



Models and the crisis

The pricing and risks of synthetic CDOs

Ákos Gyarmati¹

September 22, 2010

Abstract

We shortly discuss the mathematical modeling and its problems of one of the structured credit derivatives, the synthetic CDOs. These financial products were among the favorites prior to the crisis, and there was a general view that the mathematical models caused or at least boosted it.

We focus only on the mathematical description of these derivatives, therefore our analysis concentrates purely on the problems coming directly from the model, we ignore the related practical problems.

We find that not only the tools used in the modeling were inappropriate, but the principle used for pricing was also not correct in the framework of risk-neutral pricing. To our knowledge no one has highlighted these theoretical problems so far.

1. Introduction

When considering the financial and global crisis of 2007-2010, it is often said that the mathematical models used for pricing complex derivatives played a central role in causing or boosting it. There were articles available online that stated there was one single formula responsible for the crash of Wall Street.²

While this is obviously an exaggeration, and we stress that none of these articles can be considered scientific, they still have put the models in the spotlight. We aim to examine what were some of these complex derivatives and what models were used for their pricing. We focus on synthetic CDOs and the famous Gaussian copula model of Li (2000). We give a short introduction to synthetic CDOs and show the model itself, why it was used and the shortcomings why it was blamed afterwards. We remark that the deficiencies were known to the modelers and researchers, they did not claim that the main source of problems was the model.

¹ Department of Finance, Corvinus University of Budapest, gyarmati.akos@gmail.com

² http://www.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all

There are many known extensions to the original model (e.g. Andersen & Sidenius (2005), Gregory & Laurent (2004) or Kalemánova et. al.), we only present the main idea behind them and discuss their advantages and disadvantages.

The main point of this paper is that there are theoretical problems beside the known deficiencies of the original model. If we examine the principle used for pricing we find that it does not fit in the usual framework of risk-neutral pricing. Not only the original model, but the extensions also ignore this. To our knowledge no one has highlighted and dealt with this problem so far.

The rest of the paper is organized as follows: section 2 describes synthetic CDOs, section 3 deals with the problem of pricing, section 4 shows the risks associated with a synthetic CDO and the practical problems and extensions of the original model.

2. About synthetic CDOs

In this section we discuss what a synthetic CDO is, and why they were one of the favored products before the crisis. Our description only includes terms that are necessary to understand the basic concept of these credit derivatives and that are used in the following sections when we discuss the mathematical modeling.

2.1. About CDSs

While our goal is to get familiar with synthetic CDOs, it is necessary to understand basic credit derivatives such as CDSs. Furthermore as we will see, these are the products from which synthetic CDOs are built and why the name “synthetic” is given. We only give the definition of CDSs, we don’t deal with any details or pricing.

A CDS or Credit Default Swap is a credit derivative in which two parties, the protection buyer and protection seller, swap the credit (default) risk of a third product or party, called reference product. In the agreement they fix, that in case of default the protection seller pays for the protection buyer and in return the protection buyer makes regular payments for the protection seller until maturity or the default event. It is possible that there is a recovery on the subject of the CDS or the parties agree that they consider a recovery. In this case the protection seller pays only the amount reduced with the recovery if default happens.

The above said are illustrated in Figure 1 below.

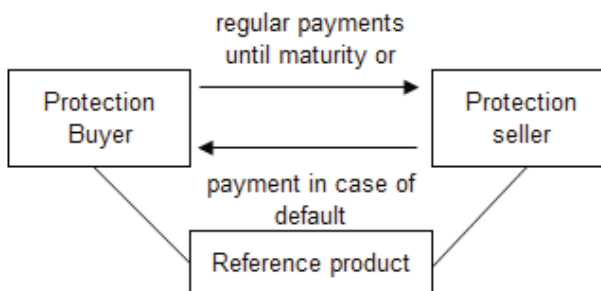


Figure 1: The basic structure of CDSs

2.2. What is a synthetic CDO?

First, the abbreviation CDO stands for Collateralized Debt Obligation. This financial product is a *basket credit derivative*, i.e. its cash flow comes from a *reference portfolio*, which contains other products that are exposed to credit risk.

The term *synthetic* comes from the fact, that in our case the reference portfolio contains CDSs (Credit Default Swaps), which are already credit derivatives, so the credit risk is represented synthetically, through the CDSs.

A synthetic CDO is a correlation product, i.e. the defaults of the CDSs in the reference portfolio are generally not independent, we have to take into account the correlation between them. What we see already is that a synthetic CDO transfers the credit risk (that was already transferred by the CDSs) of the reference portfolio. How it is done more accurately is described in the following.

A synthetic CDO is also called structured credit derivative, in the following we introduce the basic concept of its structuring:

- the reference portfolio is sliced up to so called *tranches*, which absorb the losses from the occurring defaults of the CDSs in a previously defined range. e. g. Equity tranche 0-3% means that this tranche absorbs the first 3% of losses in the reference portfolio.
- these tranches are periodically (usually quarterly) paid premium or spread from incoming cash flow of the CDSs.
 - o the premium paid for each tranche depends on its riskiness and the premium is paid for only the current notional of the tranche. e. g. tranche 0-3% is obviously riskier than tranche 12-22%, therefore it gets a greater premium. e. g. consider tranche 0-3%: if 1% of loss has already occurred, then the premium will be paid only on the remaining 2% of notional.

- in a synthetic CDO the tranches are the securities one can invest in, and the premium is considered its price.

The above described are best understood if we consider a simple, hypothetical example.

Example (trading with indices)

Consider the following structure:

Reference portfolio	Tranch name	Tranch range
A portfolio of the liquid CDSs of 125 corporation	Super senior	22-100%
	Super senior (junior)	12-22%
	Senior	9-12%
	Mezzanine (senior)	6-9%
	Mezzanine (junior)	3-6%
	Equity	0-3%

This is the synthetic CDO structure of the DJ iTraxx Europe index. Here the CDSs of the reference portfolio are equally weighted, and the premium is paid quarterly.

Suppose we invest 1 million euro into the Mezzanine (junior) tranche, which is currently traded at 300bps. For simplicity’s sake we will assume that this is now the quarterly premium. If there is no default we receive quarterly 1 mio*300bps = 30 000 euro.

Now suppose that 6 CDSs have defaulted. This is 6/125 = 4.8% of the whole portfolio, and so the losses have reached our tranche. The amount of payout: $\frac{4.8\% - 3\%}{6\% - 3\%} = 60\%$ of our total investment, i.e. we lose 600 000 euro (we would lose our complete investment if more than 6% of the reference portfolio defaulted).

From now on, if there are no more defaults, we will receive the 300bps premium only on the remaining notional of 400 000 euro, i.e. quarterly 12 000 euro.

2.3. The importance of synthetic CDOs

The issuance of synthetic CDOs grew rapidly in the year before the crisis. The motivation behind building these securities comes from both the issuer’s and the investor’s side.³

Considering the issuer, there are two main reasons:

- *regulatory capital relief*: this was the original reason for issuing CDOs. The issuer, after the securitization of the reference portfolio, is able to reduce his regulatory capital, basically because he passes on the credit risk to the investors, and therefore the only regulatory capital requirement he has to provide is for the retained pieces

³ More generally we call the buyer of a tranche protection seller, and the seller of a tranche protection buyer.

- *spread arbitrage*: this was the main reason for issuing CDOs. In practice the total spread collected from the CDSs of the reference portfolio exceeds the total amount of premium to be paid to the investors of the CDO tranches. The issuer can take advantage of this and collect the difference.

Naturally to create a successful market, the interest of the investors is also required. The advantages of CDOs for investors are:

- *broad range of risk-return profiles*: by slicing the reference portfolio into tranches we create investment opportunities from really risky to very safe, even if the reference portfolio contained only securities of mediocre risk.
- *high leverage*: as we have seen in the example above, CDOs offer a highly leveraged investment, which is especially attractive when market is booming, like prior to the crisis.

Besides these issues, as we said in the previous section *CDOs⁴ played a central role in the financial crisis of 2007-2010*. We shortly discuss the essence of their role, for a more detailed description see Gyarmati (2010).

Credit derivatives like CDOs connect two different groups of people, the investors and home-owners, through their money and mortgages. Prior to the crisis investors were looking for better investments than government bonds, because of low interest rates. On the other hand low interest rates meant cheap credit opportunities for banks. Leverage began to increase and as the profit of banks rose, the interest for credit derivatives began to grow among investors. To satisfy this demand investment banks indirectly *connected investors and home-owners through credit derivatives*. In practice this was done so that investment banks bought mortgages (or other credit derivatives), repackaged them by creating complex financial product *like CDOs* and sold them to the investors. As we saw above CDOs can satisfy the risk preferences of almost every investor even those looking for safe investments, but also *add to the complexity and reduce the transparency* of the financial system.

The demand for CDOs rose, which in the end led to the worsening of the quality of credits, as home-owner were given easier credit conditions. However the increased risk was absorbed by the increased risk-bearing of investors, they were still looking for investments like CDOs. Also (and in connection with the previous) the mathematical *models used for pricing were unable to capture the increased risk* properly, the premiums from the models did not represent the real risk of the investment. Therefore investors still thought their investments to be safe.

The raised interest rates and the passing of easy starting conditions resulted in more and more frequent defaults of bad quality credits, which eventually led to the credit and later global crisis.

As we can see there were several factors that together led to the crisis. The role of CDOs is crucial, but the models used for pricing cannot be blamed alone for everything. However a bad model does have a contribution to the crisis. To understand this properly, the model used for pricing is described in the next section. We will discuss the shortcomings and arising problems in section 4.

4 In the following description not only synthetic CDOs are considered.

3. The problem of pricing

3.1. The general approach

In this section we give the mathematical description of the above introduced synthetic CDOs and discuss their pricing. Recall that the price of a CDO tranche is the premium paid for it.

In the modeling practical issues, like the spread arbitrage opportunity of the above section or overcollateralization⁵, are generally ignored, these are very hard to incorporate and would make the models even more complicated.

We will use the following notation:

- n is the number of CDSs in the reference portfolio
- V_i is the notional of the i th CDS
- R_i the recovery on the i th CDS
- τ_i is the time of the default of the i th CDS
- $0 = t_0 < t_1 < \dots < t_N$ are the times when the premiums are paid
- $0 = K_0 < K_1 < \dots < K_m = 1$ are the so called attachment and detachment points of the tranches, i.e. the beginning and end of the range they provide protection for. e.g. tranche i provides protection between K_{i-1} and K_i . These points are given in percentage of the reference portfolio
- $B(0, t)$ is the discount factor between 0 and t
- r_i is the premium paid for tranche i , this is what we want to determine

With these notations we can define the percentage⁶ loss occurred in the reference portfolio until t , we shall denote this by $L(t)$:

$$L(t) = \frac{\sum_{i=1}^n (1 - R_i) V_i \cdot \mathcal{X}_{(\tau_i < t)}}{n V_i}$$

where $\mathcal{X}_{(A)}$ is the indicator function of the set A (it is 1 over the set and 0 otherwise). With this it is possible to determine the loss occurred to tranche i . until t .

$$L_i(t) = \max\left(\min\left[L(t); K_i\right] - K_{i-1}; 0\right)$$

In the following we will make two assumptions:

- the notional on each CDS is identical
- the recovery is also identical on each CDS

Therefore we omit the index i from now on. These may seem too strong to assume, but they are fulfilled e.g. by standardized indices, like the iTraxx Europe in the example of the previous section. Furthermore these assumptions allow us to keep the mathematical

⁵ It is basically when more collateral is put in the reference portfolio than what the tranches cover.

⁶ We will always work with percentage losses in the following.

modeling tractable. Nevertheless these are not necessary to continue, many models used in practice relax these assumptions.

The principle used today for pricing is the following argument: the premium must be determined so that initially neither the issuer (protection buyer) nor the investor (protection seller) should be able to realize one-sided profit *in expected value*. Therefore the premium should be set so that the *expected* cash flows that the investor gets and pays (the issuer pays and gets) are equal. We stress that this is not the usual no-arbitrage principle as here expected values are considered and not replicating strategies. We will discuss this issue in the next subsection in detail.

With the above defined losses we can determine both cash flows. We will call the cash flows that the investor is paid the *Premium Leg (PL)*, and what he has to pay given defaults the *Default Leg (DL)*. Recall that both PL and DL are discounted⁷, expected cash flows.

Then the PL of the investor of tranche i . can be expressed as:

$$PL_i(r_i) = \sum_{k=1}^N B(0, t_k) \cdot (t_k - t_{k-1}) \cdot r_i \cdot E_Q \left((K_i - K_{i-1}) - L_i(t_k) \right) \cdot nV$$

where $(K_i - K_{i-1}) - L_i(t_k)$ is the balance of tranche i at time t_k , E_Q denotes the expected value under a risk neutral measure used for pricing, and the multiplier nV is because of the percentage form.

Turning to the DL, we will assume that the obligations need to be fulfilled only in the discrete time points $0 = t_0 < t_1 < \dots < t_N$ not immediately on the occurrence of the default. With this the DL can be written as:

$$DL_i = \sum_{k=1}^N B(0, t_k) \cdot E_Q \left(L_i(t_k) - L_i(t_{k-1}) \right) \cdot nV$$

It is clear that $L_i(t_k) - L_i(t_{k-1})$ is what the investor of tranche i has to pay in the time interval . The price of the CDO is determined by r_i . Under the risk-neutral pricing, r_i should satisfy the following equation:

$$PL_i(r_i) = DL_i$$

From (1) and (2) we can easily express the “fair” premium of tranche i :

$$r_i = \frac{\sum_{k=1}^N B(0, t_k) \cdot E_Q \left(L_i(t_k) - L_i(t_{k-1}) \right)}{\sum_{k=1}^N B(0, t_k) \cdot (t_k - t_{k-1}) \cdot r_i \cdot E_Q \left((K_i - K_{i-1}) - L_i(t_k) \right)}$$

⁷ The discounting is done for time $t = 0$, today.

Of course this does not solve our problem, we still don't know the $E_Q(L_i(t_k))$ risk neutral expected values. It is clear that one must model the loss distribution $L(t)$ and so the joint distribution of the underlying CDSs, to calculate these. Most of the CDO pricing models in the literature focus on tackling this particular problem. We will also examine the best-known model, the Gaussian one factor copula, which was blamed the most for the crisis.

However, what most of the models don't take into account is that there are problems with the principle of pricing, already. We will discuss these issues in the next session.

3.2. Problems with the principle of pricing

As we said, the pricing principle introduced in the previous section is common in all the well-known models used for CDO pricing. However, it does not fit in the usual framework of mathematical finance used for pricing derivatives. We will argue this topic in this section. To our knowledge no one discussed these issues before.

In the above we use risk-neutral pricing, and talk about arbitrage pricing. Whether we can use the general formula of pricing⁸, is not clear however.

1. to use the formula, we should be able to reproduce the CDO cash flow *synthetically* with underlyings, we should create a *self-financing replicating portfolio*.
 - a. The *existence* and *uniqueness* of such a portfolio is not clear. We haven't found any study which discussed the replication of a synthetic CDO.
 - b. It is also not obvious what we consider underlying in this market. It should be the CDSs, but then nothing is specified about them.
2. Without the above we cannot really talk about no-arbitrage pricing.

Furthermore in the principle of the previous section there is no dynamics it is only valid for one day, so the usual *dynamic hedging* with the self-financing, replicating portfolio used in the pricing is not possible.

We haven't said anything about the *pricing measure* yet. The determination of it in the above framework is purely practical, the usual approach is to bootstrap from market data, the *Radon-Nikodym derivative* approach of mathematical finance is not used.

The main problem if we *give up the no-arbitrage principle* is that we *lose a strong economic argument* why the price should be so as the model states. Without this the whole modeling gets an ad-hoc character.

The principle used here is more like the well-known *principle of equivalence used in insurance*, but instead of real world and real probabilities, we want to use it in the risk-neutral world, with risk neutral probabilities.

The importance of the issues described in this section is that it states that when considering the crisis, not only the practical modeling of CDOs were inappropriate (as we will see

⁸ The general formula is the risk-neutral expected value of the risk-free discounted payoff.

in the following sections), but the principle used for pricing, the theoretical approach also wasn't clearly discussed.

3.3. Practical modeling – joint distributions

In this section we will briefly examine the well-known Gaussian one factor copula for modeling joint distributions. To keep track of the essence of the modeling we exclude some parts of the mathematical deductions, these can be found in the appendix (or in e.g. Eberlein et al. (2008), Gyarmati (2010)).

Remark. *According to the previous section, there is no obvious economical intuition backing up the modeling, therefore the tools used for modeling depend on the choice of the user.*

Like we said before in (3) we need to calculate the $E_Q(L_i(t_k))$ risk neutral expected values. In practice the Gaussian one factor copula was the most common to model the joint distribution of the underlying CDSs.

We define for each CDS a random variable X_i in the following way:

$$X_i = \sqrt{\rho}M + \sqrt{1-\rho}Z_i$$

where M, Z_1, \dots, Z_n are all independent and standard normally distributed.

The X_i is called state variable, and by construction X_i is also standard normally distributed, and $corr(X_i, X_j) = \rho$. We remark that this means that all pair wise correlations are equal to ρ , so according to (4) the correlation matrix of the underlying n CDSs is characterized by one number only.

There are two factors that define the state of the i th CDS: M represents the market, which affects every CDS (it is the common factor), Z_i represents the idiosyncratic factor, which describes specifically the i th CDS.

We say the i th CDS has defaulted, if its state variable sinks below a certain threshold, like in Merton (1974). These time dependent thresholds (we will use the notation $k(t)$ ⁹) can be obtained from the marginal risk neutral distribution of the default times τ_i . As in Li (2000) the usual assumption is that τ_i are exponentially distributed with λ_i intensities i.e. $Q(\tau_i < t) = 1 - \exp(-\lambda_i t)$. The default intensities can be bootstrapped from market data (see Embrechts et. al. (2005) chapter 9). We assume that the market is efficient and therefore the observed prices do not contain arbitrage. Therefore the intensities implied from market prices are considered risk neutral as they produce arbitrage-free prices. We already mentioned in the previous section that this is not an exact method, but nonetheless it is practical and the calculations can be performed.¹⁰

The idea behind (4) is that if we consider conditionally on the market factor, then by the construction defaults are independent, therefore it is easy to calculate the “conditional”

9 Formally $k(t)$ satisfies the following: $Q(\tau_i < t) = Q(X_i < k(t))$.

10 Unlike in the case of the theoretical Radon-Nikodym derivatives, where there are no such calculations.

joint distribution. From here we get the “unconditional” joint distribution by integrating over M .

More formally we can write for one CDS conditioned on M :

$$Q(\tau_i < t | M) = Q(X_i < k(t) | M) = Q\left(Z_i < \frac{k(t) - \sqrt{\rho}M}{\sqrt{1-\rho}} \mid M\right) = \Phi\left(\frac{k(t) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right) \doteq p(t | M)$$

where $\Phi(\cdot)$ is the standard normal distribution function. For the joint distribution:

$$Q(\tau_1 < t_1, \dots, \tau_n < t_n) = \int_{-\infty}^{\infty} \prod_{i=1}^n p(t_i | M = m) dF_M(m) = \int_{-\infty}^{\infty} \prod_{i=1}^n p(t_i | M = m) f(m) dm \quad (5)$$

where $f(\cdot)$ is the standard normal density function. This integral represents the so called factor Gaussian copula.

It is possible to get a closed form expression for the distribution of the relative amount of defaulted CDSs using (5) and the so called Large Homogeneous Portfolio (LHP) approximation (Vasicek (1991)). We only give the result here, the mathematical deduction can be found in the appendix.

Let D_t denote the relative amount of defaulted CDSs up to time t . For the risk neutral distribution we get:

$$F_{D_t}(h) = Q(D_t < h) = \Phi\left(\frac{\sqrt{1-\rho}\Phi^{-1} - k(t)}{\sqrt{\rho}}\right) \quad (6)$$

Using (6) we can calculate the $E_Q(L_i(t_k))$ risk neutral expected values we need to price tranche i . Because we assumed a recovery rate of R , the actual loss up to time t can be calculated as $(1 - R)L_i(t)$. Considering this, the expected values:

$$E_Q(L_i(t_k)) = (1 - R) \left(\int_{\frac{K_{i-1}}{1-R}}^1 \left(h - \frac{K_{i-1}}{1-R}\right) dF_{D_k}(h) - \int_{\frac{K_i}{1-R}}^1 \left(h - \frac{K_i}{1-R}\right) dF_{D_k}(h) \right) \quad (7)$$

since the expected loss of tranche i can be calculated as the difference of the expected losses of two “senior”¹¹ tranches (consider: $L_{[K_{i-1}, K_i]} = L_{[K_{i-1}, 1]} - L_{[K_i, 1]}$).

11 Senior tranches here are the ones that have no other tranche above them.

4. The risks of synthetic CDOs

4.1. The mark-to-market value of a CDO

In the framework we used so far, we can define the mark-to-market (MTM) value of tranche i of a synthetic CDO at time t as follows:

$$MTM_i(t) = PL_i(r_i, r^{CDS}(t), \rho(t)) - DL_i(r^{CDS}(t), \rho(t)) \quad (8)$$

Basically we take the difference of the at time t expected payoffs. Recall that according to the principle of pricing the premium is set so that the MTM value of tranche i at time 0 (at issuance) is exactly 0.

We highlighted the main sources of risk in (8). These are:

- the *change of the CDS spread*: an average spread if the CDSs are not “identical” as we assume
- the *change of the correlation*: in the model of the previous section we stressed that there is only one correlation parameter that describes the connection among the CDSs.
- *credit risk*: this was already discussed in section 2 where we described the structure and functioning of synthetic CDOs.

The effect of each of these factors can be reasoned economically. In (8) we noted that the PL depends on the premium as well, but this is determined at the beginning and does not change over the duration of the CDO.

If the CDS spread rises, it basically means that *ceteris paribus* they have become riskier, therefore one would expect a higher premium, but as we said the premium is fixed, so according to (8) the MTM value of tranche i will decrease.

In the case of correlation consider the following. If correlation is high, it means that the probability of many simultaneous defaults is high as well as the probability of no defaults at all. Conversely if correlation is low, it means that it is not likely that many defaults will occur simultaneously, but it is also unlikely that no defaults will occur at all.

The first case is bad for the investor of the senior tranche, because he could tolerate some defaults, but not many simultaneously, while it is good for the equity tranche investor as he has a greater chance to have no defaults at all and thus to “survive” (he cannot tolerate practically any defaults as he is the first one who bears the losses). In the second case the situation is the opposite: the senior investor is the favored, he doesn’t have to fear many defaults, while the equity investor has little chance to avoid losses.

Considering the above argument we have that *ceteris paribus* higher correlation makes a senior tranche investment riskier, while an equity investment safer and conversely when correlation is lower the senior investment is safer and the equity riskier. This means if correlation rises, the premium of a senior tranche should rise, and so according to (8) its MTM value decreases, while the equity premium should be lower, which means its MTM value rises. The effect on mezzanine tranches is not obvious, it has to be examined specifically.

We already discussed the case of credit risk, here we only take an interesting example: it is not necessary that the level defaults reach a given tranche to suffer from credit risk. Consider the following: defaults have just exhausted the equity tranche, but they not reached the mezzanine tranche yet. In this case the risk of the mezzanine investor is significantly higher, as now practically he is the “new” equity investor, nothing protects him from the upcoming defaults. This means his premium should rise, and so his MTM value decreases.

4.2. The positive side of the Gaussian copula

The main advantage of using the Gaussian copula is that with it the $E_Q(L_i(t_k))$ risk neutral expected values can be calculated *analytically*. In (7) we can do all the integration and we get (see e.g O’Kane & Schoegl (2001):

$$E_Q(L_i(t_k)) = (1 - R) \left(\Phi_2(f(K_{i-1}), k(t), \Sigma) - \Phi_2(f(K_i), k(t), \Sigma) \right) \quad (9)$$

where Φ_2 is the bivariate normal distribution function, $f(x) = -\Phi^{-1}\left(\frac{x}{1-R}\right)$ and is the correlation matrix.

$$\Sigma = \begin{pmatrix} 1 & -\sqrt{1-\rho} \\ -\sqrt{1-\rho} & 1 \end{pmatrix}$$

Using (9) one can easily and rapidly do the calculations based on the model:

- the calculations of the tranche premiums
- sensitivity tests for the sources of risk and hedging

We have made a set of these type of calculations to demonstrate the use of the Gaussian copula. Our data is from the DJ iTraxx Europe 5 years index series 13 at 8 April 2010.

Below in Figure 2 we see the tranche premium as the function of the CDS spread and the correlation.

These figures show exactly what we anticipated according to the previous subsection. The premiums are rising when CDS spread rises, and we see the dual effect of the correlation too.

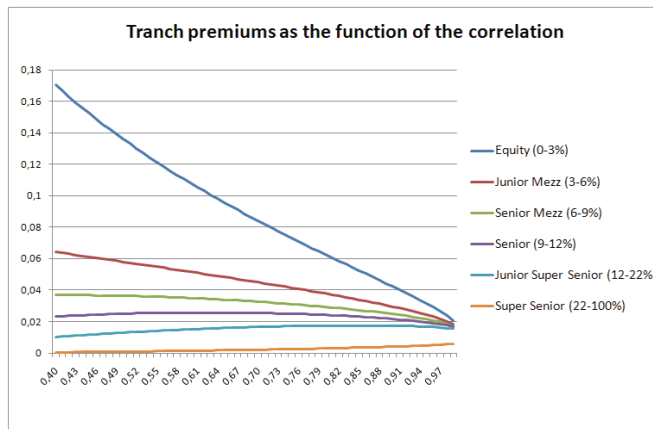
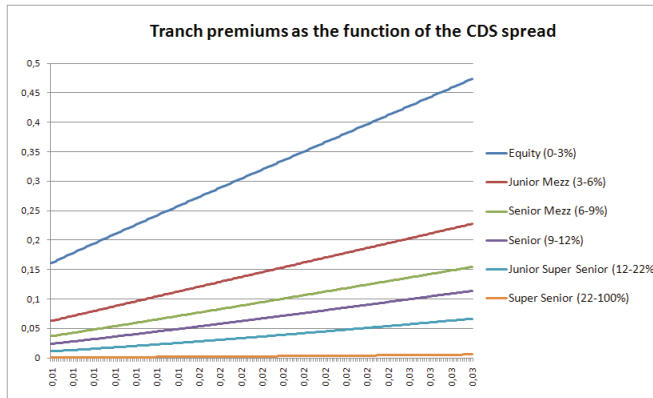
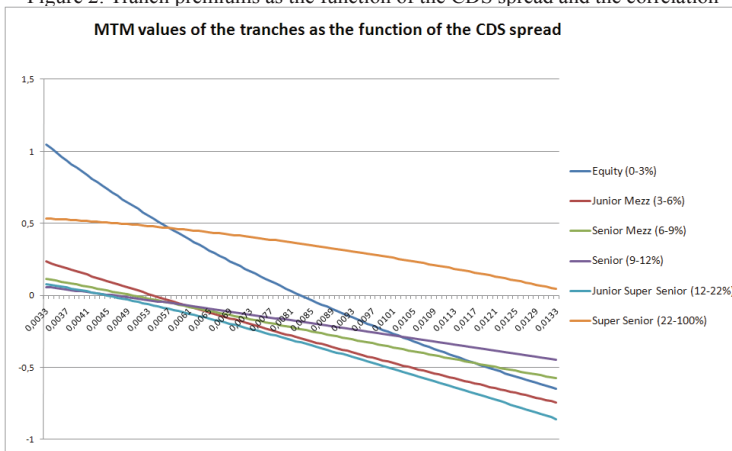


Figure 2: Tranch premiums as the function of the CDS spread and the correlation



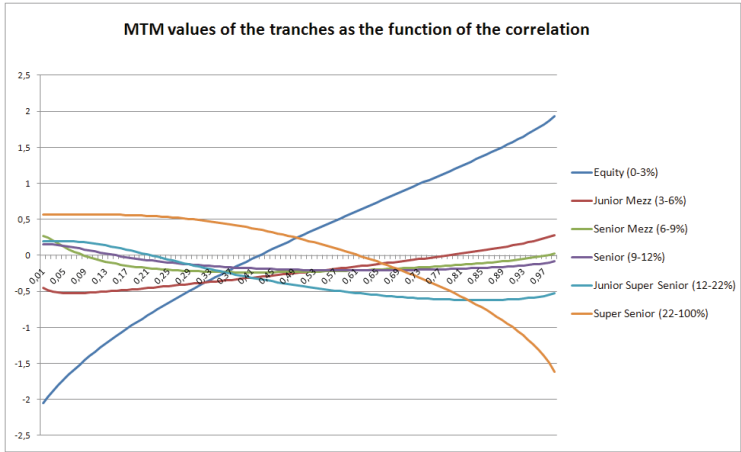


Figure 3: MTM values as the function of the CDS spread and the correlation

Considering the MTM values of the tranches, we already expect that we should see something like the opposite of the above figures as we know that the tranche premium is fixed for the duration of the CDO.

We also made the calculations of the MTM values, Figure 3 above shows the results. We can see again that the results are in line with our economical argument (and we see the “opposite” of the premium figures).

4.3. The negative side of the Gaussian copula

There are serious disadvantages of the Gaussian copula as well. From a mathematical point of view, the most important are:

- there is only one independent parameter: the correlation
- the model uses normally distributed factors, and so underestimates the real probability of joint defaults.

As a consequence with this model it is hard to fit to market data, therefore this model can lead to serious mispricing and badly constructed hedging strategies. These are the main practical problems, because of which this model was accused for causing the crisis.

Furthermore there is another, more theoretical problem with the model: if we calculate implied correlations for the tranches¹², we should get the same value over all tranches if the model was correct. However this is not the case, we get a so called correlation smile (similar to the volatility smile of the Black-Scholes model), so the model is pricing inconsistently.

In the figures below we see these issues, on our own calculations.

¹² Implied correlation is the correlation with which the model produces the observed market price.

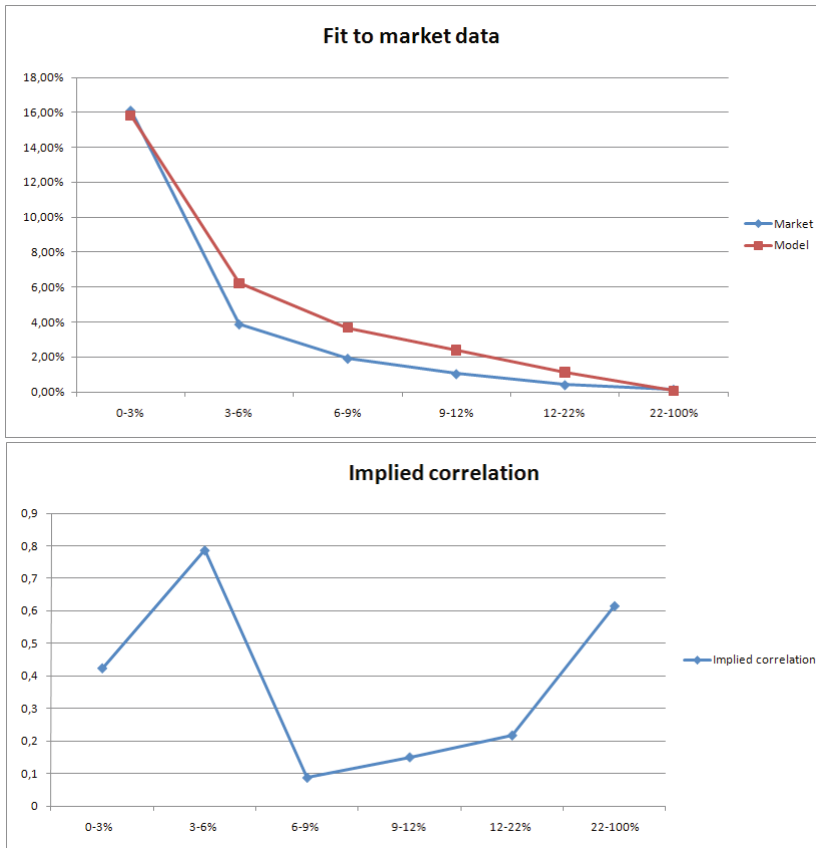


Figure 4: the fitted and real tranche premiums and the correlation smile

4.4. Possible extensions to the Gaussian copula

We consider briefly the options to overcome some of the shortcomings of the Gaussian copula. We don't examine any model thoroughly, we just give some of the best-known ideas already

present in the literature. For a more detailed description of these models see e.g. Eberlein et al. (2008), Ferrarese (2006) or Gyarmati (2010).

Basically there are two groups of possible extensions:

- *static* extensions: these models stay in the framework of the Gaussian copula, the starting point is a similar state variable equation as (4), but they use more independent variables. e.g.:
 - o models with non-normal, fat tailed factors (e.g. Normal Inverse Gaussian – NIG distributed factors)
 - o models where the correlation is also a random variable (e.g. stochastic correlation, random factor loading)

- *dynamic* extensions: these models do not use the framework of the Gaussian copula, they represent an entirely different approach. They are continuous time models, and try to model the evolution of defaults using Markov chain techniques.

The *advantage* of both categories is that they fit the market better, give more precise results than the original model.

The *disadvantage* of these models is that formulas are more complicated, and so it is more difficult to do the calculations. In most of these models an analytical solution is no more available.

As a final remark we stress, that to our knowledge the original principle described in section 3.1. remains the same in all of the now existing extensions of the model, therefore the theoretical problem of the pricing principle is still to be solved.

References

- [1] Andersen, L., & Sidenius, J. (2005), Extensions to the gaussian copula: random recovery and random factor loadings. *Journal of Credit Risk* 1, pp. 29-70.
- [2] Eberlein, E., Frey, R., & von Hammerstein, E. A. (2008), Advanced credit portfolio modeling and CDO pricing. In: *Mathematics: Key Technology for the Future*, Springer, (pp. 253-280).
- [3] Embrechts, P., Frey, R., & McNeil, A. (2005), *Quantitative risk management: Concepts, techniques, and tools*. Princeton Series in Finance, Princeton University Press
- [4] Ferrarese, C. (2006), A comparative analysis of correlation skew modeling techniques for CDO index tranches. Master's thesis, King's College London
- [5] Gregory, J., & Laurent, J.-P. (2004), In the core of correlation. *Risk*, October, pp. 87-91
- [6] Gyarmati, Á, (2010). Mathematical models and the financial crisis, the pricing and risks of CDOs, (in Hungarian), Master's thesis, Corvinus University of Budapest
- [7] Kalemanova, A., Schmid, B., & Werner, R. (2005), The normal inverse gaussian distribution for synthetic CDO pricing. Working paper, Risklab Germany
- [8] Li, D. X. (2000), On default correlation: A copula function approach. *The Journal of Fixed Income*, 6, pp. 43-54.
- [9] Merton, R. C. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*, 29, pp. 449-470.
- [10] O'Kane, D., & Schloegl, L. (2001), *Modeling credit: Theory and practice*. Lehman Brothers Analytical Research Series
- [11] Vasicek, O. A. (1991), Limiting loan loss probability distribution. KMV Corporation, technical report

Appendix

Here we show how one can get the loss distribution of (6)

Let A_k^t denote the event, that up to time t exactly k CDSs have defaulted.¹³ If condition on M , then A_k^t means that out of n independent events exactly k has happened. Therefore A_k^t is distributed binomially, with parameters $p(t | M)$ and n . To get the unconditional distribution we need to integrate over M .

To determine the loss distribution, we need the probability, that no more than k defaults occurred. We use the notation D_t for the relative amount of defaults. Then the probability we are looking for is:

$$F_{D_t}(h) = Q(D_t < h) = \sum_{k=0}^{[nh]} Q(A_k^t) = \int_{-\infty}^{\infty} \sum_{k=0}^{[nh]} \binom{n}{k} p(t | m)^k (1 - p(t | m))^{n-k} dF_M(m) \quad (10)$$

Now we will apply the LHP approximation. Let $p_t(M) \doteq \Phi\left(\frac{k(t) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right)$ as a

random variable and G_{p_t} its distribution function. With this notation and the substitution $y = p_t(u)$ we can rewrite (10):

$$F_{D_t}(h) = \int_0^1 \sum_{k=0}^{[nh]} \binom{n}{k} y^k (1-y)^{n-k} dG_{p_t}(y) \quad (11)$$

The LHP approximation means that we examine the behavior of the integrand, when $n \rightarrow \infty$. Let B_i ($i = 1, \dots, n$) denote independent Bernoulli distributed variables, with parameter y . By the law of large number we have $\bar{B}_n = \frac{1}{n} \sum_i B_i \rightarrow y$ almost surely so

for the distributions $F_{\bar{B}_n}(x) \rightarrow \chi_{[0,x]}(y)$ pointwise on $\mathbb{R} \setminus \{y\}$. The sum of Bernoulli variables is binomially distributed, so for $h \neq y$, $n \rightarrow \infty$ (let A be binomially distributed with parameter, y , n):

$$\sum_{k=0}^{[nh]} \binom{n}{k} y^k (1-y)^{n-k} = Q(A < nh) = Q\left(\sum_i B_i < nh\right) = Q(\bar{B}_n < h) \rightarrow \chi_{[0,h]}(y)$$

The sum on the left hand side is bounded by $\mathbf{1} \in L^1$, thus by Lebesgue's theorem we have from (11):

13 So by this notation the relative amount of loss is $L(t) = \frac{k}{n}(1-R)$.

$$\begin{aligned}
F_{D_t}(h) &= \int_0^1 \sum_{k=0}^{\lfloor nh \rfloor} \binom{n}{k} y^k (1-y)^{n-k} dG_{p_t}(y) \rightarrow \int_0^1 \chi_{[0,h]}(y) dG_{p_t}(y) = G_{p_t}(h) = \\
&= \mathbb{Q} \left(\Phi \left(\frac{k(t) - \sqrt{\rho} M}{\sqrt{1-\rho}} \right) < h \right) = \mathbb{Q} \left(\frac{k(t) - \sqrt{1-\rho} \Phi^{-1}(h)}{\sqrt{\rho}} < M \right) = \\
&= 1 - \Phi \left(\frac{k(t) - \sqrt{1-\rho} \Phi^{-1}(h)}{\sqrt{\rho}} \right) = \Phi \left(\frac{\sqrt{1-\rho} \Phi^{-1}(h) - k(t)}{\sqrt{\rho}} \right)
\end{aligned}$$

Where we used that M is standard normally distributed and that $1 - \Phi(x) = \Phi(-x)$.