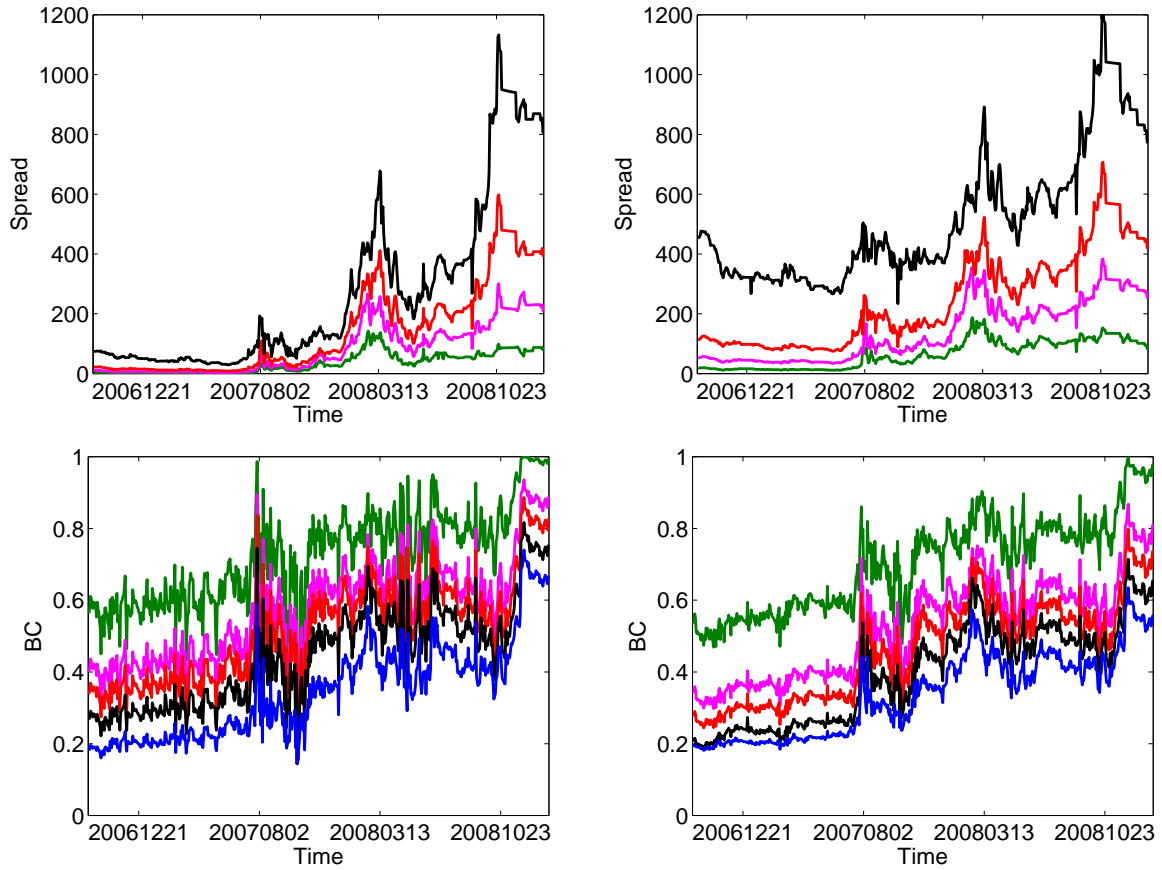


Figure 2: Market spreads (upper panel) and implied base correlations (lower panel). Series 6, maturity 5 (left) and 10 (right) years. Data from 20060920-20090202. Tranches: 1 (blue), 2 (black), 3 (red), 4 (pink), 5 (green).

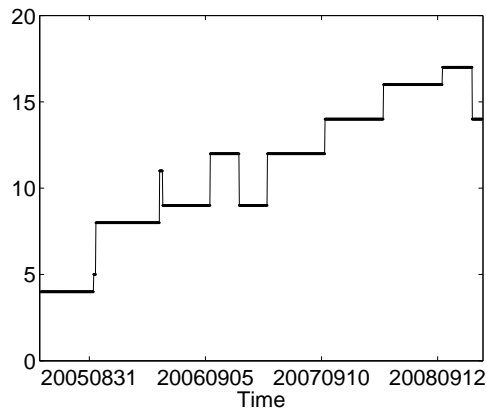


Figure 3: Daily number of curves for every surface during the period 20050330-20090202.

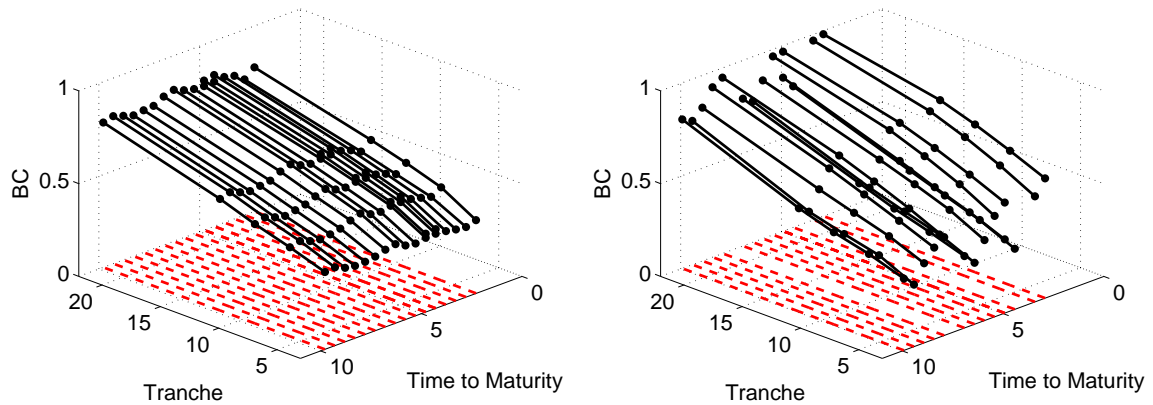


Figure 4: Base correlations of all series observed on 20080909 (left) and 20090119 (right).

2009 the iTraxx tranches became highly illiquid. Many missing data may create challenges to the econometric analysis. Because tranches with 3 years maturity were rarely traded, this maturity was excluded from our study. Tables 3 and 4 present summary statistics of

Year	3Y	5Y	7Y	10Y
2005	0	1478	715	1532
2006	181	3998	3739	4005
2007	75	5155	5170	5172
2008	232	5904	5916	5932
2009	0	260	263	263
All	488	16740	15803	16840

Table 1: Number of observed values of iTraxx tranches in the period 20050330-20090202.

Year	3Y					5Y					7Y					10Y				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
2005	100	100	100	100	100	34	34	34	34	48	5	5	5	5	5	35	34	35	34	35
2006	78	56	55	100	100	6	7	6	6	8	3	3	3	4	4	5	5	6	8	6
2007	88	99	99	100	100	3	2	2	3	3	2	2	3	2	3	3	2	3	2	2
2008	47	99	100	100	100	24	25	25	24	27	24	25	25	25	27	24	27	24	25	24
2009	100	100	100	100	100	42	42	47	42	42	42	43	43	42	42	42	43	42	42	43
All	72	93	93	100	100	16	17	17	16	20	13	14	13	13	14	16	17	16	17	16

Table 2: Percentage of missing values during the period 20050330-20090202.

	Year	Tranche 1		Tranche 2		Tranche 3		Tranche 4		Tranche 5	
Spread	2005	40.90	(13.22)	255.42	(174.34)	62.13	(41.26)	32.30	(19.49)	18.85	(12.47)
	2006	33.71	(14.76)	212.75	(173.46)	52.09	(38.71)	23.56	(16.63)	9.68	(7.02)
	2007	26.03	(12.45)	187.28	(127.72)	72.98	(55.19)	40.56	(32.85)	20.79	(18.96)
	2008	42.66	(11.22)	536.11	(230.96)	317.60	(130.18)	195.79	(74.15)	92.13	(32.29)
	2009	62.89	(8.01)	998.48	(116.96)	558.98	(80.98)	320.51	(55.32)	103.44	(11.04)
	All	35.56	(14.86)	335.27	(256.95)	162.61	(157.65)	96.36	(96.74)	44.86	(43.51)
Base Corr.	2005	0.18	(0.04)	0.26	(0.05)	0.35	(0.05)	0.43	(0.06)	0.60	(0.09)
	2006	0.16	(0.05)	0.21	(0.07)	0.28	(0.09)	0.33	(0.11)	0.48	(0.17)
	2007	0.27	(0.08)	0.36	(0.10)	0.43	(0.11)	0.49	(0.11)	0.65	(0.12)
	2008	0.43	(0.06)	0.52	(0.06)	0.58	(0.06)	0.63	(0.06)	0.79	(0.06)
	2009	0.40	(0.10)	0.48	(0.11)	0.53	(0.12)	0.59	(0.13)	0.80	(0.10)
	All	0.29	(0.13)	0.38	(0.14)	0.44	(0.15)	0.50	(0.15)	0.66	(0.16)

Table 3: Mean and standard deviation (in parentheses) of tranche spreads (UFF for the tranche 1) and implied base correlations during the period 20050330-20090202.

Year	Index	
2005	46.34	(11.08)
2006	37.70	(11.86)
2007	40.22	(16.25)
2008	121.71	(50.86)
2009	205.47	(44.04)
All	77.25	(57.60)

Table 4: Mean and standard deviation (in parentheses) of iTraxx indices' spreads during the period 20050330-20090202.

the market spreads, the base correlations, and of the iTraxx indices.

Sometimes on a particular day, for a particular tranche and a particular remaining time to maturity we observe two different spreads. As an example consider a day  $t_0$  on which a new series with 3 years maturity is issued. If 5 years earlier a series with 7 years maturity was issued, then on day  $t_0$  this series has also 3 years remaining time to maturity. In this situation we include in our data set the observation that comes from the most actual series (in the example we take the series issued on  $t_0$ ).

The base correlations (8) are implied from the market spreads using the large pool Gaussian copula model (assuming the LGD of 60%) presented in Section 2.2. The common intensity parameters are derived from iTraxx indices. The discount factors are calculated from rates of Euribor and Euro Swaps.

The structure of the equity tranche is different from the other tranches. It is quoted as an upfront payment plus 500 bp spread paid quarterly. In order to include the equity tranche in the joint analysis of all the tranches, we convert its quotes to standard spreads with zero upfront fee using the large pool Gaussian copula model.

## 4.2 DSFM Estimation Results

Since our data are positive and monotone, we convert spreads into log-spreads and for base correlations apply the Fisher's Z-transformation defined as

$$\mathcal{T}(u) = \operatorname{arctanh}(u) = \frac{1}{2} \log \frac{1+u}{1-u}.$$

It transforms the empirical Pearson's correlations between bivariate normal variables to a normally distributed variable. We will use it for the base correlations as it stabilizes their variance.

As discussed in Section 3, the number of factors  $L$ , the numbers of knots  $R_1$ ,  $R_2$ , and the orders of splines  $r_1$ ,  $r_2$  are selected according to (11). Since the design of the data in the tranche seniority dimension is fixed, we choose in this direction quadratic B-splines and five knots. Tables 5 and 6 present a proportion of the explained variation (11) by the DSFM and the DSFM without the mean factor respectively for different numbers of factors, knots and different orders of splines in the maturity dimension. Similar like in Park et al. (2009), we find that the order of splines and the number of knots have a small influence on the proportion of the explained variation. We pick the quadratic B-splines placed on 10 knots in  $\tau$  dimension for both types of data and for both DSFM models. The number of knots is close to the median number of observed strings every day. Table 7 outlines the criterion's values for the chosen tensor B-splines and for up to five factors. The information shown there reveals that two factors are sufficiently good approximation to the data. Figures 5 and 6 exhibit  $\hat{m}$  and  $\hat{Z}_t$  in the DSFM for the log-spreads and the Z-transformed base correlations respectively. Figures 7 and 8 show the DSFM without the mean factor estimates.

Number of Factors	Spline	Log-Spr				Z-BC			
	<div>Knots Order</div>	5	10	15	20	5	10	15	20
1	1	0.944	0.963	0.967	0.968	0.721	0.765	0.774	0.786
	2	0.965	0.968	0.968	0.969	0.738	0.769	0.788	0.799
	3	0.966	0.968	0.969	0.969	0.743	0.775	0.786	0.793
	4	0.966	0.968	0.969	0.969	0.740	0.774	0.792	0.797
2	1	0.960	0.978	0.982	0.983	0.813	0.871	0.814	0.894
	2	0.980	<i>0.983</i>	0.984	0.984	0.835	<i>0.883</i>	0.912	0.918
	3	0.981	0.983	0.984	0.984	0.846	0.894	0.908	0.891
	4	0.981	0.983	0.984	0.984	0.849	0.895	0.913	0.912
3	1	0.965	0.985	0.989	0.990	0.822	0.893	0.915	0.915
	2	0.986	0.989	0.990	0.990	0.869	0.915	0.928	0.932
	3	0.987	0.989	0.990	0.991	0.872	0.922	0.937	0.941
	4	0.987	0.989	0.990	0.990	0.873	0.921	0.930	0.935

Table 5: Proportion of the explained variation by the DSFM for  $L = 1, 2, 3$ , different numbers of knots and different orders of splines in the maturity dimension. The values of the selected models marked with italic.

Number of Factors	Spline	Log-Spr				Z-BC			
	<div>Knots Order</div>	5	10	15	20	5	10	15	20
1	1	0.797	0.876	0.897	0.898	0.629	0.640	0.660	0.660
	2	0.877	0.896	0.905	0.910	0.633	0.654	0.657	0.664
	3	0.867	0.898	0.906	0.908	0.638	0.650	0.662	0.664
	4	0.871	0.898	0.907	0.910	0.639	0.653	0.659	0.662
2	1	0.842	0.925	0.940	0.945	0.730	0.835	0.860	0.869
	2	0.926	<i>0.952</i>	0.961	0.954	0.781	<i>0.861</i>	0.876	0.888
	3	0.911	0.952	0.941	0.950	0.763	0.867	0.883	0.887
	4	0.917	0.956	0.947	0.954	0.783	0.870	0.881	0.886
3	1	0.858	0.940	0.959	0.973	0.746	0.854	0.888	0.898
	2	0.941	0.967	0.977	0.982	0.815	0.896	0.907	0.925
	3	0.927	0.967	0.975	0.979	0.805	0.901	0.922	0.930
	4	0.932	0.972	0.977	0.982	0.817	0.903	0.910	0.927

Table 6: Proportion of the explained variation by the DSFM without the mean factor for  $L = 1, 2, 3$ , different numbers of knots and different orders of splines in the maturity dimension. The values of the selected models marked with italic.

Number of Factors	DSFM		DSFM w/o mean f.	
	Log-Spr	Z-BC	Log-Spr	Z-BC
1	0.968	0.769	0.896	0.654
2	0.983	0.893	0.952	0.862
3	0.989	0.919	0.967	0.887
4	0.991	0.931	0.973	0.909
5	0.993	0.936	0.976	0.918

Table 7: Proportion of the explained variation by the models with  $L = 1, \dots, 5$  dynamic factors and quadratic tensor B-splines placed on  $5 \times 10$  knots.

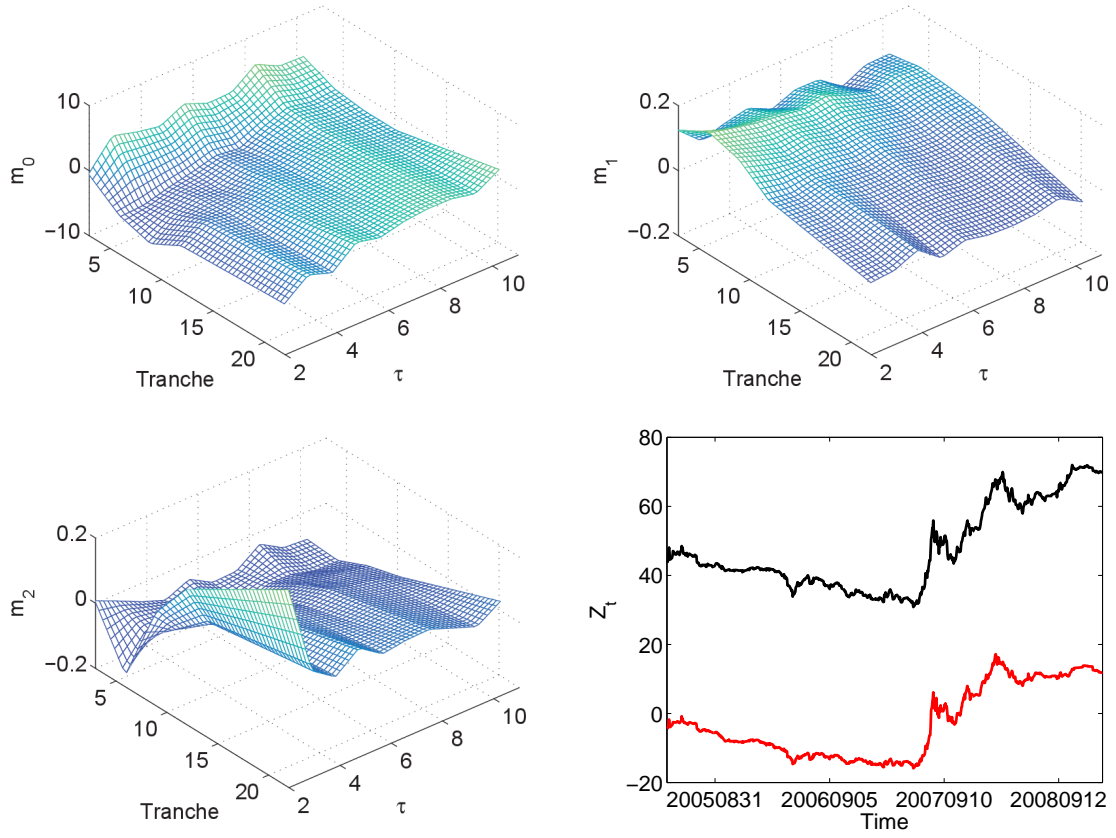


Figure 5: Estimated factors and loadings ( $Z_{t,1}$  black,  $Z_{t,2}$  red) in the DSFM for the log-spreads.

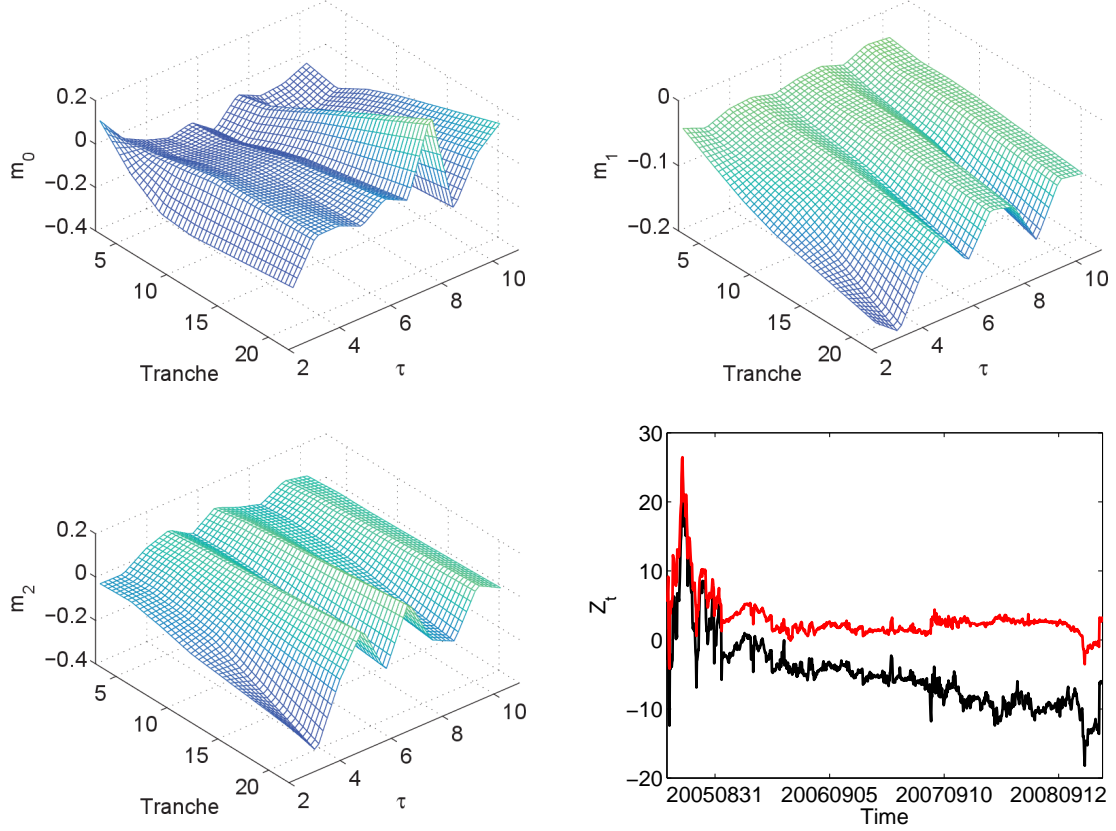


Figure 6: Estimated factors and loadings ( $Z_{t,1}$  black,  $Z_{t,2}$  red) in the DSFM for the Z-transformed base correlations.

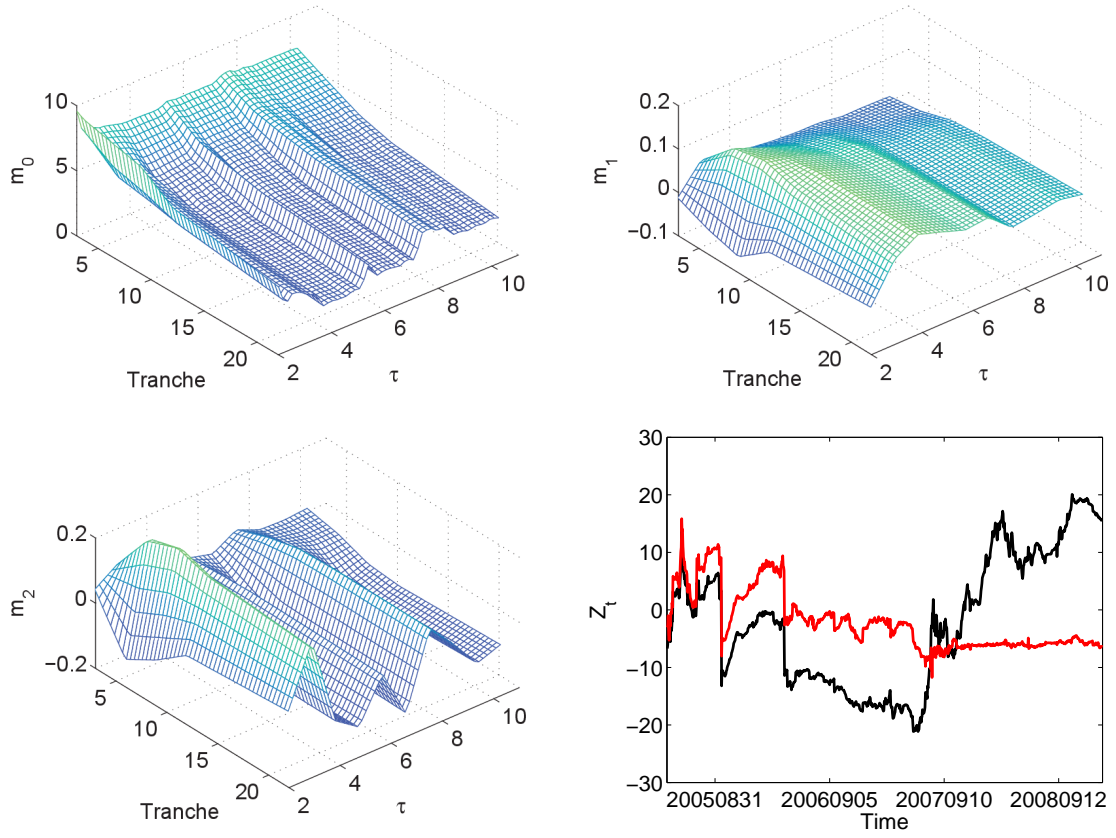


Figure 7: Sample mean, estimated factors and loadings ( $Z_{t,1}$  black,  $Z_{t,2}$  red) in the DSFM without the mean factor for the log-spreads.

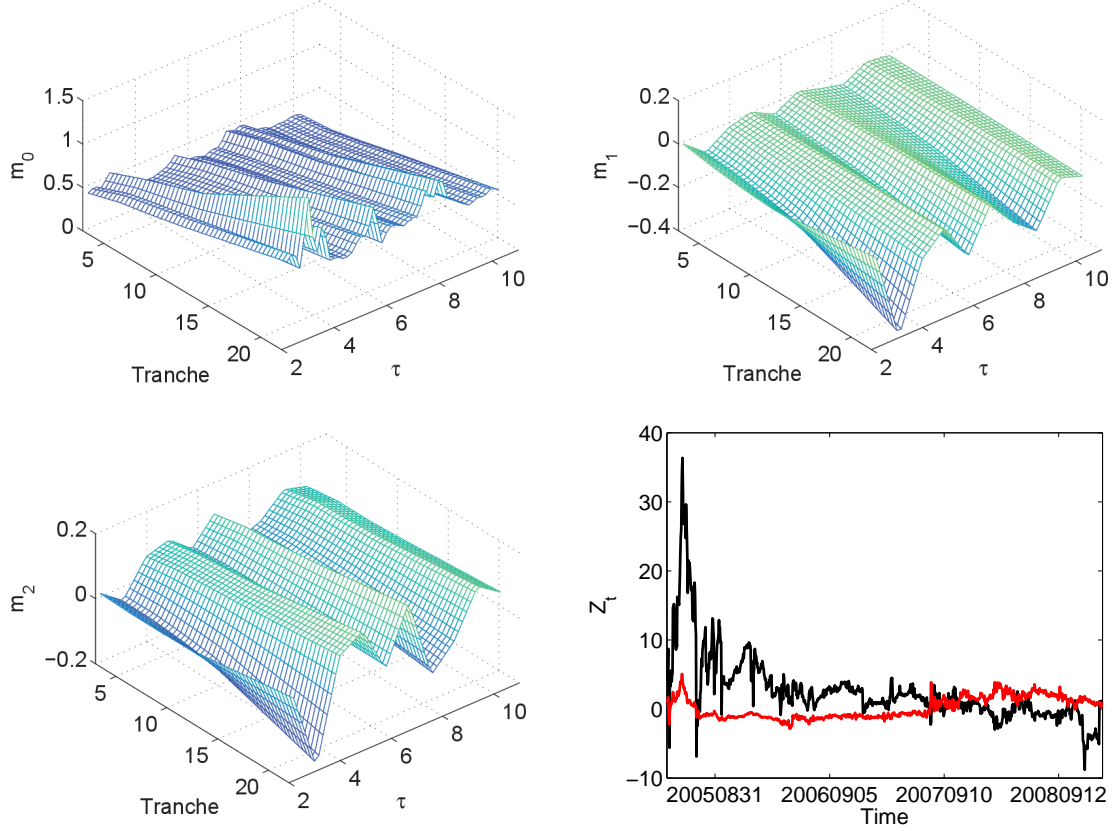


Figure 8: Sample mean, estimated factors and loadings ( $Z_{t,1}$  black,  $Z_{t,2}$  red) in the DSFM without the mean factor for the Z-transformed base correlations.



In the DSFM for both data types, the first and the second factor can be interpreted as a slope-curvature and a shift function respectively. Increasing  $\hat{Z}_{t,1}$  results in the enhancement of the surface's steepness, whereas, decreasing  $\hat{Z}_{t,1}$  implies its flattening. When we shift  $\hat{Z}_{t,2}$ , the whole surface shifts along the  $z$ -axis. In the DSFM without the mean factor for the log-spreads we observe an opposite influence of the factors. Namely,  $\hat{m}_1$  is the shift factor,  $\hat{m}_2$  is the slope-curvature factor. For the DSFM without the mean factor applied to the Z-transformed base correlations the interpretation is not so clear. When varying  $\hat{Z}_{t,1}$  and  $\hat{Z}_{t,2}$  both the slope and the curvature change. The upward shift of the surface can be a result of a decrease in  $\hat{Z}_{t,1}$  or an increase in  $\hat{Z}_{t,2}$ .

Table 8 discloses the mean squared error of the in-sample fits. The classic DSFM occurs to provide a better fit to the data than the DSFM without the mean factor. Moreover, the Z-transformed base correlations are approximated more accurately than the log-spreads for both DSFMs. Figure 9 displays the in-sample fit of the models to data on 20080909 and 20090119. The convergence of the models is typically reached after 8 cycles.

Model	Log-Spr	Z-BC
DSFM	0.016	0.004
DSFM w/o mean f.	0.045	0.006

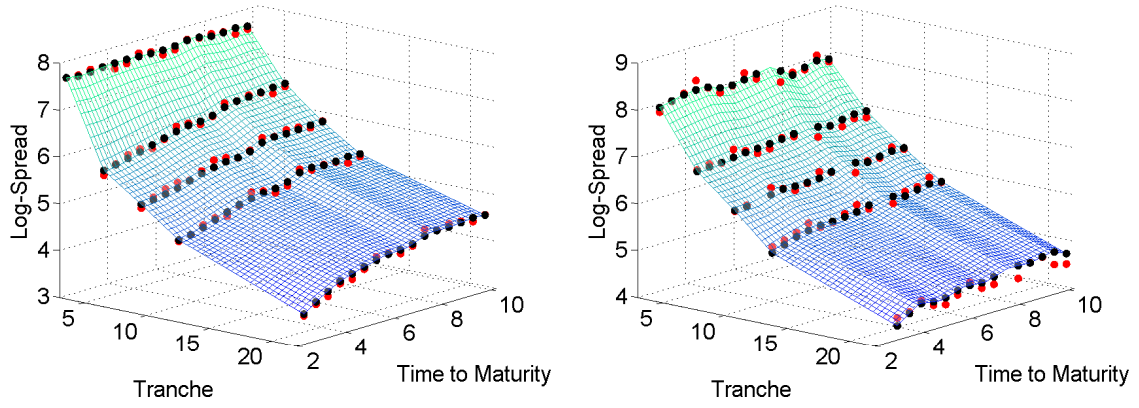
Table 8: Mean squared error of the in-sample fit.

The covariance structure of the  $\hat{Z}_t$  time series is investigated by means of VAR analysis. The augmented Dickey-Fuller test indicates that the first differences of  $\hat{Z}_t$  in both DSFMs are stationary. Figure 10 exhibits the sample partial autocorrelation functions of the residuals of the estimated VAR(1) models for the factor loadings of the DSFM without the mean factor. In general only a few autocorrelations lie slightly outside the 95% confidence interval. A similar result is observed for the classic DSFM. Thus, VAR(1) seems to be in line with the data. Certainly, one may investigate more complex multivariate time series models that account for a dynamic structure of the conditional variance-covariance and of the conditional correlation like the BEKK-GARCH or the DCC-GARCH, see Engle (2002). Since we are interested in the conditional mean process only, the VAR model appears to be sufficient. Moreover, a relatively simple out-of-sample VAR forecasting can be used in forecasting the evolution of the surfaces.

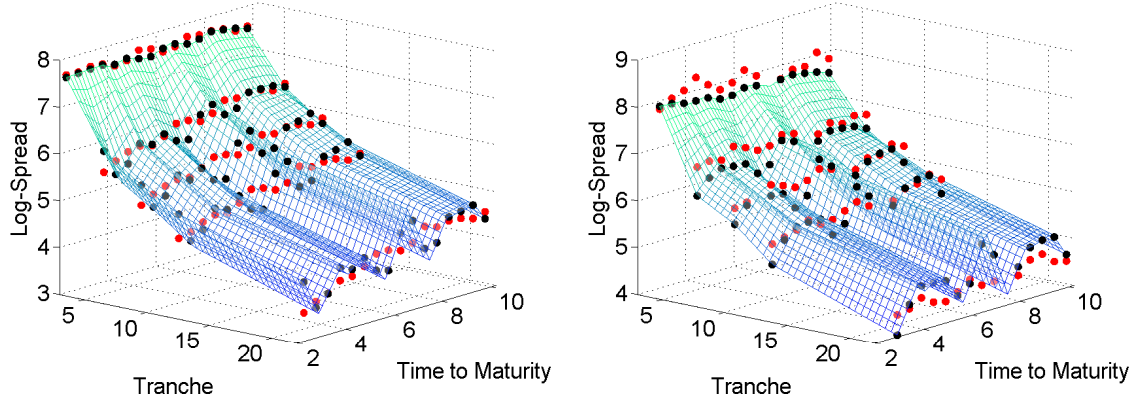
## 5 Applications in Trading

### 5.1 Curve Trades

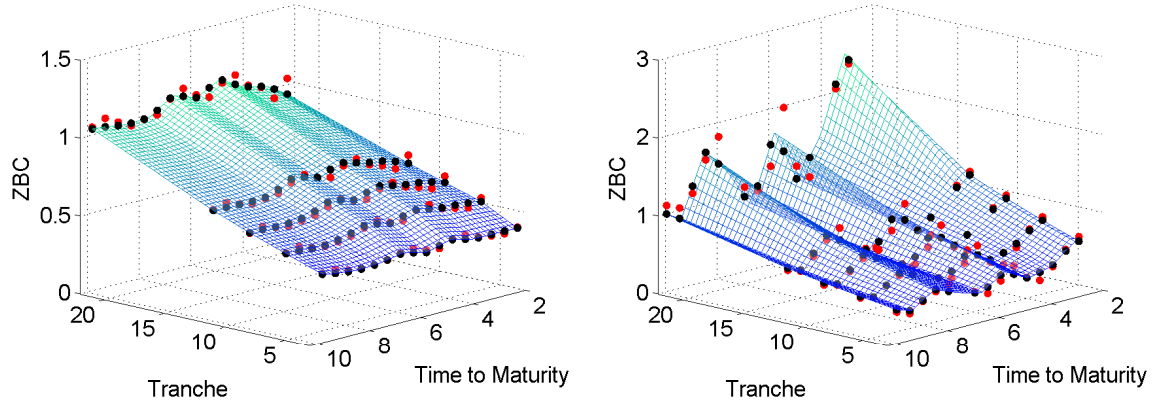
The popularity of the iTraxx market led to more liquidity in its standardized tranches allowing investors to implement complex credit positions. Here we present curve trades, namely flatteners and steepeners – strategies that combine tranches of different time to maturity, see also Kakodkar et al. (2006).



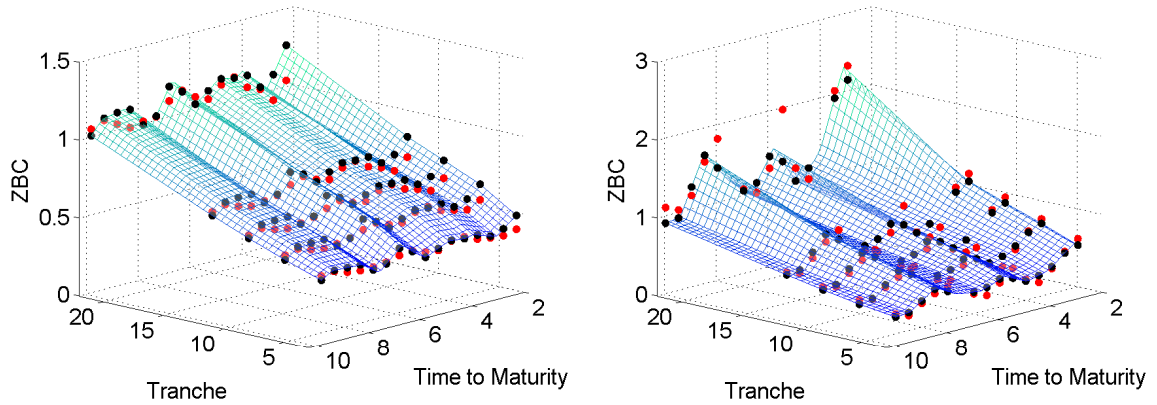
(a) DSFM for the log-spreads.



(b) DSFM without the mean factor for the log-spreads.



(c) DSFM for the Z-transformed base correlations.



(d) DSFM without the mean factor for the Z-transformed base correlations.

Figure 9: In-sample fit (black points) of the models to data (red points) on 20080909 (left) and 20090119 (right).

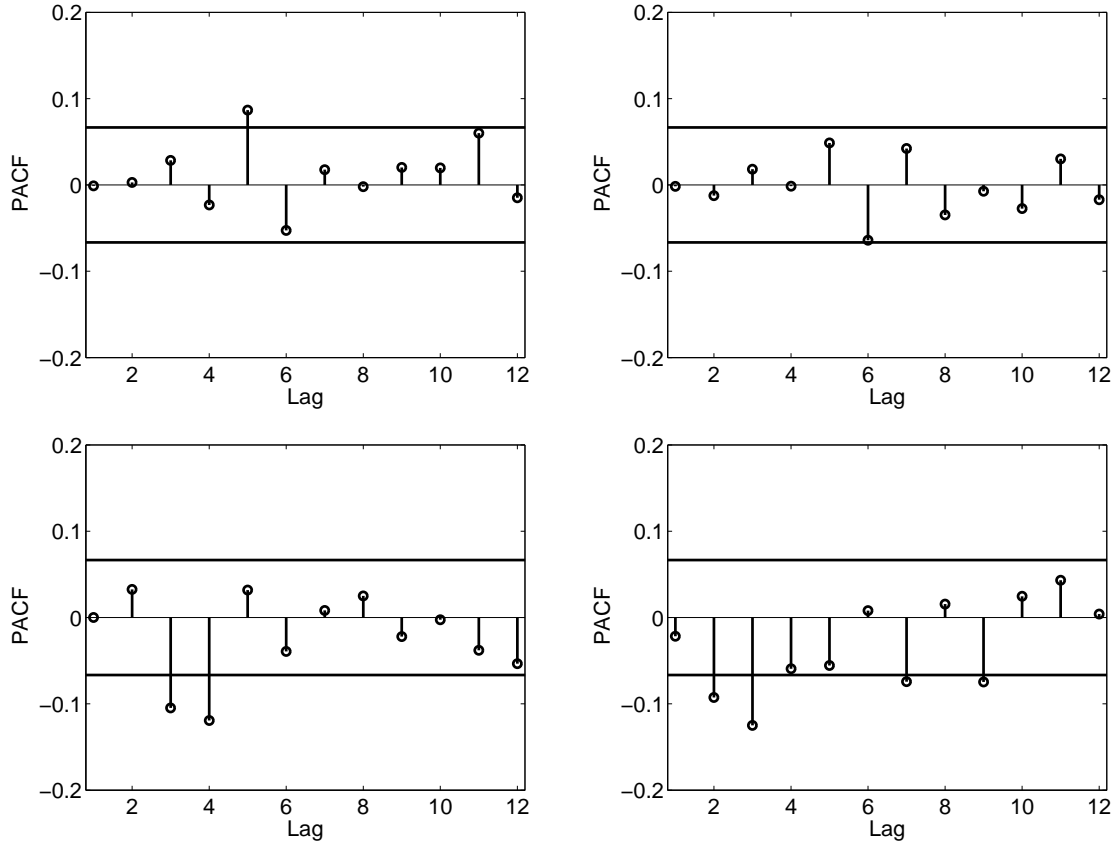


Figure 10: Sample partial autocorrelation for VAR(1) residuals of the factor loadings  $Z_{t,1}$  (left) and  $Z_{t,2}$  (right) in the DSFM without the mean factor for log-spreads (upper panel) and Z-transformed base correlations (lower panel). The solid lines correspond to the approximate lower and upper 95% confidence bounds.

A flattener is a trade that involves a simultaneous sale of a long-term tranche and a purchase of a short-term tranche. An example would be: sell 10Y 3-6% and buy 5Y 6-9%. In this trade the investor expresses a bullish long-term outlook but also a bearish short-term view on the market. The opposite trade is called a steepener. It is achieved by selling the short-term protection and buying the long-term protection. Both strategies are popular in trading CDS, credit indices, and yield curves. Credit curves got a lot of attention in May 2012 when J.P. Morgan announced a loss of \$2 billion on its flattener trade on the CDX IG 9 index. The final loss reached \$6.2 billion.

In our study both long and short term tranches have equal notional amounts. However, by adjusting the notionals, a trade can be structured so that it is risky duration neutral, carry neutral, correlation neutral, or theta (sensitivity to implied correlation changes) neutral, see Roy (2007). As recommended by Felsenheimer, Gisdakis & Zaiser (2004) we consider trades that generate no or a positive carry, i.e. the spread of the sold protection does not exceed the spread of the bought protection.

It is important to remark that very often our trades will be exposed to jump-to-default risk. For simplicity, let us consider in this paragraph flatteners only. If one buys 5Y 6-9% and sells 10Y 6-9%, then the trade is fully hedged for default only until the maturity of the 5Y tranche, i.e. any defaults that emerge from 10Y 6-9% are covered by 5Y 6-9%

till it expires. It should also be pointed out that the tranches do not have to be from the same series and there are slight differences in the composition of the collaterals of every series. Another case is if one buys 5Y 6-9% and sells 10Y 3-6%, then these tranches provide protection of different portion of portfolio risk. If there is any default in 10Y 3-6%, then we must deliver a payment obligation and incur a loss. Since we do not possess data of historical defaults in iTraxx, we cannot include the default payments in the further analysis. Consequently, in calculating the profit-and-loss (P&L) of the strategy we also do not account for the positive carry that we cumulate until the both positions are closed.

Felsenheimer et al. (2004), Kakodkar et al. (2006) and Roy (2007) consider various scenarios of flattener trades. They also assume that we do not observe any defaults in the collateral. However, the examples are not based on real data and do not investigate the performance of the trades over time.

Assume that an investor enters a curve trade and sells protection at a spread of  $s_1(t_0)$  for the period  $[t_0, T_1]$  and buys protection at a spread of  $s_2(t_0)$  for the period  $[t_0, T_2]$ . If the trade is a flattener, then  $T_1 > T_2$ . The spreads of the tranches are calculated in such a way that on the date of the trade  $t_0$  the marked-to-market (MTM) values of both positions are zero

$$\text{MTM}_\ell(t_0) = \sum_{t=t_1}^{T_\ell} \beta(t_0, t) [s_\ell(t_0)\Delta t \mathbf{E}\{F_\ell(t)\} - \mathbf{E}\{L_\ell(t) - L_\ell(t - \Delta t)\}] = 0, \ell = 1, 2.$$

Since spread values constantly vary over time, immediately after initiation of the trade,  $\tilde{t} > t_0$ , the market trades the tranches at  $s_\ell(\tilde{t})$ . In consequence, we observe a change in the MTM value of our positions

$$\text{MTM}_\ell(\tilde{t}) = \{s_\ell(t_0) - s_\ell(\tilde{t})\} \sum_{t=\tilde{t}_1}^{T_\ell} \beta(\tilde{t}, t) \Delta t \mathbf{E}\{F_\ell(t)\}, \ell = 1, 2, \quad (14)$$

where  $\tilde{t}_1$  is the first payment day after  $\tilde{t}$ .

A positive MTM means that the contract has a positive value to the protection seller. If the protection seller closes the position  $\ell$  at time  $\tilde{t}$ , then receives from the protection buyer the amount  $\text{MTM}_\ell(\tilde{t})$ .

The aim of the curve trade investor is to maximize the P&L function that equals the total MTM value

$$\text{PL}(\tilde{t}) = \text{MTM}_1(\tilde{t}) - \text{MTM}_2(\tilde{t}). \quad (15)$$

## 5.2 Empirical Results

The key decision in constructing a curve trade is which tranche to buy and which to sell. If an investor entered a flattener on 20080909, then the trade incorporated two tranches whose spreads are depicted on the left panel of Figure 1. If the investor decided to close

the positions on 20090119, then their MTM values (14) were calculated using the spread quotes exhibited on the right panel of Figure 1 and using the base correlations (need for  $E\{F_\ell\}$ ) shown on the right panel of Figure 4. Having the data displayed on Figures 1 and 4, we can compute the MTM values of all tranches that were quoted on both days. In consequence, we can easily recover those two tranches that maximize the P&L function (15). However, it is only possible if we possess the whole market information from these two points in time.

With an efficient forecasting technique, one can compute a prediction, for a given time horizon, of each point that is displayed on the left panels of Figures 1 and 4. By doing it using standard econometric methods, each tranche from every series has to be treated as an individual time series. Disregarding the fact that there are many missing values in our data, see Table 2, we have many series that do not have a long history. If an investor bought a tranche from Series 9 on 20080320, the day of its launch and decided to sell it a day or a week later, then we might not have enough past observations to fit and forecast the model.

In the DSFM modelling we do not differentiate the indices by their series number but by their remaining time to maturity. If in the past we already had observations with a very long remaining time to maturity, then we are able to price upcoming series even before they appear on the market. Moreover, we can forecast them using the DSFM.

We carry out the forecasting of log-spreads and Z-transformed base correlations in moving windows. A moving window procedure is used when only the most recent data are considered to be relevant for the estimation. We impose a static window of  $w = 250$  days. Then for every time  $t_0$  between the day  $w$  and the last day  $T$  in our data, we analyse  $\{Y_t\}_{t=t_0-w+1}^{t_0}$ . For these sequences the DSFM (9) and (13) are estimated separately. As a result, we obtain  $T - w + 1$  times the estimated factor functions  $\hat{m} = (\hat{m}_0, \dots, \hat{m}_L)^\top$  and the series of the factor loadings  $\hat{Z}_t = (\hat{Z}_{t,0}, \dots, \hat{Z}_{t,L})^\top$  of length  $w$ . Since the factor functions are fixed, the forecasting is performed only on the factor loadings. As discussed in Section 4.2, we apply VAR(1) models to compute the predictions for a horizon  $h$  of one day, one week (five days), and one month (20 days). Due to the fixed scheme of issuing the iTraxx on the market, for every time  $t$ ,  $w + h \leq t \leq T$  we know which indices are traded. Therefore, the number of points that could be observed  $K_t$  and the possible remaining times to maturity  $\tau_t$  are known. Thus, the bivariate vector  $X_{t,k}$ ,  $k = 1, \dots, K_t$ , does not have to be forecasted. The forecast  $\hat{Y}_t$  is calculated from the  $\hat{Z}_t$  forecast. Finally, a proper inverse transformation is applied to  $\hat{Y}_t$  in order to recover the values of the spreads and the base correlations.

The calculation of the expected tranche losses needs as an input a homogeneous default probability, see (7). Since the spread predictions are calculated out-of-sample, we also forecast the default probabilities. The time series of the default probabilities are derived from the time series of all the iTraxx indices considered and are forecasted with an AR(1)-GARCH(1,1) process in moving windows of  $w$  observations. For short data histories we reduce the window size or take as a predictor the last observed value. All predicted values of spreads and base correlations that lead to an arbitrage in prices, i.e. negative spreads, default probabilities and base correlations outside  $[0, 1]$ , were excluded.

Afterwards, for every predicted  $\{\hat{s}_k(t), \hat{\rho}_k(t)\}$ ,  $t = w+h, \dots, T$ ,  $k = 1, \dots, K_t$ , we compute  $\widehat{\text{MTM}}_k(t)$  according to (14) where the initial spread is the spread observed on  $t - h$ . Consequently, we create a surface of the predicted MTM values, see Figure 11. Each surface has its extremes that indicate the tranches recommended for buying and selling.

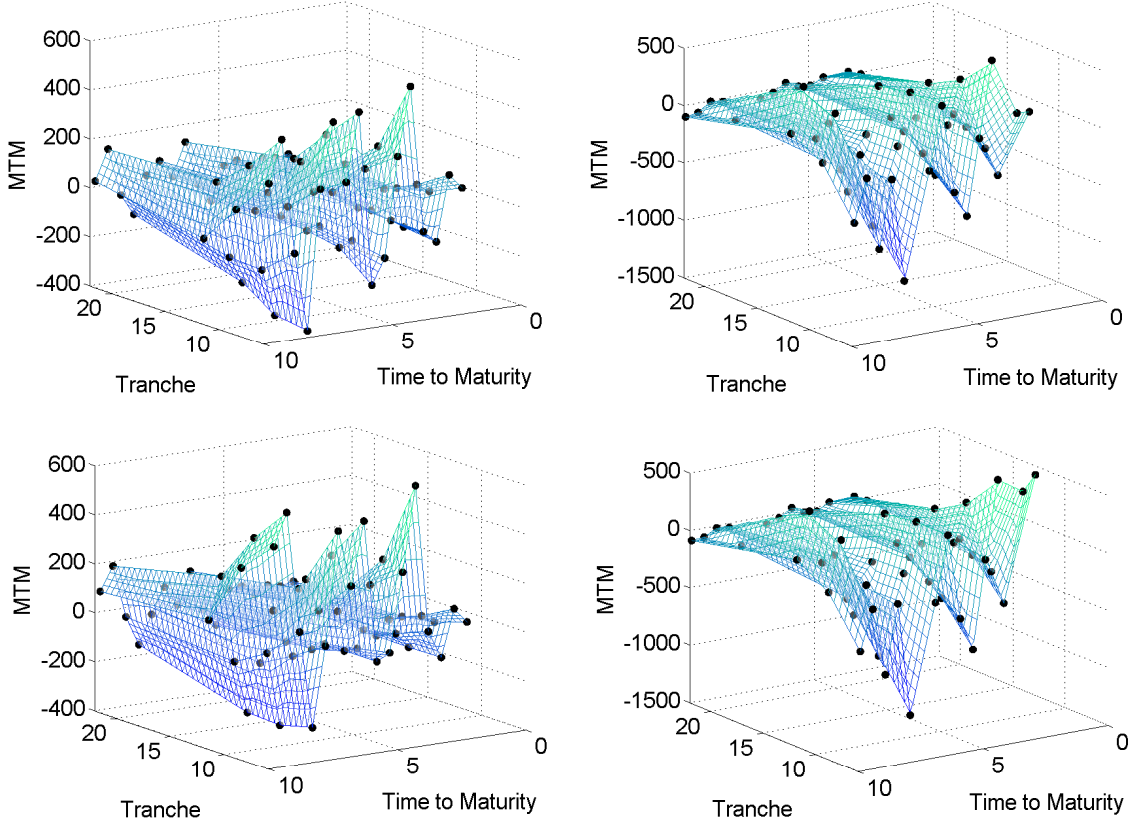


Figure 11: MTM surfaces on 20080909 (left) and 20090119 (right) calculated using one-day spread and base correlation predictions obtained with DSFM (upper panel) and DSFM without the mean factor (lower panel).

The empirical analysis of the curve trades' performance is conducted using tranches 2-5 for all dates and indices considered in Section 4.2. Since the equity tranche is quoted in percent as an upfront fee, its corresponding spread is significantly higher than the spreads of other tranches. As it causes a large skew of our spread surfaces, we excluded it from the study. In addition, we also avoid the multiple conversion of the upfront fee to the spread and back.

Buying and selling tranches involve transaction charges. However, we do not have information on trading costs neither the entire history of the bid and ask prices. We only analyse the bid-ask spreads of Series 8. Table 9 shows an average distance of the bid spread and of the ask spread from the mid spread as a percentage of the mid spread. For the investigation of the trading strategies, the tranche spread data used in this study are adjusted in the following way. The protection buyer delivers an ask spread that is calculated as a mid spread increased by a proper percent listed in Table 9. The protection seller receives a bid spread which is calculated as a mid spread reduced by this percentage.

The calculation of the spread and the MTM value of the tranche 2 requires as an input a

Maturity	1	2	3	4	5
5Y	1.88	1.78	2.52	3.77	6.28
7Y	1.49	1.65	2.31	2.97	4.87
10Y	1.41	1.66	1.83	2.52	4.09

Table 9: Average bid-ask spread excess over the mid spread as a percentage of the mid spread for Series 8 during the period 20070920-20090202

value of the base correlation of the equity tranche. Therefore, we conduct a preliminary analysis of the models (9) and (13) using all tranches and in this way obtain the forecast of the first tranche’s parameter.

For every day  $w \leq t \leq T - h$  we construct a curve trade. Namely, we fit and forecast the DSFM models and calculate  $h$ -day forecasts of the MTM surfaces. From these surfaces we recover which two tranches and from which series optimize a given strategy. We can e.g. consider a flattener that from all existing indices always buys the maturity 7Y, sells 5Y, and selects only the tranche 2. For a flattener and a steepener one can restrict the choice to a fixed tranche and fixed maturities or choose from all tranches and from all maturities. We also include a strategy that allows the investor to switch between flatteners and steepeners every day. If a strategy that combines flatteners and steepeners allows in addition choosing any tranche and any maturities, then the selected tranches are the maximum and the minimum of the forecasted MTM surface. If we consider the flatteners only, then we have to comply with the constraint  $\tau_1 - \tau_2 > 0$ . If for a particular day there are no tranches that for a given strategy return a positive P&L forecast, we assume that the investor decides not to take any action and we do not include this date in the overall summary of this strategy.

The accuracy of the predictions is evaluated by conducting a backtesting of the trades using the historical observations. For a given strategy and for the tranches selected by the DSFM forecasting procedures we check the corresponding observed market spreads, calculate the resulting MTM values, and register the realised P&L. Table 10 presents the overall means of the daily gains of different strategies given in percent. The labels in the first column should be read in the following way. Each name is composed of three parts joined with a hyphen. The first part indicates the type of the curve trade: flattener (F), steepener (S). The first five rows present results of the trades that combine flatteners and steepeners (FS). The second element of the name shows which tranches were considered: all tranches (AllT), only tranche 2 (T2), etc. The last part expresses the maturities taken: all maturities (AllM), 5Y and 10Y (510), etc. In the case of considering all maturities it is possible that the optimal strategy is composed of two tranches that have the same maturity, e.g. both are 5Y, but they come from different series and in consequence their remaining time to maturity are different. The spread predictions can alternatively be computed directly from the base correlations predictions by using (8). As a consequence it is not necessary to apply DSFMs to historical spreads. Table 10 presents also the results obtained by modelling and forecasting the Z-transformed base correlations only. Table 11 discloses the number of executed trades during the whole time period. Table 12 exhibits the corresponding Sharpe ratios calculated as the mean over the standard deviation.

To summarize the trading algorithm we enumerate its steps:

1. Consider a static rolling window of  $w$  and a forecasting horizon  $h$ .
2. For every  $t = w, \dots, T - h$  estimate the DSFMs (9) and (13) using  $\{Y_\nu\}_{\nu=t-w+1}^t$ .
3. Compute the forecast  $\hat{Z}_{t+h}$  using a VAR model.
4. Check what the possible  $X_{t+h}$  are and calculate  $\hat{Y}_{t+h}$ .
5. Transform  $\hat{Y}_{t+h}$  suitably to get  $\hat{s}(t+h)$  or  $\hat{\rho}(t+h)$ .
6. Compute the surface of forecasted MTM values with (14).
7. Select two tranches according to the strategy's restrictions.
8. Calculate the predicted P&L with (15).
9. Check the historical spread values for the selected tranches on day  $t+h$ . Imply the base correlations, calculate the realised MTM values and the realised P&L.

The results show that the highest daily gains were achieved by the strategies that invest in tranche 2 and 3. Obviously these tranches are quoted at the highest spreads but also carry the greatest risk. The steepeners for a fixed tranche and fixed maturities reveal a very good performance. However, as compared with Table 11 these strategies were rarely carried out which means that the conditions of these strategies were difficult to meet. The performance of the DSFM model (9) and (13) is comparable. The models based entirely on the predictions of the base correlations achieve better results for one-day and one-week forecasting horizon. The models that combine the spread predictions and the base correlations predictions show better results for one-month forecasting horizon. Since the forecasting for the longer time horizons is less accurate, we observe a significantly better performance of the trades designed for short term periods.

The curve trades can be also tested from the perspective of an investor that follows a certain strategy over long time horizon and constantly rebalances the trade. Assume that at  $t_0$  the investor enters an optimal curve trade for  $h$ -day horizon. At  $t_0 + h$  she either keeps the current position for the next  $h$ -days or closes it and enters a new one. In addition, we assume a margin of 10% of the notional. Every time the position is closed its gain is added or its loss is subtracted from the margin. If the loss exceeds the margin, the trade is closed. Otherwise, the investor follows her strategy for one year (250 days). If at  $\tilde{t}$  there are no positions that give a positive P&L forecast, the investor waits till  $\tilde{t} + 1$  and repeats the check.

We analyse a strategy of combined flatteners and steepeners from all tranches and all maturities using our entire data set. Figure 12 displays the final cumulated gains after one year of the investor as a function of the date on which the investor planned to finish trading (i.e. the starting date plus one year). The seasonality pattern observed on the plots appears because the strategy with  $h$ -day rebalancing constructed on day  $t_0$  and on day  $t_0 + h$  might overlap. If on  $t_0$  we started a 250-days strategy with a trade that was closed after  $h$ -days then a strategy started on  $t_0 + h$  differ from the previous one by the first and the last step only. Figure 13 shows an over time performance of the investor's trades started on 20070614 and closed on 20080529. The final cumulated profits equal 39.08%, 35.87%, and 20.35% when the account is rebalanced each day, week, and month respectively. These values are the last points depicted on the right plots on Figure 13 and they are presented on Figure 12 for the day 20080529.



Strategy	DSFM						DSFM without the mean factor					
	1 day		1 week		1 month		1 day		1 week		1 month	
	LZ	Z	LZ	Z	LZ	Z	LZ	Z	LZ	Z	LZ	Z
FS-AllT-AllM	0.29	0.35	0.10	0.13	0.05	0.04	0.30	0.30	0.11	0.13	0.04	0.03
FS-T2-AllM	0.29	0.33	0.13	0.14	0.06	0.05	0.33	0.28	0.12	0.13	0.05	0.04
FS-T3-AllM	0.19	0.22	0.07	0.08	0.03	0.02	0.18	0.23	0.07	0.07	0.02	0.02
FS-T4-AllM	0.14	0.17	0.04	0.05	0.02	0.01	0.12	0.18	0.04	0.05	0.01	0.01
FS-T5-AllM	0.09	0.11	0.04	0.04	0.02	0.01	0.08	0.11	0.03	0.04	0.01	0.01
F-T2-AllM	0.30	0.34	0.12	0.12	0.06	0.04	0.28	0.32	0.12	0.11	0.05	0.04
F-T3-AllM	0.16	0.20	0.06	0.07	0.02	0.01	0.16	0.20	0.06	0.07	0.02	0.01
F-T4-AllM	0.10	0.15	0.03	0.04	0.01	0.01	0.10	0.15	0.03	0.04	0.01	0.01
F-T5-AllM	0.09	0.10	0.03	0.03	0.01	0.01	0.08	0.10	0.03	0.03	0.01	0.01
S-T2-AllM	0.39	0.43	0.15	0.17	0.07	0.06	0.45	0.46	0.13	0.16	0.05	0.06
S-T3-AllM	0.27	0.31	0.09	0.10	0.04	0.03	0.30	0.35	0.09	0.09	0.02	0.02
S-T4-AllM	0.20	0.25	0.06	0.07	0.03	0.02	0.20	0.24	0.05	0.06	0.01	0.01
S-T5-AllM	0.12	0.15	0.04	0.04	0.02	0.02	0.12	0.16	0.04	0.04	0.01	0.02
F-AllT-105	0.20	0.21	0.07	0.09	0.03	0.02	0.19	0.21	0.06	0.08	0.02	0.02
F-AllT-107	0.22	0.26	0.07	0.08	0.03	0.03	0.25	0.25	0.08	0.08	0.03	0.03
F-AllT-75	0.15	0.15	0.04	0.06	0.01	-0.00	0.14	0.15	0.04	0.05	0.01	0.00
S-AllT-510	0.16	0.17	0.05	0.08	0.02	0.01	0.16	0.18	0.05	0.08	0.01	0.00
S-AllT-710	0.17	0.23	0.05	0.10	0.02	0.03	0.21	0.25	0.07	0.09	0.02	0.02
S-AllT-57	0.11	0.13	0.03	0.03	0.01	-0.01	0.12	0.13	0.03	0.03	0.00	-0.01
F-T2-105	0.30	0.26	0.13	0.14	0.06	0.04	0.26	0.26	0.12	0.13	0.04	0.04
F-T3-105	0.13	0.17	0.05	0.05	0.01	-0.01	0.12	0.17	0.05	0.05	0.01	-0.01
F-T4-105	0.12	0.12	0.03	0.03	0.01	-0.00	0.12	0.11	0.03	0.02	0.01	-0.01
F-T5-105	0.09	0.07	0.03	0.02	0.01	0.00	0.07	0.07	0.03	0.02	0.01	0.00
F-T2-107	0.33	0.30	0.12	0.12	0.06	0.04	0.34	0.29	0.14	0.12	0.06	0.05
F-T3-107	0.13	0.22	0.05	0.06	0.01	0.00	0.17	0.22	0.06	0.05	0.02	0.00
F-T4-107	0.11	0.13	0.03	0.03	0.01	0.00	0.12	0.13	0.03	0.03	0.01	0.00
F-T5-107	0.10	0.08	0.03	0.02	0.01	0.00	0.08	0.07	0.03	0.02	0.01	0.00
F-T2-75	0.16	0.17	0.05	0.07	0.02	0.00	0.17	0.18	0.06	0.06	0.02	0.00
F-T3-75	0.09	0.10	0.04	0.03	0.01	-0.01	0.09	0.11	0.04	0.03	0.01	-0.01
F-T4-75	0.06	0.07	0.02	0.02	0.01	-0.00	0.08	0.08	0.02	0.02	0.01	-0.00
F-T5-75	0.05	0.05	0.02	0.02	0.01	0.00	0.06	0.05	0.02	0.02	0.01	0.00
S-T2-510	0.22	0.32	0.16	0.40	0.10	0.19	0.33	0.48	0.13	0.35	0.05	0.12
S-T3-510	-0.01	0.10	0.01	0.03	0.04	-0.04	0.02	0.11	-0.03	0.04	-0.05	-0.04
S-T4-510	0.22	0.23	0.09	0.09	0.08	0.08	0.22	0.26	0.09	0.10	0.08	0.09
S-T5-510	0.13	0.13	0.07	0.07	0.04	0.05	0.09	0.13	0.07	0.07	0.04	0.05
S-T2-710	0.39	0.60	0.21	0.47	0.14	0.18	0.44	0.66	0.21	0.42	0.10	0.11
S-T3-710	0.35	0.54	0.12	0.20	0.10	0.04	0.41	0.56	0.09	0.23	0.03	0.04
S-T4-710	0.37	0.54	0.12	0.15	0.08	0.04	0.38	0.55	0.11	0.13	0.04	0.04
S-T5-710	0.25	0.28	0.10	0.12	0.05	0.05	0.24	0.29	0.12	0.13	0.05	0.05
S-T2-57	0.50	0.48	0.20	0.23	0.10	0.05	0.56	0.50	0.27	0.19	0.15	0.02
S-T3-57	0.40	0.39	0.15	0.07	0.08	0.03	0.40	0.39	0.16	0.09	0.09	0.03
S-T4-57	0.24	0.24	0.07	0.06	0.04	0.04	0.23	0.28	0.08	0.07	0.05	0.04
S-T5-57	0.20	0.18	0.07	0.05	0.03	0.02	0.20	0.18	0.07	0.05	0.03	0.02

Table 10: Mean of daily gains in percent. Calculations based on predictions of log-spreads and Z-transformed BCs marked as LZ; based only on Z-transformed BCs marked as Z.

Strategy	DSFM						DSFM without the mean factor					
	1 day		1 week		1 month		1 day		1 week		1 month	
	LZ	Z	LZ	Z	LZ	Z	LZ	Z	LZ	Z	LZ	Z
FS-AllT-AllM	754	754	750	750	735	735	754	754	750	750	735	735
FS-T2-AllM	748	751	746	748	734	735	752	753	745	748	729	735
FS-T3-AllM	752	754	749	750	735	735	754	754	750	750	735	735
FS-T4-AllM	754	754	750	750	734	735	754	754	750	750	735	734
FS-T5-AllM	741	736	735	737	735	733	741	744	735	739	735	733
F-T2-AllM	734	748	736	742	730	727	751	750	744	742	728	729
F-T3-AllM	752	754	749	750	734	727	754	754	750	750	734	735
F-T4-AllM	754	754	748	747	733	734	754	754	750	750	735	734
F-T5-AllM	741	734	733	734	734	733	741	739	733	734	735	733
S-T2-AllM	526	530	512	510	483	471	509	507	498	487	473	451
S-T3-AllM	466	433	458	430	432	406	435	423	427	412	423	401
S-T4-AllM	487	444	480	443	453	429	439	435	441	445	426	423
S-T5-AllM	512	454	500	448	494	439	474	455	459	462	472	443
F-AllT-105	695	753	694	749	701	731	698	754	700	750	701	727
F-AllT-107	731	754	731	750	726	730	736	754	730	749	726	730
F-AllT-75	722	750	728	747	723	734	724	750	714	748	717	733
S-AllT-510	736	744	726	743	709	730	723	744	711	741	708	722
S-AllT-710	750	745	741	746	729	735	726	748	716	748	717	735
S-AllT-57	745	752	742	748	721	731	747	749	740	746	717	727
F-T2-105	460	522	460	555	500	604	476	549	483	566	495	622
F-T3-105	551	706	556	698	541	680	576	718	587	705	571	690
F-T4-105	515	724	519	715	509	673	573	731	582	722	582	695
F-T5-105	502	695	497	692	508	655	587	696	572	689	577	660
F-T2-107	544	548	537	549	570	630	555	542	560	571	545	635
F-T3-107	627	701	632	696	613	708	635	704	635	703	616	712
F-T4-107	612	727	628	723	620	712	596	721	610	717	602	719
F-T5-107	582	713	570	710	576	702	604	707	590	702	595	709
F-T2-75	558	693	555	695	552	704	587	721	573	708	587	708
F-T3-75	596	728	597	734	618	714	627	727	616	732	637	718
F-T4-75	559	698	562	703	577	708	592	704	595	724	623	711
F-T5-75	645	707	642	713	637	702	650	704	645	718	644	703
S-T2-510	80	80	76	76	61	61	80	80	76	76	61	61
S-T3-510	69	65	65	65	50	50	69	61	65	57	50	50
S-T4-510	51	51	49	49	37	37	51	48	49	44	37	30
S-T5-510	50	47	49	42	36	28	49	47	48	42	35	28
S-T2-710	86	87	82	82	67	68	86	87	82	83	67	68
S-T3-710	82	84	79	77	64	61	81	84	77	78	62	61
S-T4-710	89	92	86	82	70	69	89	93	85	82	71	68
S-T5-710	120	121	116	112	101	95	122	121	119	113	102	95
S-T2-57	93	85	89	83	74	66	93	80	89	78	74	63
S-T3-57	104	91	100	82	84	72	103	91	99	83	83	64
S-T4-57	110	107	106	100	89	81	108	105	105	94	86	73
S-T5-57	121	114	120	108	97	78	109	116	108	108	85	78

Table 11: Number of executed trades. Calculations based on predictions of log-spreads and Z-transformed BCs marked as LZ; based only on Z-transformed BCs marked as Z.

Strategy	DSFM						DSFM without the mean factor					
	1 day		1 week		1 month		1 day		1 week		1 month	
	LZ	Z	LZ	Z	LZ	Z	LZ	Z	LZ	Z	LZ	Z
FS-AllT-AllM	0.20	0.24	0.20	0.25	0.31	0.23	0.20	0.20	0.21	0.25	0.24	0.21
FS-T2-AllM	0.20	0.22	0.25	0.28	0.35	0.32	0.21	0.19	0.23	0.25	0.30	0.30
FS-T3-AllM	0.19	0.22	0.23	0.26	0.28	0.17	0.18	0.23	0.23	0.23	0.21	0.19
FS-T4-AllM	0.19	0.23	0.21	0.23	0.26	0.25	0.17	0.23	0.18	0.23	0.16	0.26
FS-T5-AllM	0.24	0.29	0.32	0.34	0.40	0.28	0.24	0.29	0.26	0.33	0.31	0.29
F-T2-AllM	0.21	0.23	0.24	0.25	0.35	0.32	0.19	0.20	0.25	0.24	0.33	0.31
F-T3-AllM	0.16	0.20	0.20	0.26	0.24	0.16	0.17	0.21	0.20	0.24	0.20	0.17
F-T4-AllM	0.15	0.21	0.17	0.21	0.21	0.18	0.15	0.20	0.16	0.20	0.15	0.19
F-T5-AllM	0.25	0.27	0.32	0.33	0.34	0.26	0.23	0.28	0.26	0.33	0.27	0.27
S-T2-AllM	0.25	0.26	0.27	0.29	0.33	0.33	0.28	0.27	0.23	0.27	0.28	0.33
S-T3-AllM	0.24	0.27	0.30	0.29	0.30	0.25	0.27	0.29	0.29	0.27	0.20	0.21
S-T4-AllM	0.22	0.28	0.26	0.32	0.33	0.24	0.24	0.27	0.24	0.26	0.17	0.20
S-T5-AllM	0.27	0.32	0.32	0.35	0.33	0.34	0.28	0.32	0.30	0.34	0.32	0.34
F-AllT-105	0.17	0.20	0.18	0.21	0.18	0.17	0.16	0.19	0.15	0.20	0.11	0.15
F-AllT-107	0.18	0.23	0.16	0.19	0.19	0.19	0.21	0.22	0.17	0.18	0.15	0.22
F-AllT-75	0.14	0.14	0.11	0.15	0.04	-0.02	0.13	0.14	0.10	0.14	0.04	0.00
S-AllT-510	0.14	0.16	0.13	0.19	0.12	0.08	0.16	0.19	0.13	0.20	0.04	0.03
S-AllT-710	0.13	0.18	0.12	0.22	0.12	0.17	0.16	0.20	0.16	0.21	0.10	0.12
S-AllT-57	0.12	0.14	0.10	0.09	0.07	-0.09	0.14	0.15	0.10	0.09	0.03	-0.09
F-T2-105	0.22	0.21	0.27	0.32	0.29	0.29	0.19	0.19	0.24	0.29	0.22	0.31
F-T3-105	0.13	0.19	0.18	0.18	0.09	-0.07	0.13	0.19	0.15	0.18	0.06	-0.06
F-T4-105	0.16	0.18	0.14	0.13	0.13	-0.07	0.18	0.16	0.13	0.11	0.10	-0.08
F-T5-105	0.26	0.23	0.32	0.21	0.26	0.05	0.21	0.24	0.23	0.24	0.14	0.08
F-T2-107	0.25	0.26	0.25	0.29	0.35	0.32	0.25	0.24	0.30	0.29	0.36	0.34
F-T3-107	0.13	0.28	0.18	0.25	0.14	0.06	0.18	0.27	0.20	0.21	0.19	0.05
F-T4-107	0.16	0.19	0.18	0.17	0.15	0.03	0.18	0.19	0.18	0.16	0.19	0.03
F-T5-107	0.25	0.23	0.32	0.24	0.28	0.10	0.24	0.21	0.28	0.24	0.21	0.10
F-T2-75	0.14	0.16	0.15	0.22	0.15	0.05	0.15	0.17	0.18	0.17	0.17	0.03
F-T3-75	0.12	0.14	0.16	0.14	0.09	-0.12	0.13	0.15	0.16	0.16	0.13	-0.10
F-T4-75	0.13	0.16	0.13	0.14	0.15	-0.08	0.16	0.16	0.16	0.13	0.19	-0.07
F-T5-75	0.20	0.20	0.25	0.21	0.18	0.10	0.22	0.22	0.27	0.20	0.18	0.11
S-T2-510	0.08	0.14	0.14	0.38	0.25	0.48	0.14	0.20	0.12	0.33	0.12	0.29
S-T3-510	-0.01	0.07	0.02	0.06	0.15	-0.17	0.02	0.07	-0.07	0.06	-0.24	-0.17
S-T4-510	0.17	0.18	0.23	0.23	0.47	0.47	0.17	0.20	0.23	0.24	0.47	0.45
S-T5-510	0.23	0.22	0.42	0.40	0.60	0.75	0.18	0.22	0.40	0.40	0.60	0.75
S-T2-710	0.13	0.22	0.19	0.46	0.37	0.52	0.16	0.25	0.21	0.43	0.28	0.32
S-T3-710	0.17	0.26	0.18	0.32	0.34	0.19	0.19	0.28	0.13	0.38	0.12	0.19
S-T4-710	0.25	0.38	0.27	0.39	0.45	0.31	0.26	0.39	0.24	0.35	0.31	0.32
S-T5-710	0.40	0.42	0.54	0.64	0.74	0.66	0.38	0.44	0.60	0.66	0.70	0.62
S-T2-57	0.22	0.20	0.24	0.24	0.32	0.14	0.26	0.22	0.34	0.21	0.46	0.07
S-T3-57	0.28	0.26	0.29	0.13	0.40	0.15	0.27	0.26	0.30	0.17	0.46	0.14
S-T4-57	0.30	0.27	0.28	0.23	0.37	0.38	0.28	0.31	0.31	0.25	0.44	0.35
S-T5-57	0.49	0.42	0.64	0.46	0.71	0.55	0.45	0.42	0.66	0.46	0.72	0.55

Table 12: Sharpe ratios. Calculations based on predictions of log-spreads and Z-transformed BCs marked as LZ; based only on Z-transformed BCs marked as Z.

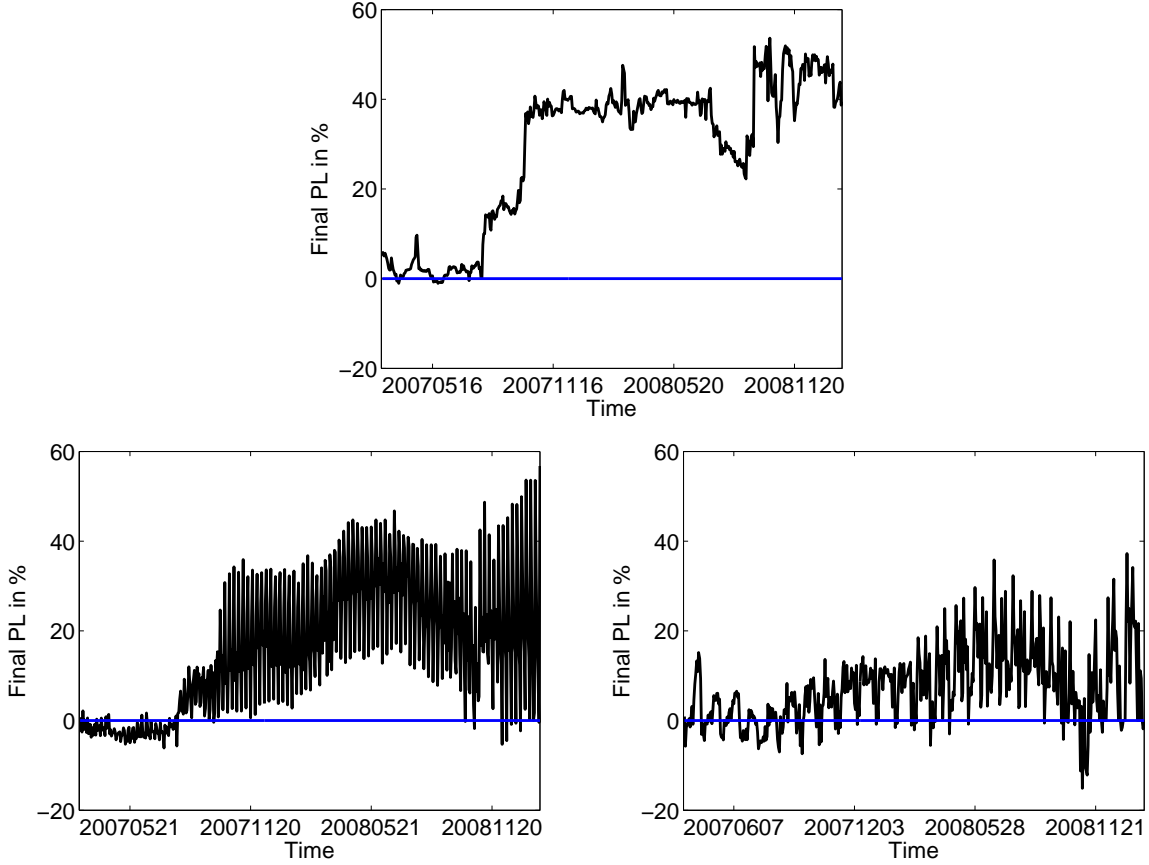


Figure 12: Closing profits of investor's strategies after one year. Rebalancing after: 1 day (upper panel), 1 week (lower panel, left), 1 month (lower panel, right). Calculations based on the DSFM predictions of log-spreads and Z-transformed BCs.

## 6 Conclusions

This work investigates dynamics of collateralized debt obligations (CDOs) by modelling the evolution of tranche spread surfaces and base correlation surfaces using a dynamic semiparametric factor model (DSFM). The empirical study is conducted using an extensive data set of 49,502 observations of iTraxx Europe tranches of Series 2 to 10 for the time period between 30 March 2005 and 2 February 2009. The base correlations are implied from spreads using the large pool Gaussian copula model. The tranche spreads and the base correlations are represented as a function of the tranche seniority and the remaining time to maturity. Every day data appear in a small number of curves that form a surface in the three-dimensional space. As time passes, the surfaces move through the space towards expiry and simultaneously change their shapes. The DSFM captures their evolution simultaneously in space and time dimensions by a small number of factors. We propose a modification of the classic DSFM and of the criterion of choosing the number of factors. The results show that both DSFMs successfully reproduce the dynamics in data. The study is completed by presenting an application in trading strategies. We show how DSFM can be used in constructing the curve trades. Based on the DSFM predictions of the spread and base correlation surfaces we calculate the predictions of the marked-to-market (MTM) surfaces for different investment horizons. We analyse the performance

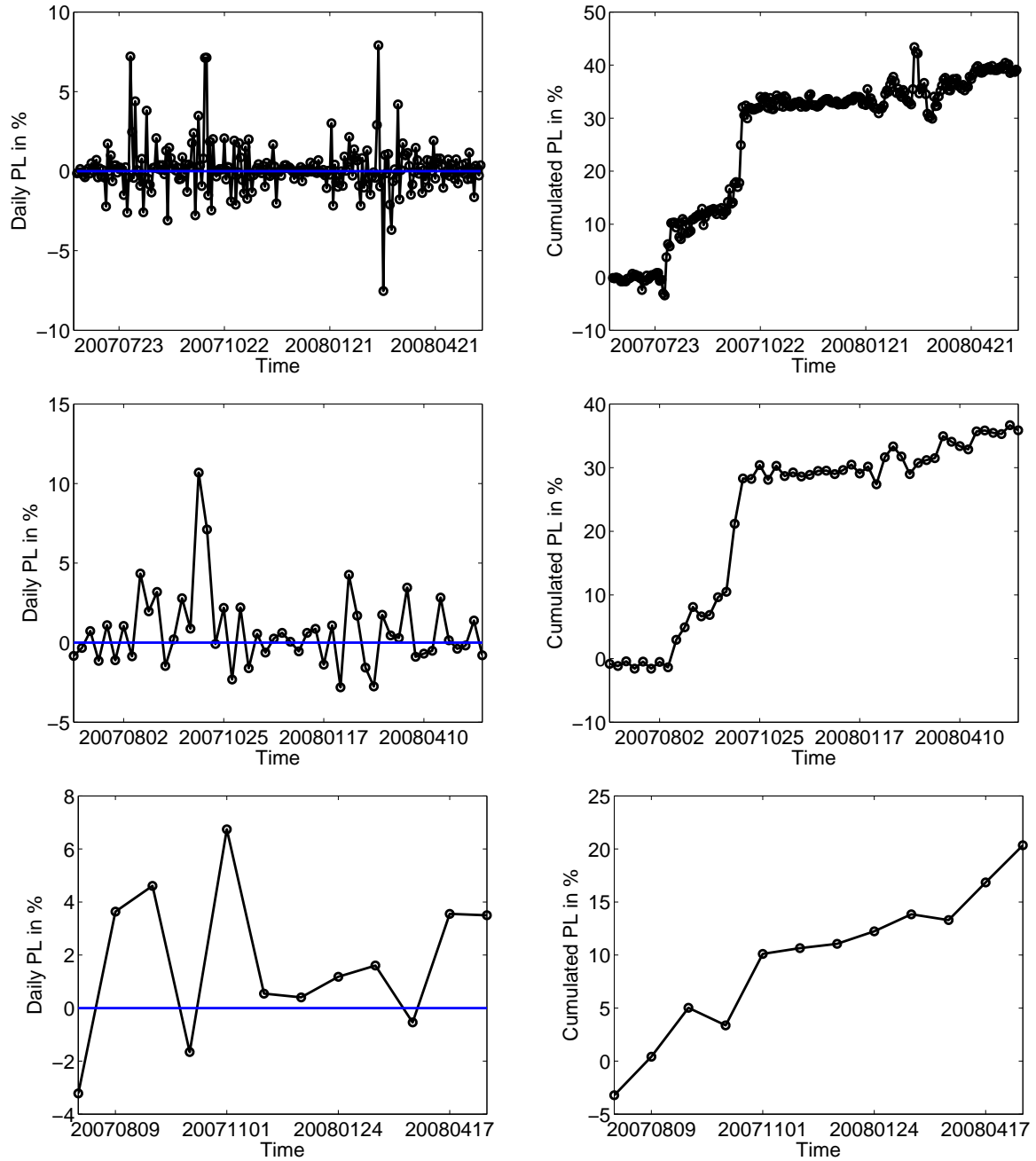


Figure 13: Investor's strategy. Start date 20070614, end date 20080529. Rebalancing after: 1 day (upper panel), 1 week (middle panel), 1 month (lower panel). Daily P&L (left) and daily cumulated P&L (right) over one year. Calculations based on the DSFM predictions of log-spreads and Z-transformed BCs.

of 43 strategies that combine different positions, tranches, and maturities. A backtesting using historical data shows that the curve trades achieve high daily gains.

## References

- Ağca, S., Agrawal, D. & Islam, S. (2008). Implied correlations: Smiles or smirks?, *The Journal of Derivatives* **16**(2): 7–35.
- Bielecki, T. R. & Rutkowski, M. (2004). *Credit Risk: Modelling Valuation and Hedging*, Springer Finance.
- Bluhm, C. & Overbeck, L. (2006). *Structured Credit Portfolio Analysis, Baskets and CDOs*, CRC Press LLC.
- Burtshell, X., Gregory, J. & Laurent, J.-P. (2009). A comparative analysis of cdo pricing models under the factor copula framework, *The Journal of Derivatives* **16**(4): 9–37.
- de Boor, C. (2001). *A Practical Guide to Splines*, Springer-Verlag.
- Detlefsen, K. & Härdle, W. K. (2013). Variance swap dynamics, *Quantitative Finance* **13**(5): 675–685.
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics* **20**(3): 339–350.
- Felsenheimer, J., Gisdakis, P. & Zaiser, M. (2004). DJ iTraxx: Credit at its best!, *Credit derivatives special*, HVB Corporates & Markets.
- Fengler, M., Härdle, W. K. & Mammen, E. (2007). A semiparametric factor model for implied volatility surface dynamics, *Journal of Financial Econometrics* **5**(2): 189–218.
- Giacomini, E., Härdle, W. K. & Krätschmer, V. (2009). Dynamic semiparametric factor models in risk neutral density estimation, *AStA - Advances in Statistical Analysis* **93**(4): 387–402.
- Gourieroux, C. & Jasiak, J. (2001). Dynamic factor models, *Econometric Reviews* **20**(4): 385–424.
- Gromenko, O., Kokoszka, P., Zhu, L., & Sojka, J. (2012). Estimation and testing for spatially distributed curves with application to ionospheric and magnetic field trends, *Annals of Applied Statistics* **6**(2): 669–696.
- Hallin, M. & Liška, R. (2007). Determining the number of factors in the general dynamic factor model, *Journal of the American Statistical Association* **102**(478): 603–617.
- Hamerle, A., Igl, A. & Plank, K. (2012). Correlation smile, volatility skew, and systematic risk sensitivity of tranches, *The Journal of Derivatives* **19**(3): 8–27.
- Härdle, W. K., Hautsch & Mihoci (2012). Modelling and forecasting liquidity supply using semiparametric factor dynamics, *Journal of Empirical Finance* **19**(4): 610–625.

- Kakodkar, A., Galiani, S., Jónsson, J. G. & Gallo, A. (2006). Credit derivatives handbook, Vol. 2: A guide to the exotics credit derivatives market, *Technical report*, Merrill Lynch.
- Kauermann, G., Krivobokova, T. & Fahrmeir, L. (2009). Some asymptotic results on generalized penalized spline smoothing, *Journal of the Royal Statistical Society, Series B* **71**(2): 487–503.
- Kneip, A. & Gasser, T. (1992). Statistical tools to analyze data representing a sample of curves, *Annals of Statistics* **20**(3): 1266–1305.
- Koopman, S. J., Lucas, A. & Schwaab, B. (2012). Dynamic factor models with macro, frailty, and industry effects for U.S. default counts: The credit crisis of 2008, *Econometric Reviews* **30**(4): 521–532.
- Li, D. X. (2000). On default correlation: a copula function approach, *The Journal of Fixed Income* **9**(4): 43–54.
- Liu, X. & Müller, H.-G. (2004). Functional convex averaging and synchronization for time-warped random curves, *Journal of the American Statistical Association* **99**(467): 687–699.
- McGinty, L. & Ahluwalia, R. (2004). A model for base correlation calculation, *Technical report*, JP Morgan.
- Myšičková, A., Song, S., Majer, P., Mohr, P., Heekeren, H. & Härdle, W. K. (2013). Risk patterns and correlated brain activities. multidimensional statistical analysis of fMRI data in economic decision making study, *Psychometrika* . Forthcoming.
- Park, B., Mammen, E., Härdle, W. K. & Borak, S. (2009). Time series modelling with semiparametric factor dynamics, *Journal of the American Statistical Association* **104**(485): 284–298.
- Roy, R. (2007). Trading credit tranches: Taking default correlation out of the black box, in A. Rajan, G. McDermott & R. Roy (eds), *The Structured Credit Handbook*, John Wiley & Sons.
- Song, S., Härdle, W. K. & Ritov, Y. (2013). High dimensional nonstationary time series modelling with generalized dynamic semiparametric factor model, *Discussion paper*, SFB 649 (2010-039), Humboldt Universität zu Berlin, invited resubmission to *The Econometrics Journal*.
- Vasicek, O. (1987). Probability of loss on loan portfolio, *Technical report*, KMV Corporation.

# SFB 649 Discussion Paper Series 2013

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

- 001 "Functional Data Analysis of Generalized Quantile Regressions" by Mengmeng Guo, Lhan Zhou, Jianhua Z. Huang and Wolfgang Karl Härdle, January 2013.
- 002 "Statistical properties and stability of ratings in a subset of US firms" by Alexander B. Matthies, January 2013.
- 003 "Empirical Research on Corporate Credit-Ratings: A Literature Review" by Alexander B. Matthies, January 2013.
- 004 "Preference for Randomization: Empirical and Experimental Evidence" by Nadja Dwenger, Dorothea Kübler and Georg Weizsäcker, January 2013.
- 005 "Pricing Rainfall Derivatives at the CME" by Brenda López Cabrera, Martin Odening and Matthias Ritter, January 2013.
- 006 "Inference for Multi-Dimensional High-Frequency Data: Equivalence of Methods, Central Limit Theorems, and an Application to Conditional Independence Testing" by Markus Bibinger and Per A. Mykland, January 2013.
- 007 "Crossing Network versus Dealer Market: Unique Equilibrium in the Allocation of Order Flow" by Jutta Dönges, Frank Heinemann and Tijmen R. Daniëls, January 2013.
- 008 "Forecasting systemic impact in financial networks" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, January 2013.
- 009 "'I'll do it by myself as I knew it all along': On the failure of hindsight-biased principals to delegate optimally" by David Danz, Frank Hüber, Dorothea Kübler, Lydia Mechtenberg and Julia Schmid, January 2013.
- 010 "Composite Quantile Regression for the Single-Index Model" by Yan Fan, Wolfgang Karl Härdle, Weining Wang and Lixing Zhu, February 2013.
- 011 "The Real Consequences of Financial Stress" by Stefan Mittnik and Willi Semmler, February 2013.
- 012 "Are There Bubbles in the Sterling-dollar Exchange Rate? New Evidence from Sequential ADF Tests" by Timo Bettendorf and Wenjuan Chen, February 2013.
- 013 "A Transfer Mechanism for a Monetary Union" by Philipp Engler and Simon Voigts, March 2013.
- 014 "Do High-Frequency Data Improve High-Dimensional Portfolio Allocations?" by Nikolaus Hautsch, Lada M. Kyj and Peter Malec, March 2013.
- 015 "Cyclical Variation in Labor Hours and Productivity Using the ATUS" by Michael C. Burda, Daniel S. Hamermesh and Jay Stewart, March 2013.
- 016 "Quantitative forward guidance and the predictability of monetary policy – A wavelet based jump detection approach –" by Lars Winkelmann, April 2013.
- 017 "Estimating the Quadratic Covariation Matrix from Noisy Observations: Local Method of Moments and Efficiency" by Markus Bibinger, Nikolaus Hautsch, Peter Malec and Markus Reiss, April 2013.
- 018 "Fair re-valuation of wine as an investment" by Fabian Y.R.P. Bocart and Christian M. Hafner, April 2013.
- 019 "The European Debt Crisis: How did we get into this mess? How can we get out of it?" by Michael C. Burda, April 2013.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".





# SFB 649 Discussion Paper Series 2013

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

- 020 "Disaster Risk in a New Keynesian Model" by Maren Brede, April 2013.
- 021 "Econometrics of co-jumps in high-frequency data with noise" by Markus Bibinger and Lars Winkelmann, May 2013.
- 022 "Decomposing Risk in Dynamic Stochastic General Equilibrium" by Hong Lan and Alexander Meyer-Gohde, May 2013.
- 023 "Reference Dependent Preferences and the EPK Puzzle" by Maria Grith, Wolfgang Karl Härdle and Volker Krätschmer, May 2013.
- 024 "Pruning in Perturbation DSGE Models - Guidance from Nonlinear Moving Average Approximations" by Hong Lan and Alexander Meyer-Gohde, May 2013.
- 025 "The 'Celtic Crisis': Guarantees, transparency, and systemic liquidity risk" by Philipp König, Kartik Anand and Frank Heinemann, May 2013.
- 026 "State Price Densities implied from weather derivatives" by Wolfgang Karl Härdle, Brenda López-Cabrera and Huei-Wen Teng, May 2013.
- 027 "Bank Lending Relationships and the Use of Performance-Sensitive Debt" by Tim R. Adam and Daniel Streitz, May 2013.
- 028 "Analysis of Deviance in Generalized Partial Linear Models" by Wolfgang Karl Härdle and Li-Shan Huang, May 2013.
- 029 "Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps" by Markus Bibinger and Mathias Vetter, May 2013.
- 030 "Can expert knowledge compensate for data scarcity in crop insurance pricing?" by Zhiwei Shen, Martin Odening and Ostap Okhrin, May 2013.
- 031 "Comparison of Methods for Constructing Joint Confidence Bands for Impulse Response Functions" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, May 2013.
- 032 "CDO Surfaces Dynamics" by Barbara Choroś-Tomczyk, Wolfgang Karl Härdle and Ostap Okhrin, July 2013.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

