Financ Mark Portf Manag (2010) 24: 289–308
DOI 10.1007/s11408-010-0136-8

Modeling the evolution of implied CDO correlations

Marius Hofert · Matthias Scherer · Rudi Zagst

Published online: 30 June 2010

Abstract CDO tranche spreads (and prices of related portfolio-credit derivatives) depend on the market's perception of the future loss distribution of the underlying credit portfolio. Applying Sklar's seminal decomposition to the distribution of the vector of default times, the portfolio-loss distribution derived thereof is specified through individual default probabilities and the dependence among obligors' default times. Moreover, the loss severity, specified via obligors' recovery rates, is an additional determinant. Several (specifically univariate) credit derivatives are primarily driven by individual default probabilities, allowing investments in (or hedging against) default risk. However, there is no derivative that allows separately trading (or hedging) default correlations; all products exposed to correlation risk are contemporaneously also exposed to default risk. Moreover, the abstract notion of *dependence* among the names in a credit portfolio is not directly observable from traded assets. Inverting the classical Vasicek/Gauss copula model for the correlation parameter allows constructing time series of implied (compound and base) correlations. Based on such time series, it is possible to identify observable variables that describe implied correlations in terms of a regression model. This provides an economic model of the time evolution of the market's view of the dependence structure. Different regression models are developed and investigated for the European CDO market. Applications and extensions to other markets are discussed.

M. Hofert (✉)
Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland
e-mail: marius.hofert@math.ethz.ch

M. Scherer · R. Zagst
HVB Stiftungsinstitut für Finanzmathematik, Technische Universität München, Parkring 11, 85748 Garching-Hochbrück, Germany

M. Scherer
e-mail: scherer@tum.de

R. Zagst
e-mail: zagst@tum.de

Keywords CDO · Implied correlation · Gaussian copula model

JEL Classification C 13 · C 52 · G 01 · G 13

1 Introduction

The market for credit derivatives grew steadily from its beginning in the early 1990s, until the recent financial crisis put an abrupt end to this trend. Yet, it is still one of the most important markets for derivatives, for several reasons: (a) the search for high-yield investments in a low-interest rate environment, (b) regulatory issues/rules that allow banks to control credit risk with derivatives, (c) the use of credit derivatives for risk management of credit portfolios, and (d) the opportunity to tap the market for new investments such as correlation risk. The creation of credit indices and a standardization of, e.g., default events and several credit derivatives, have made the market more transparent and liquid (for a survey, see, e.g., Brommundt et al. 2006). Moreover, market quotes for *credit default swaps* (CDS), *collateralized debt obligations* (CDO), and other products are now publicly available.

One type of product that has attracted a great deal of interest from both researchers and practitioners is the CDO. The important feature of a CDO is that the spread of its tranches depends to a large extent on the correlation structure among the names in the underlying credit portfolio, thus creating a new type of risk. The core idea behind a CDO is to pool multiple credits (or other credit-risky assets) and then partition this portfolio into slices with different risk profile, called *tranches*. Creating tranches with different seniority allows for senior debt as well as more risky investments within the same portfolio. Hence, it is possible to satisfy the risk appetites of a variety of investors. Crucial for pricing and risk management of CDOs is that dependence among the portfolio constituents affects the default risk (and, hence, the market price) of each tranche. However, *dependence* is not directly observable in the markets. Moreover, a great many models/dependence structures (typically represented by some copula) have been suggested to explain dependent defaults. Inspired by structural-default models that link equity to credit risk, correlation of the respective equity prices is sometimes suggested as an approximation for default correlation (see, e.g., O'Kane and Livesey 2004). Taking advantage of the current situation of relatively liquid market prices for CDO tranches, we can alternatively infer time series of implied correlations by inverting Vasicek's model.[1] This model (see Vasicek 1987; Li 2000) is often criticized for its simplifying assumptions, and several generalizations (e.g., by Andersen et al. 2003; Gregory and Laurent 2003, 2004; Hull and White 2004; Andersen and Sidenius 2005; Albrecher et al. 2007; Burtschell et al. 2007, 2009, and Kalemanova et al. 2007) have been proposed. Nevertheless, as the model continues to be used for practical applications, it is applied here to derive time series of compound and base correlations. Below, we address the question of whether (and, if so, using which variables) it is

[1] An empirical study based on historical (default) correlations (opposed to implied correlations used in the present investigation) is presented in Rösch (2003).

possible to explain implied correlations. To approach this problem, we investigate different regression models for compound and base correlations and test the significance and explanatory power of the variables used. To validate the models, we compare the resulting model spreads with a direct regression on tranche spreads.[2]

The remainder of this paper is organized as follows. After a brief introduction to CDOs in Sect. 2, the pricing of CDO tranches and the derivation of implied correlations is reviewed. The data for the statistical investigation are presented in Sect. 3. The results of the empirical study are discussed in Sects. 4 and 5. Possible extensions and applications are discussed in Sect. 6. Section 7 concludes.

2 Pricing CDOs and computing implied correlations

CDOs allow the originator to sell default risk in tranches, each of which has different seniority. Until its complete elimination, the riskiest tranche (the *equity tranche*) bears all losses within the credit portfolio. As soon as the equity tranche is eliminated, the subsequent tranche (the *junior mezzanine tranche*) is affected. This process continues up to the most senior tranche. There are two types of CDOs: *cash CDOs* (investors buy some tranche and subsequently receive interest payments) and *synthetic CDOs* (investors exchange premium payments against potential loss compensation). We focus on the latter, for which we have market quotes. Here, constructed as a swap contract, tranche spreads are found by equating the expected discounted premium and default legs and by solving this relation for the fair spread. Payments of each tranche's premium leg are typically made on a quarter-yearly schedule and depend on the remaining nominal of the corresponding tranche at the respective payment date. Default payments are due if losses affect the respective tranche. For simplicity, default payments are deferred to the next premium payment date. The CDO's tranches are specified via their lower and upper attachment points, respectively, denoted by l_j and u_j, $j \in \{1, \dots, J\}$. According to the *iTraxx Europe* convention, the portfolio under consideration is partitioned into six tranches: $[0\%, 3\%]$, $[3\%, 6\%]$, $[6\%, 9\%]$, $[9\%, 12\%]$, $[12\%, 22\%]$, and $[22\%, 100\%]$. Market quotes are available for the first five of these, indexed by $j \in \{1, \dots, 5\}$. Denoting the overall (relative) portfolio-loss process by $L_t \in [0, 1]$, the loss affecting tranche j is given by

$$L_{t,j} = \min\{L_t, u_j\} - \min\{L_t, l_j\}, \quad t \in [0, T]. \tag{1}$$

Other contractual details (e.g., payment conventions, assumptions as to recovery rates and default events) are set as close as possible to the conventions of the *International Index Company*. These specifications will be explained more thoroughly as needed throughout the paper.

Computing implied correlations, i.e., compound or base correlations, for the tranches of a CDO is related to computation of implied volatilities from observed option prices. More precisely, the correlation parameter ρ in Vasicek's model is chosen such that the model resembles the quoted spread of the considered tranche. To accomplish this, a pricing formula for CDO tranche spreads is needed. Hence, we next briefly review computation of model prices.

[2] A related study, aimed at explaining CDS spreads, is presented in Düllmann and Sosinska (2007).

2.1 The portfolio-loss distribution in Vasicek's model

Throughout this investigation, we work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The measure $\mathbb{P}$ is a pricing measure that is calibrated to market quotes in a later step. Generally speaking, the idea behind Vasicek's model (see Vasicek 1987) is to extend Merton's seminal univariate firm-value model (see Merton 1974) to a multivariate default model. Dependence between the individual default events is introduced via correlated firm-value processes. After each firm-value process in the multivariate Merton-type model is centered and scaled to follow a standard normal distribution at time t (originally, Merton used a geometric Brownian motion), dependence among the I firms is introduced via a market factor construction. More precisely, the factors $\{X_t^i\}_{i=1,\dots,I}$ are decomposed into $X_t^i := \sqrt{\rho}\, M_t + \sqrt{1-\rho}\, \varepsilon_i$. The random variable $M_t \sim \mathcal{N}(0,1)$ is interpreted as a *market factor* affecting all firms. This factor is assumed to be independent of the independent and identically distributed (i.i.d.) random variables $\varepsilon_i \sim \mathcal{N}(0,1)$ representing the *idiosyncratic residual risk* of each firm. Further simplifying assumptions, necessary for applying stochastic limit theorems at a later stage, postulate a homogeneous portfolio with respect to default probabilities, recovery rates, and portfolio weights. The default probability (up to time t) of all firms is a known input factor, denoted by $p_t := \mathbb{P}(\tau_i \le t)$. The time t default probability is set via the default threshold D, which is also identical for all firms. The crucial observation for the derivation of a *large portfolio approximation*, referred to as *conditional independence*, is that all firms are independent given the market factor M_t, and that

$$p_t(m) := \mathbb{P}\big(X_t^i \le D \,|\, M_t = m\big) = \Phi\left(\frac{D - \sqrt{\rho}\, m}{\sqrt{1-\rho}}\right), \quad m \in \mathbb{R},$$

is the conditional default probability (of all firms) up to time t given $M_t = m$. The unconditional probability of $i \in \{0, \dots, I\}$ defaults in the portfolio is computed by integrating out the market factor distribution. Moreover, assuming a sufficiently large portfolio (the *large-portfolio assumption*), the following approximation of the portfolio-loss distribution (with zero recovery), denoted L_t^0, can be derived via the central limit theorem (see Vasicek 1987; Schönbucher 2003),

$$F_{L_t^0}(x) := \mathbb{P}\big(L_t^0 \le x\big) \approx \Phi\left(\frac{\sqrt{1-\rho}\, \Phi^{-1}(x) - \Phi^{-1}(p_t)}{\sqrt{\rho}}\right), \quad x \in [0,1].$$

2.2 Pricing the tranches of a CDO

Consider a synthetic CDO, constructed using CDS on I obligors, each contributing $1/I$ to the unit nominal of the portfolio. The initial maturity (in years) of all swap contracts is denoted by T, the initial payment schedule by $\mathcal{T} := \{t_0 = 0 < t_1 < \cdots < t_n = T\}$. Because it is a swap contract, we consider two legs: the *premium leg* and the *default leg*. The former is based on the schedule $\mathcal{T}$, where quarter-yearly payments are most common. Default payments of the latter leg are deferred to the subsequent premium payment date after some default. Accrued interest, i.e., the interest accumulated between a default and the last payment date, is approximated by assuming that default payments occur in the middle of two payment dates. Pricing the tranches of a CDO corresponds to equating the present values of expected premium and default payments and solving for the spread. To simplify computations, it is common to assume a homogeneous deterministic recovery rate (set to $R := 40\%$), which gives

$$L_t := \frac{1-R}{I}\sum_{i=1}^{I}\mathbb{1}_{\{\tau_i \le t\}} = (1-R)L_t^0, \quad t \in [0,T].$$

Given L_t, the loss affecting tranche $j \in \{1,\dots,J\}$ is found from (1) and denoted $L_{t,j}$. Given the payment schedule $\mathcal{T}$ and the discount factors d_{t_k} corresponding to the time points t_k, $k \in \{1,\dots,n\}$, the fair spread of tranche j with maturity T is computed as

$$s^j = \mathbb{E}\left[\sum_{k=1}^{n} d_{t_k}(L_{t_k,j} - L_{t_{k-1},j})\right] \Big/ \mathbb{E}\left[\sum_{k=1}^{n} d_{t_k}\Delta t_k\big(N_{t_k,j} + (N_{t_{k-1},j} - N_{t_k,j})/2\big)\right], \tag{2}$$

where $\Delta t_k := (t_k - t_{k-1})$ and the remaining nominal at time t in tranche j is computed as $N_{t,j} := u_j - l_j - L_{t,j}$. The summands $(N_{t_{k-1}} - N_{t_k})/2$ account for accrued interest. An exception to the above pricing formula is made for the first tranche, for which it is standard market practice to combine a running spread of 500 basis points (one basis point is 0.0001) with an upfront payment. This upfront payment is quoted as a percentage of the nominal of the equity tranche. All the above formulas can easily be solved if discount factors are deterministic and the portfolio-loss distribution is analytically available. In the present framework it is even possible to explicitly derive expected tranche losses (see below).

Theorem 2.1 (O'Kane and Livesey 2004) *The expected tranche loss of the equity tranche in Vasicek's model is*

$$\mathbb{E}[L_{t,1}] = (1-R)\big(p_t - \boldsymbol{\Phi}\big(-\Phi^{-1}\big(\min\{u_1/(1-R),1\}\big), \Phi^{-1}(p_t), -\sqrt{1-\rho}\big)\big).$$

For more senior tranches $j \in \{2,\dots,J\}$, the expected tranche loss is given by

$$\begin{aligned}\mathbb{E}[L_{t,j}] = (1-R)\big(&\boldsymbol{\Phi}\big(-\Phi^{-1}\big(\min\{l_j/(1-R),1\}\big), \Phi^{-1}(p_t), -\sqrt{1-\rho}\big)\\ &- \boldsymbol{\Phi}\big(-\Phi^{-1}\big(\min\{u_j/(1-R),1\}\big), \Phi^{-1}(p_t), -\sqrt{1-\rho}\big)\big),\end{aligned}$$

where $\boldsymbol{\Phi}(x,y,\rho)$ denotes the cumulative distribution function of the bivariate standard normal distribution with correlation coefficient ρ.

The proof can be found in the Appendix to this paper.

2.3 Computing compound and base correlations

As explained above, computing *compound correlations* corresponds to inverting the CDO pricing formula, derived from Vasicek's model of the portfolio-loss distribution, for the dependence parameter ρ. For this purpose, necessary input variables are individual default probabilities. In our analysis, these are taken from portfolio CDS spreads, computed under the assumption of a constant homogeneous default intensity. However, tranche spreads of mezzanine tranches are not guaranteed to be monotone in the parameter ρ. Hence, the compound correlation is not necessarily unique. Therefore, we choose the smallest solution if at least one solution exists and the minimizing argument of the distance between the model spread and the market quote if no solution exists.[3]

Base correlations refer to artificial tranches covering $[0\%, x\%]$ of the portfolio, i.e., the lower attachment point is always zero. We observe that the payment streams of tranche $[x\%, y\%]$ agree with a long position in tranche $[0\%, y\%]$ combined with a short position in tranche $[0\%, x\%]$. This observation is used to construct a bootstrapping procedure that computes the base correlation of tranche j from previously computed base correlations and the market spread of tranche j. The base correlation of the equity tranche agrees with the compound correlation of this tranche. A detailed description of computing base correlations is given in O'Kane and Livesey (2004). This procedure is analogous to the well-known bootstrapping method for computing zero rates from swap par yields (see, e.g., Hull 2008, p. 80).

3 An empirical study: data and regression models

Our empirical study is based on the fifth (a calm market prior to the financial crisis) and ninth (a period during the financial crisis) series of the Markit iTraxx Europe indices. The underlying portfolio consists of CDS on 125 equally weighted European companies. We focus on CDO tranches and portfolio CDS for contracts maturing on 2011-06-20 and 2013-06-20, respectively. Moreover, data of individual CDS maturing in five years on all 125 companies are collected. Additionally, we collected all available asset values of the related companies (excluding privately owned companies and other firms without stock quotations) and risk-free interest rates for the Eurozone. The periods under consideration cover 2006-03-20 to 2007-03-15 (Series 5) and 2008-04-17 to 2008-08-15 (Series 9). Overall, the data basis is quite solid. Only a few individual CDS spreads are missing, in which case we use the last available quote. This investigation focuses on Series 5. Later in the paper, a condensed version of the results for Series 9 is provided.

3.1 Compound and base correlations

Based on the assumptions of Sect. 2 and the data described above we obtain time series of compound and base correlations (see Fig. 1).

Compound correlations show the typical *correlation smile*, i.e., they are increasing from the junior mezzanine toward the super-senior tranche and the equity tranche's

[3] This approach could lead to unrealistic compound correlations during the financial crisis. For mezzanine tranches, two completely different solutions are found: one solution close to zero, the second close to one. Given the compound correlations of the other tranches, it could be argued that using the solution close to one would be more realistic. However, this problem is not persistent in the case of base correlations. Hence, we focus on base correlations in our empirical study of the credit crisis.

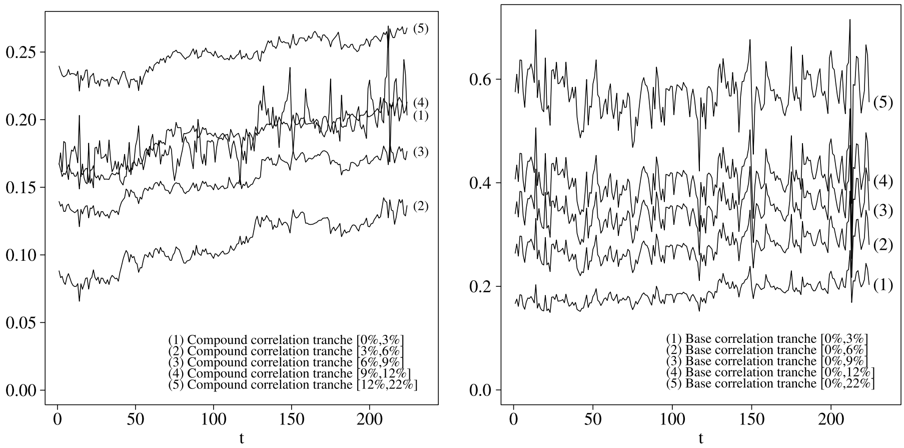

Fig. 1 Compound (*left*) and base (*right*) correlations for Series 5 in the considered period

Table 1 Pairwise correlation (in percent) of the different tranche spreads, compound, and base correlations

Series 5	Quoted spreads				Compound correlations				Base correlations			
	2	3	4	5	2	3	4	5	2	3	4	5
1	95.1	95.7	88.7	97.5	66.9	65.8	62.3	60.8	92.4	85.0	78.9	67.6
2		98.9	91.2	95.2		97.7	91.3	90.8		98.5	96.0	89.3
3			92.6	96.5			92.4	93.5			99.3	95.3
4				94.1				95.0				98.2

compound correlation is above that of the junior mezzanine tranche. O'Kane and Livesey (2004, p. 6) provide a possible explanation for this observation by attributing the smile to market prices being the result of multiple effects such as "… concerns about systemic versus idiosyncratic credit risk, fear of principal versus mark-to-market losses, liquidity effects, and supply and demand for certain tranches." Base correlations are increasing in the tranches' seniority. The standard deviations (in percent and increasing seniority) for the five time series are 2.0, 1.8, 1.5, 1.6, and 1.2 for the compound, and 2.0, 2.6, 3.1, 3.5, and 4.4 for the base correlations. Compound correlations have decreasing volatility in the tranches' seniority, while base correlations have increasing volatility in the tranches' width. Note that the time series of the different tranches show quite similar characteristics. The large pairwise correlations of these time series, especially of tranches close to each other, are reported in Table 1. As they overlap, it is not surprising that base correlations are more correlated with each other compared to compound correlations (which refer to disjoint tranches). Moreover, this explains the increasing correlation of base correlations the greater the tranches overlap.

3.2 The regressor variables

One aim of our investigation is to detect variables that can explain implied correlations and tranche spreads in the context of a regression model. Intuitively, implied correlations represent the market's view of the dependence among the firms in a credit portfolio. However, dependence is not directly observable and condensing it into a single number involves severe oversimplification. Moreover, we believe that implied correlations could be interpreted as some sort of risk premium for the risk of dependent defaults. Therefore, it seems reasonable that market participants demand a larger risk premium for correlation in times of large default probabilities, even if, theoretically, default probabilities and dependence are different types of risk. Hence, we include measures of dependence but also measures of default risk in the list of possible regressors. Inspired by Merton's firm-value model, we additionally include information from the equity market.

For each firm $i \in \{1, \dots, I\}$, $I = 125$, we first compute daily logarithmic returns (log returns) of CDS spreads. We then derive the sample variances $\sigma^2_{0,i}$ for the first 20 data points of these time series and compute the variances $\sigma^2_{t,i}$ for the remaining time points using the exponentially weighted moving average (EWMA) model with $\lambda \in (0, 1)$:

$$\sigma^2_{t,i} := \lambda\sigma^2_{t-1,i} + (1-\lambda)r^2_{t,i}, \quad i = 1, \dots, I, \tag{3}$$

(see, e.g., Alexander 1998, p. 130), where $r_{t,i}$ denotes the log return of the CDS spread of firm i at time t. Finally, for each time point t, we compute the mean value taken over all firms i to end up with the regressor "C^σ_t". Based on log returns of CDS spreads, we also derive all $I(I-1)/2$ pairwise correlations via

$$\begin{aligned} \sigma_{t,kl} &:= \lambda\sigma_{t-1,kl} + (1-\lambda)r_{t,k}r_{t,l}, \\ \rho_{t,kl} &:= \sigma_{t,kl}/\big(\sigma^2_{t,k}\sigma^2_{t,l}\big)^{1/2}, \quad k \neq l \end{aligned} \tag{4}$$

(see, e.g., Alexander 1998, p. 130), where the sample covariance $\sigma_{0,kl}$ between firms k and l is computed based on the first 20 log returns. After taking the mean over all firms for each time point as before, we obtain the regressor "C^ρ_t." Precisely the same procedure is applied to the available stock quotes: the corresponding regressors are denoted by "S^σ_t" and "S^ρ_t." Following the same procedure for portfolio CDS spreads (instead of individual CDS spreads) leads to the regressor "P^σ_t" (with the slight adjustment that no mean has to be taken). As a measure of (average) default risk for the overall credit market, we also include the mean of individual CDS spreads taken over all firms ("C_t"). Moreover, asset values are linked to default probabilities via the idea of structural-default models. To account for this observation, we include a stock index as regressor. This stock index is computed as the mean of the stock prices taken over all firms in the portfolio, where each stock is normed to unit weight by its initial value as of 2006-03-20 or 2008-04-17, respectively. This regressor is denoted by "S_t." Finally, we include the portfolio CDS spreads, "P_t," as a regressor. The intercept in all models is denoted "1."

3.3 The response variables and different regression models

The first set of response variables consists of market quotes (upfront or spread, respectively) for the five tranches of the CDO. The upfront payment of the equity tranche is quoted in percent; the spread of all other tranches is quoted in basis points. Since all quoted CDO tranche spreads s_t^j (and upfront, respectively) are positive, we use $y_t^{1,j} = f(s_t^j)$, with $f(x) = \log x$, as dependent variables. Therefore, our first set of regression models is given by

$$y_t^{1,j} = \big(1, S_t^\rho, S_t^\sigma, S_t, C_t^\rho, C_t^\sigma, C_t, P_t^\sigma, P_t\big)\big(\beta_0^{1,j}, \beta_1^{1,j}, \dots, \beta_8^{1,j}\big)^T + \varepsilon_t^{1,j},$$
$$j \in \{1, \dots, 5\}.$$

The second set of response variables consits of compound correlations, denoted by $\rho_t^{c,j}$, for each tranche $j \in \{1, \dots, 5\}$, with values in $[0, 1]$. We therefore consider $y_t^{2,j} = f(\rho_t^{c,j})$, with $f(x) = \log(x/(1-x))$, as response variables in a second regression model (using the same regressors as above).

The third set of response variables, $y_t^{3,j}$, is constructed via the same transformation f as for the second set, but with $\rho_t^{c,j}$ being replaced by the respective base correlation $\rho_t^{b,j}$.

It is convenient to assume i.i.d. zero-mean normally distributed residuals with constant variance. However, this assumption is typically violated if time series are used as response variables. We therefore additionally consider models with residuals following an autoregressive moving average (ARMA) process. Combined, this results in six regression models for each tranche. Moreover, we investigated a rolling-window procedure over the preceding 20 trading days for the computation of volatilities and correlations. This procedure can be considered an alternative to the EWMA approach. In this procedure, standard sample statistics are used for an estimation of volatilities and correlations based on the rolling window over the preceding 20 trading days.

4 Series 5: the estimated regression models

In this section, we present the fitted regression models for Series 5. To compute the EWMA processes, we investigated two smoothing parameters: $\lambda = 0.94$ (suggested by and used in *RiskMetrics*) and $\lambda = 0.33$ (pronouncing current observations). The results for the rolling-window procedure are very similar to the EWMA approach with $\lambda = 0.94$, and thus we do not present them here. The tables showing the results of the regressions contain abbreviations of the regressor variables corresponding to significant coefficients according to individual t-tests with significance level $\alpha = 5\%$. For models with i.i.d. normally distributed errors, coefficients of determination R^2 are reported as a goodness-of-fit measure. For models with ARMA residuals, the square of the sample correlation coefficients between observed and fitted values is used ("ρ^2"). Furthermore, the number of rejections of the Ljung–Box test (LB) applied to the model's residuals (according to the first 10 lags) is reported. This serves as an indicator for independence of the residuals. Other model checks that were conducted include visual tests using residual plots, QQ-plots, and plots of the autocorrelation and partial autocorrelation function of the model residuals.

The AR(1) models

According to the behavior of the partial autocorrelation and autocorrelation function computed for the empirical residuals of the models with normally distributed errors, we found AR(1) processes to fit the empirical residuals significantly better than i.i.d. normally distributed residuals. This observation holds for each tranche and choice of smoothing parameter λ, especially for regressions involving base correlations.

Due to violation of the model assumption of i.i.d. normally distributed residuals, the results for the regressions based on the smoothing parameter $\lambda = 0.94$, shown in Table 4, are presented only for AR(1) residuals. As to the different choices of λ, we find that when past observations (vs. present observations) are emphasized (which corresponds to using a large value of λ in (3) and (4)), the coefficients measuring stock and CDS correlation (S_t^ρ) and (C_t^ρ) are no longer significant, which could be due to a smoothing effect that covers large daily co-movements of stocks and CDS. A similar result holds for the rolling-window approach.

4.1 Explaining implied correlations

Table 2 sets forth the results of the regression models for the EWMA approaches with parameter $\lambda = 0.33$ for the computation of the regressors and the assumption of i.i.d. normally distributed residuals. The results for the corresponding models with residuals following ARMA processes are presented in Table 3. Table 4 presents the results for the EWMA approach with parameter $\lambda = 0.94$ and AR(1) residuals.

Explaining compound correlations

All the regression models for compound correlations have large explanatory power and many of the suggested coefficients are significant. Measures of default risk as well as measures of dependence significantly influence compound correlation. Also, different asset information and most information from the credit market are significant.

Table 2 Regression results with $\lambda = 0.33$ and i.i.d. normally distributed residuals

Series 5	Compound correlation regression			Base correlation regression		
Tranche	Significant	R^2	LB	Significant	R^2	LB
1	C^ρ, C, P^σ	61.6	10	C^ρ, C, P^σ	61.6	10
2	S, C^σ, C, P	87.8	10	C^ρ, C, P	42.9	10
3	$1, S^\rho, S^\sigma, S, C^\rho, C^\sigma, C, P$	88.3	10	S^ρ, C^ρ, C, P	34.2	10
4	$1, S^\sigma, C^\rho, C^\sigma, C, P$	81.6	10	S^ρ, C^ρ, C, P	29.7	10
5	$1, S^\rho, S^\sigma, C^\sigma, C, P$	83.6	10	S^ρ, C, P	23.7	8

Table 3 Regression results with $\lambda = 0.33$ and AR(1) distributed residuals

Series 5	Compound correlation regression			Base correlation regression		
Tranche	Significant	ρ^2	LB	Significant	ρ^2	LB
1	C^ρ, C, P^σ	81.6	0	C^ρ, C, P^σ	81.6	0
2	$1, C, P^\sigma, P$	98.9	7	C^ρ, C, P^σ	67.9	0
3	$1, C, P^\sigma, P$	98.4	0	C^ρ, C, P^σ, P	61.2	0
4	$1, S, C, P$	97.9	9	C^ρ, C, P^σ, P	57.5	0
5	$1, S, C, P$	97.3	7	S^ρ, C, P	51.7	0

Table 4 Regression results with $\lambda = 0.94$ and AR(1) distributed residuals

Series 5	Compound correlation regression			Base correlation regression		
Tranche	Significant	ρ^2	LB	Significant	ρ^2	LB
1	$1, S, C$	80.8	0	$1, S, C$	80.8	0
2	$1, C, P$	98.9	9	C, P	66.2	0
3	$1, C, P$	98.4	7	C, P	59.2	0
4	$1, C, P$	97.9	10	C, P	55.2	0
5	$1, C, P$	97.2	8	C, P	49.1	0

Explaining base correlations

The regression models explaining base correlations have less explanatory power but most fits are satisfactory. Recall that the volatility of base correlations is increasing in the tranches' width. This might explain why models explaining base correlations of higher tranches have less explanatory power. The measures of dependence (S_t^ρ) and (C_t^ρ) are more pronounced in these models (especially for models with $\lambda = 0.33$) compared to models explaining compound correlations. Moreover, other than asset return correlations (S_t^ρ), no information from the asset market is relevant in the base correlation regression models.

4.2 Explaining CDO tranche spreads

For practical applications, the most important criterion is the accuracy of the model in terms of the resulting pricing error. Moreover, high correlations between model-derived prices and market prices indicate model appropriateness. We therefore examine different methodologies for describing CDO tranche spreads with respect to these criteria. To this end, we first consider different regression models for directly describing spreads. The results of this approach are set forth in Tables 5 and 6.

Moreover, we use the regression models of the previous section to explain compound and base correlations. These model-derived correlations are then converted into CDO tranche spreads via (2). One advantage of using regression models for implied correlations (compared to a direct regression of tranche spreads) is that we can additionally compute spreads for tranches with non-standard attachment points (*bespoke CDOs*). This is not possible in a regression model for the spread of a specific

Table 5 Direct regressions on CDO tranche spreads with $\lambda = 0.33$

Tranche	Direct regression, i.i.d.			Direct regression, AR(1)		
	Significant	R^2	LB	Significant	ρ^2	LB
1	$1, S^\sigma, S, C, P$	96.6	10	S^σ, C, P	99.3	1
2	$1, S, C^\sigma, C$	95.3	10	C, P^σ, P	99.3	4
3	$1, S^\rho, S^\sigma, S, C^\rho, C^\sigma, C, P$	95.5	10	C, P^σ, P	99.1	3
4	$S^\sigma, C^\rho, C^\sigma, C, P$	87.7	10	C	98.3	9
5	$S^\sigma, S, C^\rho, C^\sigma, P$	95.0	10	$1, S^\sigma, C, P$	99.1	9

Table 6 Direct regressions on CDO tranche spreads with $\lambda = 0.94$

Tranche	Direct regression, i.i.d.			Direct regression, AR(1)		
	Significant	R^2	LB	Significant	ρ^2	LB
1	$1, S^\rho, S^\sigma, S, C^\rho, C^\sigma, C, P$	97.7	10	C, P	99.3	1
2	$1, S^\sigma, S, C^\sigma, C, P^\sigma, P$	96.3	10	C	99.3	5
3	$1, S^\rho, S^\sigma, C^\rho, C^\sigma, C, P$	96.4	10	C, P	99.0	7
4	$S^\rho, S^\sigma, C^\rho, C^\sigma, C, P$	91.6	10	C	98.2	10
5	$S^\rho, S^\sigma, S, C^\rho, C^\sigma, P$	96.8	10	$1, C, P$	99.1	6

tranche. Also, CDO tranche spreads depend not only on the dependence among the firms, but also on the future portfolio-loss distribution. Therefore, it is difficult to detect drivers of correlation using regressions on CDO tranche spreads.

An important observation that supports the latter methodology for describing CDO tranche spreads is that the different approaches yield comparable results (see Fig. 2 and Table 7). Even though the base-correlation approach yields slightly higher pricing errors, they are still far below bid–ask spreads. We thus conclude that all approaches accurately explain CDO tranche spreads.

5 Series 9: the estimated regression models (summary)

In this short discussion of Series 9, we present only the (re-)estimated models that were most appropriate for Series 5, i.e., models with AR(1) residuals and $\lambda = 0.33$. These models are also the most successful in describing Series 9. When computing implied correlations for Series 9, we find that the compound correlation of the mezzanine tranches is not well defined: one solution is close to zero, another is close to one. To circumvent this problem, we focus on base correlations.

The standard deviations for the time series of base correlations (in Series 9) are 4.5, 4.6, 4.6, 5.2, and 5.1 (in percent)—again, they are increasing in the tranches' width. However, they are much higher here than they were for Series 5, revealing a higher risk for changing default correlation. Considering pairwise correlations of quoted spreads, we find that for Series 9 the pairwise cross-correlation of spreads is much smaller than that found for Series 5, especially for the correlation of the equity

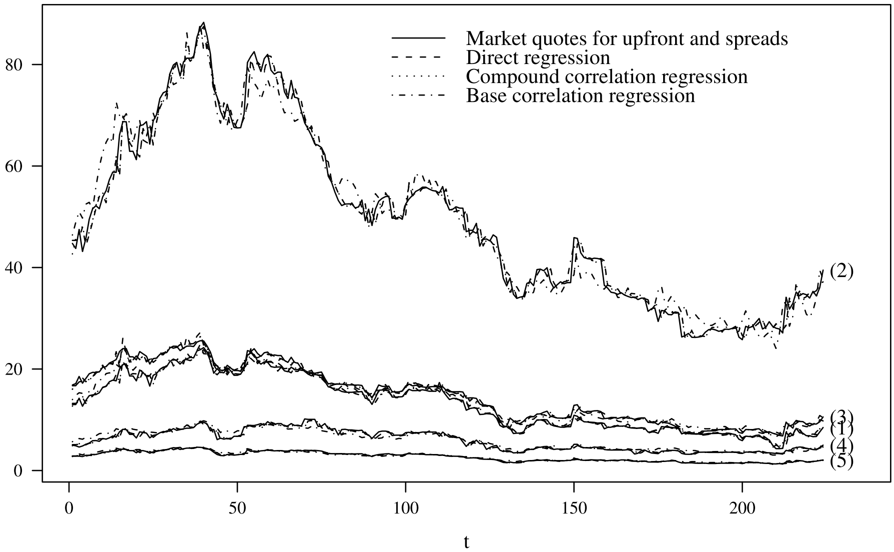

Fig. 2 Market and model-implied spreads for Series 5 for the EWMA approach with $\lambda = 0.33$ and AR(1) distributed residuals

Table 7 The fitting quality of the different methodologies in terms of an average pricing error and ρ^2 (in percent). All regression models use the EWMA approach with $\lambda = 0.33$ and AR(1) distributed residuals

Series 5	Direct regression		Compound cor. regression		Base cor. regression	
Tranche	∅ error	ρ^2	∅ error	ρ^2	∅ error	ρ^2
1	0.47 %	98.9	0.37 %	99.3	0.37 %	99.3
2	1.37 ‰	98.7	1.36 ‰	98.8	2.43 ‰	96.7
3	0.46 ‰	98.2	0.44 ‰	98.3	0.77 ‰	96.1
4	0.28 ‰	96.0	0.27 ‰	96.1	0.52 ‰	88.8
5	0.10 ‰	97.9	0.09 ‰	98.2	0.17 ‰	95.0

tranche to all other tranches. In contrast, base correlations of the different tranches are again highly correlated; their correlation is even higher than it was for Series 5, indicating a higher default correlation during the financial crisis, implying that losses in the different tranches are more likely to occur simultaneously. Also, compared to Series 5, the regression models for the different tranche spreads (and the models for the base correlations, respectively) now have the same significant coefficients (except for the intercept), indicating a higher systemic risk during the financial crisis. This is a convenient situation from a modeler's perspective, since models with the same regressor variables can be used to explain all tranches. In a direct regression on tranche spreads, the intercept, the stock index (S), and the average CDS level (C) are each significant. When base correlations are modeled, the coefficient of the portfolio CDS (P) is also significant. Moreover, the fact that the equity index is now included in all models might be seen as an indication of market convergence during financial

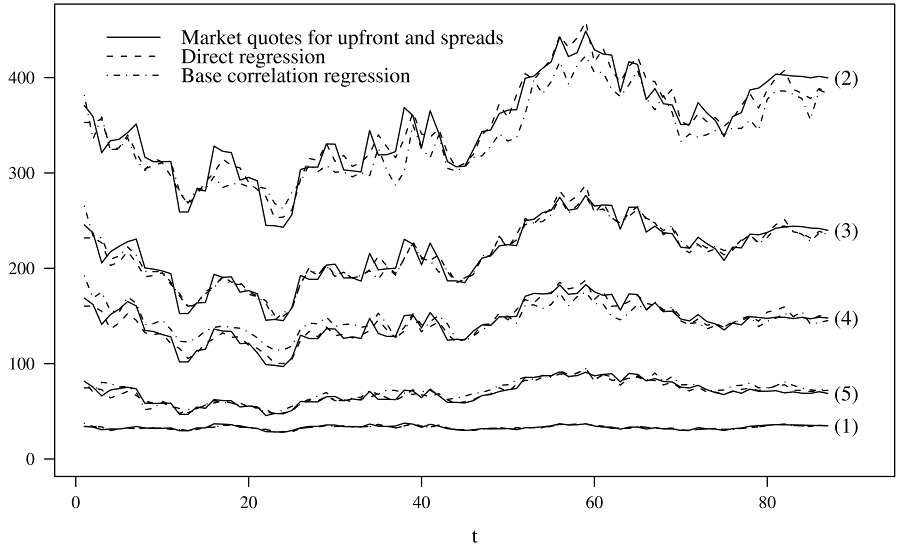

Fig. 3 Market and model-implied spreads for Series 9 for the EWMA approach with $\lambda = 0.33$ and AR(1) distributed residuals. Note that the level of spreads is up to a factor of 10 higher than for Series 5

Table 8 Pairwise correlation (in percent) of the different tranche spreads and base correlations

Series 9	Quoted spreads				Base correlations			
	2	3	4	5	2	3	4	5
1	53.0	48.4	46.4	36.9	97.4	95.1	93.8	90.3
2		98.4	92.7	91.8		99.2	98.2	95.6
3			96.8	95.5			99.6	97.9
4				97.3				99.1

distress, meaning that the credit market is now highly sensitive to signals from the equity market. Note that all significant factors are traded assets, which is important when the model is used to hedge some tranche of the CDO. We also observe that the LB test does not reject the models. Figure 3 and Table 10 show that the models again produce a very small pricing error. The overall spread level (as well as bid–ask spreads) is up to a factor of 10 higher for Series 9 compared to Series 5.

6 Possible applications of the model

6.1 Pricing fictitious CDOs

Valuation of a synthetic CDO, built from a CDS portfolio, is not easy, especially when there is insufficient market information to correctly specify the parameters of the pricing formula. While CDS spreads contain information about the firms' default

Table 9 Regression results with $\lambda = 0.33$ and AR(1) distributed residuals

Series 9	Direct regression			Base correlation regression		
Tranche	Significant	ρ^2	LB	Significant	ρ^2	LB
1	$1, S, C$	85.2	0	S, C, P	94.6	1
2	$1, S, C$	96.2	0	S, C, P	92.3	0
3	$1, S, C$	95.8	0	$1, S, C, P$	91.7	0
4	$1, S, C$	94.6	0	$1, S, C, P$	91.3	0
5	$1, S, C$	95.4	0	$1, S, C, P$	90.6	0

Table 10 The fitting quality of the different methodologies in terms of an average pricing error and ρ^2 (in percent). Both regression models use the EWMA approach with $\lambda = 0.33$ and AR(1) distributed residuals

Series 9	Direct regression		Base correlation regression	
Tranche	∅ error	ρ^2	∅ error	ρ^2
1	0.88 %	72.1	0.95 %	70.9
2	10.61 ‰	92.8	18.25 ‰	86.4
3	7.22 ‰	91.9	6.76 ‰	92.6
4	5.16 ‰	89.1	8.83 ‰	72.5
5	2.73 ‰	91.0	3.74 ‰	87.2

probabilities, the CDO pricing formula also requires a correlation number as input. The presented regression models can be used to asses (or predict) the correlation from observable quantities such as asset values and CDS spreads. As an example, let us consider the pricing of fictitious tranches of the Markit iTraxx Asia portfolio.

The Markit iTraxx Asia is a portfolio containing CDS on 50 Asian companies. However, tranches of this portfolio are not traded. While it is possible to infer the average default probability from the traded index or from individual CDS, it is difficult to find a reasonable value for the correlation parameter. The most simple approach to setting the correlations for the Asian portfolio is to use the values observed in the European market. However, by comparing the average pairwise correlations of all CDS and stock returns of the firms in the Asian portfolio, we observe this number to be significantly smaller than the European analogue, casting some doubt on the appropriateness of this approach. Alternatively, one could use the regression models for implied correlations, calibrated to the European CDO market, to compute implied correlations for Asia (given information from the Asian market). The results of both approaches are presented in Fig. 4. Using regression models for compound correlations yields spreads that consider smaller empirically observed CDS and stock correlations compared to the European market.

Please note that the results of the outlined approach are based on the assumption that the regression models can be transferred from one market to another. However, this may not be true. Also, it cannot be guaranteed that prices for different tranches obtained from this approach do not allow for arbitrage.

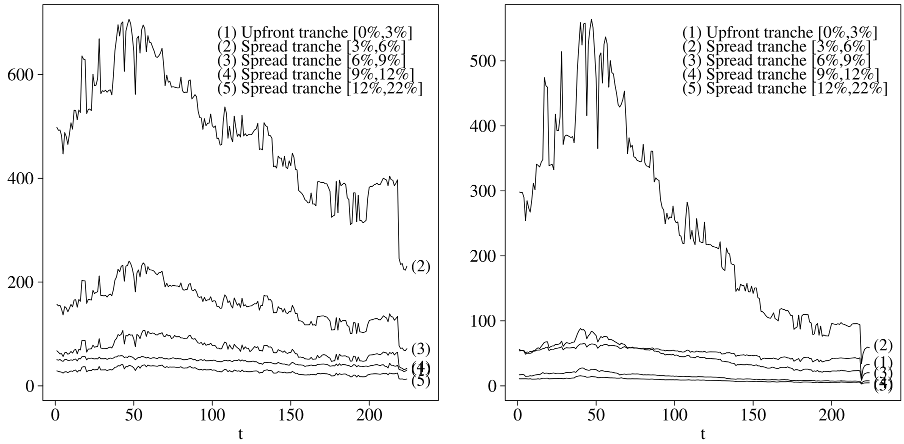

Fig. 4 Fictitious tranche spreads for the Markit iTraxx Asia portfolio for the EWMA approach with $\lambda = 0.33$ and AR(1) distributed residuals. The *figure to the left* is computed with implied correlations for the Markit iTraxx Europe on the same days; the *right-hand side* is based on predicted implied correlations for the Asian market

6.2 Scenario generation for risk management

A dynamic version of Vasicek's model is a powerful tool for risk management and portfolio optimization, but it requires an equally dynamic description of the correlation structure among the firms in the portfolio under consideration. We demonstrated above that appropriate regression models can quite accurately explain the evolution of observed CDO correlations. The fact that all regression models used in this investigation rely on observable factors (obtained from equity and CDS quotes) is especially convenient in simulation studies / economic scenario generators, where these factors are modeled as primary objects. Given the paths of simulated assets and CDS spreads, one can simulate a path of implied correlations or CDO tranche spreads from one of our suggested models, whichever is of interest. Then, prices for correlation-sensitive products can be computed.

7 Conclusion

The aim of this paper was to discover determinants of correlation-sensitive credit products and a description of implied correlations. For this purpose, we developed different regression models, which turned out to precisely explain CDO tranche spreads as well as implied correlations. We found that information from the credit market has a larger influence on implied correlations compared to information from the asset market, a result that casts some doubt on the sole use of asset information as a proxy variable for calibration of a portfolio default model. More precisely, we found that for Series 5, empirical correlations of CDS returns, their volatilities, and the absolute level of CDS spreads significantly influence implied correlations. Moreover, the absolute level and volatility of portfolio CDS spreads also significantly determine implied correlations. Considering asset information, only empirical correlations and a stock index were significant in some of the investigated regression models. For the direct description of CDO tranche spreads, several other factors were statistically significant, which is not surprising considering that tranche spreads depend on default probabilities and the dependence structure. The significant factors in the direct description of tranche spreads come from both the asset and the credit market. For the description of Series 9, important factors are a stock index, the average CDS level, and the portfolio CDS spreads. This shows that systemic factors have a special influence on the financial market during financial distress. In addition to the formulation and estimation of different regression models, we discussed their possible application for risk management and the valuation of CDO tranches in other markets.

This study is a first attempt at explaining the time evolution of the market's perception of the dependence structure of a credit portfolio and many interesting questions remain open for investigation. For instance, we focused on traded assets (and quantities derived thereof) as regressor variables. The current financial crisis was accompanied by a massive loss in liquidity. Considering this, it seems reasonable to include some sort of risk premium for liquidity in an extended regression model. Also, it would be interesting to have sufficient data to extend the investigation over a longer time horizon.

Acknowledgements We thank an anonymous referee for helpful comments and suggestions. Any remaining errors are our own.

Appendix

Proof of Theorem 2.1 First, we compute $I(a,b) := \int_a^b F_{L_t^0}(x)\,dx$ for $0 \le a < b \le 1$. This is

$$\int_a^b \int_{-\infty}^{(\sqrt{1-\rho}\Phi^{-1}(x)-\Phi^{-1}(p_t))/\sqrt{\rho}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy\,dx.$$

By substituting $\tilde{x} := -\Phi^{-1}(x)$, we obtain

$$I(a,b) = \int_{-\Phi^{-1}(b)}^{-\Phi^{-1}(a)} \int_{-\infty}^{-(\sqrt{1-\rho}\tilde{x}+\Phi^{-1}(p_t))/\sqrt{\rho}} \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left(\tilde{x}^2+y^2\right)\right) dy\,d\tilde{x}.$$

Substituting $\tilde{y} := -(y\sqrt{\rho}+\sqrt{1-\rho}\tilde{x})$ leads to

$$I(a,b) = \int_{-\Phi^{-1}(b)}^{-\Phi^{-1}(a)} \int_{\Phi^{-1}(p_t)}^{\infty} \frac{1}{2\pi\sqrt{\rho}} \exp\left(-\frac{1}{2}\frac{\tilde{x}^2+2\tilde{x}\tilde{y}\sqrt{1-\rho}+\tilde{y}^2}{\rho}\right) d\tilde{y}\,d\tilde{x}.$$

Let X and Y follow a bivariate standard normal distribution with correlation $-\sqrt{1-\rho}$. Then,

$$\begin{aligned}
I(a,b) &= \mathbb{P}\big(-\Phi^{-1}(b) < X \le -\Phi^{-1}(a), Y > \Phi^{-1}(p_t)\big) \\
&= \mathbb{P}\big(-\Phi^{-1}(b) < X \le -\Phi^{-1}(a)\big) \\
&\quad - \mathbb{P}\big(-\Phi^{-1}(b) < X \le -\Phi^{-1}(a), Y \le \Phi^{-1}(p_t)\big) \\
&= b - a + \boldsymbol{\Phi}\big(-\Phi^{-1}(b), \Phi^{-1}(p_t), -\sqrt{1-\rho}\big) \\
&\quad - \boldsymbol{\Phi}\big(-\Phi^{-1}(a), \Phi^{-1}(p_t), -\sqrt{1-\rho}\big).
\end{aligned}$$

Now, for all $j \in \{1, \ldots, J\}$,

$$\begin{aligned}
\mathbb{E}[L_{t,j}] &= \int_0^1 \big(\min\big\{(1-R)x, u_j\big\} - \min\big\{(1-R)x, l_j\big\}\big)\, dF_{L_t^0}(x) \\
&= J(u_j) - J(l_j),
\end{aligned} \tag{5}$$

where $J(z) := \int_0^1 \min\{(1-R)x, z\}\, dF_{L_t^0}(x)$, $z \in [0,1]$. If $z < 1-R$, splitting up this integral at $z/(1-R)$ and using the integration by parts formula leads to $J(z) = z - (1-R)I(0, z/(1-R))$. For $z \ge 1-R$, the same formula leads to $J(z) = (1-R) - (1-R)I(0,1)$. Therefore, both cases are comprised in the formula

$$\begin{aligned}
J(z) &= \min\{z, 1-R\} - (1-R)I\big(0, \min\big\{z/(1-R), 1\big\}\big) \\
&= (1-R)\big(\min\big\{z/(1-R), 1\big\} - I\big(0, \min\big\{z/(1-R), 1\big\}\big)\big).
\end{aligned}$$

Thus,

$$J(z) = (1-R)\big(p_t - \boldsymbol{\Phi}\big(-\Phi^{-1}\big(\min\big\{z/(1-R), 1\big\}\big), \Phi^{-1}(p_t), -\sqrt{1-\rho}\big)\big).$$

Combining this result with (5) directly leads to the result as stated. □

References

Albrecher, H., Ladoucette, S., Schoutens, W.: A generic one-factor Lévy model for pricing synthetic CDOs. In: Fu, M.C., Jarrow, R.A., Yen, J.-Y.J., Elliott, R.J. (eds.) Advances in Mathematical Finance, pp. 259–277. Birkhäuser, Basel (2007)

Alexander, C.: Risk Management and Analysis, vol. 1. Wiley, New York (1998)

Andersen, L., Sidenius, J.: Extensions to the Gaussian copula: Random recovery and random factor loadings. J. Credit Risk **1**(1), 29–70 (2005)

Andersen, L., Sidenius, J., Basu, S.: All your hedges in one basket. RISK 67–72 (November 2003)

Brommundt, B., Felsenheimer, J., Gisdakis, P., Zaiser, M.: Recent developments in credit markets. Financ. Mark. Portf. Manag. **20**, 221–234 (2006)

Burtschell, X., Gregory, J., Laurent, J.-P.: Beyond the Gaussian copula: Stochastic and local correlation. J. Credit Risk **3**(1), 31–62 (2007)

Burtschell, X., Gregory, J., Laurent, J.-P.: A comparative analysis of CDO pricing models. In: Meissner, G. (ed.) The Definitive Guide to CDOs, pp. 389–427. Risk Books, London (2009)

Düllmann, K., Sosinska, A.: Credit default swap prices as risk indicators of listed German banks. Financ. Mark. Portf. Manag. **21**, 269–292 (2007)

Gregory, J., Laurent, J.-P.: I will survive. RISK 103–107 (June 2003)
Gregory, J., Laurent, J.-P.: In the core of correlation. RISK 87–91 (October 2004)
Hull, J., White, A.: Valuation of a CDO and an n-th to default CDS without Monte Carlo simulation. J. Deriv. **12**(2), 8–23 (2004)
Hull, J.C.: Options, Futures, and Other Derivatives, 7th edn. Prentice Hall Series in Finance. Prentice Hall, New York (2008)
Kalemanova, A., Schmid, B., Werner, R.: The normal inverse Gaussian distribution for synthetic CDO pricing. J. Deriv. **14**(3), 80–93 (2007)
Li, D.X.: On default correlation: A copula function approach. J. Fixed Income **9**(4), 43–54 (2000)
Merton, R.C.: On the pricing of corporate debt: The risk structure of interest rates. J. Finance **29**(2), 449–470 (1974)
O'Kane, D., Livesey, M.: Base correlation explained. Quant. Credit Res. Q. **3** (2004)
Rösch, D.: Correlations and business cycles of credit risk: Evidence from bankruptcies in Germany. Financ. Mark. Portf. Manag. **17**(3), 309–331 (2003)
Schönbucher, P.J.: Credit Derivatives Pricing Models. Wiley, New York (2003)
Vasicek, O.: Probability of Loss on Loan Portfolio. KMV Corporation, San Francisco (1987)

Marius Hofert is a postdoc in the Department of Mathematics at ETH Zurich. He completed a Master of Science in Mathematics from Syracuse University and a Diploma in Economathematics, as well as a Doctor of Philosophy in Mathematics, from Ulm University.

Matthias Scherer is Professor (deputy) for Financial Mathematics at the Technische Universität München and coordinates the elite graduate programme "Finance and Information Management". His research focus lies on multivariate models and dependence concepts. He wrote his dissertation (Universität Ulm) on multivariate structural default models and holds a diploma in Business Mathematics (Universität Ulm) as well as a Master's degree in Mathematics (Syracuse University).

Rudi Zagst
Rudi Zagst is Professor of Mathematical Finance, Director of the Center of Mathematics and Head of the Institute for Mathematical Finance at Technische Universität München (TUM), Germany. He is also President of risklab GmbH, a German-based consulting company offering advanced asset management solutions. He is a consultant and a professional trainer to a number of leading institutions. His current research interests are in financial engineering, credit risk modeling and quantitative asset management.