

# Implied Correlation Smile

Master Thesis submitted to

**Prof. Dr. Wolfgang Härdle**

Institute for Statistics and Econometrics

CASE - Center for Applied Statistics and Economics

**Humboldt-Universität zu Berlin**



by

**Maria Feld**

(503301)

in partial fulfillment of the requirements  
for the degree of **Master of Sciences in Statistics**

July 30, 2007

## DECLARATION OF AUTHORSHIP

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

Berlin, July 30, 2007,

Maria Feld

## CONTENTS

1. <i>Introduction</i> . . . . .	8
2. <i>Modeling Correlated Defaults</i> . . . . .	10
2.1 Basic Concepts . . . . .	10
2.2 The Bernoulli Model . . . . .	13
2.3 The Poisson Model . . . . .	14
2.4 The Industrial Models . . . . .	16
2.5 One Factor Models . . . . .	18
2.6 Estimation of asset correlation . . . . .	22
3. <i>Collateralized Debt Obligations</i> . . . . .	25
3.1 Typical CDO Structure . . . . .	25
3.1.1 Cash flow structure of CDO . . . . .	26
3.1.2 Motivation for CDO transaction . . . . .	28
3.2 Valuation of CDO . . . . .	30
4. <i>Implied Correlation Smile</i> . . . . .	33
4.1 Correlation and Tranche Loss . . . . .	33
4.2 Compound Correlation . . . . .	36
4.2.1 The Data . . . . .	36
4.2.2 LPHGC Model . . . . .	39
4.2.3 Possible Explanation of the Correlation Smile . . . . .	42

4.2.4 Existence, uniqueness, and monotonicity of compound correlation . . . . .	43
5. <i>Base Correlation</i> . . . . .	47
5.1 Advantages of Base Correlation . . . . .	48
5.2 Pitfalls of Base Correlation . . . . .	51
6. <i>Conclusion</i> . . . . .	53
7. <i>Appendix A</i> . . . . .	54
7.1 How to extract the term structure of default probabilities? . .	54
8. <i>Appendix B</i> . . . . .	57
8.1 Proof of the Proposition 2.2 . . . . .	57

## LIST OF FIGURES

2.1	Portfolio Loss Distribution and Risk Measures. . . . .	12
2.2	The default probability $p(y)$ as a function of the state of the economy. . . . .	18
2.3	Mean default rate and Default rate volatility. The red and blue lines represent the historic default and the regression by exponential function fitting correspondingly. . . . .	24
3.1	CDO Scheme . . . . .	27
4.1	Portfolio Loss Distribution for different correlation parameters and fixed probability of default 10%. . . . .	34
4.2	Expected Tranche Loss as a function of default correlation. . . . .	35
4.3	Fair spread as a function of default correlation. . . . .	37
4.4	iTraxx family. Source: IIC presentation "iTraxx CDS Indices". . . . .	38
4.5	Implied Correlation Smile. . . . .	41
4.6	Non-existence of implied correlation for Mezzanine tranche;iTraxx 7S 7Y, 3 Jul 07. . . . .	44
4.7	$function(corr, CDS_{spread})$ - difference between market and model spread for mezzanine tranche, iTraxx 7S 7Y, 3 Jul 07. . . . .	45
4.8	Non-uniqueness of implied correlation for mezzanine tranche, iTraxx 5S 10Y, 1 Jun 06. . . . .	46
5.1	Compound vs Base Correlation. . . . .	49
5.2	Expected Fictive Tranche Loss as a function of default correlation. . . . .	50
5.3	Bootstrapped Base Correlations vs "True" Default Intensities. . . . .	51

## LIST OF TABLES

2.1	Moody's Corporate Bond Historic Default Frequency 1970-2004	23
2.2	Calibration Results due to exponential function fitting . . . .	24
3.1	Typical CDO tranches . . . . .	26
3.2	Unexpected Loss (% of Portfolio Notional) for homogenous portfolio ( $L \sim F_{p,\omega}$ ) w.r.t. $p$ and $\omega$ . . . . .	29
3.3	VAR (99.5 quantile) (% of Portfolio Notional) for homogenous portfolio ( $L \sim F_{p,\omega}$ ) w.r.t. $p$ and $\omega$ . . . . .	29
4.1	Example of the dataset. Source: Bloomberg. . . . .	39

## 1. INTRODUCTION

The market for credit derivatives appeared in the mid 1990s. Credit derivatives are instruments aimed at protecting debt securities investors against adverse movements in the credit quality of the borrower. Initially invented for hedging, in the beginning these instruments were privately negotiated financial contracts, purely over-the-counter traded. The development of statistic techniques for pricing credit derivatives led to the standardizing of these products, which are now liquidly traded and exhibit significant growth.

One of the most popular credit derivative products are CDO — collateral debt obligations. CDOs are asset backed securities, which have typically a loan/debt assets portfolio as a collateral securitizing a portfolio of credit-linked notes. The cash flow generated by the collateral is structured in order to meet investor's risk preferences. CDOs are used for credit risk transfer, capital relief as well as for arbitrage.

A relatively recent innovation in the credit derivatives market is the introduction of standardized CDS (credit default swaps) indices such as iTraxx. The standard tranches on these reference indices are also actively quoted. The new liquidity of the market instigated the quotation of tranching products in terms of implied correlation parameter rather than in terms of the tranche spread. This practice was inspired by the use of implied volatilities in options markets. The implied compound correlation of a tranche is the uniform asset correlation that makes the tranche spread computed by the standard market model equal to its observed market spread. This correlation is used to price off-market tranches with the same underlying or for relative value considerations (when comparing alternative investments in CDO tranches).

Applying the implied correlation concept has nevertheless a few drawbacks that are discussed in this paper. First, the tranche spreads are not necessarily monotone in correlation, and we may observe market prices that are not attainable by a choice of correlation. Moreover, implied correlations suffer from both existence and uniqueness problems. Finally, a so called correlation smile is observed when using the standard market model for pricing CDO

---

tranches. Quotations available in the market indicate that different tranches on the same underlying portfolio trade at different implied correlations. If the standard market model described market prices correctly the implied default correlation would trivially be constant over tranches. Using the data on CDS indices and tranching products we will demonstrate the pricing methodology for CDO instrument and analyze the properties of implied correlation smile.

The paper is organized as follows: the second chapter presents the theoretical background on credit risk modeling. The third chapter introduces collateral debt obligations and deals with its pricing techniques. Further we consider the concepts of compound and base correlation. The fourth section includes direct modeling of tranche spreads and finding the implied correlations observed in the market. The fifth chapter focuses on the concept of base correlations. The last chapter concludes.

## 2. MODELING CORRELATED DEFAULTS

In this chapter we present a background in the theory of credit risk modeling and ideas broadly used in the pricing of credit derivatives. Since in this work the empirical analysis bases on the use of standard credit risk models and focuses on the notion of correlation, we will proceed with covering correlation modeling approaches. These concepts have recently become popular and their comprehensive presentation could be found in many books, e. g. [?]. For more references see [11], [6].

### 2.1 Basic Concepts

We will start with defining the basic notions of credit risk theory.

Definition 2.1.1 (Loss function):

The loss fraction in case of default is called loss given default (LGD). The exposure at default in a considered time period is abbreviated as EAD. Then the loss of an obligor is defined by the following loss function:

$$\tilde{L} = EAD \times LGD \times L$$

with  $L = I_D$ . Here  $D$  stands for the default event of an obligor in a given time period, (e. g. one year).  $P(D)$  is the probability of the event  $D$ .

Definition 2.1.2 (Expected Loss):

The expected loss (EL) is defined as:

$$EL = E(\tilde{L}) = EAD \times LGD \times P(D).$$

For simplicity reasons we assume here and thereafter EAD and LGD to be deterministic therefore implying their independence from the default event. However violating these assumptions leads to a more specified and realistic model.

To calculate  $EL$  we need to find default probabilities, which could be inferred either from credit ratings or from market prices of defaultable bonds or credit derivatives. According to the first approach we use data on default frequencies for different rating classes to perform a mapping from the ratings' space into the default probabilities' space. Second approach has recently become more popular. Models for implying default probabilities from spreads of the credit default swaps are incorporated in most data systems. A detailed presentation of how this could be done is given in the Appendix A. For details on bootstrapping default probabilities please to [18], [19], [3].

The expected loss  $EL$  defines the necessary loss reserve that a bank must hold as an insurance against the default. In addition to the expected loss the bank should have a cushion to cover unexpected losses.

Definition 2.1.3 (Unexpected Loss):

The unexpected loss (UL) is defined as:

$$UL = \sqrt{\text{Var}(\tilde{L})} = \sqrt{\text{Var}(EAD \times LGD \times L)}$$

with  $\text{Var}(L) = P(D)(1 - P(D))$ .

So far we have considered the loss estimates for a single obligor. Now assume we have a credit portfolio consisting of  $m$  loans.

Definition 2.1.4 (Portfolio Loss):

The expected portfolio loss is defined by the following random variable:

$$\tilde{L}_{PF} = \sum_{i=1}^m \tilde{L}_i = \sum_{i=1}^m EAD_i \times LGD_i \times L_i$$

with  $L_i = I_{D_i}$ .

Analogously to the single obligor case we can calculate  $EL_{PF}$  and  $UL_{PF}$ :

$$EL_{PF} = \sum_{i=1}^m EL_i = \sum_{i=1}^m EAD_i \times LGD_i \times P(D_i)$$

$$UL_{PF} = \sqrt{\sum_{i,j=1}^m EAD_i \times EAD_j \times LGD_i \times LGD_j \times \text{Cov}(L_i, L_j)}.$$

It is possible to rewrite the covariance term as following:

$$\text{Cov}(L_i, L_j) = \sqrt{\text{Var}(L_i) \times \text{Var}(L_j) \times \rho_{ij}}.$$

Now we obviously face the problem of the unknown default correlations  $\rho_{ij}$ . One could assume that loss variables are uncorrelated but this severely contradicts our empirical observations: defaults are likely to happen jointly so that the correlation between obligors becomes the main driver of credit risk and the key issue in credit modeling.

The discussed above risk characteristics such as Expected and Unexpected Loss and also well-known risk measure VAR could be easily calculated given the distribution of the portfolio loss variable  $\tilde{L}_{PF}$  (see Figure (2.1)). Later we will show that finding portfolio loss distribution is essential to pricing credit derivatives.

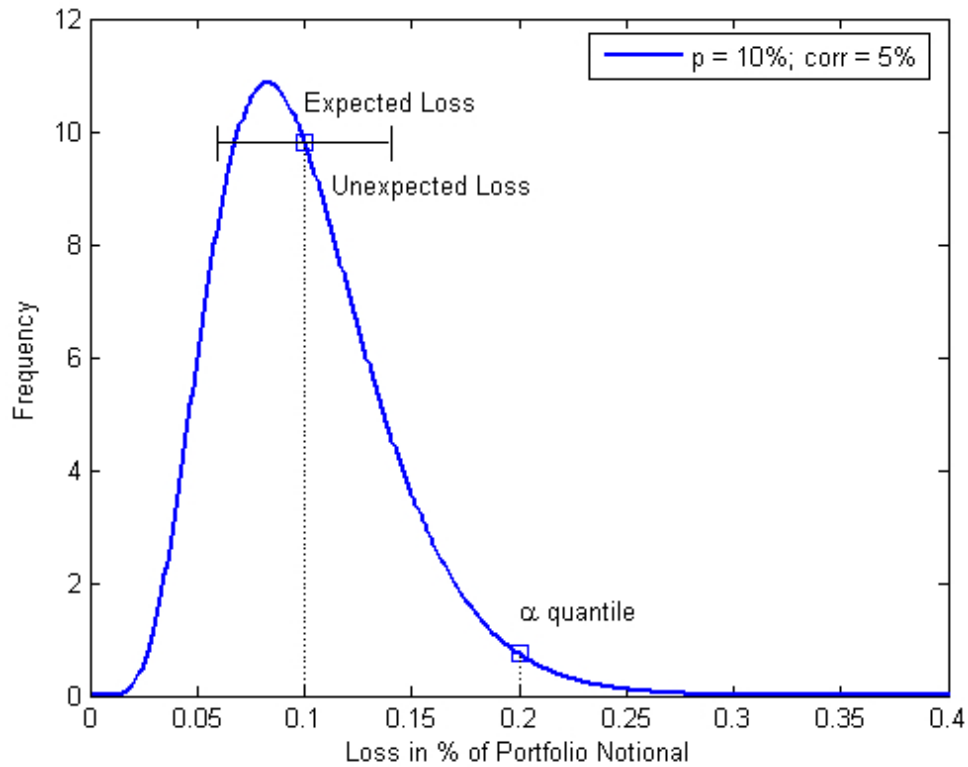


Fig. 2.1: Portfolio Loss Distribution and Risk Measures.

There are two methods to generate a loss distribution. The first solution is applying Monte Carlo simulation, the second is based on some analytical

approximation. In the Monte Carlo framework we simulate portfolio losses assuming some driving distribution of the single loss variables and correlation between them. Analytic approximation also requires correlation as an input. Further we will introduce the models which incorporate the statistical techniques for calibrating default correlations.

## 2.2 The Bernoulli Model

In the preceding section we have implicitly introduced the Bernoulli loss variable defined as  $L_i \sim B(1; p_i)$ , with  $L_i$  being the default variable of obligor  $i$ , i. e. loss is generated with probability  $p_i$  and not generated with probability  $(1 - p_i)$ . The default correlation was therefore defined as correlation between random variables, which follow Bernoulli distribution. The fundamental idea in the modeling of joint defaults is the randomization of the involved default probabilities. While in our previous analysis we considered extracted from market data or ratings default probabilities, now we assume that the loss probabilities are also random variables that follow some distribution  $F$  within  $[0, 1]^m$ :  $P = (P_1, \dots, P_m) \sim F$ .

We assume that Bernoulli loss variables  $L_1, \dots, L_m$  are independent conditional on a realization  $p = (p_1, \dots, p_m)$  of vector  $P$ . The joint distribution of the loss function is then:

$$\begin{aligned} P(L_1 = l_1, \dots, L_m = l_m) = \\ \int_{[0,1]^m} \prod_{i=1}^m p_i^{l_i} (1 - p_i)^{1-l_i} dF(p_1, \dots, p_m), \end{aligned} \quad (2.1)$$

where  $l_i \in \{0, 1\}$ . The first and second moments of the single losses  $L_i$  are:

$$E(L_i) = E(P_i), \quad \text{Var}(L_i) = E(P_i)\{1 - E(P_i)\}$$

The covariance of single losses is given by:

$$\text{Cov}(L_i, L_j) = E(L_i, L_j) - E(L_i)E(L_j) = \text{Cov}(P_i, P_j) \quad (2.2)$$

The correlation for two counterparties' default is:

$$\text{Corr}(L_i, L_j) = \frac{\text{Cov}(P_i, P_j)}{\sqrt{E(P_i)\{1 - E(P_i)\}}\sqrt{E(P_j)\{1 - E(P_j)\}}}. \quad (2.3)$$

Thus we succeeded in expressing the unknown default correlations in terms of covariances of the  $F$  distribution. Later in this chapter it will be shown how to obtain an appropriate specification for the distribution of default probabilities and consequently solve for the default correlations.

A major simplification is possible if one assumes an equal default probability  $P_i$  for all obligors. It is suitable for the uniform portfolios with loans of comparable size and with similar risk characteristics. In this case (2.1) simplifies to

$$P(L_1 = l_1, \dots, L_m = l_m) = \int_0^1 p^k (1-p)^{m-k} dF(p) \quad (2.4)$$

where  $k = \sum_{i=1}^m l_i$  is the number of defaults in the credit portfolio. Note that  $EL$  equals:

$$\bar{p} = \int_0^1 p dF(p) \quad (2.5)$$

Therefore the default correlation between two different counterparties equals:

$$\begin{aligned} \rho_{ij} &= \text{Corr}(L_i, L_j) = \\ &= \frac{P(L_i = 1, L_j = 1) - \bar{p}^2}{\bar{p}(1 - \bar{p})} = \frac{\int_0^1 p^2 dF(p) - \bar{p}^2}{\bar{p}(1 - \bar{p})}. \end{aligned} \quad (2.6)$$

Formula (2.6) shows that the higher volatility of  $P$  corresponds to the higher default correlation. Since the numerator of (2.6) equals  $\text{Var}(P) \geq 0$  the default correlation in the Bernoulli model is always positive and can not mimic negative default correlation.

### 2.3 The Poisson Model

Another widely-spread approach to joint default modeling is the assumption of Poisson-distributed loss variable  $L_i$  with intensity  $\Lambda_i$ . This means that  $L_i \sim \text{Pois}(\lambda_i)$ ,  $p_i = P(L_i \geq 1)$ ,  $L_i \in \{0, 1, 2, \dots\}$  modeling the fact that multiple defaults of one obligor  $i$  may occur. Analogously to the Bernoulli mixture model we assume not only the loss variable vector  $L$  but also the intensity vector  $\Lambda = (\Lambda_1, \dots, \Lambda_m)$  to be random:  $\Lambda \sim F$  within  $[0, \infty)^m$ . Also assume that  $L_1, \dots, L_m$  (conditional on a realization of  $\Lambda$ ) are independent.

The joint distribution of  $L_i$  is given:

$$\begin{aligned} & P(L_i = l_i, \dots, L_i = l_i) \\ &= \int_{[0, \infty)^m} e^{-(\lambda_1 + \dots + \lambda_m)} \prod_{i=1}^m \frac{\lambda_i^{l_i}}{l_i!} dF(\lambda_1, \dots, \lambda_m), \end{aligned} \quad (2.7)$$

Similar to the Bernoulli case, we have for  $i = 1, \dots, m$ :

$$E(L_i) = E(\Lambda_i)$$

$$\text{Var}(L_i) = \text{Var}\{E(L_i|\Lambda)\} + E\{\text{Var}(L_i|\Lambda)\} = \text{Var}(\Lambda_i) + E(\Lambda_i). \quad (2.8)$$

The correlation is given then:

$$\text{Corr}(L_i, L_j) = \frac{\text{Cov}(\Lambda_i, \Lambda_j)}{\sqrt{\text{Var}(\Lambda_i) + E(\Lambda_i)}\sqrt{\text{Var}(\Lambda_j) + E(\Lambda_j)}}. \quad (2.9)$$

Like in the Bernoulli Model we can express the default correlation through the covariances of the intensity vector distribution  $F$ . For the uniform portfolios we could assume a single distribution for all obligors. The analogue of (2.2) is then:

$$\text{Corr}(L_i, L_j) = \frac{\text{Var}(\Lambda)}{\text{Var}(\Lambda) + E(\Lambda)}. \quad (2.10)$$

This formula is especially intuitive if we look at it from a dispersion point of view. The dispersion of a distribution is its variance to mean ratio. The dispersion of a Poisson distribution is equal to 1. Using dispersion, we get the following formula:

$$\text{Corr}(L_i, L_j) = \frac{D(\Lambda)}{D(\Lambda) + 1}. \quad (2.11)$$

We therefore conclude: an increase in dispersion will increase the mixture effect, which strengthens the dependence between obligor's defaults.

### **Bernoulli vs. Poisson**

Comparing Bernoulli with Poisson distribution of the default risk, we see that there always exists a higher default correlation in Bernoulli distribution than in Poisson distribution. In other words even in case the mean of Bernoulli matches with the Poisson distribution, the Poisson variance will always exceed the variance of Bernoulli. The higher default correlations result in fatter tails of the corresponding loss distributions.

## 2.4 The Industrial Models

In the empirical part of this work we will apply so called Large Pool Homogenous Gaussian Copula Model to market data. This model is based in its turn on the implications of widely used industrial models, which are briefly presented below.

Two well-known factor models *CreditMetrics<sup>TM</sup>* and KMV belong to the Bernoulli class and imply only two possible outcomes — default or survive. Default of an obligor  $i$  occurs if the value of the obligor's assets  $A_T^{(i)}$  in a valuation horizon  $T$  falls below a threshold value  $C_i$ , often interpreted as the value of the obligor's liabilities.

$$L_i = \mathbb{I}_{(A_T^{(i)}) < C_i} \sim B\{1; P(A_T^{(i)} < C_i)\} \quad (2.12)$$

Thus  $A_T$  can be regarded as a latent variable, which drives the default event implicitly replacing the notion of default correlation for the asset correlation. How is the correlation matrix of the latent variables defined? The answer lies in the basic assumption of both models, according to which the asset value dynamics relate to the changes in some common factors reflecting economic issues. Therefore asset correlations between obligors are induced exclusively by the correlation between the respective composite factors denoted by  $Y_i$ . In the typical model parametrization the latent variables are presented in the form of standardized asset log-returns:

$$r_i = \frac{\tilde{r}_i - \mathbb{E}(\tilde{r}_i)}{\sqrt{\text{Var}(\tilde{r}_i)}} \quad \text{with} \quad \tilde{r}_i = \log(A_T^{(i)}/A_0^{(i)}).$$

Suppose that the standardized log return of the asset value can be written as:

$$r_i = R_i Y_i + \varepsilon_i. \quad (2.13)$$

Here  $Y_i$  represents a weighted sum of many industry and country indices (composite factor). From the simple regression analysis we conclude that  $R_i^2$  defines how much the volatility of  $r_i$  can be explained by the volatility of  $Y_i$  and therefore it stands for the systematic risk of the obligor  $i$ . Respectively  $\varepsilon_i$  is the firm-specific effect.

The core assumption of *CreditMetrics*<sup>TM</sup> and KMV models is the multivariate normal (Gaussian) distribution of the latent variables  $r_i$ :

$$\begin{aligned} r_i &\sim N(0, 1) \\ Y_i &\sim N(0, 1) \\ \varepsilon_i &\sim N(0, 1 - R_i^2) \end{aligned}$$

In this case we can rewrite (2.12) as:

$$L_i = \mathbb{I}_{r_i < c_i} \quad (2.14)$$

where  $c_i$  is the threshold corresponding to  $C_i$  after replacing  $A_T$  for the standardized log returns  $r_i$ . Using (2.13) we can rewrite the threshold condition  $r_i < c_i$  as  $\varepsilon_i < c_i - R_i Y_i$ . Because  $r_i \sim N(0, 1)$ , from  $p_i = \mathbb{P}(r_i < c_i)$  we obtain

$$c_i = \Phi^{-1}(p_i).$$

After standardizing of  $\varepsilon_i$  the threshold condition changes to:

$$\frac{\varepsilon_i}{\sqrt{1 - R_i^2}} < \frac{\Phi^{-1}(p_i) - R_i Y_i}{\sqrt{1 - R_i^2}}. \quad (2.15)$$

On the right hand side of (2.15)  $Y_i$  is the only stochastic element. We therefore obtain (conditional on  $Y_i = y$ )

$$p_i(y) = \Phi\left(\frac{\Phi^{-1}(p_i) - R_i y}{\sqrt{1 - R_i^2}}\right). \quad (2.16)$$

Transforming this into the Bernoulli mixture setting yields

$$\begin{aligned} P(L_1 = l_1, \dots, L_m = l_m) \\ = \int_{[0,1]^m} \prod_{i=1}^m q_i^{l_i} (1 - q_i)^{1-l_i} dF(q_1, \dots, q_m). \end{aligned}$$

Now we are able to specify the probability distribution function:

$$F(q_1, \dots, q_m) = N_m(\mu, \Gamma)$$

where  $\mu = (p_1^{-1}(q_1), \dots, p_m^{-1}(q_m))^T$  and  $\Gamma$  is the asset correlation matrix of the log returns  $\tilde{r}_i$ .

The described modeling framework belongs to the KMV model. Though being based on the same assumptions, *CreditMetrics*<sup>TM</sup> differs from the KMV mainly in two issues: it uses equity instead of asset value process and it incorporates slightly different approach to defining composite factors. For further information on the model please refer to *CreditMetrics*<sup>TM</sup> Technical Document ([13]).

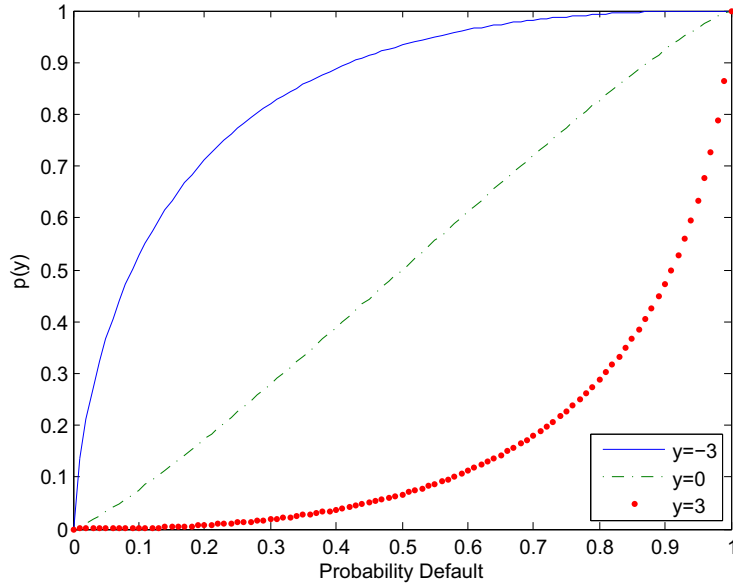


Fig. 2.2: The default probability  $p(y)$  as a function of the state of the economy.

## 2.5 One Factor Models

The multiple factor model has been introduced with equation (2.13). A one factor model simplifies the analysis since there is only one driving factor common to all obligors:  $Y \sim N(0, 1)$ . In our discussion we concentrate on the KMV Model (for references see [?]). In a one factor setup we model the (standardized) log returns:

$$r_i = \sqrt{\omega}Y + \sqrt{1 - \omega}Z_i \quad (2.17)$$

with idiosyncratic  $Z_i \sim N(0, 1)$ . The uniform asset correlation is denoted  $\omega$ . As before  $Z_i$  is assumed to be independent from the factor  $Y$ . Given a single factor and identical for all obligors  $\omega$ , we can rewrite equation (2.16) as:

$$p_i(y) = \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\omega}y}{\sqrt{1 - \omega}}\right) \quad (2.18)$$

In order to demonstrate the dependence of  $p_i(y)$  on the default probability given  $y$  values let us fix  $\omega = 20\%$  and  $y \in \{-3, 0, 3\}$ . The variable  $y \sim N(0, 1)$  can be interpreted as the state of the economy,  $y = -3$  corresponding to a bad state,  $y = 0$  meaning a typical state and  $y = 3$  indicating a good state of the economy. See Figure (2.2).

The joint default probability is given in the following proposition.

PROPOSITION 2.1:

In a one-factor portfolio model with uniform correlation  $\omega$  and loss statistics  $L_1, \dots, L_m$  with  $L_i \sim B(1, p_i)$ , where  $p_i$  is defined as in (2.18), the joint default probability (JDP) of two obligors is:

$$JDP_{ij} = P(L_i = 1, L_j = 1) = \Phi_2 \{ \Phi^{-1}(p_i), \Phi^{-1}(p_j); \omega \},$$

where  $\Phi_2[\cdot, \cdot; \omega]$  denotes the bivariate normal cumulative distribution function with correlation  $\omega$ .

*Proof.* The joint default probability can be calculated as

$$P(L_i = 1, L_j = 1) = P(r_i < \Phi^{-1}(p_i), r_j < \Phi^{-1}(p_j)).$$

By construction, the correlation between the asset value log-returns and  $r_i, r_j \sim \Phi(0, 1)$  is  $\omega$ . This proves the proposition.

As we want to derive the analytical approximation for portfolio loss distribution, the next step is to prove that with increasing portfolio size in terms of the number of obligors  $m$ , the portfolio loss distribution converges to a closed-form limit distribution. Here we allow for random *LGD* but deterministic (fixed) *EAD*. Also we do not exclude the dependence of *LGD* on the state of economy  $Y$ . This dependence between the default and the recovery rate on the same underlying factor is reasonable as empirical evidence shows that recovery rates tend to decrease while default rates rise. The framework used is Bernoulli mixture model, such that the counterparties are modeled by random variables:

$$\begin{aligned} L_i &\sim B(1, p_i) \\ Y &\sim \Phi(0, 1) \end{aligned}$$

$$((LGD_i \times L_i)|_{Y=y})_{i=1, \dots, m} \quad \text{independent} \quad (2.19)$$

where we assume that all involved random variables are defined on a common probability space.

The last condition in (2.19) means that we assume conditional independence of losses rather than independence of default indicators. The derivation of the limit loss distribution does not however depend on the particular distribution of the factor  $Y$ . It is sufficient that  $Y$  and the residual variables  $\varepsilon_1, \varepsilon_2, \dots$  are random variables in  $\mathbb{R}$  defined on some probability spaces  $(\Omega_Y, F_Y, P_Y)$ ,

$(\Omega_1, F_1, P_1)$ ,  $(\Omega_2, F_2, P_2)$  etc. Then the suitable probability space for Proposition (2.1) is the product space

$$(\Omega, F, P) = (\Omega_Y, F_Y, P_Y) \otimes (\Omega_1, F_1, P_1) \otimes (\Omega_2, F_2, P_2) \otimes \dots$$

because the variables  $Y, \varepsilon_1, \varepsilon_2, \dots$  are always assumed to be independent. For every  $\omega = (y, \epsilon_1, \epsilon_2, \dots) \in \Omega$  the loss variables  $L_i(\omega)$  are given by latent variable indicators evaluated with regard to the realization of  $\omega$ :

$$L_i(\omega) = \mathbb{I}_{[\sqrt{\omega}y + \sqrt{1-\omega}\epsilon_i < c_i]}.$$

For a portfolio of  $m$  obligors, the portfolio loss relative to the portfolio's total exposure is given by:

$$L^{(m)} = \sum_{i=1}^m w_i LGD_i L_i, \quad w_i = \frac{EAD_i}{\sum_{j=1}^m EAD_j}. \quad (2.20)$$

Following [?] we will introduce a technical assumption in order to infer the closed-form limit loss distribution. Consider an infinite number of loans with exposure  $EAD_i$ . Assume that the following holds:

$$\sum_{i=1}^m EAD_i \nearrow \infty \quad (m \rightarrow \infty), \quad (2.21)$$

$$\sum_{m=1}^{\infty} \left( \frac{EAD_m}{\sum_{i=1}^m EAD_i} \right)^2 < \infty. \quad (2.22)$$

Condition (2.21) states that the total exposure of the portfolio strictly increases to infinity with increasing number of obligors. Condition (2.22) implies that the exposure weights compress very rapidly with increasing number of obligors. These assumptions are necessary to make sure that the exposure share of each entity in the reference portfolio tends to zero.

**PROPOSITION 2.2:**

Assumptions (2.21), (2.22) are sufficient to guarantee that in the limit the percentage portfolio loss  $L^{(m)}$  defined in (2.20) and the conditional expectation  $E(L^{(m)}|Y)$  are equal almost surely, such that

$$P \left[ \lim_{m \rightarrow \infty} (L^{(m)} - E(L^{(m)}|Y)) = 0 \right] = 1. \quad (2.23)$$

See proof of the proposition (2.2) in the Appendix B.

In the case that  $(LGD_i L_i)_{i \geq 1}$  are not only conditionally independent but also identically distributed, Proposition (2.2) can be reformulated as follows:

There exists some measurable function  $p : \mathbb{R} \rightarrow \mathbb{R}$  such that for  $m \rightarrow \infty$  the portfolio loss  $L^{(m)}$  converges to  $p \circ Y$  almost surely. Moreover,  $p \circ Y$  equals  $E(LGD_i L_i | Y)$  almost surely.

Because  $E(L^{(m)} | Y)$  is by definition  $\sigma(Y)$ -measurable, where  $\sigma(Y)$  denotes the  $\sigma$ -Algebra generated by  $Y$ , there exists some measurable function  $p : \mathbb{R} \rightarrow \mathbb{R}$  with  $E(L^{(m)} | Y) = p \circ Y$ . Combined with the Proposition (2.2) and the assumption that all losses are identically distributed this concludes the proof of the above statement.

Thus it is shown that for  $m \rightarrow \infty$  the randomness of the portfolio loss  $L^{(m)}$  solely depends on the randomness of the factor  $Y$ : by increasing the number of obligors in the portfolio, the specific risk is completely removed and only the systematic risk arising from the volatility of the common factor remains in the portfolio. Assuming uniform default probabilities  $p_i$  for all obligors  $i$  and applying KMV framework to our analysis we infer:

$$E(L^{(m)}) = \sum_{i=1}^m w_i E(L_i | Y) = \Phi \left\{ \frac{\Phi^{-1}(p_i) - \sqrt{\omega} y}{\sqrt{1 - \omega}} \right\} =: p(Y),$$

such that from (2.2) it follows that

$$L^{(m)} \xrightarrow{m \rightarrow \infty} p(Y) = \Phi \left( \frac{\Phi^{-1}(p_i) - \sqrt{\omega} y}{\sqrt{1 - \omega}} \right) \quad \text{almost surely.} \quad (2.24)$$

We have obtained that for sufficiently large portfolios the percentage of defaulted loans given a certain state of economy  $Y = y$  is approximately equal the conditional default probability  $p(Y)$ . Now we want to derive the cumulative distribution function of the limit loss variable  $p(Y)$  and thus define the loss distribution. Denote here the limit of  $L^{(m)}$  by  $L$ . For every  $0 \leq x \leq 1$  we have then:

$$\begin{aligned} P(L \leq x) &= P(p(Y) \leq x) & (2.25) \\ &= P \left( -Y \leq \frac{1}{\sqrt{\omega}} (\Phi^{-1}(x) \sqrt{1 - \omega} - \Phi^{-1}(p)) \right) \\ &= \Phi \left( \frac{1}{\sqrt{\omega}} (\Phi^{-1}(x) \sqrt{1 - \omega} - \Phi^{-1}(p)) \right). \end{aligned}$$

Now we can calculate the corresponding probability distribution function by taking the derivative of (2.25):

$$f_{p,\omega}(x) = \sqrt{\frac{1-\omega}{\omega}} \exp\left(\frac{1}{2}(\Phi^{-1}(x))^2 - \frac{1}{2\omega}(\Phi^{-1}(p) - \sqrt{1-\omega}\Phi^{-1}(x))^2\right) \quad (2.26)$$

It is also possible to derive the expression for  $m$ -moment of portfolio loss distribution:

$$E(L^{(m)}) = \Phi_m[(\Phi^{-1}(p), \dots, \Phi^{-1}(p)), \Gamma_\omega] \quad (2.27)$$

Obviously we need a factor model to define asset correlation  $\omega$  and some market data to calibrate the default probability  $p$ . Next we will show how to estimate asset correlation from historic default frequencies using one factor model.

## 2.6 Estimation of asset correlation

Our first step is to calibrate default probabilities. Table (2.1) presents Moody's historic corporate bond default frequencies from 1970 to 2004. For each rating class  $R_i$  we calculate the mean and the standard error of the historic default frequencies. Then we use simple regression to fit the mean by an exponential function. As a result we obtain fitted default probabilities  $\mu_1, \dots, \mu_6$  for all rating classes (see table (2.2)). Analogously we fit the volatilities of the default frequencies.

The second step includes calculating the asset correlations. We refer to the formula (2.18) from the one-factor model, in which we replace true default probability  $p_i$  with the fitted mean default rate  $\mu_i$ . It could be shown that the following expression is true for the considered model:

$$\text{Var}(P(Y)) = \Phi_2\{\Phi^{-1}(p), \Phi^{-1}(p); \omega\} - p^2 \quad (2.28)$$

where we again replace the true unknown variance of default rate for the fitted default volatility  $\sigma$ . Thus, the asset correlation  $\omega$  is the only unknown parameter in (2.28). The calibrated correlations are showed in the last column of table (2.2).

Year	Aaa	Aa	A	Baa	Ba	B
1970	0.00%	0.00%	0.00%	0.27%	4.12%	20.78%
1971	0.00%	0.00%	0.00%	0.00%	0.42%	3.85%
1972	0.00%	0.00%	0.00%	0.00%	0.00%	7.14%
1973	0.00%	0.00%	0.00%	0.46%	0.00%	3.77%
1974	0.00%	0.00%	0.00%	0.00%	0.00%	10.00%
1975	0.00%	0.00%	0.00%	0.00%	1.02%	5.97%
1976	0.00%	0.00%	0.00%	0.00%	1.01%	0.00%
1977	0.00%	0.00%	0.00%	0.28%	0.52%	3.28%
1978	0.00%	0.00%	0.00%	0.00%	1.08%	5.41%
1979	0.00%	0.00%	0.00%	0.00%	0.49%	0.00%
1980	0.00%	0.00%	0.00%	0.00%	0.00%	4.94%
1981	0.00%	0.00%	0.00%	0.00%	0.00%	4.49%
1982	0.00%	0.00%	0.26%	0.31%	2.72%	2.41%
1983	0.00%	0.00%	0.00%	0.00%	0.91%	6.31%
1984	0.00%	0.00%	0.00%	0.36%	0.83%	6.72%
1985	0.00%	0.00%	0.00%	0.00%	1.40%	8.22%
1986	0.00%	0.00%	0.00%	1.33%	2.03%	11.73%
1987	0.00%	0.00%	0.00%	0.00%	2.71%	6.23%
1988	0.00%	0.00%	0.00%	0.00%	1.24%	6.36%
1989	0.00%	0.61%	0.00%	0.59%	2.98%	8.95%
1990	0.00%	0.00%	0.00%	0.00%	3.35%	16.18%
1991	0.00%	0.00%	0.00%	0.27%	5.35%	14.56%
1992	0.00%	0.00%	0.00%	0.00%	0.30%	9.03%
1993	0.00%	0.00%	0.00%	0.00%	0.56%	5.71%
1994	0.00%	0.00%	0.00%	0.00%	0.24%	3.82%
1995	0.00%	0.00%	0.00%	0.00%	0.69%	4.81%
1996	0.00%	0.00%	0.00%	0.00%	0.00%	1.44%
1997	0.00%	0.00%	0.00%	0.00%	0.19%	2.12%
1998	0.00%	0.00%	0.00%	0.12%	0.63%	4.26%
1999	0.00%	0.00%	0.00%	0.10%	1.01%	5.85%
2000	0.00%	0.00%	0.00%	0.38%	0.89%	5.49%
2001	0.00%	0.00%	0.16%	0.19%	1.57%	9.36%
2002	0.00%	0.00%	0.16%	1.21%	1.54%	4.97%
2003	0.00%	0.00%	0.00%	0.00%	0.95%	2.66%
2004	0.00%	0.00%	0.00%	0.00%	0.19%	0.65%

Tab. 2.1: Moody's Corporate Bond Historic Default Frequency 1970-2004

Rating	Mean	Stand.Dev.	$\mu$	$\sigma$	$\rho$
Aaa	0.000%	0.000%	0.0124%	0.1326%	42%
Aa	0.017%	0.103%	0.0413%	0.2546%	37%
A	0.922%	3.766%	0.1374%	0.4890%	31%
Baa	0.168%	0.319%	0.4565%	0.9390%	23%
Ba	1.170%	1.267%	1.5171%	1.8031%	15%
B	6.271%	4.643%	5.0417%	3.4627%	11%
Mean	1.425%	1.683%	1.2011%	1.1802%	26%

Tab. 2.2: Calibration Results due to exponential function fitting

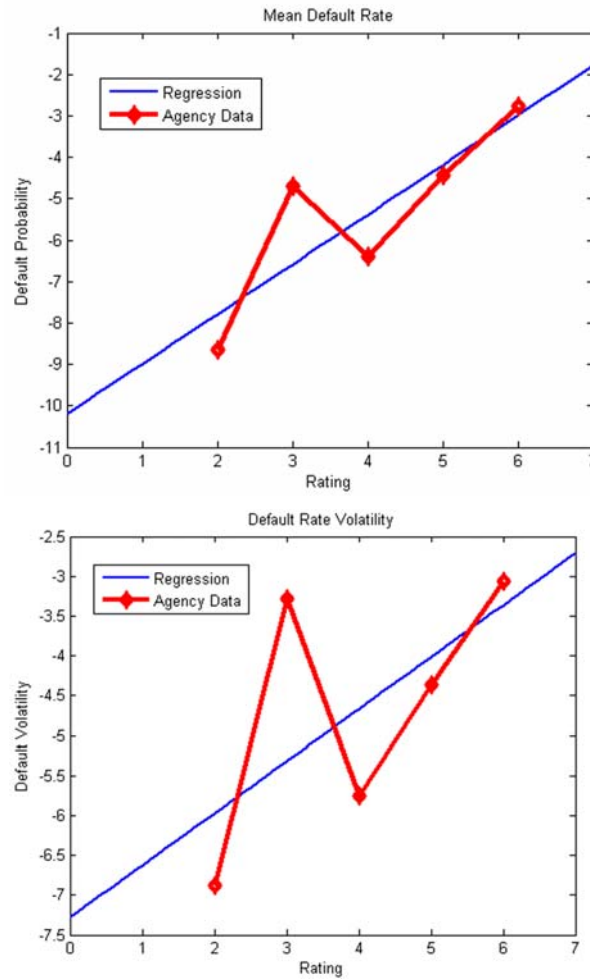


Fig. 2.3: Mean default rate and Default rate volatility. The red and blue lines represent the historic default and the regression by exponential function fitting correspondingly.

## 3. COLLATERALIZED DEBT OBLIGATIONS

### 3.1 *Typical CDO Structure*

Collateralized debt obligation (CDO) is a structured financial product that securitises a diversified portfolio of debt assets (e./ g. bank loans) by transferring the credit risk to the external investors. The initial debt portfolio is called the collateral. The idea behind CDO is to sell bonds of different seniority parceled in tranches and backed by the assets in the collateral. In fact the initiator of CDO performs the repackaging of the original risk profile by offering investors with different risk-return preferences a choice of respective tranches. The seniority of the tranches reflects the order, in which the losses in the collateral affect the tranche investors. Each tranche is defined by the percentage of losses in the collateral that it carries. For example, the most subordinated tranche (called equity) suffers from the first 3% losses in the collateral. If losses are for example 5% of the collateral notional, the equity investors carry the first 3% (thus losing all their investment), and the next 2% are carried by those who invested in the junior tranche. The senior tranche investors suffer only if the total collateral portfolio loss exceeds 22% of its notional value. The details of the risk transfer mechanism incorporated in the typical CDO structure and its advantages are presented further in this chapter. For references see [3], [1].

The first prerequisite for the CDO transaction is a pool of credit risky assets on the balance sheet of the originator. It is often that banks purchase such pool intentionally for CDO, which is in that case motivated by the arbitrage spread opportunities. The credit risk can also be artificially created by purchasing a pool of CDS<sup>1</sup> (Credit Default Swap), thus selling an insurance

---

<sup>1</sup> Credit Default Swap is a contract, which references the debt assets of a particular firm. In case these assets default, the seller of the swap (protection seller) pays compensation to the swap buyer (protection buyer). Typically the compensation implies buying-out the debt assets at their face value. In return for this, protection buyer pays a periodic fee until default occurs or until the swap matures. This fee is termed the CDS spread and it is fixed at the initiation of the contract.

on the default event. In this case the transaction is called synthetic CDO. Next step in the CDO scheme is the transfer of the collateral assets to a so

Tranche	Rating	Premium (bp)
Super senior (22-100%)	Unrated	1.5
Super senior(junior) (12-22%)	AAA	4
Senior (9-12%)	AAA	9
Mezzanine (senior) (6-9%)	AAA	18
Mezzanine (junior) (3-6%)	BBB	63
Equity (0-3%)	N.A.	34% (upfront fee)

Tab. 3.1: Typical CDO tranches

called SPV (Special Purpose Vehicle) - a company set-up specially for the transaction. The main reason for founding a new company is the condition of bankruptcy remoteness of the SPV. That means SPV's own bankruptcy risk is minimized as it will not default on its obligations because of the insolvency of its originator. In terms of cash flow the transfer of the collateral to the SPV can be a complete sale, though the originator often keeps the administration of the asset pool. The cash proceeds of the originator are the principal of the collateral. The SPV becomes the owner of all the cash flows arising from the asset pool. Any fixed-rate loans in the collateral pool are hedged by SPV against interest rate risk through interest-rate (fixed-to-floating) swaps. To pay out to the originating company for the debt assets SPV issues a securities - structured notes backed by asset pool on its balance. The total notional of the notes equals the principal balance of the collateral pool. Interests and principal of the notes are paid from the cash proceeds and principal of the pool. As it was already said the notes are structured into tranches of different risk classes.

### 3.1.1 Cash flow structure of CDO

Principal (repayment/amortization of debt securities) and interest are distributed sequentially top-down to the note investors in the order of seniority. There is a deleveraging mechanism in the CDO cash flow waterfall represented by the overcollateralization (O/C) and interest coverage (I/C) tests. O/C tests make certain that the principal coverage of the collateral are sufficient for a certain (over)coverage of the premium to be paid out to the

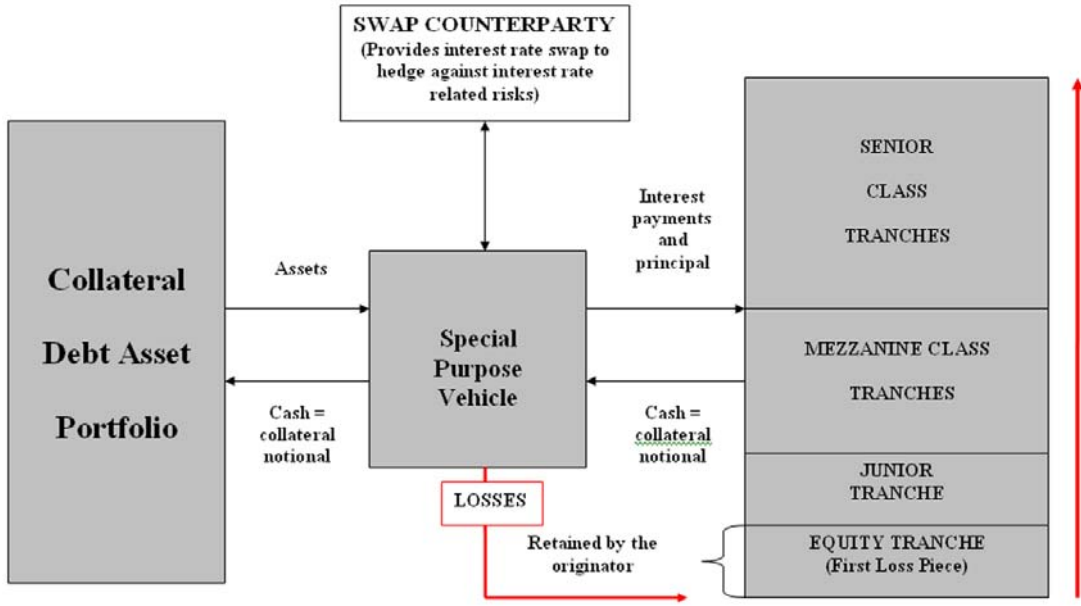


Fig. 3.1: CDO Scheme

note holders. For example, first O/C test is to be taken for the most senior tranche (A):

$$(O/C)_A = \frac{PV_{pool}}{PV_A},$$

where  $PV_A$  is the par value of the class A (senior) notes. This test is passed if  $(O/C)_A \geq 120\%$ . The O/C test for the class B notes is passed if:

$$(O/C)_B = \frac{PV_{pool}}{PV_A + PV_B} \geq 110\%.$$

An I/C tests take care that the interest proceeds from the collateral are sufficient for paying any expenses and coupons on the liability side of the structure and due to any counterparties involved. An example of I/C test for class A notes is given below:

$$(I/C)_A = \frac{PV_{pool} \times WAC \times 0.5 - FEES \times 0.5}{PV_A \times C_A \times 0.5},$$

where  $FEES$  are required annual fees to any counterparties (e.g. hedge counterparty),  $WAC$  is weighted average coupon of the collateral pool,  $C_A$  — the coupon on class A notes, and the factor 0.5 reflects that the interest is calculated w.r.t. a semiannual horizon, covering two (quarterly) payment

periods. This test is passed if

$$(I/C)_A \geq 140\%.$$

If any of these tests is broken, cash is typically redirected in a way trying to bring the broken test in line again. In this way, the interest stream to the tranche investors is restructured backwards (from the bottom to the top) thus mitigating losses. Most senior note holders receive their payments first and it is respectively the lowest coupon. More junior note investors receive payments only if more prioritized payments are in line with the documentation of the structure. The more junior tranche is, the higher coupon it carries, reflecting the higher risk borne by these notes. The equity tranche (also called "first loss piece") carries no promised coupons. Instead equity investors obtain the excess spread of the structure in every payment period, i. d. the cash left-over after payments of coupons on classes A to D. Typically the originator of the CDO keeps some part of the equity tranche in order to participate in the excess spread of the interest waterfall. This serves as a "guaranty" to the market showing that the arranger itself trusts the structure and the underlying credits as he retains part of the credit risk.

The timing of defaults is crucial to first loss piece investors. If defaults occur at the end of the CDO lifetime (backloaded structure), equity holders have enough time to collect the excess spread. In this case though the probability that 3% of the collateral assets default is relatively high, the equity holders are rewarded by an attractive overall return even if they loose a substantial part of their investment. On the other hand, the frontloaded transaction with defaults taking place at an early stage of the deal presents a bad scenario for subordinated investors, because they bear not only the first loss loosing their investment but also miss the spread, as excess cash is redirected and no longer distributed to them.

### 3.1.2 Motivation for CDO transaction

As it was already mentioned, one of the possible reasons for initiating CDO is the spread arbitrage. The arbitrage exists when the total spread collected on credit risky instruments at the asset side of the transaction exceeds the total "diversified" spread to be paid to investors on the tranced liability side. The first loss piece investors benefit from the possible arbitrage. The second motivation is the credit risk transfer. This is typically the case with CLO - collateralised loan obligations (with bank loans in the collateral portfolio).

The securitisation has an impact on the loss distribution of the underlying

reference pool: it divides the loss distribution in two parts. The first loss segment refers to losses carried by the originator by retaining the equity piece. The excess loss of the first loss piece is taken by the CDO investors. Thus the upper boundary of the first loss piece is an effective loss cap for the originator. Even if the first loss is higher than the expected loss on the original asset pool, the CDO transaction leads to the risk transfer because it protects the originator against the unexpected loss. Formally, there is a risk transfer if  $P(L > FLP) > 0$ , where  $FLP$  stands for "first loss piece". The risk measures for different values of expected loss could be found in tables (3.2), (3.3).

EL = p (in %)	Correlation parameter $\omega$					
	1%	5%	10%	20%	30%	60%
0.1	0.04	0.08	0.14	0.32	0.4	1
0.5	0.14	0.38	0.5	0.98	1.4	2.6
1	0.27	0.73	1	1.4	2.3	4.1
3	0.69	1.56	2.2	3.5	4.7	10
5	1.1	2.15	3.05	4.64	6.18	11.3
10	1.77	3.84	5.13	8.12	9.76	16.26

Tab. 3.2: Unexpected Loss (% of Portfolio Notional) for homogenous portfolio ( $L \sim F_{p,\omega}$ ) w.r.t.  $p$  and  $\omega$

EL = p (in %)	Correlation parameter $\omega$					
	1%	5%	10%	20%	30%	60%
0.1	0.22	0.49	0.82	1.5	2.24	4.17
0.5	0.99	2	3.17	5.57	8.19	18
1	1.88	3.6	5.6	9.46	13.7	30
3	5.14	9	13	20.7	28.72	57.2
5	8.16	13.6	19	29.1	39	71
10	15.2	23.5	31.1	44.2	56.14	87

Tab. 3.3: VAR (99.5 quantile) (% of Portfolio Notional) for homogenous portfolio ( $L \sim F_{p,\omega}$ ) w.r.t.  $p$  and  $\omega$

The third motivation is the regulatory capital relief. In general, loan pools require a financial institution hold regulatory capital in amount of 8% of the

---

risk-weighted assets of the reference pool. After the CDO transaction the originator needs to hold capital only for the retained equity piece of the CDO structure. However, there are opportunity costs for the capital relief. The originator has to cover upfront expenses (rating agencies, lawyers, structuring and underwriting costs) and bear the ongoing asset pool administration costs. Finally, the CDO deal leads to the cash in-flow as a result of the assets' sale to an SPV. The advantage of the refinancing by means of the securitization is that resulting costs are mainly related to the credit quality of the transferred assets (collateral) and not so much to the rating of the originator. Thus, CDO can be a way out for companies with fast declining rating when the funding from other sources becomes too expensive for them.

### 3.2 Valuation of CDO

Valuation of a CDO implies finding the fair spread of each tranche. The fair spread is by definition a spread, with which the mark-to-market value of the contract is zero. The issuer of credit linked notes pays the fair spread to tranche investors, if the present value of the fee payments is equal to the present value of the contingent payments, using risk-neutral valuation. We assume here that there exists a risk-neutral martingale measure  $Q$ , under which all price processes that are discounted under interest rate process are martingales. All expectations in the following model are with respect to this measure.

The fee payments are called "protection leg" and refer to cash flow that covers losses affecting the specific tranche and are paid out to the protection buyer (i. e. CDO initiator). The contingent payments are called "premium leg" and is generally paid out quarterly in arrears to the protection seller (i. e. tranche investor). On order to evaluate protection and premium legs we need to specify the expected losses of each CDO tranche, how they are determined by the expected portfolio (collateral) loss. Thus the tranche spread is derived in terms of expected losses.

For convenience we define the following variables:

$K_j$	upper attachment point of tranche $j$
$S_j$	spread of tranche $j$ paid per year
$F_j(t)$	face value of tranche $j$ at time $t$
$L(t)$	portfolio loss at time $t$
$L_j(t)$	loss of tranche $j$ at time $t$
$DF(t)$	discount factor at time $t$
$T$	time of maturity of CDO, as a fraction of year

As it was already stated a tranche suffers a loss only if the total portfolio loss (in % of its notional) exceeds the lower attachment point of this tranche. The maximum loss of a tranche  $j$  is the tranche's size, which could be defined as  $K_j - K_{j-1}$ . Then it is possible to express the loss of tranche  $j$  in terms of total portfolio loss  $L(t)$ :

$$L_j(t) = \min[\max(0; L(t) - K_{j-1}); K_j - K_{j-1}] \quad (3.1)$$

Then the present value of the contingent (protection leg) paid out, given the following payment dates

$$0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T$$

can be calculated taking the expectation with respect to the risk-neutral probability measure:

$$PV_j^{protection} = E \left[ \sum_{i=1}^n (L_j(t_i) - L_j(t_{i-1})) DF(t_i) \right] \quad (3.2)$$

Assuming a continuous time payment equation (3.2) turns to the following expression:

$$PV_j^{protection} = E \left[ \int_0^T DF(t) dL_j(t) \right] \quad (3.3)$$

Premium payments to the investors depend on the face value at time  $t$  of the tranche  $F_j(t)$ , i. e. of the expected survival of the tranche, which can be written as

$$F_j(t) = (K_j - K_{j-1}) - L_j(t) \quad (3.4)$$

From (3.4) we obtain the expected present value of the fee payments (premium leg):

$$PV_j^{premium} = S_j \cdot E \left[ \sum_{i=1}^n \Delta t_i DF(t_i) F_j(t_i) \right] \quad (3.5)$$

---

Equalizing (3.3) and (3.5) we infer the fair spread  $S_j^*$ :

$$S_j^* = \frac{\mathbb{E} \left[ \int_0^T DF(t) dL_j(t) \right]}{\mathbb{E} [\sum_{i=1}^n \Delta t_i DF(t_i) F_j(t_i)]} \quad (3.6)$$

We conclude that in order to price a CDO it is essential to calculate the cumulative portfolio loss distribution. Possible approach to finding portfolio loss distribution is described in Chapter (2). For references to this section see [4], [10], [12].

## 4. IMPLIED CORRELATION SMILE

### 4.1 *Correlation and Tranche Loss*

In the recent time market for credit derivatives boomed. The liquidity in trading standardized CDO indices led to the market convention of pricing tranches in terms of implied correlation rather than in terms of spread. Market participants are interested in correlation as it can facilitate a comparison of prices across tranches. This practice was inspired by the use of Black-Scholes implied volatilities of equity markets. Even though Black-Scholes model is simplifying reality, the implied volatilities provide a common benchmark for comparison of options across maturities and strikes. A direct extension from Black-Scholes implied volatilities to the CDO market is termed compound correlations that are by definition obtained through inverting a standard pricing model for tranching products. The implied correlation is found by matching the model generated prices to market quoted spreads.

This chapter introduces the concept of implied (compound) correlation, provides analysis of the dependence between correlation and tranche spreads, and presents the pitfalls and weaknesses of this concept that are necessary to know when applying it on practice. For references to this chapter please see [25],[17], [1], [9].

As we have already seen the prices of the CDO tranches depend on the perceived likelihood of the joint defaults of the underlying pool (collateral portfolio). The loss distribution of the portfolio is in turn influenced by the parameter of default correlation. The figure below demonstrates this dependence. It is obvious that higher correlation between defaults (more correct: correlation between the random variables  $L_i = I_D$ , where  $L_i = 1$  when  $i$  asset defaults) leads to higher probability of joint default occurrence. In other words, high correlation implies that there will be either few defaults or many. On the other hand, low correlation between defaults of entities in the reference portfolio means that the probability that many securitites default at the same time is also relatively low keeping marginal default probabilities

fixed. This is demonstrated on the Figure 4.1 (Please, note that the scale for y axis was chosen different in every case for convenience of the reader, the x axis represents loss as a percentage of exposure).

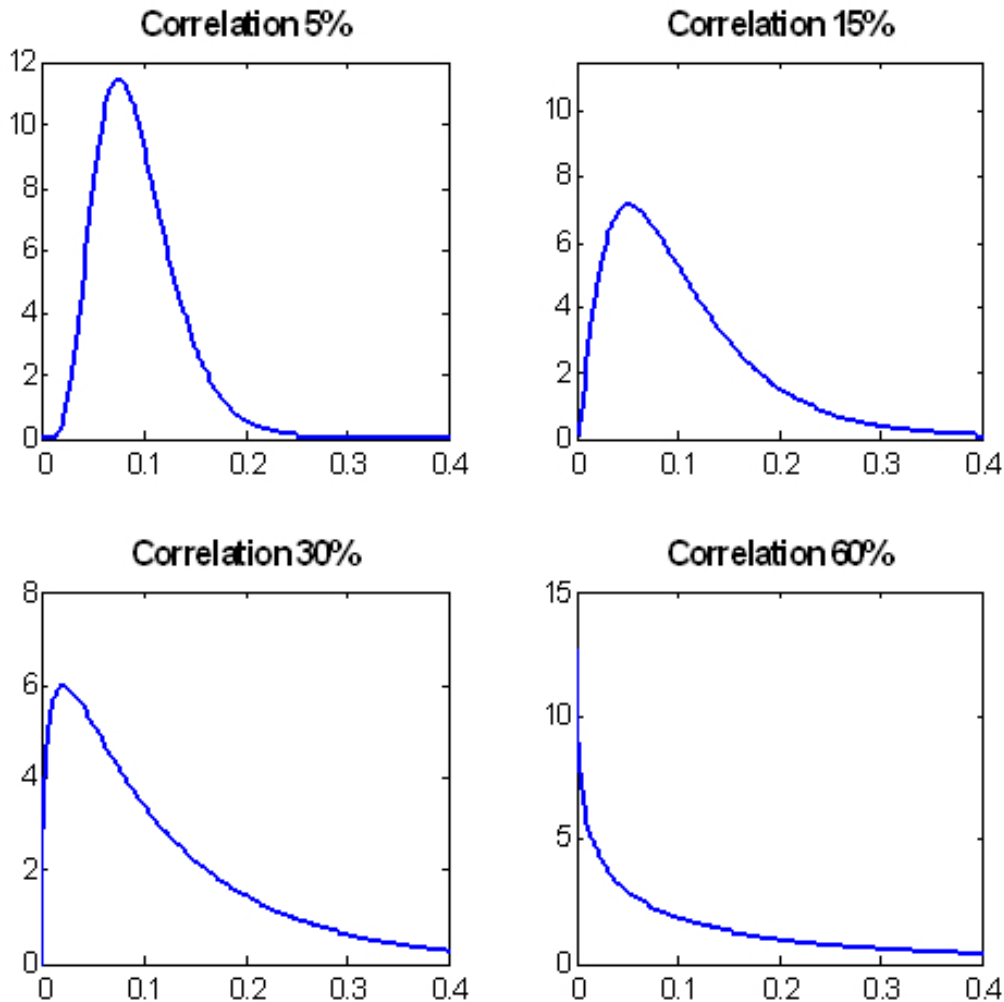


Fig. 4.1: Portfolio Loss Distribution for different correlation parameters and fixed probability of default 10%.

Further we provide intuition on how default correlation affects tranche spreads through its influence on collateral loss distribution. First, let us look at the equity tranche, that suffers the first losses, absorbing any losses below its upper attachment point of 3%. Because of this upper limit on losses, the equity tranche is not much influenced by occurrence of many defaults while equity investors are better off with few defaults occurred as payments to them are not redirected and their investment is not totally lost. This reduces the

expected loss of the tranche and therefore reduces the fair spread paid to its holders. Analogously, when correlation parameter increases, the expected tranche loss and the respective fair spread also go up.

In contrast senior tranche investors are better off with default correlation at a low level. Senior notes holders are affected only when losses in the collateral exceed 22% of the pool notional. Many defaults should occur in order to make such loss possible. The probability of this event is higher when correlation increases. Thus the expected tranche loss and the fair spread increase monotonically with correlation. These intuitively driven conclusion is summarized on the Figures 4.2, 4.3.

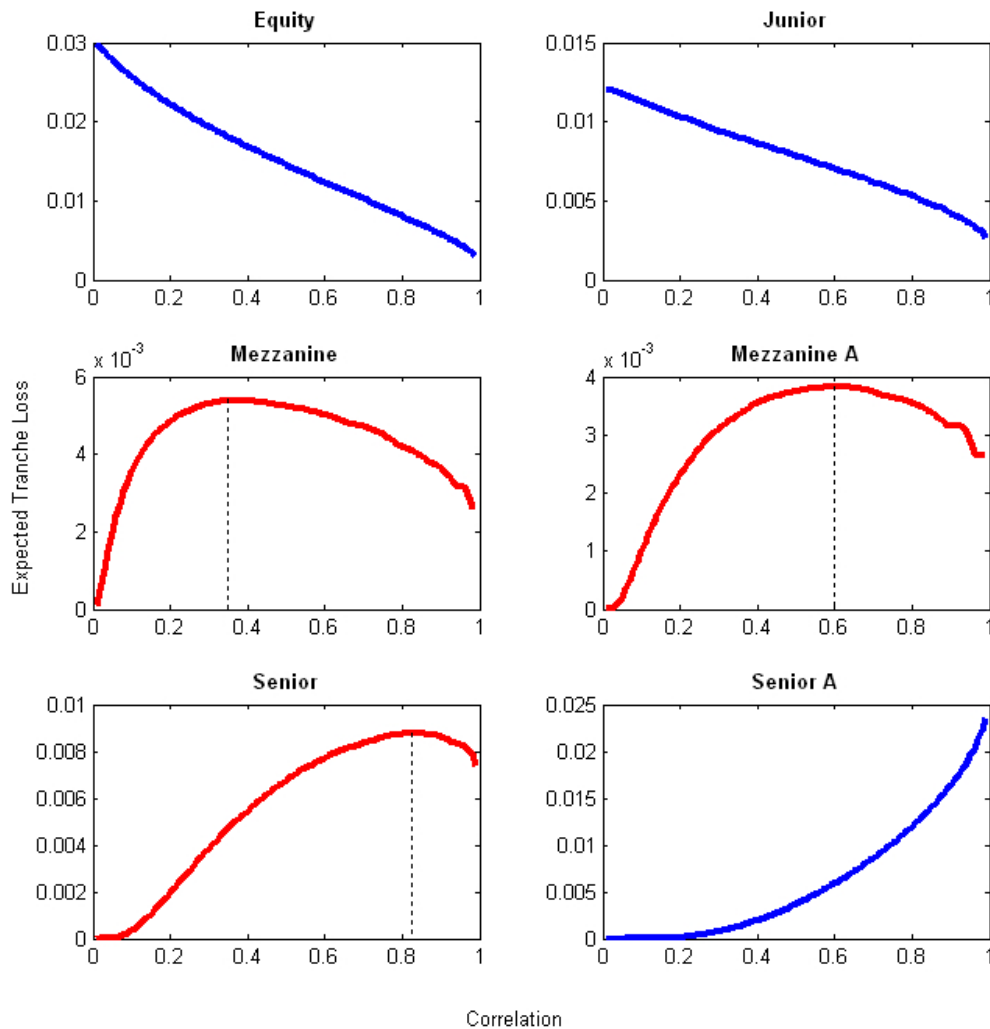


Fig. 4.2: Expected Tranche Loss as a function of default correlation.

While for the most subordinated and senior tranches the dependence between correlation parameter and expected tranche loss is monotonic, though with a different sign, for mezzanine tranches it is not the case. The tranche loss is defined (omit the time subscript indices for simplicity)

$$L_{Mezz} = \min(L, K_U) - \min(L, K_L), \quad (4.1)$$

where  $L$  is portfolio loss variable,  $K_U$  — upper attachment point for mezzanine tranche, and  $K_L$  - lower attachment point respectively.

There are three possible outcomes:

- $L < K_L$ , then  $L_{Mezz} = 0$ ;
- $K_L < L < K_U$ , then  $L_{Mezz} = L - K_L$ ;
- $L > K_U$ , then  $L_{Mezz} = K_U - K_L$ .

It can be seen now that for both summands in the equation (4.1) the expected value is decreasing in the correlation characteristic of the portfolio loss distribution. Since these components enter the equation with opposite signs, we are not able to assume that tranche loss is monotonic in correlation. This is well observed on the middle graphs in Figure 4.2. Analogous picture could be seen on the diagram with correlation plotted against fair spread (Figure 4.3). Modeled spread for equity tranche rises dramatically as correlation goes to zero, because equity investors are the first to suffer any losses.

## 4.2 *Compound Correlation*

In this section we will use market data to infer the implied correlation. First, we present the assumptions of the standard market model used and give the description of the dataset. Second, we will explore the main properties of compound correlation.

### 4.2.1 *The Data*

The market data we take as input consists of prices of CDS Europe iTraxx index as well as of prices of iTraxx derivative product - tranching iTraxx. ITRAXX Europe is one of the most popular CDS indices representing a

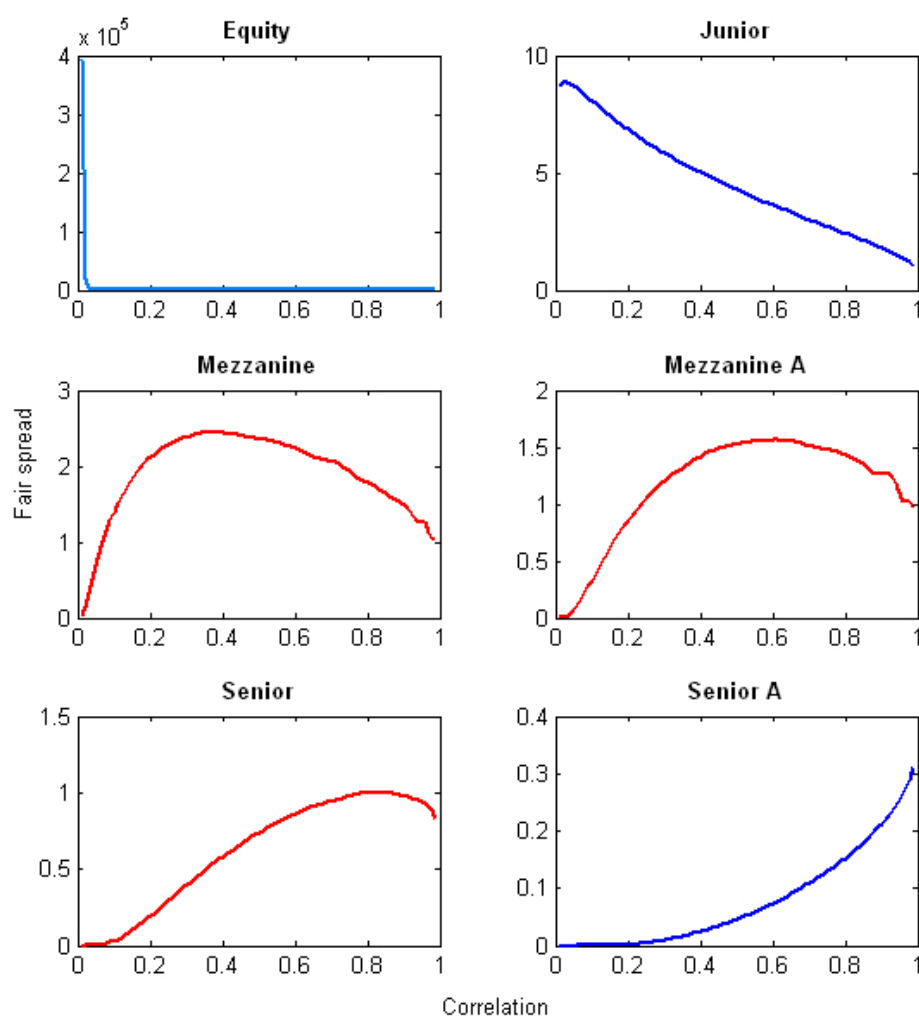


Fig. 4.3: Fair spread as a function of default correlation.

portfolio of 125 equally-weighted default swaps on European names. It is split into traded sector indices (Autos, Consumer, Energy, Financial senior and subordinated, Industrials, TMT and Non-Financial) and a HiVol index of the 30 widest spread non-financial names. A Crossover index comprising the 45 most liquid sub-investment grade non-financials is also traded. Figure 4.4 demonstrates the whole family of iTraxx products. The iTraxx CDS indices typically trade 5 and 10-year maturities and a new series is issued every 6 months in March and September. The Europe iTraxx also trade 3 and 7-year maturities. More detailed information on iTraxx could be found on the website of International Index Company Limited ([www.indexco.com](http://www.indexco.com)).

BENCHMARK ISSUES	SECTOR INDICES		DERIVATIVES
<b>iTRAXX EUROPE</b> 125 credit traded entities	<b>NON-FINANCIALS</b> 100 entities	<b>INDUSTRIALS</b> 20 entities	<b>TRANCHED iTRAXX</b> five standardised tranches of iTraxx Europe 0-3% 3-6% 6-9% 9-12% 12-22%
<b>iTRAXX EUROPE HiVol</b> 30 highest spread entities from iTraxx Europe	<b>FINANCIALS SENIOR</b> 25 entities	<b>ENERGY</b> 20 entities	
<b>iTRAXX EUROPE CROSSOVER</b> 45 sub-investment grade entities	<b>FINANCIALS SUB</b> Same 25 as Financials Senior	<b>CONSUMERS</b> 30 entities	<b>iTRAXX OPTIONS</b>
	<b>TMT</b> 20 entities	<b>AUTOS</b> 10 entities	<b>iTRAXX FUTURES</b> iTraxx Europe exposure traded as a futures contract being planned
First to Default baskets: Autos, Consumer, Energy, Financial (Sen/Sub), Industrials, TMT, HiVol, Crossover, Diversified			

Fig. 4.4: iTraxx family. Source: IIC presentation "iTraxx CDS Indices".

The tranching iTraxx investor who is essentially the seller of protection, is responsible for all default losses on an underlying index portfolio of default swaps in excess of a respective tranche attachment point up to the detachment point. Thus, a tranching index has the same risk characteristics as a collateralized debt obligation. In return for covering the losses, protection sellers receive fees quoted in two parts: an upfront-fee and a running spread. Both of these are quoted in turn as a fraction of maximum amount of loss for a tranche. Once default occurs, the notional amount upon which the running spread is charged is reduced, dollar for dollar, with losses. All tranches have an assigned ("promised") running spread, and only the equity tranche has an upfront fee. Unlike other tranches, the "First Loss Piece" tranche has a contractually set running spread of 500 basis points, and the upfront fee is negotiated in the market. The prices of CDS Index are used to bootstrap the default probabilities, while the market quoted spreads and the equity upfront fee for tranching iTraxx are applied to CDO pricing function when inverting it for implied correlation. An example of data for one day is presented in the table 4.1.

iTraxx Europe 7S 10Y 25 May 2007		
Tranche	Running Spread (in bp)	Upfront Fee (in %)
22-100%	4.04	0
12-22%	12.7	0
9-12%	41.2	0
6-9%	89	0
3-6%	304	0
0-3%	500	36

Tab. 4.1: Example of the dataset. Source: Bloomberg.

#### 4.2.2 LPHGC Model

We use standard pricing model, described in section 3.2, which relies on the following assumptions:

- The default of a reference entity in the portfolio is triggered when its asset value falls below a barrier, the asset value of the portfolio is driven by a common, standard normally distributed factor  $Y$  (the framework of *CreditMetrics*<sup>TM</sup> one-factor model).
- There is an equally weighted homogeneous portfolio of credit risky assets, the number of entities  $m$  tends to infinity, which effectively cancels the effect of a single name's performance on tranche loss.
- The assets in the collateral are assumed to have the same default probability  $p$ .
- The uniform default probability is bootstrapped from the mid market spread on equivalent CDS index. It would be more consistent to use individual spreads for all the credits in the reference portfolio to account for single name blow ups, but this would require data that is not available instantly.
- Loss-given-Default ( $LGD = 1 - R$ ) is assumed to be constant and therefore independent on default. The simple approach is to take the recovery rate  $R$  equal to the average historical recovery rate on senior unsecured bonds for US corporations (40%).

- The model assumes identical constant pairwise default correlation  $\omega$ .
- The timing of defaults for the  $m$  assets over the lifetime of the contract can be calculated from a joint default probability distribution, which is assumed to be multivariate normal distribution (in other words, Gaussian copula). The construction of this analytical approximation is presented in section 2.5.

It is obvious that the Large Pool Homogeneous Gaussian Copula Model is simplistic, and because of its transparency and replicability it became a market standard. Besides, it requires few inputs, which are easy to agree on. Details on practical implementation could be found in [22], [24].

Next, we took data on market quotations and invert the pricing formula to find implied correlations for each tranche. This procedure of implying correlation differs for the first equity tranche, because we don't have a market quoted running spread for this tranche, we need to apply a recursive technique in order to account for upfront fee and contractually fixed running spread of 500 bp. The formal procedure is as follows:

1. Find a vector of expected survival ( $Esur$ ) for equity tranche for all premium payment dates as a function of default correlation

$$Esur = f(corr)$$

2. Denote  $spread(\cdot)$  a function that gives a model spread using CDO pricing model (e.g. LPHGC). This is essentially a function of default correlation.
3. Define a function  $UpfrontFee(\cdot)$  that finds the upfront fee given annual market spread and market convention of 500 bp p.a.

$$UpFee = \frac{\frac{MarketSpread}{4} - \frac{500}{4}}{DF}, \quad (4.2)$$

where  $DF$  presents the sum of effective discount factors for each payment date. Effective discount factor incorporates the discount rate and expected survival rate as the running spread is paid on the effective tranche size, which decreases in time with losses occurring. Omitting discount rate (set it equal to zero) for simplicity (it was proved to have minor influence on the spread value) we define  $DF = \sum_{t=1}^T Esur_t$ . As expected survival depends on correlation,  $DF$  could be also represented as a function of correlation.

Further, denoting  $DF = DF(corr)$  and plugging

$$MarketSpread = spread(corr)$$

in equation (4.2) we obtain

$$UpFee = \frac{\frac{spread(ccorr)}{4} - \frac{500}{4}}{DF(corr)}, \quad (4.3)$$

4. Thus we expressed upfront fee as a function of correlation. Given the quoted upfront fee  $MarketUpFee$  we apply numerical inversion to equation (4.3) and find implied correlation for equity tranche  $corr_{Equity}$  such that

$$UpFee(corr_{Equity}) = MarketUpFee.$$

As a result we get the following implied correlation graph for 6 tranches(see Figure 4.5):

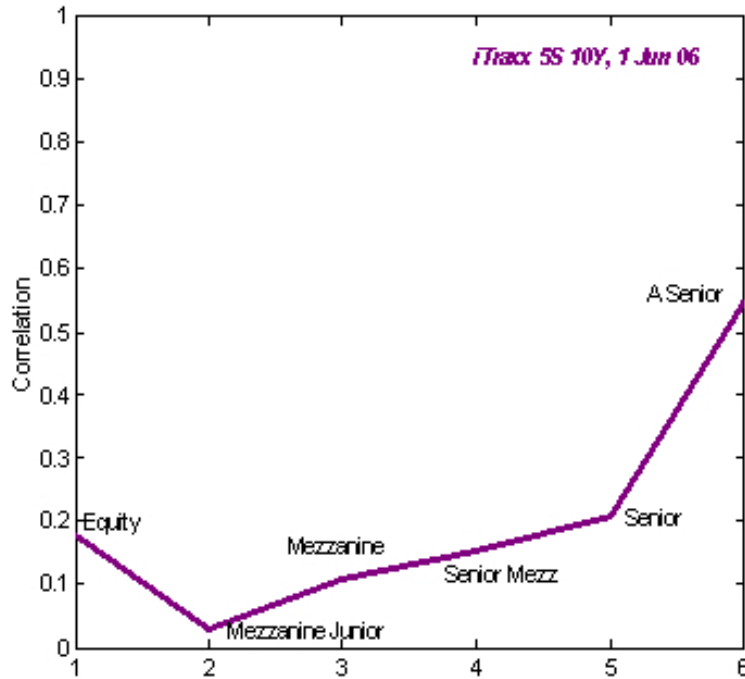


Fig. 4.5: Implied Correlation Smile.

This shape is known as "correlation smile". Though the LPHGC model uses only one single parameter to summarize all correlations among the various borrowers defaults, market tranche spreads imply that different tranches

---

on the same underlying portfolio trade at various correlations. Correlation smile reveals that the mezzanine (junior) tranches typically show a lower compound correlation than equity or senior tranches. For senior and super senior tranches a higher correlation is necessary as compared to the equity tranche. As shown in the section 2.6 typically mean asset correlation varies within the bounds of 0–30%. That means that market data implies that the model underprices the senior tranches - a bigger input correlation is needed for model to compensate increasing risk of default with higher spreads that match market quotes. It means essentially that Gaussian Copula model fails to model fat tails of the portfolio loss distribution, underestimating the chance of observing a very high or very low number of defaults. Other reasons for why we observe a skew in implied correlation are given below.

#### 4.2.3 *Possible Explanation of the Correlation Smile*

The existence of the implied correlation smile (skew) is not quite obvious. Though standard Large Pool Homogeneous Gaussian Copula model is simplifying as it assumes unrealistic uniform default correlation, nevertheless it doesn't overlook any factors which might possibly influence correlation parameter as well as any of specific tranche characteristics. Thus, correlation doesn't depend on tranche priced in the model. Consequently, one would expect an almost flat implied correlation. As there is an analogy between implied correlation and implied volatility, it is also possible to apply similar explanations to the fact that both parameters demonstrate a smile. The existence of volatility skew is usually explained by two lines of argument. The first consideration refers to the general market supply and demand conditions, while the second approach focuses on the difference between the asset's return distribution assumed in the Black-Scholes option pricing model (lognormal distribution) and the one implicit in the market quotes. The correlation smile can be explained analogously. Three possible explanations for this phenomena are presented below. For further details please refer to [1], [20], [21].

- *Demand and supply conditions.* The smile could stem from different preferences of the market (in other words, of the arbitrage seeking investors) in selling protection on some tranches. Implied correlation skew typically denotes a lower correlation for mezzanine tranches as compared to junior tranches. The strong demand for mezzanine and senior tranches leads to the situation when spreads on respective tranches decrease relative to those of equity tranches unrepresentative of the

underlying risk. On the other hand the demand for equity tranches is usually lower than for other notes, because it represents unrated securities which are sold in the first place as financial institutions free up the regulatory capital under the new Basell II capital requirement rules.

- *Segmentation among investors.* Different investors hold varying views about default correlations. The views of sellers of protection on equity tranches may differ from those of sellers of protection on mezzanine tranches. However, it is not clear why this difference would be systematic.
- *Model weakness.* The main assumptions of the standard market model used to imply correlation are too simplistic. The implicit portfolio loss distribution has fatter tails than the Gaussian copula, recovery rates are not fixed but can be stochastic, moreover, the recovery rate and the default variable are not independent as it is assumed.

#### 4.2.4 Existence, uniqueness, and monotonicity of compound correlation

A natural question to ask next is: can we always find compound correlation and if such a correlation exists, is it unique? Inverting pricing formula under LPHGC model for spreads of tranching on different dates, series, and maturities shows that for a mezzanine tranche the existence of compound correlation is not always the case. Figure 4.6 demonstrates such situation. The function  $f(x)$  on y axis is the following:  $f(corr) = ModelSpread(corr) - MarketSpread$ .

The function  $f(corr)$  doesn't cross the zero line that means that any possible correlation input in the model produces a lower spread than quoted in the market, thus underpricing the mezzanine tranche. To investigate this problem, we look at parameters in the model other than correlation that influence the fair spread value. The other two parameters are default probability and discount rate. The spread proved to be insensitive to large changes in discount rate, but it is strongly dependent on default probability, which is in turn driven by the market average spread of the CDS index.

The figure 4.7 reveals  $f(x)$  as a function of both correlation and CDS spread. The read plane represents the surface, on which  $f(\cdot, \cdot) = 0$ , thus equalizing market and model spread for mezzanine tranche. The blue surface shows the true value of  $f(\cdot, \cdot)$  for most possible values of correlation and CDS average spread. Consequently, the border between red and blue surfaces on the graph contains those combinations of correlation and CDS spread values, for which

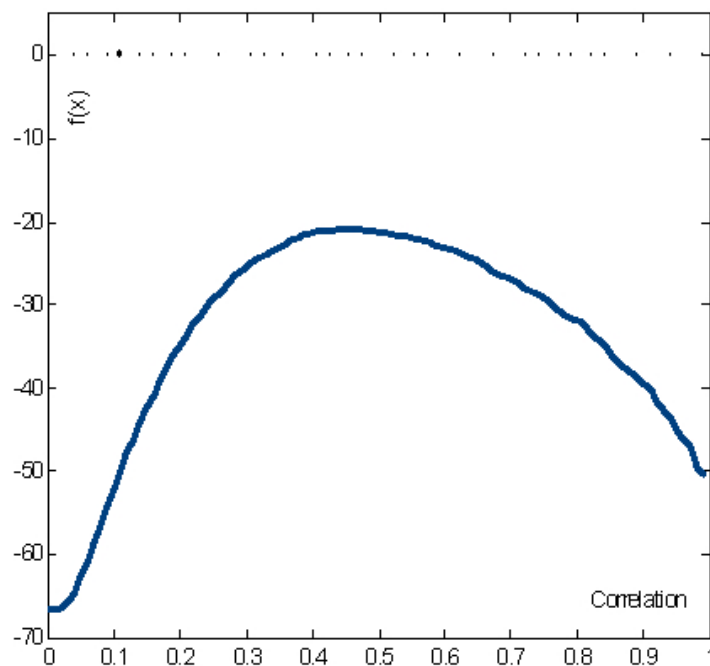


Fig. 4.6: Non-existence of implied correlation for Mezzanine tranche; iTraxx 7S 7Y, 3 Jul 07.

it market and models spreads are equal. It is clear now, that it is not possible to imply correlation for any values of CDS spread except for those that constitute the border.

The second question concerns uniqueness of compound correlation. Figure 4.8 demonstrates the case when there are two distinct correlations that give the same par spread. Such a lack of uniqueness is typical for mezzanine tranche and this property makes an implied correlation concept less applicable as it is difficult to interpret correlation parameter and even more difficult to use in CDO hedging. The non-uniqueness of implied correlation is the direct corollary of the fact that "middle" tranches spreads are not monotonic in correlation (see Figure 4.3).

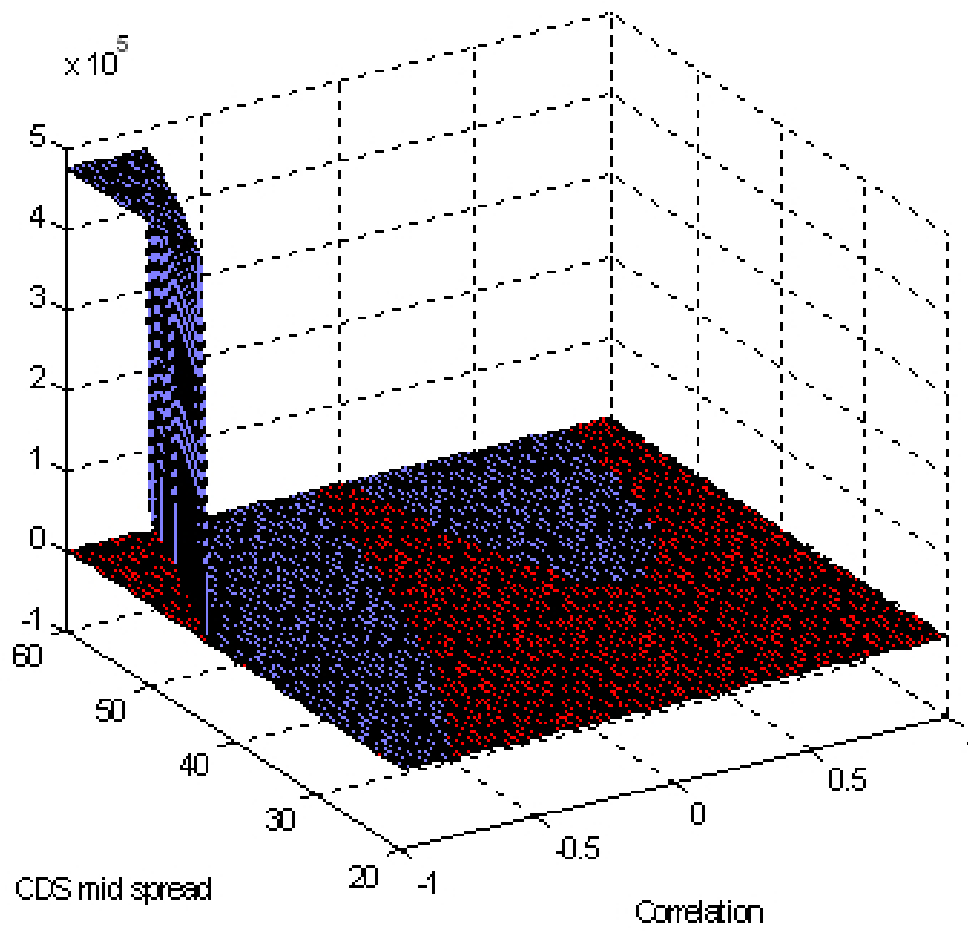


Fig. 4.7:  $function(corr, CDS_{spread})$  - difference between market and model spread for mezzanine tranche, iTraxx 7S 7Y, 3 Jul 07.

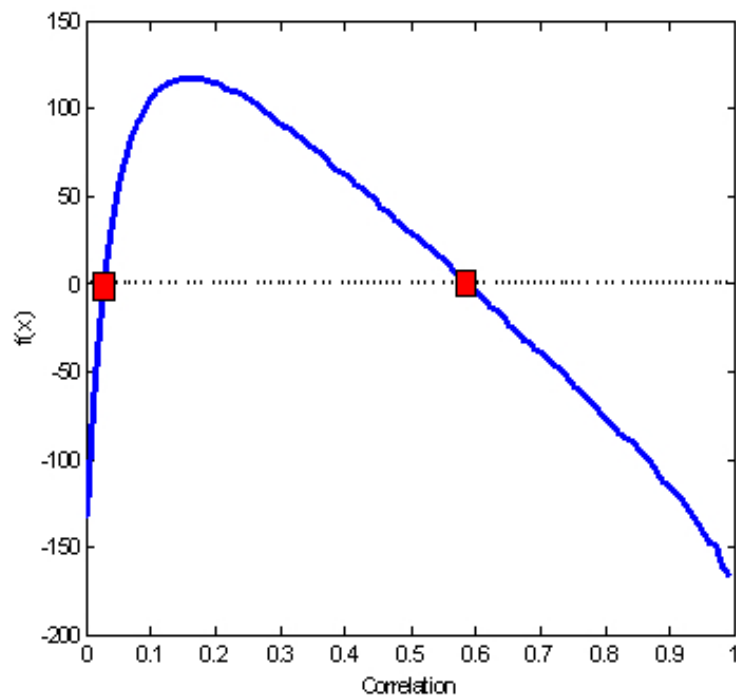


Fig. 4.8: Non-uniqueness of implied correlation for mezzanine tranche, iTraxx 5S 10Y, 1 Jun 06.

## 5. BASE CORRELATION

The concept of base correlations was proposed in 2004 by McGinty, Beinstein, Ahluwalia, and Watts([20]) from JP Morgan and it was designed to overcome the limitations of the compound correlation. For references please see also [?], [22], [27], [15], [23]. The simple idea behind this concept is following: instead of implying correlation on regular (so called "interest") tranches, we define virtual tranches, which are often called "first loss tranches", as they all have the same lower attachment point of 0% and are therefore similar in that to standard equity tranche. A  $j$ -tranche is defined by  $0 - K_j\%$ , where a vector of upper attachment points  $K = (3\%, 6\%, 9\%, 12\%, 22\%, 100\%)$ . Further, we apply some CDO pricing model (e. g. LPHGC model) for pricing these fictive constructed tranches and use numerical inversion to find implied correlation, termed base correlation as opposed to compound. It is straightforward to infer the expected tranche loss for these virtual first loss pieces once we have calculated the expected tranche loss for regular tranches. A loss on  $j$  tranche is decomposed into losses on two "equity" tranches. From this it follows:

$$E(L[0, K_j]) = E(L[0, K_{j-1}]) + E(L[K_{j-1}, K_j]) \quad (5.1)$$

The first virtual tranche therefore coincides with equity tranche ( $0 - 3\%$ ), and the following pieces ( $[0 - 6\%]$ ,  $[0 - 9\%]$ ,  $\dots$ ,  $[0 - 100\%]$ ) are bootstrapped using formula (5.1).

Since premium spread quotes are for regular tranches only, the implied base correlation of a tranche cannot be obtained directly. To calculate the base correlation a recursive method is used.

- The base correlation for the first equity tranche ( $0 - 3\%$ ) is simply the implied compound correlation for this piece.
- Given the unique correlation for the equity piece we fix the expected loss in this tranche  $E(L[0, 3\%])$ .
- Given the market spread for the next tranche ( $3 - 6\%$ ) we iterate over the correlation parameter  $\omega$ , which generates an expected loss of fictive

tranche (0 – 6%) such that the expected loss of the (3 – 6%) tranche via equation 5.1 implies the given spread.

- This correlation is denoted base correlation for 6% detachment point, and the iteration continues, extracting base correlations for all detachment points.

Formally, the procedure is executed as follows:

$$ETL_{3-6\%} = f(corr),$$

where  $ETL$  is expected tranche loss.

$$f(corr) + ETL_{0-3\%}(corr_B^1) = ETL_{0-6\%}(corr),$$

where  $corr_B^1$  is base correlation for 0 – 3% tranche;

$$ModelSpread_{3-6\%} = g(ETL_{3-6\%}) = g(f(corr))$$

$$ModelSpread_{3-6\%} = g(ETL_{0-6\%}(corr) - ETL_{0-3\%}(corr_B^1)) := h(corr) \quad (5.2)$$

Define a function  $F(corr)$ :

$$F(corr) = MarketSpread_{3-6\%} - ModelSpread_{3-6\%} \quad (5.3)$$

Plug equation 5.2 in 5.3:

$$F = MarketSpread_{3-6\%} - h(corr) \quad (5.4)$$

Find the zero of function  $F$ , solving equation 5.4 for  $corr$  by numerical inversion, then denote the solution  $corr = corr_B^2$ . Proceed with iterations for next tranches. The base correlation graph is presented on figure 5.1.

### 5.1 Advantages of Base Correlation

The base correlation approach seeks to exploit the monotonicity of equity tranches in order to solve the problem of non-uniqueness of implied correlations for mezzanine tranches. No matter how big the detachment point of the

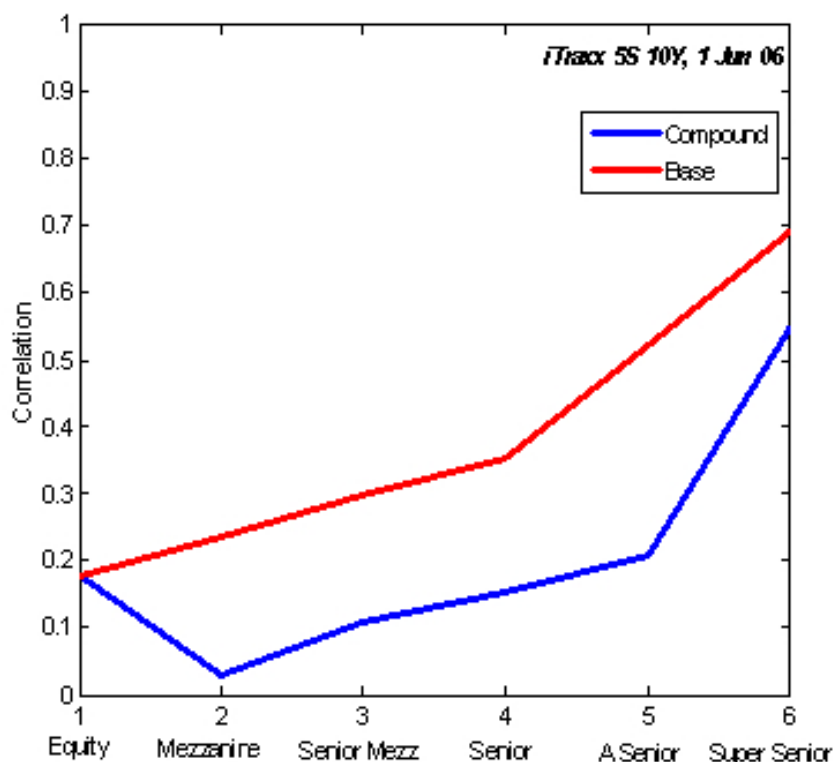


Fig. 5.1: Compound vs Base Correlation.

first loss tranche is - 3% or 100%, it is equally susceptible to any losses occurring in the reference portfolio. That is why the virtual equity tranches are monotonic in correlation. This is shown on Figure 5.2. Figures 5.2 and 4.2 reveal also one more difference between regular and fictive tranches - for the latter expected tranche loss is always decreasing with correlation, implying that fair spread is also negatively dependent on correlation for all tranches.

Though the base correlation values can seem to be very high compared with compound correlation, the base skew on figure 5.1 is in fact meaningful that the skew of compound correlation as it reflects the difference between real world pricing and the LPHGC model assumptions. The reason for this skew is that one-factor models place very low probabilities on losses for very senior tranches (e. g. 22–100% tranche). Therefore the fair spread for these tranches inferred from the model is zero basis points. In reality the market participants won't take any risk for nothing and thus the market charges a few basis points. This skews the implied correlation substantially for the remainder of the capital structure.

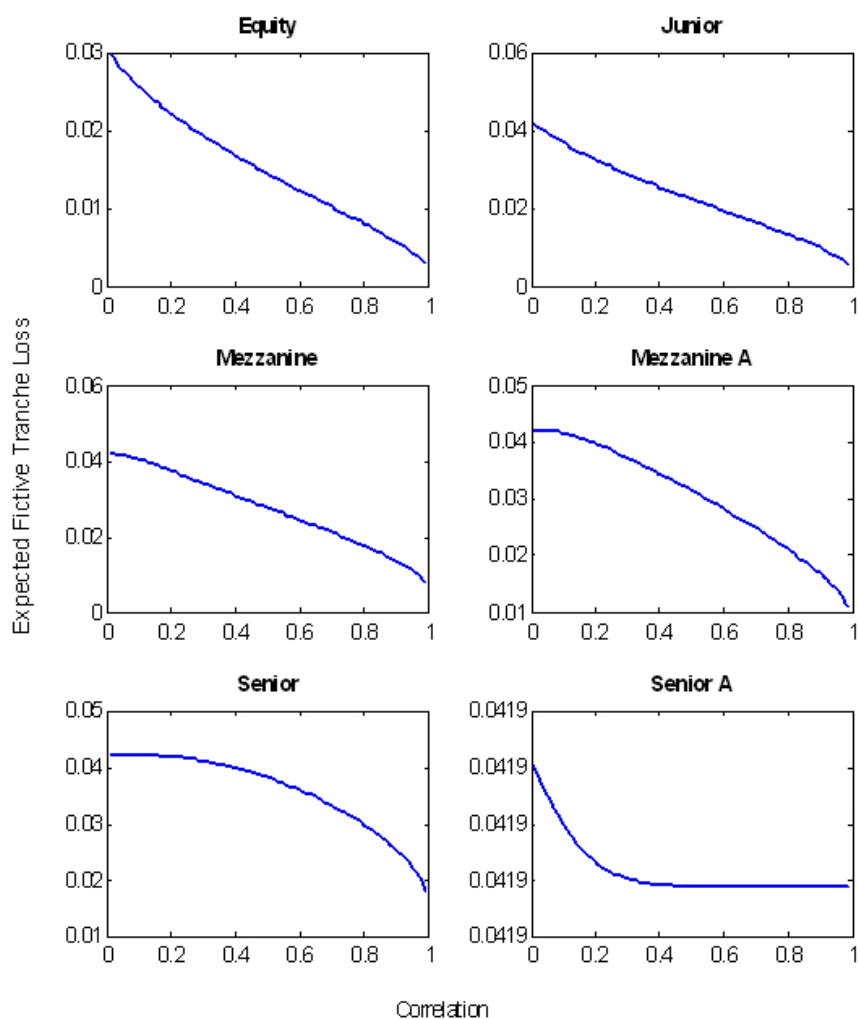


Fig. 5.2: Expected Fictive Tranche Loss as a function of default correlation.

The next thing that belongs to the advantages of base correlation is that it can be used for pricing off-market CDO tranches with the same collateral pool. Any kind of interpolation process from the compound correlation smile is compromised by the fact that these compound correlations are function of two base correlations. Using the base correlation framework, we can use the standard market tranches to calibrate the model for base correlation inputs and then interpolate from these base correlations to value a CDO tranche that is not actively traded. Base correlations are thus used to produce consistent implied correlations for non-standardised tranches, which greatly increases the risk-return horizon for investors trying to check the quoted prices of tranches with non-standard attachment points.

## 5.2 Pitfalls of Base Correlation

Willeman [27] examined fair tranche spreads, base correlations and their dependence on the assumed default correlation. It was shown that even if the intensity correlation increases, base correlations for senior tranches may actually decrease. Also it was proven that base correlations at a given attachment point depend on the placement of all prior attachment points.

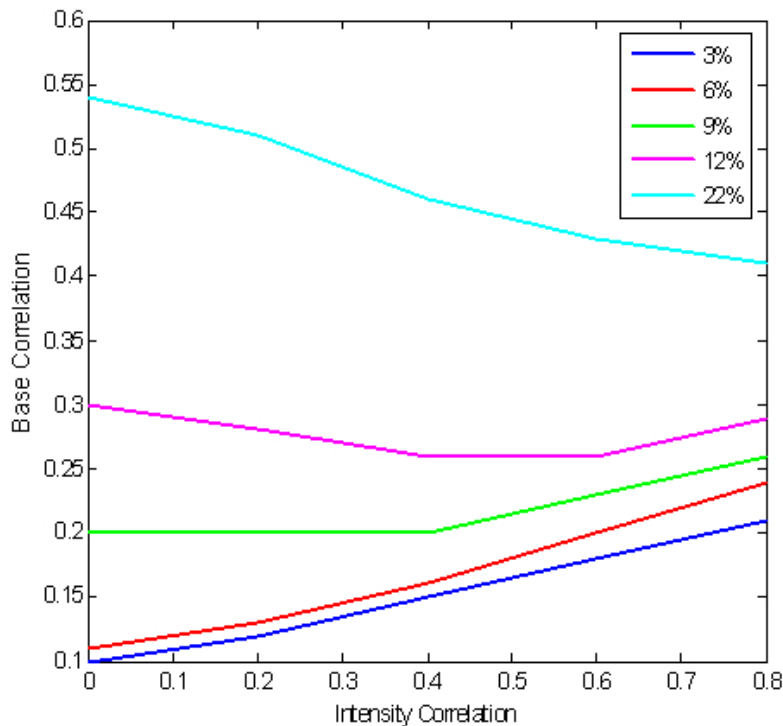


Fig. 5.3: Bootstrapped Base Correlations vs "True" Default Intensities.

Figure 5.3 shows base correlations as a function of assumed "true default correlation" of first loss pieces with different detachment points. For a given correlation, the tranche spreads decrease with seniority and for a given tranche there is a monotonic relationship between fair tranche spreads and correlation (see Figure 4.3). Turning to the base correlations, we see that for a fixed intensity correlation, there is the expected relationship that the base correlation is increasing with seniority as observed in the market. But for a given fixed detachment point, we see that, for more junior tranches, the base correlation goes up when the intensity correlation goes up, as expected. However, for the more senior tranches the base correlation of the senior tranches

---

may decrease with increasing default correlation. This means that even if spreads change as to reflect an increased correlation, the base correlations seem to imply that the correlation actually has gone down. To conclude, one should be careful when making judgements on true correlation changes by the changes in implied base correlation. For more details on this topic please refer to [23], [25].

## 6. CONCLUSION

Over the past years the CDO market exhibited rapid growth accompanied with increasing liquidity and market improvement in transparency of pricing framework. Publicly quoted CDS indices such as iTraxx offer an opportunity to calibrate the correlation parameter in pricing models. Correlation is now market observable and becomes itself a subject for arbitrage trading placing further importance on its correct calibration and valuation of tranching products. In this work we applied market standard model with analytical loss distribution approximation and assumption of flat correlation structure to find implied correlation. We have analyzed the implications of the selection of parameters and examined the skew observed in implied correlation over various tranches.

The approach presented here is by no means comprehensive. There are many more models developed for pricing CDO and other correlation-dependent products. Among them are models that use Monte Carlo simulation instead of analytic approximation to find portfolio loss distribution, models with use more realistic assumptions (fatter tailed distributions, stochastic recovery rate etc.), as well as models with dynamic intensity process for default times. These models allow for modeling correlation smile and claim to produce more consistent results though at cost of speed.

In this work we have presented proofs from recent data that quoting CDO tranches in terms of implied correlation — either compound, or base — does suffer in both cases from inconsistency, making it difficult to draw conclusions about relative attractiveness of two tranches as well as using implied correlation for pricing non-standard tranches.

## 7. APPENDIX A

### 7.1 How to extract the term structure of default probabilities?

The rating approach we used to calibrate default probabilities in chapter 2 is rarely applied in practice. Market based approach reflects the market-agreed perception about the evolution of the market in future, while historical information reflects past development, which is not necessarily stable over long periods in time. Below it is shown how to obtain a so called credit curve — default probability as a function of maturity.

Let  $\tau_i$  be a continuous random variable that measures the length of time for security  $i$  until default occurs. In this framework, the current time could be assumed to be the starting point. The time scale is defined in terms of years for continuous models or number of periods for discrete version. The meaning of default could be defined as rating D according to Moody's system. Let  $F(t) = P(\tau_i \leq t)$  be the distribution function of time-to-default. Assume  $F(0) = 0$ . We can also define the probability density function:

$$f(t) = F'(t) = \lim_{\Delta \rightarrow 0^+} \left( \frac{P(t \leq \tau_i \leq t + \Delta)}{\Delta} \right)$$

Denote  $p_x(t)$  the conditional probability that the security  $i$  will default within the next  $t$  years conditional on its survival in  $x$  years:

$$p_x(t) = P(\tau_i - x \leq t | \tau_i > x), \quad t \geq 0.$$

Plugging  $t = 1$  we obtain marginal default probability, conditional on its survival to the beginning of the period:

$$p_x = P(\tau_i - x \leq 1 | \tau_i > x).$$

A credit curve in a discrete world therefore can be expressed as a sequence of  $p_0, p_1, \dots, p_n$ .

Now we employ the concept of hazard rate function to get a convenient representation for marginal default probabilities. Hazard rate function gives the instantaneous default probability for a security that has attained age  $x$ . For notational simplicity we omit further the subscript  $i$ .

$$P(x < \tau \leq x + \Delta x | \tau > x) = \frac{F(x + \Delta x) - F(x)}{1 - F(x)} \approx \frac{f(x)\Delta x}{1 - F(x)} \quad (7.1)$$

The hazard rate function is given then

$$h(x) = \frac{f(x)}{1 - F(x)}.$$

We can now express distribution and probability density functions in terms of  $h(x)$ :

$$F(x) = 1 - e^{-\int_0^x h(s) ds} \quad (7.2)$$

$$f(x) = h(x) \cdot e^{-\int_0^x h(s) ds}.$$

Understanding the first arrival time  $\tau$  as associated with a Poisson arrival process, the constant mean arrival rate  $h$  can be interpreted as default intensity. Changing from a deterministically varying intensity to stochastic intensity we obtain the following expression for conditional default probability

$$p_x(t) = 1 - E_s \left[ e^{-\int_0^t h(x+s) ds} \right],$$

where  $E_s$  denotes expectation given all information available at time  $s$ . Nevertheless a typical assumption here is that the hazard rate is constant. Also we need to put constraints of positiveness on  $h(x)$ , so that the probability  $p_x(t)$  is less than 1. Given that we obtain

$$f(t) = he^{-ht}, \quad t > 0, h > 0 \quad (7.3)$$

which shows that the density function follows the exponential distribution with parameter  $h$ . Under this assumption the default probability over the time interval  $[0, t]$  is

$$p_x(t) = 1 - e^{-\int_0^t h(x+s) ds} = 1 - e^{-ht} = (p_x)^t.$$

Now we are interested in how to estimate the hazard rate  $h$ . If  $h$  is continuous then  $h(t)\Delta t$  approximately equals the probability of default between  $t$  and  $\Delta t$  conditional on survival to  $t$ . As far as by construction market-based

---

probabilities are risk-neutral, we can calculate an approximation for  $h(t)\Delta t$  in the following way. The link between spreads of traded credit derivatives and default probabilities is analogous to the link between interest rates and discount factors in fixed income markets. Thus if  $spr$  represents a spread over the risk-free rate, then we get an expression for risk-neutral default probability:

$$DP^* = \frac{1 - \frac{1}{1+spr}}{1-R} \approx \frac{spr}{1-R},$$

where  $R$  is assumed recovery rate.

Then taking in account time period, we obtain

$$p_t(x) = 1 - e^{-(\frac{spr}{1-R})t}, \quad (7.4)$$

where we use so called clean spread ( $\frac{spr}{1-R}$ ) for credit default swap of respective maturity.

## 8. APPENDIX B

### 8.1 Proof of the Proposition 2.2

*Proof.* For fixed  $Y = y \in \mathbb{R}$  define the conditional probability measure  $P_y$

$$P_y(\cdot) = P(\cdot | Y = y)$$

Consider the random variable

$$X_k = EAD_k(LGD_i L_k - E(LGD_i L_k | Y))$$

With respect to  $P_y$ , the random sequence  $(X_k)_{k \geq 1}$  is independent due to 2.19 and centered by definition. Let us define

$$\eta_m = \sum_{i=1}^m EAD_i,$$

such that  $(\eta_m)_{m \geq 1}$  is a positive sequence strictly increasing to the infinity due to assumptions 2.21, 2.22. If we could prove that

$$\sum_{k=1}^{\infty} \frac{1}{(\eta_k)^2} E(X_k^2) < \infty, \quad (8.1)$$

then a version of the strong law of large numbers would yield

$$\lim_{m \rightarrow \infty} \frac{1}{\eta_m} \sum_{k=1}^m X_k = 0 \quad P_y - \text{almost surely.} \quad (8.2)$$

This version of the law of the large numbers is based on Kronecker's Lemma, stating that whenever  $(x_k)_{k \geq 1}$  and  $(\eta_k)_{k \geq 1}$  are sequences with the latter being positive and strictly increasing to infinity, such that

$$\sum_{k=1}^{\infty} \frac{x_k}{\eta_k} \quad \text{converges,}$$

we obtain:

$$\lim_{m \rightarrow \infty} \frac{1}{\eta_m} \sum_{k=1}^m x_k = 0.$$

The next step is to prove the statement 8.1. From assumptions 2.21, 2.22 we get

$$\sum_{k=1}^{\infty} \frac{1}{(\eta_m)^2} \mathbb{E}(X_k^2) \leq \sum_{k=1}^{\infty} \frac{4 \cdot EAD_k^2}{\eta_k^2} < \infty \quad (8.3)$$

due to the uniform boundedness of  $(LGD_k L_k - \mathbb{E}(LGD_k L_k | Y))$ . Thus we have proved 8.2 for every  $y \in \mathbb{R}$ .

We can write now

$$\mathbb{P} \left[ \lim_{m \rightarrow \infty} (L^{(m)} - \mathbb{E}(L^{(m)} | Y)) = 0 | Y = y \right] = 1 \quad (8.4)$$

for every  $y \in \mathbb{R}$ . Then almost sure convergence also holds unconditionally and we obtain

$$\begin{aligned} & \mathbb{P} \left[ \lim_{m \rightarrow \infty} (L^{(m)} - \mathbb{E}(L^{(m)} | Y)) = 0 \right] = \\ & = \int \mathbb{P} \left[ \lim_{m \rightarrow \infty} (L^{(m)} - \mathbb{E}(L^{(m)} | Y)) = 0 | Y = y \right] dP_Y(y) = 1. \end{aligned}$$

The proposition is proved.

## BIBLIOGRAPHY

- [1] Amato, J.D., Gyntelberg, J. (2005) CDS index tranches and the pricing of credit risk correlations. *Bank for International Settlements Quarterly Review*, March
  
- [2] Bluhm, C., Overbeck, L. and Wagner, C. (2003). *An Introduction to Credit Risk Modeling*, Chapman Hall/CRC Financial Mathematics Series.
  
- [3] Bluhm, C. (2003) CDO Modeling: Techniques, Examples, and Applications, *Working Paper*, HypoVereinsbank, Group Credit Portfolio Management, 10 December.
  
- [4] Duffie, D., Gârleanu, N. (2001). Risk and Valuation of Collateralized Debt Obligations, *Working Paper*, Graduate School of Business, Stanford University.
  
- [5] Embrechts, P., McNeil, A. and Straumann, D. (1999). *Correlation and Dependence in Risk management: Properties and Pitfalls. Correlation, Risk Management: Value at Risk and Beyond*, Cambridge University Press, Cambridge.
  
- [6] Finger, C. (2000). A Comparison of Stochastic Default Rate Models, *RiskMetrics Journal*, Volume 1.
  
- [7] Frey, R., McNeil, A.J., Nyfeler, M.A. (2001). Copulas and Credit Models. *Working Paper*,

- 
- [8] Frey, R., McNeil, A. (2001). Modeling Dependent Defaults, *Preprint, ETH Zürich*.
- [9] Friend, A., Rogge, E. (2005). Correlation at first sight, *Economic Notes* 34, No. 2, 155-183.
- [10] Jackson, K., Kreinin, A., Ma, x. (2007) Loss Distribution Evaluation for Synthetic CDOs, *Working Paper, Natural Sciences and Engineering Research Council*, Department of Computer Science, University of Toronto.
- [11] Gordy, M. (2000). A Comparative Anatomy of Credit Risk Models, *Journal of Banking and Finance*, 24 January, 119-149.
- [12] Greenberg, A., O’Kane, D., Schloegl, L. (2004). LH+: A Fast Analytical Model for CDO: Hedging and Risk Management, *Quantitative Credit Research Quarterly, Lehman Brothers*.
- [13] Gupton, G.M., Finger, C.C., Bhatia, M. (1997). CreditMetrics - Technical Document, *Risk Management Research, Morgan Guaranty Trust Company*.
- [14] Hull, J., White, A. (2004). Valuation of a CDO and an n-th to Default CDS without Monte Carlo Simulation. *Journals of Derivatives*.
- [15] Hager, S., Schöbel, R. (2006). A Note on the Correlation Smile, *Working Paper*, Eberhard-Karls-University of Tübingen.
- [16] Franke, J., Hädler, W. and Hafner, C. (2004). *Statistics of Financial Markets*, Springer-Verlag, Heidelberg.
- [17] Lehnert, N., Altrock, F., Rachev et. al. (2005). Implied Correlation in CDO Tranches, *Working Paper*.

- 
- [18] Li, D. (1998) Constructing a credit curve, *Credit Risk, a Risk Special Report*, November, 40-44.
- [19] Li, D. (2000) On Default Correlation: A Copula function approach, *Journal of Fixed Income*, 9, March, 43-54.
- [20] McGinty, L., Beinstein, E., Ahluwalia, R., Watts, M. (2004) Introducing Base Correlations, *Credit Derivatives Strategy JP Morgan*.
- [21] McGinty, Ahluwalia, R. (2004a) Credit Correlation: A Guide. *Credit Derivatives Strategy JP Morgan*.
- [22] McGinty, Ahluwalia, R. (2004b). A Model for Base Correlation Calculation, *Credit Derivatives Strategy JP Morgan*.
- [23] Parcell, E., Wood, J. (2007) Wiping the smile off your base (correlation curve), *Derivative Fitch*, 21 June.
- [24] Schönbucher, P.J., Rogge, E. (2003) Modelling dynamic portfolio credit risk, *Working Paper*.
- [25] Torresetti, R., Brigo, D., Pallavicini, A. (2006). Implied Correlation in CDO tranches: a Paradigm to be handled with care. *Credit Models, Banca IMI, Milan*.
- [26] Nelsen, R. (1998) *An Introduction to Copulas*, Springer-Verlag, New York.
- [27] Willeman, S. (2004). An Evaluation of the Base Correlation framework for synthetic CDOs, *Working Paper*, Aarhus School of Business.