Valuation of Synthetic and Cash CDO in Practice
and related issues

Gerhard Schweimayer*

Working Paper

Last Version March 2008
This Version January 2010

Abstract

This paper explains some of the existing approaches to value CDO and NtD instruments. This is done to such an extend that enables the practitioner with some mathematical background to implement a valuation model quickly. Thereby, not only unfunded CDOs but also Cash CDOs are considered. There is no need to consult the huge variety of literature on the pricing of unfunded CDOs, as the underlying concepts are sufficiently explained. While the valuation of unfunded CDOs is covered well in the literature there is to the knowledge of the author not much to find on how to approximate the sophisticated waterfall structures of Cash CDOs analytically to a sufficient extend. As concerns how the cash flows out of the underlying pool pour out into the waterfall structure there is also only a few sources that treat this issue. Within this article the focus is not only on how to circumvent the use of term structure models, but also the use of a term structure model to estimate future cash flows of a given pool is outlined. In addition, the paper investigates the role of pricing formulas, credit spreads and recovery rates in bootstrapping of default probabilities. Thereby, the implied recursive formula to calculate the risk neutral default probabilities is discussed, as marginal default probabilities are the basic building block of the needed joint distributions of default times. Moreover, the relationship between hazard rates and credit spreads is quantified. It turns out that pricing assumptions can have substantial impact on the resulting default probabilities especially when recovery rates and credit spreads are high. It is also shown how pricing assumptions for CDS influence the replication of a CDS with bonds.

Key Words: Simulation, Factor Model, Structural Model, Loss Distribution, Base Correlation, Waterfall, Hazard Rate, Bootstrapping, Black Derman Toy

*Dr. Gerhard Schweimayer, gerhard@schweimayer.de.

Contents

1 Introduction

This paper was written to provide a practical guide through the various valuation approaches for unfunded CDOs and NtD and to give a glue of how the involved risk parameters influence the valuation. In addition the issue of how to apply these concepts to Cash CDOs is tackled. The paper intends to enable the practitioner with some mathematical background to implement a valuation model quickly.

A *NtD* is a n-th to default swap which has two legs, the default leg that pays the protection buyer the loss given the n-th default within a basket of reference obligations and the premium leg that pays the protection seller a quarterly premium the so called credit spread. In case of a default between two payment dates also the accrued premium is considered, and the protection seller is obliged to pay the loss given default on the defaulted member of the basket. A *Cash CDO* is a collaterlized debt obligation that usually has a basket of bonds (CBO) or loans (CLO) as its reference, it is distinguished between static and managed baskets of collateral. A *Synthetic CDO* instead is not based on physically owned bonds or loans but rather on a portfolio of single name CDS, possibly standardized by the iTRAXX index of European corporates. While cash CDO are usually *funded*, that is, the protection seller buys a bond for initially paying the nominal that is reduced by future defaults, synthetic CDO are usually *unfunded*, which means that there is no initial payment because portfolio losses are directly compensated in case of default. A *Single Tranche CDO* is unfunded and restricts the possible losses to the intervall $[a\%, b\%]$ of the total nominal. The percentage premium is always paid on the outstanding nominal that is reduced by past defaults that exceed the lower bound of the tranche. Single tranche CDOs are usually synthetic and unfunded on the liquid iTraxx or CDx indices that represent European, North American baskets of reference entities respectively.[1] From the protection buyers point of view (the issuer) this loss due to defaults that exceeds the lower bound of the interval is what he receives. This payment increases with further losses until the upper bound is reached. For a depiction of a CDO be referred to figure 4 in the appendix.

In order to arrive at the set out goal section 2 explains the theory of correlated default times, and then the evaluation of the joint default distribution as well as the role of expected loss and further loss related distributions is outlined by section 3. This section concludes with a short explanation of correlation trading. Section 4 shows how these distributions are applied to price NtD and CDO using the flexible approach of Monte-Carlo simulation, followed by the (semi-)analytical factor approach for these products and a brief outline of the correlation crisis in summer 2005 and implied correlation. This section concludes briefly with the calculation and

[1]The DJ iTraxx Europe covers the 125 most liquid European investment grade CDS issues. All members are equally weighted. It is thus an CDS index. The CDx is similar but for North American CDS.

importance of sensitivities. Finally, 5 details the bootstrapping of marginal default probabilities and related issues.

2 Correlated Default-Times

When valuing a CDO, as for other asset backed securities, it is crucial to model the default correlation properly.[2] The valuation is based on the independent marginal default-time-distributions of each member of the basket, that have to be coupled properly. Below we want to give an accessible introduction into the general theory of how to obtain a correlated default model of these single distributions, as well as the important special case of the factor model. Finally, trading practice for correlation is briefly explained.

2.1 General Theory

Let us assume a basket with M members and respective random default times $T_1, ..., T_M$, with given distributions $F_1(t_1), ..., F_M(t_M)$, e.g. $F_m(t) = 1 - e^{-\lambda_m t}$ in a reduced form model with constant hazard rate λ_m. How to derive these distributions from market data is explained in section 5.[3]

From Sklar's Copula-Theorem (cf. Li (2000), Nelson (1999)) we know that given a copula $C(*)$ we receive the respective joint distribution of default times

$$F(t_1, ..., t_M) = C(F_1(t_1), ..., F_M(t_M)), \tag{1}$$

whereby the copula determines the correlation structure and distributional behavior of $F(*)$. In other words, the copula tells us how to couple the marginal distributions. In practice the Gaussian Copula

$$F(t_1, ..., t_M) = \Phi_\Sigma(\phi^{-1}(F_1(t_1)), ..., \phi^{-1}(F_M(t_M))) \tag{2}$$

is often used. Where ϕ is the standard normal distribution function and Φ_Σ the multivariate extension with correlation matrix Σ.

In order to obtain dependent default times and inspired by formula (2), whose right hand side refers to standard normal variables, we define dependent standard normal variables $Y_m \sim N(0,1)$ and set

$$\begin{aligned} Y_m = \phi^{-1}(F_m(T_m)) \qquad & \forall m \in \{1, ..., M\} \\ \Longleftrightarrow T_m = & F_m^{-1}(\phi(Y_m)). \end{aligned} \tag{3}$$

[2] Default Correlation is defined via the correlation between stochastic default (stopping) times, cf. Li (2000).

[3] More precisely, credit spread market data for liquid single name CDS.

Which yields the dependent Default Times T_m. Thereby, asset-return correlations are transformed into default correlations, more precisely, into correlations of default times, if we interpret the Y_m in terms of the structural Merton model. In general, default correlations are less than asset or equity correlations.

In practice $F_m(t)$ is rarely assumed a homogeneous poisson process but an inhomogeneous generalization that is able to match the term structure of default rates better. If we derived, given n CDS spreads, respective probabilities $F_{m,k}$ with $k \in \{1, ..., n\}$ then an inhomogeneous process with hazard rates $(\lambda_{m,1}, ..., \lambda_{m,n})$ can be calibrated iteratively with n equations

$$F_m(t_k) = 1 - exp\left(-\sum_{i=1}^{k} \lambda_{m,i} \cdot (t_i - t_{i-1})\right) = F_{m,k} \quad \forall k. \tag{4}$$

This is done via bootstrapping based on market CDS spreads of different terms of the reference entity m.

Dependent variables can either be obtained by performing the conditional probability approach as outlined in Nelson (1999) that can be used for any well defined copula, or the Cholesky decomposition that is specific to multidimensional normal distributions. Using the Cholesky decomposition we need to know which pairwise correlations make up the correlation matrix Σ. Li (2000) shows that applying the Gaussian copula is in fact equivalent to the Credit Metrics approach that uses the correlations of equity returns.[4] Credit Metrics slices a standard normal distribution such that its default probability equals the historically or market given default probability. Thus the above Y_m can be viewed as log-asset-returns that have been transformed to obey the standard normal distribution, as they are used by Credit Metrics.

In order to engage the Cholesky decomposition we search for an upper triangular matrix Q such that

$$Q^T \cdot Q = \Sigma \tag{5}$$

holds, and transform independent standard normal variables Z_m into the dependent Y_m via

$$Y = Q^T \cdot Z \in \mathbb{R}^M. \tag{6}$$

2.2 Factor Approach

Factor approaches avoid the Cholesky decomposition by calculating two independent standard normal variables V, and Z_m for each Y_m and set

$$Y_m = \rho_m \cdot V + \sqrt{1 - \rho_m^2} \cdot Z_m, \tag{7}$$

[4]Cf. Li (2000), pp. 16-17.

which could be used within simulation via (3) implying a restricted correlation structure. The correlation of the Y_m with their factor V in common determines the correlation structure, as $\rho_m = Corr(Y_m, V)$ implying $Corr(Y_i, Y_j) = \rho_i \cdot \rho_j$. It is easily verified that $Y_m \sim N(0, 1)$ and

$$(Y_m \mid V) \sim N(\rho_m V, 1 - \rho_m^2) \quad (8)$$

holds.[5]

The factor approach not only reduces the specification of the correlation matrix from $M(M-1)/2$ many correlation entries to the specification of $\rho_1, ..., \rho_M$ within the Monte-Carlo simulation but its main benefit is to enable one to find semi-explicit or even analytical formulas for pricing, because the $(Y_m \mid V)$ are independent. This will be shown in section 3.3.

2.3 Trading Practice

Credit default correlation can be liquidly traded in the market via a long/short position in a NtD or single tranche CDO while spread-delta hedging the respective basket with single name CDS contracts, a basket of CDS respectively. A liquidly traded example of these is the *Linear iTraxx*, that represents a portfolio of single name European CDS. The single name CDS are priced independently and thus bear no correlation but only spread risk, i.e. the present value of a linear iTraxx is simply the sum of present values of CDS in the basket, whereas the NtD or CDO bear both correlation and spread risk. Hence, by investing in a CDO on the basket (selling protection) and selling the spread risk by buying linear iTraxx protection one is left with a long position in correlation risk. A short position on the other hand can be obtained by doing the investment the other way round.

Based on model assumptions, market participants can perform a different kind of correlation trade where they are long the equity tranche of a synthetic unfunded CDO and short the mezzanine tranche. As all trances are delta positive in their present value to spread increase of the underlying pool an initial delta neutral position can be achieved even though the mezzanine tranche has lower delta than the equity tranche.[6] As the mezzanine tranche is insensitive to changes in correlation, that is to say, its correlation delta is very low, but the equity tranche exhibits a positive correlation delta, the investor takes a long position in correlation and is compensated for that risk by what the excess spread exceeds the mezzanine premium payments.[7]

[5]Note, the correlation here is the correlation of normalized log-asset returns as explained in detail by section 3.2.

[6]cf. Mahadevan et al. (2006), p. 67-68. In contrast to the trances the assumption of equal default probability over time of the pool-assets implies equal spread deltas for these underlying assets.

[7]cf. Mahadevan et al. (2006) p. 78.

3 Evaluation of needed Distributions

3.1 Monte-Carlo-Simulation

We will not treat the evaluation of distributions via simulation in depth here, because the issue is naturally incorporated in section 4. But it is very intuitive to obtain the distribution of the number of defaults up to time t via simulation. We denote it by $P(N(t) = k)$ with $N(t) = \Sigma_{m=1}^{M} \mathbf{1}_{(T_m<t)}$.

3.2 Structural Model

As it is not clear from scratch why one is able to use normalized log-asset returns to determine the probability of default for a given time horizon $[0, t]$ the structural model is briefly explained, before we address the derivation of the needed distributions under the factor approach.[8] We will first consider the joint distribution of default times, and then the distribution of the number of defaults, afterwards the linear loss distribution leading to valuation based on the expected loss principle is outlined.[9] Finally, the use of the HLPGC in practice as well as implied correlation is discussed.

A structural model postulates a stochastic process for the asset values $(A_t)_t$ as well as a deterministic behavior for liabilities $(x_t)_t$ over time. A default occurs when the asset value falls below the liabilities, i.e. $A_t \leq x_t$. Thereby x_t is called the default barrier. It is easily observed that the criteria $A_t \leq x_t$ is equivalent to

$$Y_t = \frac{ln\left(\frac{A_t}{A_0}\right) - \mu_{R_t}}{\sigma_{R_t}} \leq \frac{ln\left(\frac{x_t}{A_0}\right) - \mu_{R_t}}{\sigma_{R_t}} = y_t, \tag{9}$$

for the given investment horizon (time bucket) $[0, t]$, where $R_t = ln(A_t/A_0)$ is the log-asset return and $\mu_{R_t} = E(R_t)$. When $(A_t)_t$ is explained by a Geometric Brownian Motion we are left with $Y_t \sim N(0, 1)$. Expected value and standard deviation of R_t are obtained if one discretizes

$$dA_t = \mu \cdot A_0\, dt + \sigma \cdot A_0\, dW_t \Leftrightarrow^{Ito} dlnA_t = (\mu - \frac{1}{2}\sigma^2)\, dt + \sigma\, dW_t$$

and solves analytically to receive

$$\begin{aligned} lnA_t &= lnA_0 + (\mu - \frac{1}{2}\sigma^2) \cdot t + \sigma\sqrt{t} \cdot \epsilon_t \\ \Leftrightarrow ln\left(\frac{A_t}{A_0}\right) &= \left(\mu - \frac{1}{2}\sigma^2\right) \cdot t + \sigma\sqrt{t} \cdot \epsilon_t, \end{aligned} \tag{10}$$

[8]The following has been partly inspired by N.N. (2004).

[9]It is distinguished between the linear loss distribution that refers to the collateral pool and does not consider structural features, and the loss distribution that refers to the tranche under consideration.

with $\epsilon_t \sim N(0,1)$. Thus $\mu_{R_t} = (\mu - \frac{1}{2}\sigma^2) \cdot t, \;\; \sigma_{R_t} = \sigma\sqrt{t}$.

Having for each member m of the basket $\{1, ..., M\}$ of reference assets this normalized log-asset returns $Y_{t,m}$ and the respective default barriers $y_{t,m}$ we consider also the normalized return $V_t \sim N(0,1)$ of some market factor in common as well as $Z_m \sim N(0,1)$ to establish a correlation structure between the asset returns, which is formula (7)

$$Y_{t,m} = \rho_m \cdot V_t + \sqrt{1 - \rho_m^2} \cdot Z_m. \tag{11}$$

The common factor V_t accounts for systematic risk and the Z_m imply specific risk. The derivation shows that even the time horizon $[0,t]$ is not decisive for Y_m, V and Z_m, as they are all standard normal, but the time horizon may influence the correlations and is crucial for the default barrier.[10]

When implementing the factor approach one is able to work around knowing the common factor as it is standard normally distributed, and the formulas just integrate over its whole range to calculate the unconditional distributions, cf. formula (25).

3.3 Joint Distribution of Default Times

In this section we want to derive the joint distribution of default times using the Factor Approach. An explicit representation of the Gaussian copula as given by (2) requires the use of an M-dimensional integral, whereas the factor representation allows to simplify this to an parsimonious 1-dimensional integral-representation. Even though this representation will not be used here for valuation purposes we will derive it in the following because some of the intermediate results are basic building blocks of valuation under the Factor Approach.

Assume for example the single survival probabilities $P(T_m > t)$ as being explained by a *Reduced Form Model* with constant default intensities,[11] that is, created by a Poisson process with constant intensities λ_i:

$$P(T_m > t) = q_t^m = e^{-\lambda_m t} = 1 - p_t^m = 1 - P(T_m < t), \tag{12}$$

where p_t^m is the unconditional probability of default prior to time t and q_t^m is the probability of survival. In order to bring these distributions into dependence the one-factor approach uses equation (11) and relates

$$\phi(y_m) = p_t^m \Longleftrightarrow P(Y_{t,m} < y_{t,m}) = P(T_m < t), \tag{13}$$

via the Structural Model. Thereby, ϕ is the standard normal distribution function of $Y_{t,m}$ and $P(*)$ is the probability. Formula (13) is how the normalized log-asset-

[10] Cf. section 3.3, as $y_{t,m} = \phi^{-1}(p_t^m)$ increases with p_t^m. This can not be seen from right hand side of formula (9), because the role of x_t is not clear.

[11] Term structures of credit Spread, default intensities and default probabilities are equivalent concepts.

returns of the structural model are transformed into default times, and the asset-return correlations are again converted into correlations of default times that are a bit below correlations of asset returns.[12] Thereby, $y_{t,m}$ signifies the normalized default barrier. Market practice calibrates the default barrier to fit the term structure of risk neutral default probabilities $(p_t^m)_t$ that are implied by the credit spreads of traded single name CDS, via $\phi^{-1}(p_t^m) = y_{t,m}$. Section 5 shows how single name risk neutral default probabilities are derived by bootstrapping. Moreover, one can easily recognize that as $p_t^m = 1 - e^{-\lambda_m t}$ increases with t also the default barrier $y_{t,m}$ increases with t.

It is important to note: the variables $(Y_{t,m} \mid V_t)$ and also the default times conditioned on the common factor V_t are independent. Due to (8) we can express the conditional default probability via the standard normal distribution ϕ as

$$P(Y_{t,m} < y_{t,m} \mid V_t) = \phi\left(\frac{y_{t,m} - \rho_m V_t}{\sqrt{1-\rho_m^2}}\right) \tag{14}$$

$$\Longleftrightarrow p_t^{m|V_t} = P(T_m < t \mid V_t) = \phi\left(\frac{\phi^{-1}(p_t^m) - \rho_m V_t}{\sqrt{1-\rho_m^2}}\right) \tag{15}$$

$$\Longleftrightarrow q_t^{m|V_t} = 1 - \phi\left(\frac{\phi^{-1}(p_t^m) - \rho_m V_t}{\sqrt{1-\rho_m^2}}\right). \tag{16}$$

The conditional independence allows us to write the joint conditional distribution of default times as

$$P(T_1 < t_1, ..., T_M < t_M \mid V_t) = \prod_{i=1}^{M} P(T_i < t_i \mid V_t) = \prod_{i=1}^{M} p_t^{i|V_t}. \tag{17}$$

The iterated expectations theorem

$$\begin{aligned} E_x(X) &= E_y(E_x(X \mid Y)) \\ \Longrightarrow P(A) = E(\mathbf{1}_A) &= E_y(E(\mathbf{1}_A \mid Y)) = E_y(P(A \mid Y)) \end{aligned} \tag{18}$$

allows to express the searched unconditional distribution of default times $F(t_1, ..., t_M)$ by means of its conditional counterpart and via the standard normal density f_V of the common factor:

$$\begin{aligned} P(T_1 < t_1, ..., T_M < t_M) &=^{it.exp.} E_V(P(T_1 < t_1, ..., T_M < t_M \mid V)) \\ &=^{cond.indp.} E_V(\prod_{i=1}^{M} p_{t_i}^{i|V_t}) \\ = F(t_1, ..., t_M) &= \int_{-\infty}^{\infty} \prod_{i=1}^{M} p_{t_i}^{i|V} \cdot f_V(x)\ dx. \end{aligned} \tag{19}$$

[12]See appendix for how the conversion works.

3.4 Distribution of the Number of Defaults

Let $P(N(t)=k)$ with $N(t)=\Sigma_{m=1}^{M}\mathbf{1}_{(T_m<t)}$ denote the distribution of the number of defaults up to time t. This, or alternatively the linear loss distribution, is the key distribution on which the valuation of NtD and CDO is based. From a theoretical point of view the distribution of the number of defaults is equivalent to the linear loss distribution - given a homogeneous basket. But its mathematical representation suits more to baskets with only few members (say up to 10) and is thus rather used for N-th to default swaps than CDOs whose baskets cover around 150 members. We consider the moment generating function

$$g_{N(t)}(x)=E(x^{N(t)})=\sum_{m=1}^{M}P(N(t)=k)\,x^k. \tag{20}$$

Again, we can rewrite this by the theorem of iterated expectations, cf. formula (18), as

$$\begin{aligned}
E(x^{N(t)}) &= E_V(E(x^{N(t)}\mid V)) = E_V(E(x^{\Sigma_{m=1}^{M}\mathbf{1}_{(T_m<t)}}\mid V)) \\
= E_V(E(\prod_{m=1}^{M}x^{\mathbf{1}_{(T_m<t)}}\mid V)) &=^{CondIndp} E_V(\prod_{m=1}^{M}E(x^{\mathbf{1}_{(T_m<t)}}\mid V)) \\
&= E_V\left(\prod_{m=1}^{M}\left((1-p_t^{m|V})+p_t^{m|V}\cdot x\right)\right) \qquad (21) \\
&= \int_{-\infty}^{\infty}f_V(v)\prod_{m=1}^{M}((1-p_t^{m|V})+p_t^{m|V}\cdot x)\;dv \qquad (22)
\end{aligned}$$

Hereby f_V is the standard normal density of the common factor, and equation (21) follows from (15) and

$$\begin{aligned}
E(x^{\mathbf{1}_{(T_m<t)}}\mid V) &= P(\mathbf{1}_{(T_m<t)}=0\mid V)\cdot x^0+P(\mathbf{1}_{(T_m<t)}=1\mid V)\cdot x^1 \\
&= P(T_m>t\mid V)\cdot 1+P(T_m<t\mid V)\cdot x=(1-p_t^{m|V})+p_t^{m|V}x.
\end{aligned}$$

The initial representation of the moment generating function uses a polynomial in x of degree M. In order to obtain a tractable representation of $P(N(t)=k)$ we recognize that $\Pi_{m=1}^{M}\left[(1-p_t^{m|V})+p_t^{m|V}x\right]$ is also a polynomial in x of degree M and can be thus written as

$$\sum_{k=0}^{M}\xi_k(p_t^{1|V},...,p_t^{M|V})\cdot x^k, \tag{23}$$

with the factors $\xi_k(p_t^{1|V},...,p_t^{M|V})$ made up of the conditional default probabilities. We can now rearrange (22) to become

$$\sum_{k=0}^{M}\left[\int_{-\infty}^{\infty}f_V(v)\cdot\xi_k(p_t^{1|V},...,p_t^{M|V})\;dv\right]\cdot x^k. \tag{24}$$

This finally yields the desired distribution for the number of defaults up to time t as

$$P(N(t) = k) = \int_{-\infty}^{\infty} f_V(v) \cdot \xi_k(p_t^{1|V}, ..., p_t^{M|V})\ dv, \tag{25}$$

which is called the semi-analytical, semi-implicit, expression respectively.

As mentioned at the end of section 3.2 we do not need to know about the common factor V, because we just integrate over its whole range. As regards the asset correlations, market practice uses either *equity-return correlations* (with the need to specify V anyway) or *implied correlations* that meet market spreads. A further possibility to avoid the determination of the common factor is to use the average correlation between equity returns directly. Unfortunately, often this has to be corrected to match the *compound correlation* of a given tranche.[13] For tailor made baskets that underly so called *Bespoken CDO* this is not possible due to a lack of liquidity. A first simple approach to solve this problem is to combine the existing standard baskets of Europe, iTRAXX, and the US, CDx, to match the tailor made basket as good as possible, and calculate the respective weighted average, more precisely: if the bespoken basket comprises of α% EUR and β% USD reference entities and thus weight $\alpha \cdot corr_{iTraxx} + \beta \cdot corr_{CDx}$. For a more elaborated analysis of this issue please be referred to Huddart/Picone (2007).

3.5 Linear Loss Distribution

We emphasize the difference between the *linear loss distribution* as of the underlying assets and its derivative that is transformed by the location of the tranche, waterfall characteristics, respectively. The latter is denoted *loss distribution.* The loss distribution matters directly to the investor while the linear loss distribution characterizes the underlying collateral pool.

We concentrate on an analytical approach proposed by Kalemanova et al. (2005) and which is based on the Large Homogeneous Portfolio (LHP) assumption. Note, there are also semi-analytical approaches as proposed by Hull/White (2004) to obtain the distribution of the number of defaults as well as the linear loss distribution. These require iterative procedures but relax some of the assumptions of LHP.

Within the factor approach the LHP assumes the correlations of the asset returns with the market, as well as the default barriers and recovery rates, to be equal for all members of the basket. Formula (14) shows that this assumption implies an identical term structure of default probabilities for all members of the basket. In addition, the basket is assumed to consist of a large number of members that are equally weighted, in order to make the central limit theorem applicable. Using the above assumptions we have equal $p_t^V = p_t^{m|V}$ $\forall m$ and receive the conditional loss

[13] For an explanation of compound and base correlation please be referred to section 4.2.3.

distribution via the binomial distribution that models a drawing with dropping back, because the probabilities do not change given a default or a draw. The cumulative loss due to k defaults is $l_k = k \cdot \frac{N}{M}(1-R)$, where N is the nominal of the whole CDO and M is the number of elements in the basket. The conditional probability for such a loss at time t is then

$$P(cL_t = l_k \mid V) = \begin{pmatrix} M \\ k \end{pmatrix} \cdot (p_t^V)^k \cdot (1 - p_t^V)^{M-k}. \tag{26}$$

The conditional distribution can be extended to cover different default probabilities, i.e. different correlations to the common factor and/or different default barriers, but still with identical nominals and recovery rates:

$$P(cL_t = l_k \mid V) = \sum_{I_k} \prod_{j=1}^{k} p_t^{i_j|V} \cdot \prod_{j=1}^{M-k} (1 - p_t^{h_j|V}), \tag{27}$$

where $I_k = \{\{i_1, ..., i_k\} : \{i_1, ..., i_k\} \subset \{1, ..., M\}\}$ is the index set of all sets out of the basket that comprise of k different members, and $h_j \in \overline{I}_k$ with $\overline{I}_k = \{1, ..., M\} \setminus I_k$. Admittedly, the last representation is rather complex and calls for a numerical estimation of the distribution function. Nevertheless, it is an alternative to the approach to evaluate the distribution of the number of defaults in the previous section.

Integrating over the density of the return of the market factor yields the unconditional linear loss distribution of the underlying pool

$$P(cL_t = l_k) = \int_{-\infty}^{\infty} f_V(x) \cdot P(cL_t = l_k \mid x)\, dx = E_V(P(cL_t = l_k \mid V)). \tag{28}$$

4 Applications

Within this section we want to show how simulation can be used to value NtD and unfunded CDOs easily even with inhomogeneous baskets. Clearly, simulation is also suited to value cash CDOs with their sophisticated waterfall structures. The proposed mapping of the waterfall is applicable to a wide range of cash CDOs and we believe it to cover the most significant features of a waterfall that make up the loss distribution. Afterwards, the factor approach is used to find computationally less demanding analytical pricing formulas for NtDs and synthetic unfunded CDOs.

4.1 Simulation

Based on a realization $(t_1, ..., t_M)_k$ of risk neutral default times we can calculate a deterministic Pv_k under the given scenario k, realization of default times, respectively, when we assume a deterministic risk free interest rate. Having created n such scenarios and Pv's the present value is obtained as

$$Pv = \frac{1}{n} \cdot \sum_{k=1}^{n} Pv_k. \tag{29}$$

The above principle is now outlined in examples.

4.1.1 NtD

Let t_{mat} be the time to maturity of the NtD at valuation and $t_{m,k}$ the m-th default time in the k-th scenario of our simulation. Further $N_{m,k}$ the nominal of the asset that causes the m-th default in the k-th scenario,[14] and $L_{m,k}$ the loss given default of asset (m, k), i.e. the LGD could vary over the scenarios. Form the protection sellers point of view the present value of the k-th scenario is given by

$$Pv_k = Pv_{CumPremium,k} + Pv_{Accrued,k} - Pv_{Default,k}. \tag{30}$$

The present value of the **default leg** of the m-th asset at time of default is

$$Pv_{Default}(t_{m,k}) = \begin{cases} e^{-r_f \cdot t_{m,k}} \cdot N_{m,k} \cdot L_{m,k} & if \quad t_{m,k} \leq t_{mat} \\ 0 & if \quad t_{m,k} > t_{mat} \end{cases}. \tag{31}$$

In case of default the protection seller pays the loss given default $(N_{m,k} \cdot L_{m,k})$ which is discounted by the risk free rate and otherwise nothing.

As regards the **premium leg**, and if we denote by $t_p \leq t_{m,k}$ the time to the last premium payment before the m-th default time, and by cs the premium spread or fair credit spread, we have

$$Pv_{Premium}(t_{m,k}) = \begin{cases} \sum_{i=1}^{p} e^{-r_f \cdot t_i} \cdot cs_i \cdot N & if \quad t_{m,k} \leq t_{mat} \\ \sum_{i=1}^{l} e^{-r_f \cdot t_i} \cdot cs_i \cdot N & if \quad t_{m,k} > t_{mat} \end{cases}, \tag{32}$$

where l is the total number of premium payments. In case of $t_p < t_{m,k} < t_{p+1}$ the **accrued premium** is given by

$$cs_{p+1} \cdot N \cdot \frac{t_{m,k} - t_p}{t_{p+1} - t_p}. \tag{33}$$

---

[14] From notation we see the that no homogeneous nominals are required.

4.1.2 Synthetic Unfunded CDO

In contrast to the NtD we just treat the **default leg**. If we had a homogeneous CDO with lets say 100 names and identical nominals and we want to price the tranche between 6% and 15% its present value was the sum of present values of a 6-th NtD up to a respective 15th NtD.

For a more flexible approach we consider a $a\%$ to $b\%$ CDO tranche with m-many underlying assets and different nominals $N_1, ..., N_M$ that are still outstanding, i.e. $N = \Sigma N_m$ and $N_{Tr}(0) = (b/100 - a/100) \cdot N(0)$ being the principal of the tranche at deal inception. For the ease of notation we will neglect the time index for the starting nominals in the sequel. The by-product of our MC-simulation will be the transition distribution of the cumulative loss $P(cL(t) \leq l)$.

In the k-th scenario the cumulative linear loss $cL_k(t)$ at time $t \leq t_{mat}$ is calculated by the single name losses given default $L_{m,k}$ as[15]

$$cL_k(t) = \sum_{m=1}^{M} L_{m,k} \cdot \mathbf{1}_{(t_{m,k} \leq t)} \tag{34}$$

The cumulative loss that has to be borne by the tranche at time t in scenario k is then

$$cL_k^{[a,b]}(t) = \begin{cases} 0 & if \quad cL_k(t) \leq a/100 \cdot N \\ N_{Tr} & if \quad cL_k(t) \geq b/100 \cdot N \\ cL_k(t) - (a/100) \cdot N & if \quad cL_k(t) \in (a/100 \cdot N, b/100 \cdot N) \end{cases} . \tag{35}$$

If the sum of all losses given default at time t is less than the lower bound of the tranche the tranche bears no losses. On the other hand, if cumulative losses exceed the upper bound the whole nominal of the tranche is paid to the protection buyer. In case the losses are in between the range of the tranche only what is above the lower bound has to be paid. The currently remaining nominal of the tranche $[a, b]$ equals $N_{Tr}(t) = N_{Tr} - cL_k^{[a,b]}(t)$. Over all scenarios the density of the loss distribution can be obtained via kernel estimation over the $cL_k^{[a,b]}(t)$ for an arbitrary fixed point in time. Given some densities for selected points in time the transition density can be obtained via interpolation.

With $cL^{[a,b]}(t)$ the average loss over all scenarios the present value of the default leg becomes

$$Pv_{Default} = \int_0^{t_{mat}} e^{-r_f \cdot t} \cdot cL^{[a,b]}(t) \; dt. \tag{36}$$

[15] Usually, the loss given default is fixed over the scenarios, e.g. $L_m = 40\%$, here we allow the loss given default to vary over the scenarios.

4.1.3 Cash CDO

This section shows how an investor is able to treat cash CDOs with sufficient precision, without increasing the necessary effort to an extend that would endanger the application in practice under restricted resources. In spite of this simplification, the considered method will most likely be only engaged if there are doubts as concerns the spread approach, as it is currently observed during the sub-prime crisis. The reason is that the spread with its mapping on a publicly observable risk factor is only able to account for systematic risk.

A cash CDO usually goes through 3 phases. During the *ramp up period*, that lasts 6 to 12 months, the collateral manager successively establishes the underlying pool of collateral. Afterwards, the *reinvestment period* follows, which takes 4 to 6 years, where principal proceeds, recovery and the protecting share of excess spread are reinvested into new collateral assets. During the reinvestment period the collateral manager actively trades the the assets. The assessment of the collateral manager by the investor is thus of pivotal importance for the investment decision if it is to be invested in a managed CDO. Finally, the trade enters the *amortization period* during which the notes are redeemed, and recovery as well as the curing part of excess spread are no longer reinvested but increase the redemption top down.

In a cash CDO the linear loss distribution is transformed by the waterfall into the loss distribution. A *waterfall* defines in which order the interest rate and principal payments out of the collateral pool are poured out over the rated trances. In most cases interest rate swaps transform the variable interest rate payments of the pool into fixed rates. The rest, which is left to the equity holder after the rated trances have received their cash flows, is called *excess spread*, cf. figure 1. It can thus be interpreted as some kind of dividend. The waterfall also determines what happens to the payments if defaults occur. Due to the immanent complexity of the waterfall a Cash CDO is predestined for a simulation approach.

The *credit enhancement* implied by the waterfall can be distinguished into *subordination*, *excess spread* and *trigger values*, cf. figure 1. Usually, OC triggers determine the sequential redemption of tranche notes beginning with the most senior. IC triggers determine as to when interest rate cash flows from the collateral are used to first pay interest to the senior tranche and then to redeem this tranche until the IC test is met again, followed by the next subordinated tranche. A further structural feature are *deferred interest* accounts, on which it is booked when the remaining cash flow is insufficient to meet the contractual interest rate of the tranche. In the next period, the structure needs first to serve the interest claim booked on the deferred interest account before it is able to pay the ordinary interest rate cash flow. Excess spread is used in conjunction with supplement OC tests that are triggered before OC tests are breached and imply the reinvestment of excess spread instead

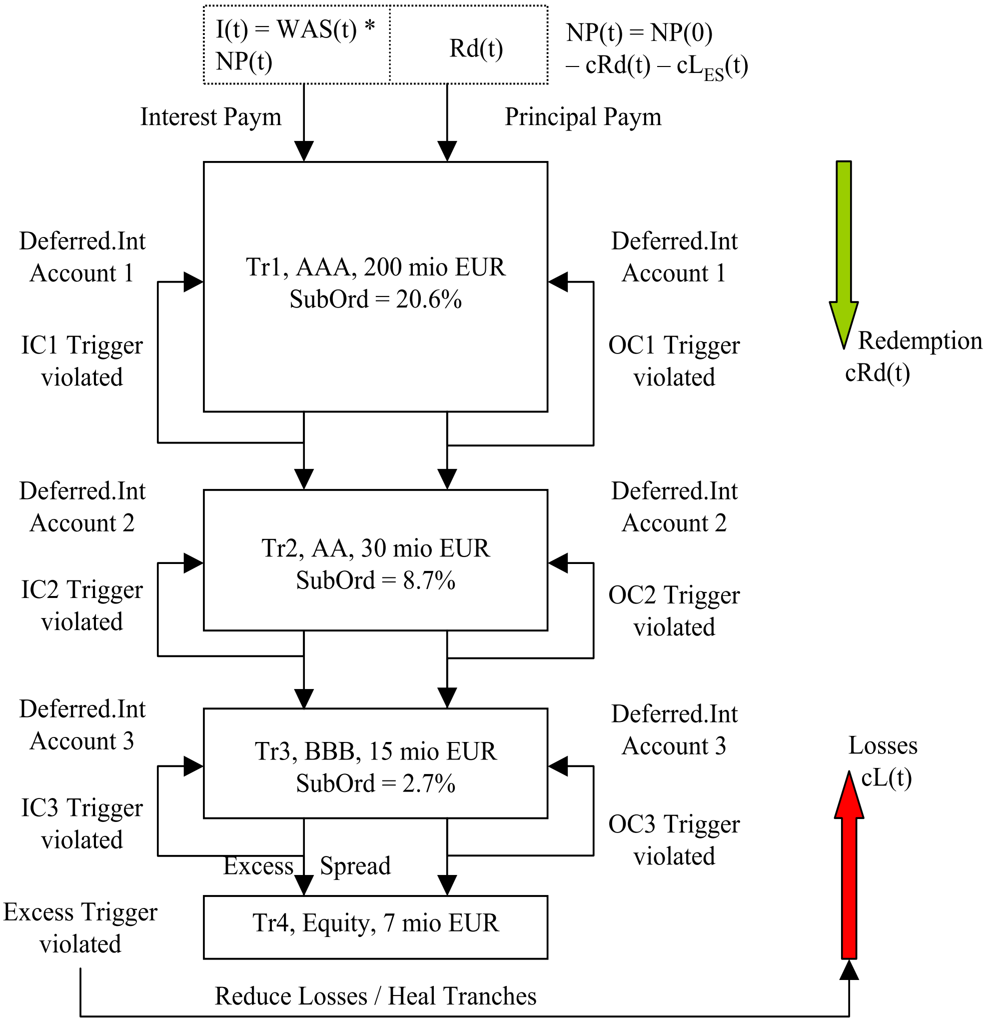

Figure 1: Waterfall Structure

of redemption. This mechanism is intended to avoid early amortization.[16] These triggers are activated before the OC/IC trigger and direct the reinvestment of excess spread which has a similar impact as redemption, that is, it increases the trigger values. OC trigger divide the current nominal of the pool by the nominal of the respective tranche plus the nominal of all superior trances. Hence, except for the lowest tranche, reinvestment increases the numerator but not the denominator. As regards redemption, we need to distinguish two cases. Redemption by interest rate cash flows reduces the denominator but not the numerator of trigger values, while redemption through principal cash flows or recovery not only diminishes the numerator but also the denominator. As the numerator is higher than the denominator and thus the reduction of the denominator dominates the reduction of the numerator, redemption too increases the trigger value. Note thereby, redemption by principal cash flows and recovery tend to reduce the interest rate cash flow available to the lower trances, because the lost WAS is higher than the saved interest rate cash flow to the most superior tranche. Accordingly, this kind of redemption increases the IC trigger of high and middle trances, and reinvestment has a similar impact.

The sequential process outlined above coincides with the vast majority of traded cash CDOs. But there exist some other structures at the market that utilize fully or in part the pro-rata-principle, that is, they distribute losses equally over the trances. For such structures the proposed model is inappropriate. A substantial simplification results given a bullet redemption, because then the amortization period collapses and we are left with the reinvestment period. This special case will be discussed afterwards, and is a more conservative approach for investment grade trances under stress scenarios, because the early redemption is a main part of the credit enhancement. In addition, it has a negative impact on the present value, when valuation spreads tend to increase with decreasing time to maturity.

Before we investigate the values of the waterfall that are specific for a given tranche the variables that influence the whole structure are outlined. Let $cL(t) = \sum_{i=1}^{t}(L(i) - R(i))$ be the cumulative net loss at the end of period t, then the the cumulative net loss reduced by excess spread can be written recursively as

$$cL_{es}(t) = cL_{es}(t-1) + Max\left(L(t) - R(t) - \alpha \cdot es(t-1), 0\right) \tag{37}$$

with $cL_{es}(1) = cL(1)$. In this representation the contractually available excess spread $\alpha \in [0,1]$ to heal losses is reinvested after the reinvestment period instead of paying down the trances with recovery and ordinary redemption. This assumption simplifies the mathematical representation, *because amortizing excess spread reduces the trances but not the collateral pool.* In order to avoid double counting when calculating the current nominal $NP(t)$ of collateral assets, the recovery in (37)

[16] Cf. Mahadevan et al. (2006), pp. 18.

reduces the gross loss also after the reinvestment period, but increases

$$cRd(t) = \sum_{i=1}^{t} \left(Rd(t) + R(i) \cdot \mathbf{1}_{(i>ReInv)} \right), \tag{38}$$

the cumulative redemption out of the collateral pool until period t. That is to say, the recovery will not be reinvested after the end of the reinvestment period. For the current nominal of the collateral pool we thus receive

$$\begin{aligned} NP(t) &= NP - cRd(t) - cL_{es}(t), \text{ and} \\ I(t) &= WAS(t) \cdot NP(t) \end{aligned} \tag{39}$$

is the implied interest rate payment, where we neglect management fees and other costs. As in the previous section we abstain from adding the time index to the starting nominal.

Based on these global structural features we now focus on the tranche-specific parameters. We start with the cumulative redemption, the cumulative net loss and the implied remaining nominal of tranche Tr:

$$\begin{aligned} cRd_{Tr}(t) &= Min\left(Max\left(cRd(t) - \sum_{j=1}^{Tr-1} N_j, 0 \right), N_{Tr} \right), \\ cL_{Tr}(t) &= Min\left(Max\left(cL_{es}(t) - \sum_{j=Tr+1}^{\#Tr} N_j, 0 \right), N_{Tr} \right) \text{ and} \\ N_{Tr}(t) &= N_{Tr} - Min\left(cRd_{Tr}(t) + cL_{Tr}(t), N_{Tr} \right). \end{aligned} \tag{40}$$

Thereby, N_{Tr} is the original nominal value of the tranche. The last minimum-function is not necessary in this setting, but if excess spread was used to redeem during the amortizing period, or if trigger values redirected interest rate cash flows into the early redemption of trances, then, the sum of the nominal amounts of all trances would no longer equal the nominal of the collateral pool but would be less. Thus, the condition $cRd_{Tr}(t) + cL_{Tr}(t) \leq N_{Tr}$ no longer holds for all trances Tr. Given this model and under the assumption of a purely sequential waterfall the interest rate cash flow to tranche Tr at the end of period t is

$$I_{Tr}(t) = Min\left((br(t) + cs_{Tr}) \cdot N_{Tr}(t), Max\left(I(t) - \sum_{j=1}^{Tr-1} I_j(t), 0 \right) \right). \tag{41}$$

The term $br(t)$ refers to the base rate, that is for example the 3M-EURIBOR interest rate, and cs_{Tr} is the emission spread of the tranche. Formula (41) is recursive in the number of superior trances. If the waterfall is not to be modelled completely within our model it is possible to approximate the superior payments by $I_j(t) = (br(t) + cs_j) \cdot N_j(t)$. The interest rate cash flow $I(t)$ out of the collateral pool, reduced by the interest rate payments that are consumed by superior trances feeds the interest rate claim of the tranche under consideration. But this interest rate can not be negative, which is ensured by the inner maximum-function. The outer minimum-function accounts for the fact that even though the tranche has a contractual claim on the base rate enhanced by the emission-spread, the available interest rate cash flow may not suffice to feed the contractual claim completely due to losses.

The above framework covers some critical values. Under a complete model of the waterfall the excess spread is given endogenously. As the waterfall is not modelled completely here we need to address the impact of our simplifying assumptions on excess spread. Our model increases the excess spread. It thus increases the expected time to maturity in two ways. First, it does not accelerate the redemption of investment grade trances, and second, it increases the excess spread that heals low trances. Hence, this is a conservative approach for investment grade investors, while it is the other was round for lower rated trances. Within our model the excess spread is given by $es(t) = Min(I(t) - \sum_{j=1}^{\#Tr-1} I_j(t), 0)$. This relationship together with formula (41) shows that the structure of superior nominal values $N_j(t),\ j \in \{1, ..., Tr-1\}$ can not be neglected when evaluating a given tranche.

In order to figure out exactly the $WAS(t)$ that determines the interest rate cash flow, the investor were to know how the collateral pool is made up at each point in time during a simulation. A possible circumvent of this problem is to assume that the credit structure of the collateral pool will not change during life time, that is, $WAS(t) \equiv WAS$. For *managed pools* this is a feasible assumption. For *static pools* that do not replace the high yield components, that are likely to default first, the weighted average spread will decrease during the term.

If the model does not employ the simplifying assumption of a bullet redemption the investor needs a function of time which tells about the behavior of scheduled redemption $sRd(t)$ and prepayments $uRd(t)$ during the term. The latter, and thus its cumulative version too, depends on the level of interest rates as well as macroeconomic factors, like inflation and unemployment. Hence, its modelling can lead to sophisticated approaches that cover term structure models. The prepayment risk not only concerns *pass-through fonds*, after the reinvestment period also *pay-through structures* are concerned, on which this article puts its emphasis. In any case, the evaluation of interest rate and principal cash flows requires an exact analysis of the collateral assets over time, even for planned amortization. As some investors do not have the necessary IT and human resources at their disposal to accomplish this we discuss a simple approximation. This proxy aggregates the two kinds of principals payments above and lets them increase in a concave manner, more precisely:

$$sRd(\tau) + uRd(\tau) = Rd(\tau) = \quad (42)$$

$$\begin{cases} 0 & if \quad \tau \in [0, \#t_{ReInv}] \\ \frac{NP - 2\cdot cL_{es}(\tau)}{(\#t_{WAL} - \#t_{ReInv})^2} \cdot (\tau - \#t_{ReInv}) + \frac{cL_{es}(\tau)}{(\#t_{WAL} - \#t_{ReInv})} & if \quad \tau \in (\#t_{ReInv}, \#t_{WAL}] \end{cases} .$$

The basic idea of this method is how a loss after period $\tau^* \in [\#t_{ReInv}, \#t_{WAL}]$ must reduce the slope of a redemption, which has been linearly increasing until this loss occurred, as such to ensure that the remaining nominal is not repayed before t_{WAL}. Hereby, $(\#t_{WAL} - \#t_{ReInv})$ is the number of periods between the end of the reinvestment phase t_{ReInv} and the expected maturity date t_{WAL} of the structure, and τ labels the number of periods that have been passed since valuation date. The shown representation assumes the valuation date to lay within the reinvestment period. The model does not try to calculate the weighted average life (WAL) based on the above mentioned sophisticated models by itself but takes it as an exogenous input, e.g. from Bloomberg. For $\tau \in [0, \#t_{ReInv}]$ we receive $Rd(\tau) = 0$ from (43), and for $\tau = \#t_{WAL}$ the formula yields $Rd(\#t_{WAL}) = (NP - cL_{es}(\#t_{WAL}))/(\#t_{WAL} - \#t_{ReInv})$. Formula (43) too rests on an excess spread that is reinvested not only during the reinvestment phase. Due to the fact that $Rd(\tau)$ is strictly increasing the proposed calibration implies an extended time to maturity, as it was targeted. In any case, as shown in figure 3, the user is able to adjust the functional behavior by varying the parameter $\#t_{WAL}$ to match the scenario of a professional system.

The outlined concepts can be directly used to assess the impact of predefined stress scenarios of defaulting collateral assets over time on the cash flows left to the tranche under consideration. This is often done by investors to check what *constant annual default rate* (CADR) the tranche they are interested in can bear. In addition, as in the previous section, one can simulate the cumulative linear loss $cL_k(t)$ according to formula (34) with $k \in \{1, ..., n\}$ n times. For each scenario k this yields the needed interest rate and principal cash flows. Discounted with appropriate swap rates the present value of this scenario follows. Finally, the statistical average over all scenarios yields the searched present value of the tranche. For a real world structure figure 3 compares the resulting interest rate and principal cash flows of an extended version of the outlined model that takes OC/IC trigger values into account with the output of a professional model. Thereby, for the behavior of the cumulative loss, the scenario of a CADR of 8% has been assumed. Note, the two sums of all cash flows in both models are equal, but with different timing. The difference in timing can be particularly recognized at the principal payments of tranche A, which shows what difference it makes when each component with its redemption and maturity profile is known and influences the model.

We consider the proposed model as a quick to implement approach. It is applicable to a broad scope of structures in a conservative way while being as simple as possible. One advantage is that it avoids the need to model the whole structure, waterfall respectively, as it is done by more sophisticated approaches.

If it is allowed for stepwise amortization of rated notes after the reinvestment period or for early amortization it can be stated that percentage subordination would increase because trances are redeemed sequentially, while equity is kept constant. As concerns the excess spread early redemption of the low interest rate senior notes reduces the excess spread. Moreover, deleveraging or repaying notes reduces the granularity of a portfolio and its diversification. Despite the last two aspects, it can be concluded that improved subordination through early amortization is the main driver for better default protection. Thus the assumption of bullet redemption and the reinvestment approach are conservative.

4.2 Factor Approach

While the explained Monte-Carlo-Approach is very flexible when the impact of correlations is to be modelled precisely or an inhomogeneous CDO is to be valued, there are as discussed analytical and semi-analytical (iterative) procedures that avoid simulation, cf. Hull/White (2004). Some apply a special case of Hulls linear loss distribution (used for CDO) and his distribution of the number of defaults (used for NtD), where the correlation matrix is aggregated into an average correlation factor, assuming that correlations do not vary much around their average.[17]

4.2.1 NtD

We focus on a homogeneous NtD, implying equal nominals for the underlying assets $N_i = N_j = \widehat{N} \;\forall\, i, j \in \{1, ..., M\}$. We moreover assume equal loss given default $\widehat{L}$.

The present value of the **default leg** of a homogeneous NtD is readily obtained as[18]

$$Pv_{Default} = \int_0^{t_{mat}} e^{-r_f t} \cdot \widehat{L} \cdot \widehat{N} \;\; dP(N(t) \geq k). \tag{43}$$

An inhomogeneous NtD would cause a problem because we do not know what asset defaults, but different assets can have different nominals and loss given default in case of an inhomogeneous NtD. For this reason one needs to model the distributions of each asset to cause the k-th default.

The above formula can be transformed into an analytically more tractable form using integration by parts

$$\int_a^b f \cdot g' = [f \cdot g]_a^b - \int_a^b f' \cdot g. \tag{44}$$

For this we define the distribution function of the k-th default as $F^{(k)}(t) = P(N(t) \geq$

[17] In its most flexible version the approach uses a set $\{V_1, ..., V_q\}$ of explaining indicators and their correlations with the assets $A_1, ..., A_M$. For each V_k of these explaining factors its asset correlations $\rho_{k,1}, ..., \rho_{k,M}$ with the reference assets are used and the correlation between two assets A_i, A_j becomes $\rho_{1,i}\rho_{1,j} + ... + \rho_{q,i}\rho_{q,j}$.

[18] Note, $dP(N(t) \geq k) = P(N(t + dt) \geq k) - P(N(t) \geq k)$, gives the probability that the k-th default occurs exactly at time t within the integral.

k) and extend the above formula trivially by dt

$$\begin{aligned}
Pv_{Default} &= \int_0^{t_{mat}} e^{-r_f t} \cdot \widehat{L} \cdot \widehat{N} \cdot \frac{dF^{(k)}(t)}{dt} \; dt \\
=^{44} e^{-r_f t_{mat}} \cdot F^{(k)}(t_{mat}) &- \int_0^{t_{mat}} F^{(k)}(t) \cdot \widehat{L} \cdot \widehat{N} \; d(e^{-r_f t}) \qquad (45)\\
= e^{-r_f t_{mat}} \cdot F^{(k)}(t_{mat}) &+ \int_0^{t_{mat}} F^{(k)}(t) \cdot \widehat{L} \cdot \widehat{N} \cdot r_f \cdot e^{-r_f t} \; dt.
\end{aligned}$$

If the k-th default could only occur at the payment dates of the premium (credit spread cs), the present value of the **premium leg** was

$$Pv^{\diamondsuit}_{Premium} = \sum_{i=1}^{P} P(N(t_i) < k) \cdot cs_{t_i} \cdot \widehat{N}. \qquad (46)$$

The possibility of intermediate defaults increases this present value by the expected discounted value of the **accrued**

$$Pv_{Accrued} = \sum_{i=1}^{P} \int_{t_{i-1}}^{t_i} e^{-r_f t} \cdot \frac{t - t_{i-1}}{t_i - t_{i-1}} \cdot cs_{t_i} \cdot \widehat{N} \; dP(N(t) \geq k), \qquad (47)$$

that linearly increases from the beginning of a payment period and declines sharply to zero at the next premium payment giving the shape of a sawtooth. The probability $P(N(t_i) < k)$ stands for the fact that the premium is only received if less then k defaults occurred up to the payment date t_i.

4.2.2 Synthetic Unfunded CDO

A closer look at the formulas based on MC-Simulation (29), (31) and the distribution of the number of defaults used in (43) clarifies that all concepts rely on the expected loss, the loss distribution, respectively. Thus one could rewrite the formulas in terms of the expected loss. For a single tranche CDO with initial nominal N_T between attachment point a and detachment point b we have

$$Pv_{PremiumLeg} = \sum_{i=1}^{mat} cs \cdot max(0, N_T - EL_{[a,b]}(t_i)) \cdot e^{-r_f \cdot t_i} \qquad (48)$$

$$\begin{aligned}
Pv_{DefaultLeg} &= \sum_{i=1}^{mat} e^{-r_f \cdot t_i} \; (EL_{[a,b]}(t_i) - EL_{[a,b]}(t_{i-1})) \qquad (49)\\
&= \int_0^{t_{mat}} e^{-r_f \cdot \tau} \; dEL_{[a,b]}(\tau) \\
&=^{IntByParts} e^{-r_f \cdot t_{mat}} \cdot EL_{[a,b]}(t_{mat}) + \int_0^{t_{mat}} r_f \cdot e^{-r_f \cdot \tau} \cdot EL_{[a,b]}(\tau) \; d\tau,
\end{aligned}$$

with credit spread cs, expected loss $EL_{[a,b]}(t)$ on the tranche at time t,[19] and time to maturity t_{mat}. The nominal and the premium on it reduce with increasing losses on the tranche.[20] On the other hand, the value of the default leg increases successively with cumulative losses. Given the linear loss distribution $P(cL_t = l_k)$ in (28) we find the expected loss of the loss distribution at time t as

$$EL_{[a,b]}(t) = \sum_{k=1}^{M} P(cL_t = l_k) \cdot max(min(l_k, \widehat{b}) - \widehat{a}, 0), \tag{50}$$

with $l_k = k \cdot \frac{N}{M}(1-R)$, the loss due to the k-th default and $\widehat{b} = b \cdot N$, $\widehat{a} = a \cdot N$.

The representation of the integral in formula (28) as an expectation shows that $P(cL_t = l_k)$ can be obtained by simulating the standard normal V and taking the statistical average of the implied conditional probability. As this representation imposes high computational effort, at least when implemented in VBA for Excel, we decided to approximate by numerical integration, cf. figure 6,

$$P(cL_t = l_k) \approx \sum_{j=0}^{n-1} f_V(\xi_j) \cdot P(cL_t = l_k \mid \xi_j) \cdot \Delta, \tag{51}$$

with $\Delta = \frac{c_2 - c_1}{n}$ and $\xi_j = c_1 + j \cdot \Delta$, where we substituted the infinite integration borders by sufficiently large $c_1 << 0, c_2 >> 0$, and n. The LHP proposes an even faster method as it observes that

$$P(cL_t \leq l) = \phi\left(\frac{\sqrt{1-\rho^2} \cdot \phi^{-1}(\frac{l}{Nom(1-R)}) - \phi^{-1}(p_t)}{\rho}\right) = F_{\rho,p_t}(\omega) \tag{52}$$

holds for $M \longrightarrow \infty$, $\frac{k}{M} = \frac{l_k}{Nom(1-R)} \longrightarrow \frac{l}{Nom(1-R)} = \omega$. Thus for sufficiently large baskets we can use the approximation

$$P(cL_t \leq l_k) = \phi\left(\frac{\sqrt{1-\rho^2} \cdot \phi^{-1}(\frac{k}{M}) - \phi^{-1}(p_t)}{\rho}\right), \tag{53}$$

and in addition by using

$$\begin{aligned} P(cL_t = l) &= dF_{\rho,p_t}(\omega) \\ \Longrightarrow P(cL_t = l_k) &\approx P(cL_t \leq l_k) - P(cL_t \leq l_{k-1}) \end{aligned} \tag{54}$$

we obtain a computational less demanding formula at the cost of the additional assumption of a very large portfolio, such that $\frac{k_t}{M} \longrightarrow p_t^V$ for $M \to \infty$.

---

[19]Cf. (35).

[20]The loss distribution, the expected loss respectively, need to take into account that losses touch the N_T only from the attachment point on, in contrast to the linear loss distribution.

Unfortunately, as shown by figure 6 in the appendix, the normal distribution puts too much probability on low losses and too less probability on high losses. On top of this, it is not able to model *lower tail dependence*. Some authors thus heuristically replace the normal distribution by a t-distribution with appropriate degrees of freedom that puts more emphasis on the tails.[21]

In practice, the LHP is also called *HLPGC*, homogeneous large portfolio gaussian copula. The assumptions of this model were violated in summer 2005 when General Motors and Ford were substantially downgraded. At this time two kinds of structured trades were hit by this downgrade. Firstly, hedge fonds were invested long the equity tranche of iTraxx CDO while short the mezzanine tranche.[22] Secondly, capital structure speculators were short in shares but long in credits of GM and Ford. Both investment strategies created a positive carry, while they were hedged under the respective model assumptions applied in pricing.

The first strategy rests on the assumptions of the HLPGC model, that is, the spreads within the basket can not move differently and one unique correlation exists. But in the given case only GM and Ford changed substantially and thus the spread dispersion increased. This not only devalued the equity tranche relatively to the mezzanine tranche but also (further) biased the implied correlations of the different tranches which should be equal under the assumptions of HLPGC. Thus the theoretical hedge pretended by the HLPGC did not work. The second strategy is based on the assumptions of the Merton model instead, which excludes the possibility of a change in the capital structure of a company. But this was what happened, instead of default a restructuring took place, what is usually beneficial to the equity holders at the expense of debt holders. Thus, spreads went up while the share price held its level. In theory the share price would have been expected to fall with the downgrade.

Even though this simple model is not very suitable for pricing it is often used to calculate the implied correlation to have an *analog to the implied volatility* when comparing different tranches and trades. As it is an important notion for credit basket derivatives we now introduce compound correlation and base correlation.

4.2.3 Implied Correlation

For the understanding of implied correlation which is also called compound correlation it is crucial to understand the impact of higher/lower correlation within a given basket of credits on the joint default distribution.[23] Higher correlation makes

[21]Kalemanova et al. (2005). This is not only done for the marginal distribution, but also for the coupling function, because only then lower tail dependence is achieved. More precisely, the coupling distribution causes lower tail dependence or not, cf. [???].

[22]Even though Ford and GM are not itself members of iTraxx.

[23]Cf. Willemann (2004).

a high number of defaults and thus losses as well as a low number of defaults more likely, that is, it shifts mass to the tails, and lower correlation does the other way round. The *Compound Correlation* of a given tranche is the correlation that yields the quoted market spread, usually under the HLPGC model. From the other point of view, if we increase the correlation then mass is shifted to the tails and thus smaller losses decrease the spread for the equity tranche,[24] and higher losses increase the spread for the senior tranche. But matters are not clear as concerns the mezzanine tranches, and in fact the compound correlation does not increase monotonically over these tranches. More precisely, the compound correlation behaves like a smiling quadratic function over the mezzanine tranches, what leads to the term *Correlation Smile*. For this reason the notion of *Base Correlation* was introduced. Given the equity tranche $[0, a]$, mezzanine $[a, b]$ and senior $[b, 1]$of a CDO the base correlation is the implied correlation of a fictive equity tranche $[0, b]$ such that the spread implied by the expected loss $EL[a, b] = EL[0, b] - EL[0, a]$ equals the quoted market spread for the mezzanine tranche.[25] Thus the concept of base correlation yields an implied parameter that is monotonous with the credit spread, as implied correlation of equity tranches are monotonous. This allows to interpolate and derive correlations for off-market tranches and different tranches can be compared. In addition, market practice uses the base correlations together with HLPGC to evaluate the present value of the mezzanine tranche by $Pv[a, b] = Pv[b, 0] - Pv[0, a]$.

4.3 Sensitivities

With the valuation formula of the factor approach in place it is possible to compute analytical as well as numerical sensitivities that do not require MC simulation. Importantly, this applies not only to the usual credit spread or hazard rate sensitivity. The one-factor approach with single average correlation allows to compute a sensitivity with respect to changes in the average correlation too. Even though the analytically exact representation of this sensitivity is possible as the correlation is explicitly covered by the $p_t^{m|V}$, we will not state it here because it depends on the number of assets involved (the degree of the polynomial) and is thus sophisticated. We rather state the market practice via basis points:

$$\Delta_\rho = \frac{dP_{2bp}(\rho)}{d\rho_{2bp}} \approx \frac{P(\rho + 1bp) - P(\rho - 1bp)}{2bp}$$

$$\gamma_\rho = \frac{d\frac{dP_{1bp}(\rho)}{d\rho_{1bp}}}{d\rho_{1bp}} \approx \frac{P(\rho + 1bp) + P(\rho - 1bp) - 2 \cdot P(\rho)}{(1bp)^2}, \qquad (55)$$

[24] The possible loss of the equity tranche is capped by the attachment point and thus the increased probability of very high losses does not matter to the equity tranche.

[25] Given the correlation: formulas (15), (26), (28) and (50) yield the expected loss, and with the expected loss formula (49) yields the present value of the default leg that must equal the present value of the premium leg. The spread then follows from equation (48).

the use of a different distance for γ has the advantage that we do not need to calculate additional present values to obtain the second greek. The driving market parameters of basket CDS for which it is therefore beneficial to have sensitivities to verify quickly movements of present values are: recovery rates, correlations, and credit spreads. Credit spreads are used to determine the single name risk neutral default probabilities via bootstrapping which requires to determine the correct recovery rate in advance. So, not only the model but also the calculation and estimation of the relevant input parameters can imply different view points of OTC counterparts or between trading and risk control.

As concerns credit spread sensitivities the fair market spread is a value that not only aggregates adjustments of the underlying spreads but also correlation and recovery rates. The *Fair Spread* is the percentage premium that sets the present value of a CDO or CDS to zero. While a sensitivity based on the average spread of the whole basket proved to be unable to explain present value fluctuations of the tranche due to spread adjustments the fair spread explained movements of present value well.

5 Risk Neutral Default Probabilities

This section is concerned about how to extract single name risk neutral default probabilities from credit spreads given by the CDS or bond market, and finally it clarifies the relation between credit spreads and default intensities.

5.1 Default Intensity

Firstly, we introduce the concept of *Hazard Rates*, default intensities, respectively. Consider the conditional probability of a default in the interval $[t, t + dt]$ given survival until time t,

$$\begin{aligned} P(t \leq T \leq t + dt \mid T \geq t) = \frac{P(T \geq t) - P(T \geq t + dt)}{P(T \geq t)} &= \frac{F(t + dt) - F(t)}{1 - F(t)} \\ &= \frac{dF(t)}{1 - F(t)} = \frac{f(t)}{1 - F(t)}\ dt. \end{aligned} \tag{56}$$

Thus, to receive the conditional density of instantaneous default, $\lambda(t)$:

$$\begin{aligned} \lambda(t) = \frac{P(t \leq T \leq t + dt \mid T \geq t)}{dt} = \frac{f(t)}{1 - F(t)} = \frac{dF(t)/dt}{1 - F(t)} &= \frac{d(1 - S(t))/dt}{S(t)} \\ \Longrightarrow \lambda(t) &= -\frac{S'(t)}{S(t)}. \end{aligned} \tag{57}$$

The solution to this ordinary differential equation is

$$S(t) = e^{-\int_0^t \lambda(x) dx}, \tag{58}$$

the probability to survive until time t. If $\lambda(x)$ is a deterministic function this leads to inhomogeneous Poisson processes, and if $\lambda(x)$ is stochastic, but integratable, a Cox process follows. From the equation above we also learn *the hazard rate is a conditional default density function*. The relation between the process of default intensity $\lambda(t)$ and credit spreads is shown in section 5.3.

5.2 Bootstrapping

The preceding paragraphs assumed the single name default probabilities as given and just brought them together within a basket. We now want to shade some light on how to obtain them from given market data.

5.2.1 CDS Spreads

For the bootstrapping of the term structure of default probabilities the curve of par CDS spreads is used in practice. These are spreads of CDSs with various maturities, each with a net present value of zero, and same credit risk. The net present value of zero ensures the spread to reflect current credit risk.

Let $AI_P(t)$ be the accrued interest function of the premium payment at time t, and $AI_U(t)$ the accrued interest of the reference obligation of the CDS. Both functions behave like a sawtooth during the lifetime of the CDS, i.e. they increase from zero to the next premium payment and decline then sharply to zero and start to increase again. The *recovery rate plays a crucial role* in the process of bootstrapping. In case of default the protection buyer receives a payment that increases the remaining market value M_t of the bond to reach the claim of the bond holder. If the bond holder claims nominal plus accrued interest, $N + AI_U(t)$, the recovery rate is obtained as $R = M_t/(N+AI_U(t))$, and $R = M_t/N$ if accrued is neglected. During the discussion of the bootstrapping we will investigate the impact of using $R \cdot (N + AI_U(t))$ or the proxy $R \cdot N$ among the aspects of how different CDS pricing formulas influence the induced default probability. If bondholders apply the best-claim-assumption the protection seller needs only to pay $N - R(N + AI_U(t))$ instead of $N - RN$, and the pricing formula is[26]

$$
\begin{aligned}
Pv_{CDS} = \int_{t_0}^{t_k - t_0} & (N - R(N + AI_U(t))) \cdot dfac(t)\ dF(t) \\
& - \sum_{i=1}^{k} cs_k \cdot N \cdot dfac(t_i) \cdot (1 - F(t_i)) \\
& - \int_{t_0}^{t_k - t_0} AI_P(t) \cdot dfac(t)\ dF(t),
\end{aligned} \tag{59}
$$

allowing an default event at every time until maturity. In the first term we can interpret $dF(t)$ as the probability of default at time t. The second term is the present value of the discrete time premium payments (usually quarterly) given survival, and the third term accounts for the present value of accrued premium in case of default. The last term is of only little impact because the relatively little accrued figures are in addition weighted by the small probabilities of default. If we neglect the impact of accrued interest and refine in view of the bootstrapping to a valuation at deal inception, i.e. $t_0 = 0$, but keep the accrued premium, the pricing simplifies to

$$
\begin{aligned}
Pv_{CDS} = \int_{0}^{t_k} (1 - R)N \cdot dfac(t)\ dF(t) - \sum_{i=1}^{k} cs_k \cdot N \cdot dfac(t_i) \cdot (1 - F(t_i)) \\
- \int_{0}^{t_k} AI_P(t) \cdot dfac(t)\ dF(t).
\end{aligned} \tag{60}
$$

If we assume in addition, that default can only occur at the premium payments the discrete version often used in practice is

$$
Pv_{CDS} = \sum_{i=1}^{k} (1 - R)N \cdot dfac(t_i) \cdot (F(t_i) - F(t_{i-1})) - \sum_{i=1}^{k} cs_k \cdot N \cdot dfac(t_i) \cdot (1 - F(t_i)). \tag{61}
$$

Now we consider quoted (current par) CDS spreads $(cs_1, ..., cs_n)$ of a given reference obligation, comparable credit risk, respectively, for CDSs with times to maturity $t_1 - t_0, ..., t_n - t_0$ and $t_0 = 0$. The respective PVs are zero and hence the quoted par spreads reflect current credit risk. Using (61) for each $k \in \{1, ..., n\}$ we end up with n equations that can be successively solved for the $F(t_i)$ starting with $k = 1$

$$
\begin{aligned}
dfac(t_1) \cdot (1 - R) \cdot N \cdot F(t_1) - dfac(t_1) \cdot cs_1 \cdot N \cdot (1 - F(t_1)) = 0 \\
\Longleftrightarrow F(t_1) = \frac{cs_1}{1 - R + cs_1}.
\end{aligned} \tag{62}
$$

[26]We denote the original time to maturity with t_k in preparation for the recursive formula derived in section 5.2.2, and cs_k is the deal premium which at inception ($t_0 = 0$) equaled the par spread for a duration of $t_k - 0$. The valuation takes place after $t_0 \geq 0$ and $t_k - t_0$ is the remaining time to maturity.

Formula (61) implies that in case of default until t_1 no premium payment is done. This is the *first case*. If we assume in contrast that the premium has to be paid in case of default the 1-period default probability becomes

$$\begin{aligned} dfac(t_1) \cdot (1-R) \cdot N \cdot F(t_1) - dfac(t_1) \cdot cs_1 \cdot N = 0 \\ \iff F(t_1) = \frac{cs_1}{1-R} \end{aligned} \tag{63}$$

instead, which is the *second case*. And if we take the coupon into account, i.e. bond holders claim $N \cdot (1+c)$ instead of only N, together with a premium in case of default the formula becomes[27]

$$\begin{aligned} dfac(t_1) \cdot (N - RN \cdot (1+c)) \cdot F(t_1) - dfac(t_1) \cdot cs_1 \cdot N = 0 \\ \iff F(t_1) = \frac{cs_1}{1-R-R(r_f+cs_1)}, \end{aligned} \tag{64}$$

the *third case*. In the *last case* bond holders again claim $N(1+c))$ but no premium is paid under default

$$\begin{aligned} dfac(t_1) \cdot (N - RN \cdot (1+c)) \cdot F(t_1) - dfac(t_1) \cdot cs_1 \cdot N(1-F(t_1)) = 0 \\ \iff F(t_1) = \frac{cs_1}{1-R-R(r_f+cs_1)+cs_1}, \end{aligned} \tag{65}$$

Before we provide the recursive formula that bootstraps the risk neutral default probabilities in Section 5.2.2, we discuss how $F(t_1)$ behaves under the different assumptions above.

In contrast to the first case we find the second case to have a premium payment in case of default. Hence, all other things held equal, the default probability is increased to give the insurance payment a higher weight, this is accomplished by omitting the credit spread in the nominator. Compared to the second case, the third case reduces the insurance payment in case of default by Rc and thus the default probability must increase further by additionally subtracting Rc in the nominator. In contrast, the last case reduces the default probability compared to the third case, because no premium is payed in case of default. In summary, we find that how the premium payment is treated determines whether the credit spread appears in the nominator or not, and the applied claim determines whether Rc appears in the nominator or not.

It is important to note, even though all 4 cases should work in practice, only one of them yields formulas that are consistent with stochastic theory even for extreme values of R. Let $R \to 0$, $R \to 1$ respectively, then in the first case $F(t_1) \in [cs/(1+cs), 1]$ which is consistent, but for the second case we have $F(t_1) \in [cs, \infty)$, allowing for impossible probabilities. The third case yields $F(t_1) \in [cs, -cs/(r_f + cs)]$, allowing for

[27] Thereby, the reference obligation is assumed to have the same payment frequency for its coupon payments as the premium payments of the CDS, which is usually quarterly. Then $N + AI_U(t)$ becomes $N + cN = (1+c)N$ where $c = r_f + cs$ is the coupon of the defaultable bond.

negative probabilities, and the forth case too with $F(t_1) \in [cs/(1+cs), -cs/r_f]$ implies negative probabilities. Case two disqualifies also for $cs \to \infty$ with $F(t_1) \to \infty$, while all others yield $F(t_1) \to 1$.

5.2.2 Recursive Formula for Bootstrapping

We now want to discuss how the actual bootstrapping can be performed. With formula (61) we neglect the impact of the accrued premium. In order to account for its effect in a heuristic but easy to implement manner we extend this formula by the term

$$-\sum_{i=1}^{k} \frac{1}{2} \cdot cs_k \cdot N \cdot dfac(t_i) \cdot (F(t_i) - F(t_{i-1})). \tag{66}$$

That is, we estimate the integral in (60) by an approximation which assumes that in case of a default between two premium payments this default will be exactly in the middle.[28] Alternatively, we could have assumed an inhomogenous Poisson process, implying $F(t_k) = 1 - exp(-\sum_{i=1}^{k} \lambda_i \cdot (t_i - t_{i-1}))$, cf. (4), and evaluate the two integrals under this assumption "exactly".

We further note that formula (61) and its extended version are to be solved for each $k \in \{1, ..., n\}$ for a given spread curve $(cs_1, ..., cs_n)$.[29] With the extension above (writing F_i instead of $F(t_i)$) we obtain for the consistent first case (62) and $k = 1$

$$F_1 = \frac{cs_1}{1 - R + \frac{1}{2} \cdot cs_1}, \tag{67}$$

that serves as a starting point for the recursive formula

$$\begin{aligned} F_k = & \frac{(1-R) - \frac{1}{2} \cdot cs_k}{(1-R) + \frac{1}{2} \cdot cs_k} \cdot F_{k-1} + \frac{cs_k}{(1-R) + \frac{1}{2} \cdot cs_k} \\ & + \frac{A_{k-1}}{(1-R) \cdot dfac(t_k) + \frac{1}{2} \cdot cs_k \cdot dfac(t_k)}, \end{aligned} \tag{68}$$

which is obtained if we rearrange the extended version at a given k for F_k. The term A_{k-1} is found as (difference between conditional leg and premium leg)

$$\begin{aligned} -\sum_{i=1}^{k-1} (1-R) \cdot dfac(t_i) \cdot (F_i - F_{i-1}) + \sum_{i=1}^{k-1} cs_k \cdot dfac(t_i) \cdot (1 - F_i) \\ + \sum_{i=1}^{k-1} \frac{1}{2} \cdot cs_k \cdot dfac(t_i) \cdot (F_i - F_{i-1}), \end{aligned}$$

[28] In order to be consistent we would need to extend the sum of conditional payments from n many summands to $2n - 1$ many summands, i.e. to refine the approximation of the respective integral. We omit this here intentionally.

[29] The credit spreads to be used in the formula are quarterly spreads. Thus, annualized spreads need to be divided by 4.

and can be interpreted as a *correction that allows for a non-flat credit spread term structure*. We will outline this issue in more detail below. If we do not take into account the accrued premium the recursive formula simplifies to

$$\begin{aligned} F_k = & \frac{1-R}{(1-R)+cs_k} \cdot F_{k-1} + \frac{cs_k}{(1-R)+cs_k} \\ & + \frac{A_{k-1}}{(1-R) \cdot dfac(t_k) + cs_k \cdot dfac(t_k)} \end{aligned} \tag{69}$$

with a starting value of

$$F(t_1) = \frac{cs_1}{1-R+cs_1}, \tag{70}$$

and term A_{k-1} is then found as

$$-\sum_{i=1}^{k-1}(1-R) \cdot dfac(t_i) \cdot (F_i - F_{i-1}) + \sum_{i=1}^{k-1} cs_k \cdot dfac(t_i) \cdot (1-F_i). \tag{71}$$

Which is just the recursive previous equation that is satisfied by F_{k-1}. As already mentioned above, the last term of the recursive formulas can be interpreted as the correction implied by a non-flat credit spread curve. This can be seen when a flat credit spread term structure is assumed. Then the term A_{k-1} vanishes, because in this case the spread cs_k not only satisfies its equation but all preceding equations, including the equation $A_{k-1} = 0$ with given $(F_1, ..., F_{k-1})$, yielding the simple formula

$$F_k = \frac{1-R}{(1-R)+cs_k} \cdot F_{k-1} + \frac{cs_k}{(1-R)+cs_k}. \tag{72}$$

To put it another way: $(F_1, ..., F_k), (cs_1, ..., cs_k)$ satisfy $A_k = 0 \forall k \in \{1...n\}$.

In the next section we consider an alternative way to bootstrap the default probabilities using the coupons of traded par bonds. Par bonds are used because for them $c - r_f$ equals the current credit spread. We will see that the different assumptions above imply different assumptions on coupon payments and claim amount of the risk free bond, defaultable bond respectively, that make up the replicating portfolio.

5.2.3 Bond Spreads

Even though in practice most often par credit spreads from liquidly traded single name CDS and the respective pricing formula are used, we want to show the bootstrapping of risk neutral default probabilities using par bonds too. The reason for that is threefold. Firstly, we can show how to replicate and price a CDS with a portfolio that consists of a defaultable bond and a credit risk free bond of identical maturities. Secondly, we can show that credit spreads from defaultable bond prices should equal the quoted credit spreads for single name CDS with matching reference obligation. Thirdly, we gain an impression of how the default probabilities react to different pricing assumptions.

As for the CDS, we assume that a default can only occur at the coupon dates $t_1, ..., t_n$, and we do not incorporate the impact of accrued interest.

We pick up the *first case* of the previous section on CDS spreads where it is assumed that no premium is paid in case of default, and that the bond holders only claim N instead of $N(1+c)$ as a proxy for the market value. The position of a protection buyer can be replicated by buying a credit risk free bond and short selling a defaultable bond. To see this we consider the cash flows prior to default and the closing payments under default of this replicating portfolio. Prior to default the replicating portfolio receives r_f and pays $c = r_f + cs$ to the holder of the short sold defaultable bond, implying a net cash flow of cs, which exactly what a CDS protection buyer pays as a premium. In case of default the risk free bond is sold at par, because r_f is constant over time, and $R \cdot N$ is paid to the holder of the defaulted bond, implying a closing net cash flow of $(1 - R)N$, which is again exactly what a protection buyer would receive from a CDS under the first assumptions. Note, no risk free coupon and no risky coupon is paid here in case of default, which is in line with the simplified formula (61) and implies the valuation of the replicating portfolio to be

$$\begin{aligned} Pv = 0 = -\sum_{i=1}^{n} dfac(t_i) \cdot R \cdot N \cdot (F(t_i) - F(t_{i-1})) - dfac(t_n) \cdot N \cdot (1 - F(t_n)) \\ -\sum_{i=1}^{n} dfac(t_i) \cdot c \cdot N \cdot (1 - F(t_i)) \\ +dfac(t_n) \cdot N + \sum_{i=1}^{n} dfac(t_i) \cdot r_f \cdot N \cdot (1 - F(t_i)). \end{aligned} \tag{73}$$

The first term is the expected Pv of nominal in case of default prior to maturity. The second term is the expected Pv of nominal given no default. The third term in the second row is the expected Pv of coupon payments, having a coupon that has been scaled to the periodicity of the coupon payments. The forth term is the present value of the nominal of the risk free bond, and *the last term, in contrast to the present value formula of a risk free bond, weights the coupon payments by the probability of survival, because we assumed that in case of default no coupon is payed.* In a 1-period framework the above formula simplifies to

$$\begin{aligned} 0 = -dfac(t_1) \cdot R \cdot N \cdot F(t_1) - dfac(t_1) \cdot N \cdot (1 - F(t_1)) \\ -dfac(t_1) \cdot c \cdot N \cdot (1 - F(t_1)) \\ +dfac(t_1) \cdot N + dfac(t_1) \cdot r_f \cdot N \cdot (1 - F(t_1)), \end{aligned} \tag{74}$$

and further

$$\begin{aligned} 0 = -R \cdot F(t_1) - (1 - F(t_1)) - c \cdot (1 - F(t_1)) + 1 + r_f \cdot (1 - F(t_1)) \\ \Longleftrightarrow F(t_1) = \frac{cs}{1 - R + cs}, \end{aligned} \tag{75}$$

which is exactly what we derived for the respective CDS in formula (62). It can be concluded that the assumptions for the pricing of the CDS and their implications are made clear by the replication. In order to arrive at the *second case* where the premium of the CDS is paid under default our replication needs to assume that in case of default not only the risk free coupon but also the coupon of the defaulted bond is is still paid, yielding

$$\begin{aligned} 0 &= -R \cdot F(t_1) - (1 - F(t_1)) - c + 1 + r_f \\ &\iff F(t_1) = \frac{cs}{1-R}. \end{aligned} \tag{76}$$

The *third case* can be replicated by a portfolio in which the recovered value is $R(1+c)N$ instead of RN, and where again not only the risk free but also the coupon of the defaulted bond is paid

$$\begin{aligned} 0 &= -R \cdot (1+c) \cdot F(t_1) - (1 - F(t_1)) - c + 1 + r_f \\ &\iff F(t_1) = \frac{cs}{1-R-Rc}. \end{aligned} \tag{77}$$

The *forth case* follows from our portfolio if we assume that bond holders claim $N(1+c)$, but neither the risk free nor the coupon of the defaulted bond is paid under default, implying

$$\begin{aligned} 0 &= -R \cdot (1+c) \cdot F(t_1) - (1 - F(t_1)) - c \cdot (1 - F(t_1)) + 1 + r_f \cdot (1 - F(t_1)) \\ &\iff F(t_1) = \frac{cs}{1-R-Rc+cs}. \end{aligned} \tag{78}$$

For a graphical comparison of the resulting 1-period default probabilities $F_1(t_1), ..., F_4(t_1)$ we let $R \in [0, ..., 0.6], cs \in [0, 0.1]$ and set $r_f = 0.04$. In addition we slice this 3-dim depiction at $R = 0.45$ for a 2-dim depiction, cf. Figure 7 in the appendix.

We used a defaultable and a risk free bond to replicate the CDS. An alternative procedure is to apply an asset swap, see appendix. If one wants to buy credit risk only with an asset swap he buys a bond from the counterpart while entering into an interest rate swap that pays a floating rate and receives the swap rate. The bond is pledged as collateral to the RePo counterpart that pays the nominal minus a haircut and receives a floating rate for this repurchase agreement.

It was outlined above that in theory spreads from the CDS and the bond market should equal. We now want to discuss why this relationship does not exactly hold in practice.[30]

5.2.4 Default Swap Basis

Before we go into details it should be stressed that the theoretical equality rests on arbitrage arguments that require market participants to act in both market segments

[30] Cf. Zhu (2004).

- the cash market and the CDS market. At this point we introduce the notion of the *Default Swap Basis* (=DSB). If we denote by cs_b the par bond spread and by cs_p the premium of a par CDS then the default swap basis is defined by

$$DSB = cs_p - cs_b. \tag{79}$$

As concerns the iBoxx and iTraxx indices the DSB became negative up to 10bp during January 2006, indicating an increasing segmentation of the market. In our case the Market Segmentation Theory states that "different groups of credit players focus on specific instruments. While the whole structured community [...] is playing only in the CDx arena, the majority of [...] asset managers is still skewed to cash bonds.".[31] Each segment has its own supply-demand equilibrium. Thus a lack of arbitrage can also increase the gap between the two spread types.[32]

A further reason for a non-zero DSB is that due to a lack of par bonds the market also uses bonds above and below par. If a bond is above par then $c - r_f > cs_p$ yields a too high default probability, and the other way round for bonds below par. A higher default probability also narrows the DSB. Second reason is because of the fact that in case of default the accrued premium is usually to be paid by the protection buyer of a CDS. In contrast, the replicating portfolio is likely to yield the risk free rate but not the interest payment of the defaulted bond. This decreases the CDS permium compared to the bond spread and decreases DSB in practice. On top of this, the counterpart risk of the protection seller is likely to become a problem in case of default. The premium thus decreases relatively to a bond spread.

5.3 Hazard Rate and Credit Spread

It is often mentioned that the credit spread can be viewed as a hazard rate or default intensity when speaking of the reduced form model. But it should be stressed that this statement holds only for the continuously compounded credit spread and only under certain conditions. It does not hold for the kind of ordinary credit spread discussed in section 5.2. In order to bring the continuously compounded credit spread of a bond cs_b^c in we also consider the continuously compounded risk free rate r_f^c and assume that in case of default $(1+c)\cdot R\cdot N$ is claimed with respective recovery rate R, and $c = r_f + cs_b$ the coupon made up by the ordinary values. In this case the valuation formula of a 1-period bond becomes

$$Pv_1 = e^{-r_f^c \cdot t_1} \cdot (F(t_1) \cdot (1+c) \cdot R \cdot N + (1 - F(t_1)) \cdot (1+c) \cdot N) \,. \tag{80}$$

[31] Cf. Felsenheimer et al. (2006), p. 2.

[32] When arbitrage works and the DSB is negative one short sells the DSB, that is to say, he buys a cash bond (receives $cs_b + r_f$) and hedges its credit risk with a CDS (paying cs_p). Assuming his funding cost for the bond price via RePo agreement is r_f his arbitrage is $cs_b - cs_p$ which is the negative DSB.

We now replace Pv_1 by $e^{-(r_f^c+cs_b^c)t_1} \cdot (1+c) \cdot N$, i.e. we equate the risk neutral with the standard approach of pricing, and simplify

$$e^{-cs_b^c \cdot t_1} = F(t_1) \cdot R + 1 - F(t_1) \Longleftrightarrow 1 - e^{-cs_b^c \cdot t_1} = (1-R) \cdot F(t_1). \tag{81}$$

Taking into account the reduced form representation of $F(t_1) = 1 - e^{-\lambda t}$ with hazard rate λ it can be concluded that

$$1 - e^{-cs_b^c \cdot t_1} = (1-R) \cdot (1 - e^{-\lambda t_1}), \tag{82}$$

and thus for $R = 0$ the continuously compounded credit spread can be seen as an hazard or default intensity.

6 Term Structure Models

In section 4.1.3 we introduced a simple model to circumvent the need of term structure models to determine the principal cash flows from the underlying pool. Due to the vocal role of term structure models in financial engineering in general and its importance when modelling the time structure of cash flows out of the underlying pool of debt of a Cash CDO in particular the Black Derman Toy spot rate model is outlined subsequently.

6.1 A Modification of the Classical Tree Structure

The classical tree representation as it is well known from the binomial tree of stock options is replaced by a modification that allows for an easier calibration and a more elegant analytical treatment. In contrast to the classical tree at each time i the possible rate can be defined independently from the previous rates in the tree. They depend only on the minimum spot rate s_{i0} and the *Black Volatility* σ_i, implying that the sequences $(s_{i0})_i$ and $(\sigma_i)_i$ with $i \in \{1, ..., T\}$ need to be choosen as such that the resulting tree matches the current volatility structure as well as the current term structure of interest rates. Each possible short rate s_{ij} is characterized by its time i and its state of the world j.

The shown binomial process describes the future behavior of the instantaneous short rate $s(t)$. With respect to the real world expectation operator for this rate

$$E(s(t)) = f(t) \overset{\Delta \to 0}{=} f(t, t+\Delta) \tag{83}$$

holds. Thereby, $f(t, t+\Delta)$ is the observed forward rate today for investments at time $t > 0$ with a duration of length Δ in years. In practice we consider approximations for $\Delta \in \{\frac{1}{250}, \frac{1}{52}, \frac{1}{12}, \frac{1}{4}, \frac{1}{2}\ \}$, that is daily, weekly, monthly, quarterly or half-yearly spot rates. Instead of the real world measure above the model under investigation

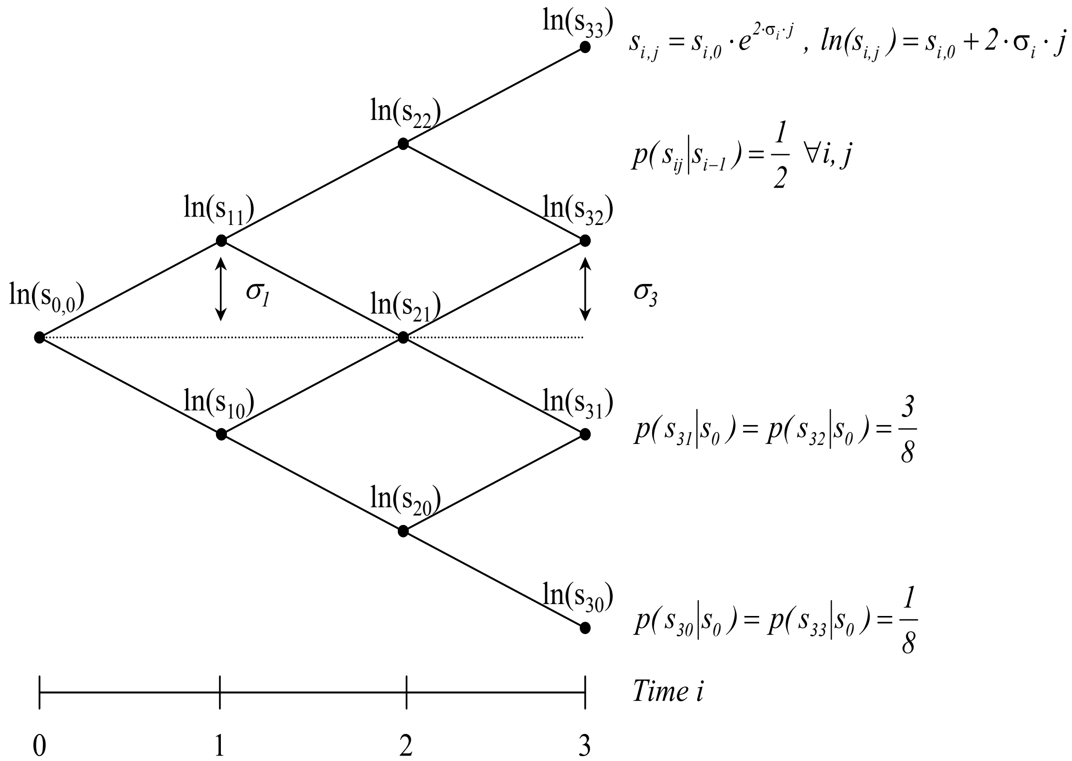

Figure 2: BDT Binary Tree

applies $E^A(s(t))$, which is the expected value with respect to the risk neutral measure A and for which the above equation does not hold. We are thus not able to calibrate by just fitting the first two moments of the process. While it is possible to fit the volatility directly, the expected value requires an indirect calibration that adjusts the $s_{i,0}$ until the valuation of selected bonds matches their current market value.

6.2 Risk Neutral Process

Risk neutral valuation distinguishes between two measures: the *classical risk neutral measure* and the *forward risk measure*. With the risk neutral measure today's value of a financial instrument equals the expected value of the discounted future value of this instrument, that is to say

$$Pv(0) = E_0^A(e^{-\int_0^t s(\tau)d\tau} \cdot Fv(t)). \tag{84}$$

As $s(t)$ is stochastic at time 0 it can not be drawn out of the expectation operator. Under the forward risk neutral measure, on the other hand, the present value of a financial instrument equals the discounted expected value of this instrument

$$Pv(0) = e^{-\int_0^t f(\tau)d\tau} \cdot E^B(Fv(t)). \tag{85}$$

Both approaches have its merits, one specific of the risk neutral measure is

$$\begin{aligned} e^{-\int_t^{t_1} f(\tau)d\tau} &= E^A\big(e^{-\int_t^{t_1} s(\tau)d\tau}\big) \\ \Rightarrow e^{-f(t)} &= E^A\big(e^{-s(t)}\big) \end{aligned} \tag{86}$$

The instantaneous short can be written as

$$e^{-s(t)\cdot\Delta} = \lim_{\Delta\to 0} B(t, t+\Delta) \Leftrightarrow s(t) = -\lim_{\Delta\to 0} \frac{ln(B(t, t+\Delta))}{\Delta}. \tag{87}$$

The non-linear relationship shows that $s(t)$ and its one-day approximation can not be linearly replicated by tradeable instruments. This influences the structure of the risk neutral process. In the classical case, that is for a tradeable underlying, not only the volatility function is invariant under risk neutralization but also the drift component and even the whole process is uniquely determined by the volatility function. The drift of the binomial tree of the short rate, on the other hand, needs to be defined as such to help the risk neutral process to fit volatility and interest rate term structure, and is not determined by the volatility function alone. From another point of view, this is due to the fact that with the above limit to zero some information on the term structure is lost. That is to say, the risk neutral process of the short rate is not only specified by the volatility function $\sigma_i, i \in \{0, ..., \infty\}$ but also the drift s_{i0} is to be calibrated. This implies that one starts to calibrate the risk neutral process directly instead of calibrating its real world equivalent and transforming it into a risk neutral form. The difference can also be seen from the treatment of the risk neutral probabilities. While in the classical case the martingale measure, i.e. the values $P(s_{ij}|s_{i-1})$, are choosen as to yield a martingale under discounting with the market-risk free rate r_f ($E^A(s(t) \cdot e^{-r_f \cdot t}) = s_0$), in our case we deviate from this approach by setting the risk neutral probabilities arbitrarily to 1/2 and appropriately determine the drift instead.

6.3 Specific Key Values of the Binomial Tree

6.3.1 Expected Value of the log-rate

A heuristic analysis of the binomial tree, which could be proofed formally via complete induction, shows the unconditional probability to be

$$P(s_{ij}) = \binom{i}{j} \cdot \frac{1}{2}^i, \quad i \geq 0, \quad j \in \{0, ..., i\}. \tag{88}$$

Due to

$$ln(s_{ij}) = ln(s_{i0}) + 2 \cdot \sigma_i \cdot j \tag{89}$$

this yields

$$E^A(ln(s_i)) = \sum_{j=0}^{i} \binom{i}{j} \cdot \frac{1}{2}^i \cdot (ln(s_{i0}) + 2 \cdot \sigma_i \cdot j)$$

$$= ln(s_{i0}) + \sigma_i \underbrace{\sum_{j=0}^{i} \binom{i}{j} \frac{1}{2}^i \cdot 2j}_{=i} = ln(s_{i0}) + \sigma_i \cdot i \qquad (90)$$

6.3.2 Volatility

With the expected value given the unconditional variance follows analogously

$$V(ln(s_i)) = V(ln(s_i|s_0)) = \sum_{j=0}^{i} \binom{i}{j} \cdot \frac{1}{2}^i \cdot (ln(s_{i0}) + 2 \cdot \sigma_i \cdot j - ln(s_{i0}) - \sigma_i \cdot i)^2$$

$$= \sigma_i^2 \cdot \underbrace{\sum_{j=0}^{i} \binom{i}{j} \cdot \frac{1}{2}^i \cdot (2j - i)^2}_{=i} = \sigma_i^2 \cdot i, \qquad (91)$$

and the conditional variance with time lag one is

$$V(ln(s_i|s_{i-1})) = \sigma_i^2 \qquad (92)$$

As the variance is independent of the measure we have omitted this in the notation of variance. For a as simple as possible calibration of the model on annualized swaption volatilities $\sigma^a_{ex,ex+\Delta_f}$ of a swaption that expires in ex and whose swap has a time to maturity of Δ_f years, as they are published by Reuters and Bloomberg, we consider the generalized conditional variance of the log Δ_f-period swap rate with $\Delta_f = k \cdot \Delta_s$

$$V(ln(s_{ex+k \cdot \Delta_s}|s_{ex})) = \sigma_{ex}^2 \cdot k, \qquad (93)$$

where k is how often the periodicity Δ_s of the tree fits into the time to maturity of the swap.

6.4 Calibration of Volatility

The most secure way to calibrate the model is to adjust volatility and drift parameters during an iterative procedure as such to match quoted swaption and bond prices.

In order to circumvent the first part we want to motivate how and why one is able to calibrate the model directly to quoted swaption volatilites. For this we consider a swaption with expiry date ex on a swap whose time to maturity is Δ_f in years. The respective underlying rate is then the forward rate $f(ex, ex + \Delta_f)$. The annualized short rate $s(t)$ is infinitesimal i.e. it has periodicity dt. The corresponding discount factor for the time frame $[ex, ex + \Delta_f]$ is

$$df(ex, ex + \Delta_f) = e^{-\int_{ex}^{ex+\Delta_f} s(t)dt} \tag{94}$$

To derive an analogous discount factor with respect to the annualized forward swap rate $f(ex, ex + \Delta_f)$, which will be denoted by f_{ex,Δ_f}, we again refer to periodicities $h_f, h_s \in \{250, 52, 12, 4, 2\}$ i.e. the periodicity in days, weeks, months, ... and more importantly in years $\Delta_{f,y}, \Delta_{s,y} \in \{\frac{1}{250}, \frac{1}{52}, \frac{1}{12}, \frac{1}{4}, \frac{1}{2}\}$. The discount factor as a discretization of the above integral is

$$df_s(ex, ex + \Delta_f) = e^{-\sum_{i=1}^{n} s_{ex+i\cdot\Delta_s}\cdot\Delta_s}, \quad n = \Delta_f \cdot h_s \overset{\Delta_f=\Delta_s}{=} 1 \tag{95}$$

That is if period Δ_s is shorter than period Δ_f it fits n-times into time to maturity of the swap. We now consider the common discounting with the annualized forward rate and the corresponding continuous compounding as a possible approximation:

$$\begin{aligned} df_f(ex, ex + \Delta_f) &= \left(\frac{1}{1 + f_{ex,\Delta_f}}\right)^{\Delta_f} \\ \underset{k\geq 1}{\approx} \left(\frac{1}{\left(1 + \frac{f_{ex,\Delta_f}}{k}\right)^k}\right)^{\Delta_f} &\underset{k\to\infty}{=} e^{-f_{ex,\Delta_f}\cdot\Delta_f} \end{aligned} \tag{96}$$

If we equate these two discount factors, $df_f = df_s$ we arrive at

$$\begin{aligned} e^{-f_{ex,\Delta_f}\cdot\Delta_f} &= e^{-\sum_{i=1}^{n} s_{ex+i\cdot\Delta_s}\cdot\Delta_s} \\ \Leftrightarrow f_{ex,\Delta_f} = \frac{\Delta_s}{\Delta_f}\cdot\sum_{i=1}^{n} s_{ex+i\cdot\Delta_s} &= \frac{1}{n}\cdot\sum_{i=1}^{n} s_{ex+i\cdot\Delta_s} \end{aligned} \tag{97}$$

As the short rates, given the previous short rate, are independent during time the variance $V(f_{ex,\Delta_f})$ is easy to calculate for arbitrary $n \geq 1$, becoming

$$V(f_{ex,\Delta_f}) = \frac{1}{n^2}\cdot\sum_{i=1}^{n} V(s(ex + i\cdot\Delta_s | ex + (i-1)\cdot\Delta_s)). \tag{98}$$

Unfortunately, we need volatilities for $ln(f_{ex,\Delta_f})$ instead, and restrict thus to $n = 1$. With $n = 1$ we choose $\Delta_s = 1/4$ or $\Delta_s = 1/2$ that is a 3-month or 6-month periodicity and can thus directly calibrate to the corresponding caplet or swaption volas.

This means, that in order to ease the calibration to volatilities as much as possible we accept that the periodicity of our model is dictated by the available swaption volas.

6.5 Calibration of Drift

As we have already mentioned above, given the correct volas, one would usually calibrate the model to quoted bond prices. Based on this approach we propose an alternative way to calibrate the model to the yield curve or swap curve. To accomplish this, we consider artificial Zero Bonds whose prices stand for the respective swap rates. The drift of the model is then adjusted to match the prices of these artificial Zero Bonds. One starts with the bond shortest time to maturity and correspondingly adjusts the drift until this time to maturity. While the drift until this date is kept constant in the following procedure one takes the next higher time to and adjusts the drift again until the next higher time to maturity, and so on, cf. figure .

7 Appendix

7.1 Proof

From section 3.3 we know $p_t = P(T \leq t) = P(Y_t \leq y_t) = \phi(y_t)$ and $y_t = \phi^{-1}(p_t), p_t = e^{-\lambda t}$. If $F(t) = e^{-\lambda t}$ is the distribution of default times and $u \in (0, 1]$ we have $F(t) = \phi(y_t) = u$ and $F^{-1}(u) = t, \phi^{-1}(u) = y_t$. We can therefore create random variables $T = F^{-1}(U), Y_t = \phi^{-1}(U)$ by the same seed U, and after having substituted U we receive $T = F^{-1}(\phi(Y_t)) = -\frac{1}{\lambda} \cdot ln(\phi(Y_t))$. The correlation of two default times is thus $Corr(T_1, T_2) = Corr(-\frac{1}{\lambda} \cdot ln(\phi(Y_{1,t})), -\frac{1}{\lambda} \cdot ln(\phi(Y_{2,t})))$. If we apply calculation rules for correlation we recognize that this equals $Corr(ln(\phi(Y_{1,t})), ln(\phi(Y_{2,t})))$.

7.2 Implementation

Required market data are:

1. Payment dates: $t_1, ..., t_{mat}$
2. Implied correlation that calibrates to market quotes: ρ
3. Term structure of average over single name default probabilities: $p_{t_1}, ..., p_{t_{mat}}$
4. Number of members in the basket: M
5. Percentage borders of tranche: a, b (attachment, detachment point)
6. Nominal of tranche: $N_T = (b - a) \cdot N \Leftrightarrow N = \frac{N_T}{b-a}$
7. Initial fair spread: cs

8. Swap curve: $r_{t_1}, ..., r_{t_{mat}}$

and then following formulas need to be implemented:

1. Formula (15)
2. Formula (26) and (28) or just (54) when $M \to \infty$
3. Formula (50) with $l_k = k \cdot \frac{Nom}{M} \cdot (1 - R)$
4. Formula (48), (49)

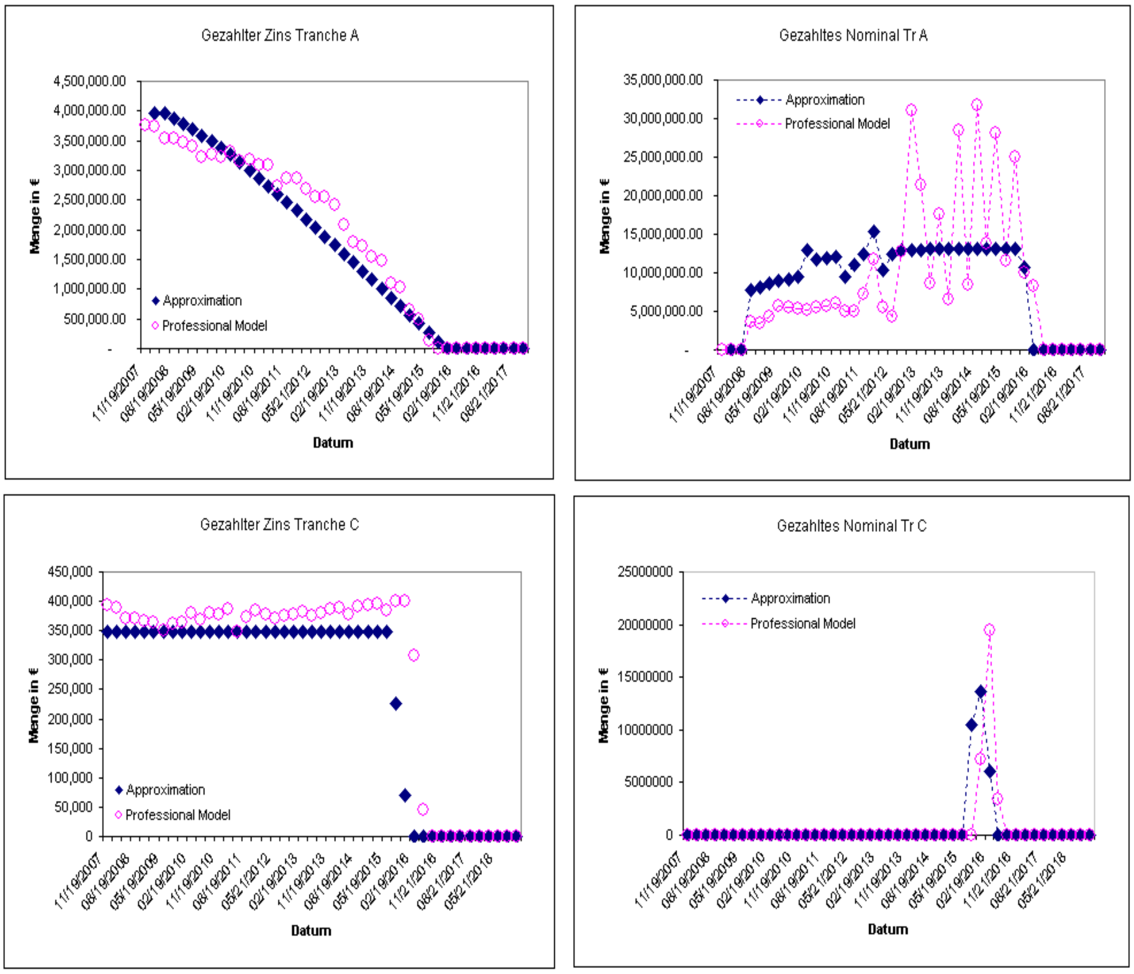

Figure 3: Comparison with Professional Model

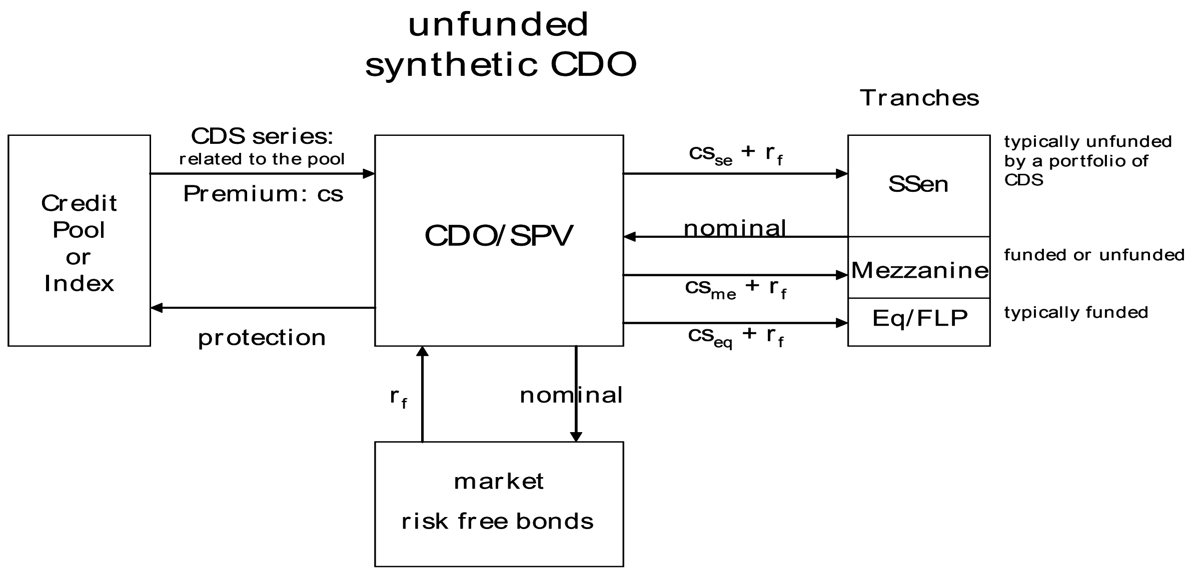

Figure 4: Synthetic CDO

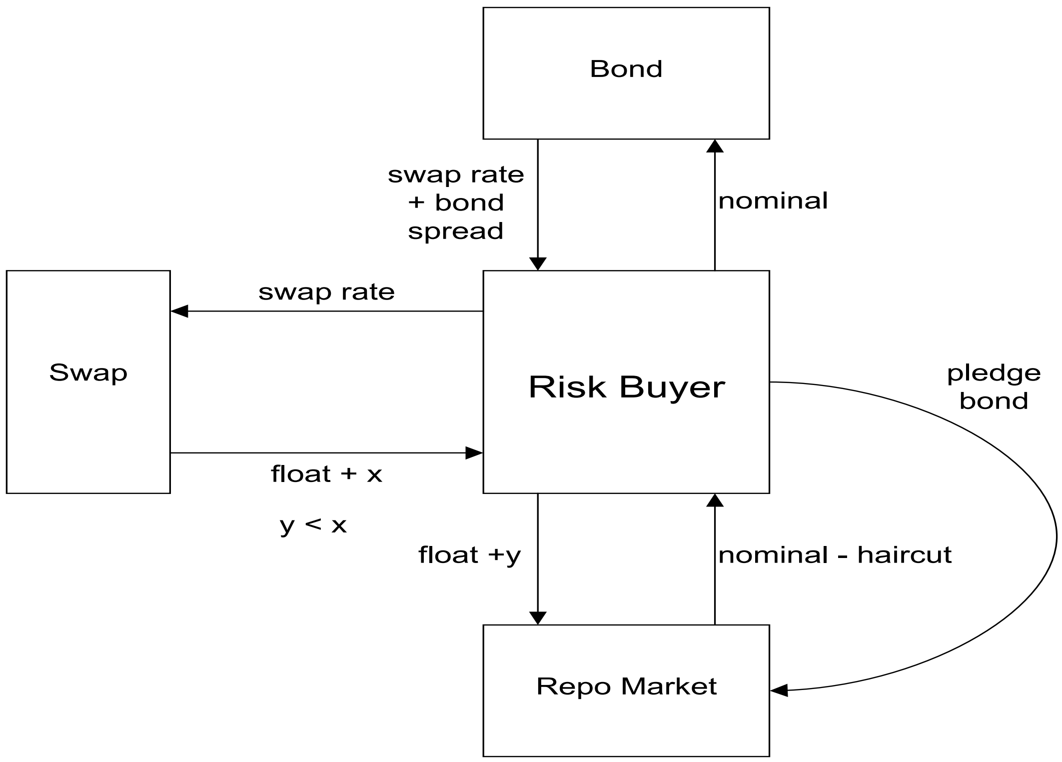

Figure 5: Asset Swap

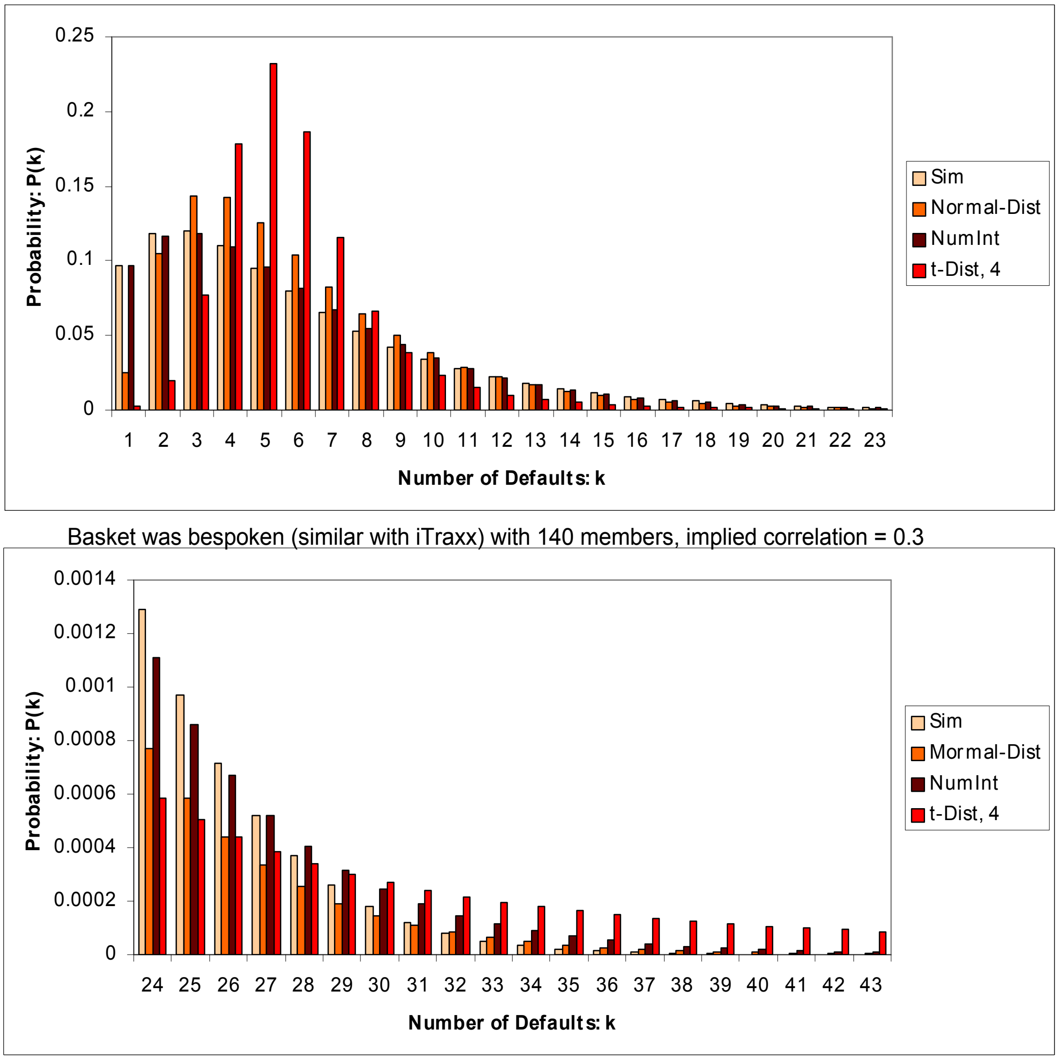

Figure 6: Comparison of Distributions

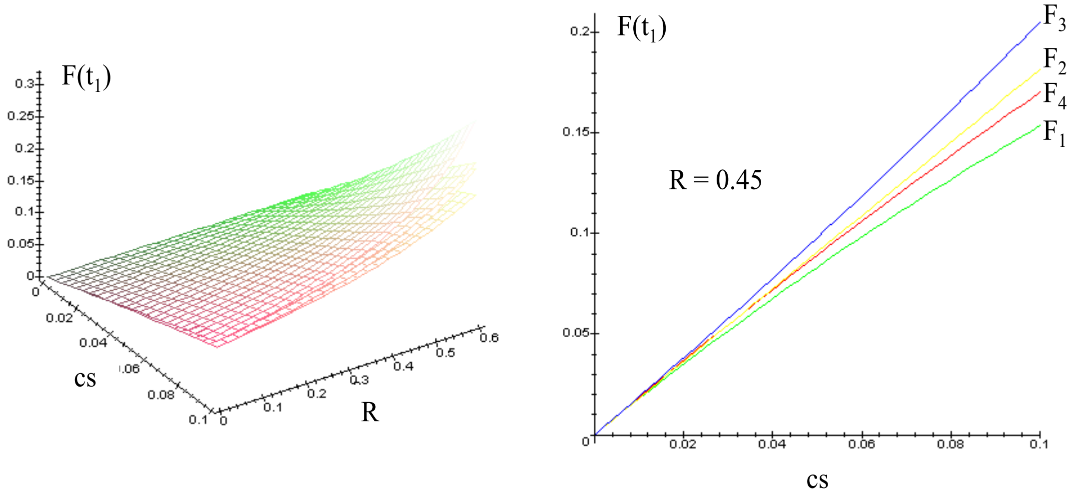

Figure 7: Bootstrapped Probabilities

References

Felsenheimer, J./Gisdakis, P./Zaiser, M. (2006), Euro credit pilot. Regular publication of HVB Credit Research, www.hvb.de/valuepilot, Ferbruary 2006.

Huddart, A./Picone, A. (2007), Synthetic CDOs: Made to measure - a guide to valuing bespoke and non-standard tranches. Dresdner Kleinwort Structured Credit Research, http://www.dresdnerkleinwort.com/research, 2007, Online-Article.

Hull, J./White, A. (2004), Valuation of a CDO and n-th to default CDS without monte carlo simulation. Working Paper, University of Toronto.

Kalemanova, A./Schmid, B./Werner, R. (2005), The normal inverse gaussian distribution for synthetic CDO pricing. Working Paper, Technical University of Munich.

Li, X. (2000), On default correlation - a copula function approach. Working Paper, RiskMetrics, New York, www.riskmetrics.com (Dec 2005).

Mahadevan, S./Polanskyj, P./Tirupattur, V./Onur, P./Sheets, A. (2006), *Structured Credit Insights - Instruments, Valuation and Strategics*, 2nd edn, Morgan Stanley.

Nelson, T. (1999), *An Introduction to Copulas*, Springer, New York.

N.N. (2004), Valuation of nth to default swaps. Masters Thesis in Mathematical Finance, University of Oxford.

Willemann, S. (2004), An evaluation of the base correlation framework for synthetic CDOs. Working Paper.

Zhu, H. (2004), An empirical comparison of credit spreads between the bond market and the credit default swap market. BIS Working Paper, No. 160.