Modeling credit risk and credit derivatives

Leijdekker, V.J.G.
In the previous chapter we have investigated a large data set of CDO tranche prices. In particular, we have looked at the performance of the, market standard, base correlation framework and a number of interpolation, or mapping, schemes applied to this framework. As the base correlation framework uses different correlation values for each CDO tranche, one has to resort to interpolation, in order to value tranches with non standard attachment and detachment points. In this chapter we investigate a number of models as alternatives to the base correlation framework, which all belong to the class of one factor copula models, which has been discussed briefly in Chapter 1. Further, for ease of reading, we present some background of credit derivatives and the base correlation framework.

5.1 Introduction

In recent years, the market for credit derivatives has undergone an enormous growth. The market for credit default swaps has grown from 14 trillion USD in 2005, to over 500 trillion USD in 2007. Further, in 2004 CDO contracts were issued for a total notional of 157 billion USD, which grew to 552 billion USD in 2007. This development has led to the need for models which can be calibrated to the available market quotes. Due to its simplicity, the one factor Gaussian copula has become market standard. As this model is not able to calibrate to all market quotes with a single correlation parameter, the base correlation framework was introduced by McGinty, Beinstein, Ahluwalia, and Watts [MBAW04]. The ideas behind this framework are similar to the implied volatility in the equity derivatives market, as different correlations are used for different CDO tranches.
The main disadvantage of the base correlation framework is that it is not an actual model of default, it merely is an advanced interpolation scheme. Therefore, many authors have suggested different models for the valuation of CDO tranches, with the goal to calibrate to CDO quotes with a single set of parameters. Hull and White [HW04] consider the student t-distribution instead of the Gaussian distribution in their factor copula. Kalemanova et al. [KSW05] consider the normal inverse Gaussian distribution. Other types of extensions have been considered by e.g. Andersen and Sidenius [AS04], who stick to the Gaussian distribution, but they let the correlation depend on the state of the economy. This reflects that correlations tend to increase when the state of the economy deteriorates. Another approach, where the correlation is modeled stochastically, is the mixture factor copula, where a mixture of gaussian copulas is considered. This idea has been discussed by Burtschell et al. [BGL05].

Other models in the factor copula setup that have been proposed are by [Moo06], who considers the variance gamma distribution for the factors. This distribution has been introduced by [MS90] into financial mathematics, and it has been used to match the volatility smile of equity options by [MCC98]. The alpha-stable distributions are considered by [PS06], and [vdV07] considers an, additional, external factor to explain for external causes of default. A brief overview of several of the models above is given by [BGL05].

Next to the models in the factor copula framework, many other models have been considered. Whereas the factor copula models belong to the class of bottom-up models, as one first models the marginal distributions and then the dependence structure, other authors have considered a top-down approach, where one focusses directly on the loss process. Giesecke et al. [GGD05] and [EGG09] consider the general setup of such top-down models, as well as some detailed examples. Brigo et al. [BPT06] and [BPT07] consider the sum of Poisson processes to model the loss process.

In this chapter we investigate the performance of a number of the models described above, where we calibrate the models to a large set of (daily) CDO tranche prices, from December 2004 up to November 2006. We consider the calibration errors for each model and each date at which we calibrate to the market data. Further, we discuss the stability of the optimal model parameters, which could indicate if a model is ‘over fitting’ to the market data.

This chapter is structured as follows. In Section 5.2 we briefly define the two most popular credit derivatives, the credit default swap and the collateralized debt obligation, and in Section 5.3 we consider the valuation of such derivatives. Thereafter, in Section 5.4 we describe the one factor copula models that we investigate in this Chapter. Section 5.5 describes the approach we have taken to compare the models, and we describe the set of market data that has been used. The results from the tests that we have performed are presented in Section 5.6, where first the calibration errors are compared. Thereafter, we compare the model quotes to the market quotes, to see how well the models match the market data. In Section 5.7 we summarize the results and we conclude with some remarks about the obtained results.
5.2 Credit Derivatives

In this section we briefly describe two of the most liquid credit derivatives currently traded. These are the credit default swap (CDS) and the collateralized debt obligation (CDO).

5.2.1 Credit Default Swaps

The credit default swap is a contract between two parties providing protection against the default of a third party. The party that buys the protection is called the protection buyer, and he is compensated for the losses incurred on the default of the third party, which is called the reference credit. The compensation is paid by the protection seller who in turn receives (periodic) premium payments from the protection buyer. This premium is usually defined in basis points per year on the notional value of the contract. The level of the premium is set such that the contract has zero value at the start of the contract, and it is paid until the maturity of the contract or until the reference credit defaults.

The default event can be defined in several ways, such as bankruptcy, restructuring or failure to pay back a loan. In case of default the settlement of the CDS can be in physical delivery, which means that the protection buyer delivers bonds of the defaulted company and in return it receives the full nominal value of the bonds. Alternatively the settlement can be settled in cash, where the value of the defaulted bonds is determined, the recovery rate, and the protection buyer receives the nominal value of the bonds reduced by this recovery rate.

Similar to the CDS the index CDS provides protection against defaults. In case of the CDS there is only one reference credit but the index CDS has a basket of reference credits as underlyings, each with its own notional value, but usually all underlyings have the same notional value. The defaults of the reference credits results in losses which are paid by the protection seller. In return the protection buyer pays a periodic premium over the remaining notional value of the reference credits which have not defaulted.

5.2.2 Collateralized Debt Obligations

Another credit derivative that provides protection on a basket of reference credits is the collateralized debt obligation (CDO). Whereas the index CDS provides the protection against all defaults during the life of the contract, the CDO provides protection on a part, or tranche, of the basket, which is defined by the attachment and detachment points. As soon as the sum of the losses on the basket reaches the level of the attachment point, the protection buyer start to receive loss payments. He stops receiving loss payments, when the losses reach the detachment point or when the maturity of the contract is reached. In return for this possible loss payments the protection seller receives a periodic premium. In some cases a part of the premium is
paid up front at the start of the contract. The CDO can be defined on several types of reference credits, such as loans, bonds or CDSs. In case the underlying basket consists of credit default swaps, we speak of a synthetic CDO. In the analysis that we perform in this chapter we only focus on this latter type of CDO structures.

5.2.3 iTraxx and CDX

In recent years the liquidity in credit derivatives have increased enormously. Currently CDSs on almost all large companies are traded, with maturities of three and six months, and one, two, three, four, five, seven, ten and twelve years. The contracts have also been standardized, such that the maturity always falls on the twentieth of March, June, September or December.

Also the market for CDOs has become much more liquid, which has led to standardization of contracts and to the existence of the iTraxx and CDX indices. These are baskets consisting of (CDSs on) 125 equally weighted European or North-American companies, respectively. On these baskets index CDSs are traded, for which daily quotes are available. Further CDO tranches are traded on these baskets. In case of the European iTraxx basket market quotes are available on five different tranches, where the first tranche, or equity tranche, covers the first 3% of the basket notional. A part of the premium on this tranche is paid upfront together with a periodic premium of 500 basis points per year. The next tranche, or junior mezzanine tranche, covers the next 3% of the basket notional and only a periodic premium is paid. The mezzanine tranche covers the 6% up to 9% region of the basket notional, the senior mezzanine tranche covers the 9% up to 12% region and the (super) senior tranche covers the 12% up to 22% region of the basket loss. Sometimes quotes are available for a sixth tranche, which covers the remaining part of the basket. These latter quotes have not been considered.

For the North-American CDX index, quotes for five year CDO tranches are given in the market. The equity tranche covers the first 3% of the basket notional and a part of the premium is paid upfront and there is a periodic premium of 500 basis points. The junior mezzanine tranche covers the 3% up to 7% region of the basket notional, the mezzanine tranche covers the 7% up to 10% of the losses and the senior mezzanine tranche covers the 10% up to 15% of the notional. The (super) senior tranche covers the 15% up to 30% of the basket notional. Again, quotes for a sixth tranche, covering the remaining part of the basket, can be given.

5.3 Pricing Credit Derivatives

In the previous section we have briefly explained two of the most traded credit derivatives. In this section we give some formulas to determine the value of a trade.
5.3.1 Pricing a Credit Default Swap

We start with the pricing of credit default swaps and we discuss the model that is used in the analysis in the present chapter. In order to price the CDS we need to value the two legs, which are the two payment streams of the swap. The first leg is the default or protection leg (DL) and the second leg is the premium leg (PL). Further we want to determine the fair spread $s$, which is the spread in basis points per year such that the contract has zero value at the start of the contract.

We write $\tau \geq 0$ for the default time of the reference credit, $T$ for the maturity of the contract and we write $D(t)$ for the risk free discount factor, which we assume to be deterministic. The expected recovery rate is given by $R$ and the notional amount of the contract is given by $N$. The expected present value of the default leg can be written as

$$DL = N \mathbb{E} \left[ D(\tau) (1 - R) 1_{\{\tau \leq T\}} \right]$$

$$\approx N \sum_{i=1}^{M} D(t_i^*) (1 - R_{\text{fix}}) \mathbb{P} (\tau \leq T). \quad (5.1)$$

The premium is paid periodically, usually every three months at the end of the period. In case the reference credit defaults, an accrued premium payment has to be paid over the period from the previous premium payment date up to the default date. When we denote the premium payment dates by $T_1 < T_2 < \ldots < T_N = T$, and we write $T_0 = 0$, we can write the present value of the fixed leg as

$$PL = N \sum_{i=1}^{N} (D(T_i) 1_{\{\tau > T_i\}}(t_j - t_{j-1}) + D(\tau)1_{\{t_{j-1} < \tau \leq t_j\}}(\tau - t_{j-1})). \quad (5.2)$$

The value of both legs is obtained by taking the expectations of the present values in (5.1) and (5.2). We write $F_{\tau}$ for the distribution function of $\tau$. The value of the two legs of the credit default swap are given by

$$\mathbb{E}DL = (1 - R) N \int_0^T D(t) dF_{\tau}(t) \quad (5.3)$$

$$s \mathbb{E}PL = s N \sum_{j=1}^{m} \left( D(t_j)(t_j - t_{j-1})(1 - F_{\tau}(t_j)) + \int_{t_{j-1}}^{t_j} D(t)(t - t_{j-1}) dF_{\tau}(t) \right). \quad (5.4)$$

We choose to model the default time $\tau$ as the first jump of an inhomogeneous Poisson process, with a deterministic piecewise constant intensity. The levels of the intensity allow us to match the market quotes for CDSs with different maturities, by bootstrapping the intensity one maturity at a time, starting with the lowest maturity. The
distribution function of \( \tau \), or default probability function, is given by \( \mathcal{F}_\tau(t) = p(t) = 1 - \exp(- \int_0^t \lambda(s) \, ds) \). In order to evaluate the integrals in (5.3) and (5.4) we consider a discretization of the interval \([0, T]\) into time points \(0 = T_0 < T_1 < \ldots < T_n = T\), and we approximate \( D(t) \) on \([T_i, T_{i+1})\) by \( D \left( \frac{T_i + T_{i+1}}{2} \right) \) and by \( D \left( \frac{t_j + t_{j+1}}{2} \right) \) on \([t_j, t_{j+1})\). This leads to the formulas

\[
\mathbb{E} DL \approx (1 - R) N \sum_{i=1}^{n} D \left( \frac{T_{i-1} + T_i}{2} \right) (p(T_i) - p(T_{i-1}))
\]

\[
s \mathbb{E} PL \approx s N \sum_{j=1}^{m} \left( D(t_j)(t_j - t_{j-1})(1 - p(t_j)) + D \left( \frac{t_{j-1} + t_j}{2} \right) \frac{t_j - t_{j-1}}{2} (p(t_j) - p(t_{j-1})) \right).
\]

The valuation of the index CDS proceeds along the same lines as the valuation of the CDS above. An index CDS can be viewed as a portfolio of CDSs where each CDS has the same spread. Using this observation one can simply obtain the fixed and credit leg of an index CDS by summing the fixed legs and premium legs of the CDSs, respectively.

### 5.3.2 Pricing a Collateralized Debt Obligation

In order to price a CDO tranche we again have to value the credit and fixed leg. We write \( L(t) \) for the cumulative loss process of the underlying basket consisting of \( K \) reference credits:

\[
L(t) = \sum_{k=1}^{K} (1 - R_k) N_k 1_{\{\tau_k \leq t\}},
\]

where \( R_k \) is the recovery rate of reference credit \( k \) and \( N_k \) is its notional value and its default time is given by \( \tau_k \). When we denote the attachment point of the tranche by \( a \) and the detachment point by \( d \) then we define the tranche loss at time \( t \) as \( L_{a,d}(t) \) and its value can be obtained from the loss process by

\[
L_{a,d}(t) = \min(\max(L(t) - a, 0), d - a).
\]
The present value of the two legs can be written in terms of the tranche loss process. This gives

\[ DL_{a,d} = \int_0^T D(t) dL_{a,d}(t) \]

\[ sPL_{a,d} = s \sum_{j=1}^m D(t_j) \left( ((d - a) - L_{a,d}(t_j))(t_j - t_{j-1}) \right. \]

\[ \left. + \sum_{k=1}^K 1_{\{t_{j-1} < \tau_k \leq t_j\}}(\tau_k - t_{j-1}) \Delta L_{a,d}(\tau_k) \right) . \]

Here \( \Delta L_{a,d}(\tau_k) \) is the change in the loss amount on the CDO tranche \((a, d)\) due to a default of name \( k \) at time \( \tau_k \).

The values of the legs are obtained by taking the expectation of both expressions. The values can be approximated using the same discretization and assumptions as in the previous section. A straightforward calculation shows that

\[ EDL_{a,d} \approx \sum_{i=1}^n (E L_{a,d}(T_i) - E L_{a,d}(T_{i-1})) D \left( \frac{T_i + T_{i-1}}{2} \right) \]

\[ sEPL_{a,d} \approx s \sum_{j=1}^m (t_j - t_{j-1}) \left( 1 - \frac{E L_{a,d}(t_j) + E L_{a,d}(t_{j-1})}{2} \right) D(t_j) \]

From formulas (5.5) and (5.6) it is easy to see that in order to value both legs of a CDO tranche one only has to calculate expected tranche losses at the premium payment dates \( t_j \) and at the time points \( T_i \). In the next sections we discuss techniques and models which can be used to determine these expected losses.

### 5.4 Factor Copula Models

In the previous section we have seen that to value CDO tranches, one has to calculate expected tranche losses at a number of points in time. In order to calculate these expected losses one has to model the dependence structure between the different reference credits in the basket. In the analysis in this chapter we consider several factor copula models which can exactly do this. The advantage of (factor) copula models is that they allow one to model the dependence structure and marginal distributions separately. They also allow one to condition on a common factor, by which default times become independent. We start with a general description of the factor copula models and we consider a technique that can be used to determine the loss distribution. In the remainder of the section we consider the standard model and a number of alternatives to this model.
5.4.1 General Setup

We define the factor copula models through a set of random variables, which determine the dependence structure. For each of the $K$ companies in the basket underlying the CDO we define a random variable $X_k$ by

$$X_k = Z + E_k, \quad k = 1, \ldots, K,$$

where $Z$ and the $E_k$’s are independent and $Z \sim F_Z$ and $E_k \sim F_{E_k}$. Further, we write $F_{X_k}$ for the distribution of $X_k$. This factor setup has been introduced to the pricing of CDOs by [Li00], where the normal distribution is considered for all factors. The distribution function $F_{X_k}$ can be obtained from the distribution function of $Z$ and $E_k$, by conditioning on the factor $Z$ which results in

$$P(X_k \leq x) = \int_{\mathbb{R}} F_{E_k}(x - Z) dF_Z(x).$$

The factor $Z$ is used to model the global state of the economy. A low value corresponds to a bad state of the economy, and a high value corresponds to a good state of the economy. The factor $E_k$ models the state of the individual company. Again, a low value represents a bad state and a high value represents a good state of the company.

In order to link the static factor copula structure to the marginal distributions, that were obtained in Section 5.3.1, we assume that the company $k$ defaults as soon as a non-decreasing time-dependent barrier $\chi_k(t)$ reaches the level $X_k$. This results in default probabilities

$$P(\tau \leq t) = P(X_k \leq \chi_k(t)).$$

We, naturally, want both probabilities to be equal and this implies the value of the barrier. Using $p_k(t)$ to denote the marginal default probability of company $k$ then we have

$$\chi_k(t) = F_{X_k}^{-1}(p_k(t)).$$

In order to calculate the expected (tranche) losses we need to know the loss distribution. Therefore we want to compute probabilities of the form $P(L(t) \leq \ell)$, for certain loss levels $\ell$. We discretize the loss range by introducing $\delta_{\ell}$, which described the minimum loss step. This step is such that $(1 - R_k)N_k = n_k \delta_{\ell}$, for some integer $n_k$. With this discretization $L(t)$ has a discrete distribution on the points $n \delta_{\ell}$, with $n \in \mathbb{N}$.

We can build the loss distribution using ideas explored by [ASB03] and [HW04]. We first condition on the factor $Z$ and then use the independence of the factors $E_k$. The loss distribution can be obtained using the following algorithm, where we add one company at a time to the distribution of $L(t)$.

1) Start with $P^{(0)}(L(t) = n\delta_{\ell}|Z) = 1_{\{n=0\}},$ for $n = 0, 1, 2, \ldots$. 
2) Assume that we are at step $k$ and we want to determine $P^{(k+1)}(L(t) = n\delta_{\ell}|Z)$ by adding company $k+1$ to the distribution. Then we have

$$P^{(k+1)}(L(t) = n\delta_{\ell}|Z) = P^{(k)}(L(t) = n\delta_{l}|Y)(1 - F_{E_{k+1}}(\chi_{k+1}(t) - Z))$$

$$+ P^{(k)}(L(t) = (n - n_{k+1})\delta_{l}|Y)F_{E_{k+1}}(\chi_{k+1}(t) - Z)).$$

Repeat this step for $k = 1, 2, \ldots, K$.

3) Obtain the conditional loss distribution by

$$P(L(t) = n\delta_{\ell}|Z) = P^{(K)}(L(t) = n\delta_{\ell}|Z).$$

The unconditional loss distribution is obtained by integrating out over factor $Z$,

$$P(L(t) = n\delta_{\ell}) = \int_{D_Z} P(L(t) = n\delta_{\ell}|Z = x) dF_Z(x)$$

With this loss distribution we can easily calculate expected (tranche) losses as

$$\mathbb{E}L(t) = \sum_{k=1}^{[K/\delta_{\ell}]} k\delta_{\ell}P(L(t) = k\delta_{\ell})$$

$$\mathbb{E}L_{a,d}(t) = \sum_{k=1}^{[K/\delta_{\ell}]} \min(\max(k\delta_{\ell} - a, 0), d - a)P(L(t) = k\delta_{\ell})$$

### 5.4.2 Standard Model

With the formulas to determine the loss distribution in place, we can focus on specific models which fall into the general setup of formula (5.7). In this section we consider the standard model where $Z$ and the $E_k$’s have a normal distribution. To be more precise, we let

$$Z \sim N(0, \rho^2), \ E_k \sim N(0, 1 - \rho^2),$$

for a correlation parameter $0 \leq \rho \leq 1$. We can then rewrite (5.7) to the equivalent formulation,

$$X_k = \rho Y + \sqrt{1 - \rho^2} \varepsilon_k,$$

where $Y, \varepsilon_k \sim N(0, 1)$ and hence $X_k \sim N(0, 1)$. The resulting factor copula model is often referred to as the standard model or the one factor Gaussian copula. The one factor Gaussian copula has first been suggested by [Li00].

The main drawback of this model is that it is not able to match the market quotes on the European iTraxx and North-American CDX indices. This is illustrated in Table 5.1
Table 5.1: Market and Model Quotes for CDO tranches on the iTraxx index on valuation date 13 December 2004 with a maturity of five years. The correlation parameter has been chosen such that the upfront premium on the equity tranche is matched.

<table>
<thead>
<tr>
<th>Tranche</th>
<th>Market Quote</th>
<th>Model Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0%,3%)</td>
<td>23.00 %</td>
<td>23.00 %</td>
</tr>
<tr>
<td>(3%,6%)</td>
<td>131.0 bp</td>
<td>187.26 bp</td>
</tr>
<tr>
<td>(6%,9%)</td>
<td>43.5 bp</td>
<td>45.25 bp</td>
</tr>
<tr>
<td>(9%,12%)</td>
<td>27.25 bp</td>
<td>12.18 bp</td>
</tr>
<tr>
<td>(12%,22%)</td>
<td>14.25 bp</td>
<td>1.43 bp</td>
</tr>
</tbody>
</table>

for a set of CDO tranches on the iTraxx index. In this table we can see that, if we choose the correlation such that the upfront premia is matched, the standard model overestimates the premium on the junior mezzanine tranche and it underestimates the premium on the more senior tranches.

A solution to this problem is given through the base correlation framework, where a CDO tranche \((a, d)\) is split up into two tranches \((0, d)\) and \((0, a)\). Starting with the equity tranche, one can try to find a correlation value such that the upfront premium on this tranche is matched. Next one iterates over the remaining tranches by finding correlation values such that the annual premia are matched. At each step one uses the correlation from the previous step to price the tranche \((0, a)\) and the tranche \((0, d)\) is priced with the new value of the correlation parameter. The problem of this approach is that it is not a model of default times. Furthermore it is not clear what correlation parameters one should use to value a nonstandard CDO tranche, and one has to resort to interpolation, which could lead to arbitrage opportunities, as the expected loss on base tranches, as function of detachment point, might not be increasing. A more detailed description of the base correlation approach can be found in [AM04] and [ABM04].

Due to the deficiencies of the standard model and the base correlation several alternative models have been proposed. In the analysis in this chapter we restrict ourselves to models that fit the general setup of Equation (5.7). In the remainder of this section we describe the alternative models that we consider in this chapter.

### 5.4.3 Double t-Copula

The first alternative model that we consider is the double t-copula. This model is obtained by assuming a student \(t\)-distribution for the factors \(Z\) and \(E_k\). The density of random variable \(V\) with a \(t_\nu\) distribution, with \(\nu > 0\) degrees of freedom, is given by

\[
f_\nu(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu \pi} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}
\]

\[
\Gamma(x) = \int_0^\infty t^{x-1}e^{-t} \, dt
\]
In case $\nu > 1$ the expectation of $V$ is zero and, in case $\nu > 2$ the variance of $V$ is $\nu/(\nu - 2)$. For low values of $\nu$ The student $t$-distribution has much fatter tails than the normal distribution, which makes extreme outcomes more likely.

Using $t$-distributions for both factors has been considered by [HW04], where the following description is considered.

$$Z = \sqrt{\frac{\nu_Y - 2}{\nu_Y}} \rho Y$$
$$E_k = \sqrt{\frac{\nu_\varepsilon - 2}{\nu_\varepsilon}} (1 - \rho^2) \varepsilon_k,$$

where $Y \sim t_{\nu_Y}$ and $\varepsilon \sim t_{\nu_\varepsilon}$, with $\nu_Y, \nu_\varepsilon > 2$ and $0 \leq \rho \leq 1$. The factors are scaled by the inverse of their variance, such that $X_k$ has zero expectation and a variance of one.

In the beginning of this section we have seen that for the Student $t$-distribution we have $\nu > 0$. The description given above only allows for $\nu > 2$, which excludes the use of other $t$-distributions. Therefore we consider a different description by removing the scaling factors in front of both factors, which results in the double-$t$ factor copula.

$$Z = \rho Y$$
$$E_k = \sqrt{1 - \rho^2} \varepsilon_k,$$

where we still have $Y \sim t_{\nu_Y}$ and $\varepsilon \sim t_{\nu_\varepsilon}$, only in this case we have $\nu_Y, \nu_\varepsilon > 0$. This description does allow us to use a wider range of degrees of freedom, as the range from 0 up to 2 is available as well. In the analysis in this chapter we consider this latter description because of the wider parameter range. This wider range is important as tests show that in some cases one finds optimal parameter values where the number of degrees of freedom are below two.

5.4.4 Normal Inverse Gaussian Factor Copula

Similar to the student $t$-distribution we consider the normal inverse Gaussian (NIG) distribution, where we, compared to the standard model, also consider different distributions for the factors. The NIG distribution has the density

$$f_{\alpha, \beta, \mu, \delta}(x) = \frac{\delta \alpha \exp(\delta \gamma + \beta(x - \mu))}{\pi \sqrt{\delta^2 + (x - \mu)^2}} K_1\left(\alpha \sqrt{\delta^2 + (x - \mu)^2}\right)$$
$$\gamma = \sqrt{\alpha^2 - \beta^2}$$
$$K_1(\omega) = \frac{1}{2} \int_0^\infty \exp\left(-\frac{1}{2} \omega \left(t + t^{-1}\right)\right) \, dt,$$
where $0 < |\beta| < \alpha$ and $\delta > 0$. The expectation and variance of a random variable $V$ following an $NIG(\alpha, \beta, \mu, \delta)$ distribution are given by

$$
E(V) = \frac{\mu \gamma + \beta \delta}{\gamma},
$$

$$
Var(V) = \frac{\delta \alpha^2}{\gamma^3}.
$$

Further the NIG distribution satisfies the properties

$$
V \sim NIG(\alpha, \beta, \mu, \delta) \Rightarrow cV \sim NIG\left(\frac{\alpha}{c}, \frac{\beta}{c}, c\mu, c\delta\right)
$$

(5.8)

$$
V_i \sim NIG(\alpha, \beta, \mu_i, \delta_i), \text{ for } i \leq M \Rightarrow \sum_{i=1}^{M} V_i \sim NIG\left(\alpha, \beta, \sum_{i=1}^{M} \mu_i, \sum_{i=1}^{M} \delta_i\right)
$$

(5.9)

We obtain the normal inverse Gaussian factor copula from the general description (5.7) setting

$$
Z = \rho Y
$$

$$
E_k = \sqrt{1 - \rho^2} \varepsilon_k,
$$

where we assume that the common factor $Y$ is distributed as

$$
Y \sim NIG\left(\alpha, \beta, -\frac{\alpha \beta}{\sqrt{\alpha^2 - \beta^2}}, \alpha\right)
$$

and the idiosyncratic factor $\varepsilon$ as

$$
\varepsilon_k \sim NIG\left(\frac{\sqrt{1 - \rho^2}}{\rho} \alpha, \frac{\sqrt{1 - \rho^2}}{\rho} \beta, -\frac{\sqrt{1 - \rho^2}}{\rho} \frac{\alpha \beta}{\sqrt{\alpha^2 - \beta^2}}, \frac{\sqrt{1 - \rho^2}}{\rho} \alpha\right).
$$

From Equation (5.8) and (5.9) it follows that

$$
X_k \sim NIG\left(\frac{\alpha}{\rho}, \frac{\beta}{\rho}, -\frac{\alpha \beta}{\rho \sqrt{\alpha^2 - \beta^2}}, \frac{\alpha}{\rho}\right).
$$

The parameters are chosen such that $X_k$ has zero expectation. Using the normal inverse Gaussian distribution in this form for the pricing of CDO tranches has been proposed by [KSW05]. In this paper the authors price CDO tranches in the large homogeneous pool setup, where it is assumed that the underlying basket is homogeneous and its size is infinite. In this setup one ignores the different characteristics of the companies in the underlying basket, but it allows for (much) faster calculations.

Just as the student $t$-distribution, the NIG distribution has fatter tails than the normal distribution, and thus extreme events, such as a large number of losses, should be more likely.
5.4.5 Random Correlation

The previous two models changed the distribution of $X_k$ by changing the distributions of the factors. An alternative approach is to make the correlation random and maintaining the normal distribution of both factors. We consider two different models where the correlation is random.

**Mixture Factor Copula**

The first model with random correlation that we consider is a mixture of one factor Gaussian copulas. We assume that the correlation can take on $n$ different levels $\rho_1, \ldots, \rho_n$ with probability $p_1, \ldots, p_n$, with $\sum_{i=1}^{n} p_i = 1$. We arrive at the mixture factor copula from the general description (5.7) by setting

$$Z = \sum_{i=1}^{n} 1\{B=i\} \rho_i Y$$

$$E_k = \sum_{i=1}^{n} 1\{B=i\} \sqrt{1 - \rho_i^2} \varepsilon_k,$$

where $B$ is a discrete random variable that can take the values $i = 1, \ldots, n$ with probabilities $p_i$, and $Y, \varepsilon_k \sim N(0, 1)$. In the tests in Section 5.6 we will use the mixture factor copula with $n = 3$.

This simple extension of the standard model represents the uncertainty about the correlation state. Typical values of the correlation will include a very low correlation, which can be associated with an economy that is in a good state, and a very high correlation value, which corresponds to a bad state of the economy. The remaining correlation parameters is expected to lie somewhere in the middle, representing a ‘normal’ state of the economy.

The mixture factor has been considered by, amongst others, [BGL07], where also the three state version is considered. Test results for a small number of valuation dates in this paper suggest that the mixture factor copula can be calibrated well to market data. Further they provide some figures with calibrated parameters values over a larger data set, without providing the quality of the fits. In the current chapter we investigate the calibration on a larger data set and we consider both the iTraxx and CDX indices, where [BGL07] only consider the iTraxx index on a large data set.

**Random Factor Loadings Copula**

Another way to introduce randomness in the correlation parameters is by making the correlation parameter dependent on the factor $Y$. As we already briefly mentioned above, correlations tend to by higher when the market is in a bad state and they tend to be low when the market is in a good state. We choose $Z$ and $E_k$ such that they incorporate this behavior, and such that $X_k$ has zero expectation and a variance of
one.

\[ Z = \rho(Y)Y \] (5.10)

\[ E_k = \sqrt{1 - \text{Var}(\rho(Y)Y)} \varepsilon_k + \mathbb{E} \rho(Y)Y \] (5.11)

\[ \rho(y) = \alpha + \beta \frac{\exp(-\delta(y - \gamma))}{1 + \exp(-\delta(y - \gamma))}. \] (5.12)

Again we let \( Y, \varepsilon_k \sim N(0, 1) \) and further we let \( \alpha, \beta \geq 0, \alpha + \beta \leq 1, \delta > 0 \). The function to let the correlation depend on state of the economy is chosen such that it always lies between zero and one, and it is strictly decreasing. Further we can adjust the function easily through the four scaling and shifting parameters.

The general setup of the random factor loading factor copula is suggested by [AS04]. In their paper they use a much simpler description for the function \( \rho(y) \), which can take on only two different values. With the formulation (5.12) we let the correlation depend on the state of the economy in a continuous way. In a paper by [BGL07] a similar model is considered, where the correlation also depends on the state of the economy. The idiosyncratic term is defined differently though, as the term in front of the \( \varepsilon_i \) varies with the correlation function as well, instead of in our description, where we use the variance of \( \rho(Y)Y \) and its expectation. Further the authors in [BGL07] try to directly calibrate the shape of the function for the correlation as a function of \( Y \), where we have parameterized this function. Some tests with both models, with the same description for the correlation function, have shown a better performance for the random factor loading copula as described in (5.10) - (5.12). Therefore we have chosen to consider this description in the analysis in this chapter.

## 5.5 Approach

In this chapter we compare the models discussed in Sections 5.4.2 up to 5.4.5. First we want to see if the extensions can outperform the standard model, which is highly likely as the standard model can be obtained from the alternative models. In case of the Double-\( t \) factor copula we can let the degrees of freedom go to infinity, such that the distribution of the factors equals the normal distribution. For the mixture factor copula we can simply let all correlation values be equal and for the random factor loadings factor copula we set \( \beta = 0 \). The standard model cannot be obtained from the NIG factor copula.

Next we want to see if we can reproduce market data by calibrating the models to the market data such that the data is matched as closely as possible. Further we are interested if the performance of the models changes over the different points in the data set.

In this section we will mainly focus on the different techniques that we use to find these optimal parameters. Further we discuss some (possible) issues one might encounter when trying to find these optimal parameters. We conclude this section by
describing the market data that we use for the analysis. The results of this analysis are discussed in the next section.

5.5.1 The Sum of Squared Errors

The market data for one day includes CDS spreads for each company in the basket underlying the CDO tranches. Further we have quotes for the five CDO tranches, either on the European iTraxx basket or on the North-American CDX basket. The CDS spreads are matched by choosing the levels of the piecewise constant intensity, starting with the first period.

In order to match the CDO tranche data as closely as possible for a given day we need to define the distance between the market data and the fair upfront or the fair spread which results from the alternative model. Therefore we consider the sum of squared errors between the market quotes and the fair spreads obtained by an alternative model.

\[
S(p) := \sum_{i=1}^{5} w_i (Q_i - \bar{s}_i(p))^2 ,
\]

where \(Q_i\) represents the market quote for tranche \(i\) and \(\bar{s}_i\) represents the fair spread for tranche \(i\) obtained with the set of parameters \(p\), which is different for each model. Further we have introduced weights \(w_i\). In case \(i = 1\) we consider the equity tranche which in our analysis always has attachment point 0% and detachment point 3%, and the market quote and the fair spread are in this case an upfront percentage of the tranche notional. Together with this upfront premium a spread of 500 basis point has to be paid per year on the outstanding notional. In case \(i > 1\), \(Q_i\) and \(\bar{s}_i\) represent the premium in basis points which has to be paid per year on the outstanding notional of the tranche.

Before we are able to compute the sum of squared error \(S(p)\) we need to define the weights that will be used. We use two different kinds of weights. First we consider the errors relative to the market quotes, which means that we choose \(w_i = Q_i^{-2}\). An alternative set of weights that we consider is given through \(w_i = Q_i^{-1}\). The former set of weights favors market quotes with low values, which are observed for the most senior tranches. The latter set of weights, compared to the former, puts more weight on market quotes with higher values, as one divides by a smaller number and hence the contribution to the SSE becomes larger. Another possible choice of \(w_i\) is \(w_i \equiv 1\), which measures absolute errors. This choice puts high weights on tranches with high spreads, and very low weight on tranches with very low spreads. Therefore we do not consider this choice in the analysis in this chapter.

5.5.2 Finding the Minimum of the SSE

In order to fit the alternative models to the market data, we need to find the set of optimal parameters such that \(S(p)\) reaches its minimum value. When finding this
minimum one needs to keep parameters restrictions in mind. It is thus clear that we have a constrained minimization problem.

Before we try to solve this problem, we first transform the constrained minimization problem into a unconstrained minimization problem by defining transformations from $\mathbb{R}$ to the (possibly) bounded region for each parameter. The transformations are described in Section 5.5.2. After we have obtained the unconstrained problem we apply the Levenberg-Marquardt algorithm, which is designed for unconstrained minimization problems, where the objective function is of the form (5.13). In the appendix we briefly describe this algorithm. Using the Levenberg-Marquardt algorithm we are able to obtain real valued optimal parameters for the unconstrained problem. By applying the parameter transformations to these optimal values we obtain the optimal parameter values for the constrained problem.

### Parameters Transformation

As we discussed in the previous section, the Levenberg-Marquardt algorithm only works on real valued parameters. In our analysis the parameters are usually restricted to some interval or to a part of the real line. Further, parameters may depend on the value of other parameters. We denote the set of valid parameters values $\mathbf{p}$ by $\Theta \subset \mathbb{R}^k$. In this section we consider some transformations to map $\Theta$ to $\mathbb{R}$ and vice versa. Some restriction on these transformation are that they should be bijective and continuous. The first transformation maps the interval $(a,b)$, with $-\infty < a < b < \infty$ to $\mathbb{R}$. This transformation can be applied to the correlation which should be in the interval $(0,1)$. Together with the transformation we also provide its inverse.

$x \mapsto \log \left( \frac{x - a}{b - x} \right) : (a, b) \rightarrow \mathbb{R}$

$y \mapsto a + \frac{b - a}{1 + \exp(y)} : \mathbb{R} \rightarrow (a, b)$.  

These mappings are bijective and continuous. If we a have parameters which bounds are determined by the value of other parameters, e.g. $g(p_1) < p_2 < h(p_1)$ we can still apply this transformation as soon as the value of this parameter $p_1$ is known. We just have to set $a = g(p_1)$ and $b = h(p_1)$.

Another transformation that we need is a transformation that maps the interval $(a, \infty)$ onto $\mathbb{R}$. This can easily be done by the transformations

$x \mapsto \log(x - a) : (a, \infty) \rightarrow \mathbb{R}$

$y \mapsto a + \exp(y) : \mathbb{R} \rightarrow (a, \infty)$.  

Again it is directly clear that the transformations are bijective and continuous.

We only consider transformation for open intervals. In case the parameters are restricted to a closed interval we can safely ignore the parameter values at the boundaries since the objective functions depend continuously on the parameters, and thus we can get arbitrarily close to the boundaries.
Choice of Initial Parameters and Numerical Issues

Above we have shown how we can apply the Levenberg-Marquardt algorithm to the constrained optimization problem (5.13). In order to start this algorithm an initial choice for the parameters has to be made. With the initial choice of parameters the algorithm starts looking for a new set of parameters where the sum of squared errors is smaller, by calculating derivatives with respect to the parameters in the starting point. As we do not have an analytical expression it is not possible to determine an analytical expression for the derivative with respect to the parameters. Therefore we have to resort to numerical derivatives by changing the parameter by a small amount, e.g. $10^{-5}$.

In order to obtain an accurate approximation to the minimum of the SSE we need to make a good initial guess. If we choose the ‘wrong’ initial condition, the SSE might be very high, and the algorithm could end up in a local minimum that is significantly higher than the global minimum. We deal with this problem by randomly calculating initial parameters until the SSE is lower than a certain level. This level is obtained as a fraction of the global minimum obtained with the standard model. In this way we are guaranteed to end up in a minimum that is lower than the minimum obtained with the standard model. We thus avoid that the Levenberg-Marquardt algorithm chooses parameters such that the alternative model is equivalent to the standard model. On the other hand we still might not end up in the global minimum.

Another problem is that the global minimum could be in an almost flat region, by which we could end up very close to the minimum with parameters that are far from the optimal parameters. This problem can also be caused by numerical integration routines where we approximate integrals up to a certain accuracy. Due to small errors in this integration the flat region might not be smooth enough, which results in a many local minima in this flat region, all of them having approximately the same value.

5.5.3 Market Data

The analysis in this chapter is performed on a data set of CDO tranche quotes on iTraxx as well as CDX, which was provided by ABN AMRO bank. The data ranges from December 2004 up to November 2006. For a large number of days in this set we have CDS quotes from which we can obtain marginal distributions. Further we have CDO tranche quotes for maturities of five years to which we calibrate the alternative model. Further the data set includes discount curves which are used to determine the present value of future payments.

For some companies in the iTraxx or CDX indices we have some incomplete data, where CDS quotes are only available for a limited number of days. If we calibrate the models on such a day, we use the CDS data on the day that is closest to the valuation date. Further, for some days in the data set we have multiple quotes for
each CDO tranche. In these cases we compare the quotes and pick the set that seems most reasonable compared to the days before and after this date.

As this data set ranges from December 2004 up to November 2006 we do not have data for credit derivatives during the credit crisis which started in July 2007. During this period quotes have increased significantly for CDSs as well as for CDO tranches, especially the more senior tranches. The current data set does contain data from May 2005, when Ford’s and General Motors’ credit rating were downgraded. This caused some shocks in the market and the structure of the quotes changed which resulted in a steeper base correlation curve, which indicates a different correlation structure. When calibrating the alternative models it is thus interesting to see if models fit better before or after May 2005.

In the next section we discuss the results from the calibration of the alternative models to the data described above.

5.6 Results

In the previous section we have discussed the numerical techniques and the data for the tests that we have performed. In this section we discuss the results of these tests and we look which model calibrates best to the market data. For this we use several measures such as the sum of squared errors and the actual differences between market and model quotes. Further we look at the stability of the parameter values and the stability of the sum of squared errors.

5.6.1 The Sum of Squared Errors

First we consider the sum of squared errors (SSE) which is the objective function of the calibration. In case the model fits perfectly to the market data this function has zero value. In any other case it has a positive value. As we discussed in Section 5.5.1 we consider two different sets of weights \( w_i \) in the sum of squared errors, that is defined in Equation (5.13). We calibrate the models with \( w_i = Q_i^{-2} \) as well as \( w_i = Q_i^{-1} \).

For each valuation day in the data set we calibrate the models the market data, which results in a set of SSEs. In Table 5.2 we present the average, minimum, maximum and standard deviation of these SSEs corresponding to the weights \( w_i = Q_i^{-2} \). The corresponding CDO tranches have the iTraxx index as underlying and they have a five year maturity.

From this table we can see that the alternative models all perform much better than the standard model. On the other hand we see that the models, except for the mixture factor copula are not able to provide a perfect fit to the market data. In the case of the mixture factor copula we observe the perfect fit for only one day in the data set. Further we see that the differences between the averages are not very large.
Table 5.2: The average, minimum, maximum and standard deviation of the SSE for the standard model, the double-$t$ factor copula (Double-$t$), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the iTraxx index with five year maturities. The weights in the SSE (5.13) are given by $w_i = Q_i^{-2}$.

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Double-$t$</th>
<th>NIG</th>
<th>Mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg</td>
<td>2.257</td>
<td>0.0672</td>
<td>0.0574</td>
<td>0.0457</td>
<td>0.0565</td>
</tr>
<tr>
<td>min</td>
<td>0.734</td>
<td>0.0043</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0002</td>
</tr>
<tr>
<td>max</td>
<td>3.38</td>
<td>0.366</td>
<td>0.317</td>
<td>0.322</td>
<td>0.310</td>
</tr>
<tr>
<td>stdev</td>
<td>0.69</td>
<td>0.062</td>
<td>0.051</td>
<td>0.053</td>
<td>0.048</td>
</tr>
</tbody>
</table>

The same analysis has been performed for CDO tranches on the North-American CDX index, again with maturities of five years. In Table 5.3 the results are presented corresponding to the weights $w_i = Q_i^{-2}$.

Table 5.3: The average, minimum, maximum and standard deviation of the SSE for the standard model, the double-$t$ factor copula (Double-$t$), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the CDX index with five year maturities. The weights in the SSE (5.13) are given by $w_i = Q_i^{-2}$.

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Double-$t$</th>
<th>NIG</th>
<th>Mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td>2.159</td>
<td>0.0642</td>
<td>0.0622</td>
<td>0.0747</td>
<td>0.0573</td>
</tr>
<tr>
<td>minimum</td>
<td>0.666</td>
<td>0.0009</td>
<td>0.0003</td>
<td>0.0002</td>
<td>0.0007</td>
</tr>
<tr>
<td>maximum</td>
<td>3.50</td>
<td>0.271</td>
<td>0.194</td>
<td>0.537</td>
<td>0.211</td>
</tr>
<tr>
<td>st.deviation</td>
<td>0.70</td>
<td>0.056</td>
<td>0.044</td>
<td>0.106</td>
<td>0.048</td>
</tr>
</tbody>
</table>

From this table we can draw similar conclusions as from Table 5.2, although the performance of the mixture factor copula is slightly worse in this case. This is mainly due a small number of dates on which the mixture factor copula calibration performs poorly.

In Tables 5.4 and 5.5 below, the same statistics are presented, but in this case we use the weights $w_i = Q_i^{-1}$. With this choice the resulting SSEs are larger than in the previous case, since we have $Q_i^{-2} < Q_i^{-1}$ as long as $Q_i > 1$, which is (almost) always the case. Further we expect that we fit better to tranches with high spreads, and worse to tranches with low spreads.

In these two tables, we observe that the calibration errors are substantially larger than in Tables 5.2 and 5.3, which was to be expected, as the pricing errors are divided by the quote instead of the squared quote. Further we again observe that alternative models performs better than the standard model. The mixture factor copula again shows a number of large calibration errors, which results in the higher average.

Tables 5.2 - 5.5 only give a general view on the performance of the models. The statistics in these tables can be influenced by some outliers. It could be the case that
Table 5.4: The average, minimum, maximum and standard deviation of the SSEs, calculated on each valuation date for the standard model, the double-\(t\) factor copula (Double-\(t\)), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the iTraxx index with five year maturities. The weights in the SSE (5.13) are given by \(w_i = Q_i^{-1}\).

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Double-(t)</th>
<th>NIG</th>
<th>Mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg</td>
<td>41.9</td>
<td>1.61</td>
<td>1.47</td>
<td>1.50</td>
<td>1.33</td>
</tr>
<tr>
<td>min</td>
<td>17.1</td>
<td>0.107</td>
<td>0.026</td>
<td>0.000</td>
<td>0.004</td>
</tr>
<tr>
<td>max</td>
<td>143</td>
<td>8.75</td>
<td>17.7</td>
<td>27.2</td>
<td>8.57</td>
</tr>
<tr>
<td>stdev</td>
<td>19.5</td>
<td>1.72</td>
<td>2.11</td>
<td>3.25</td>
<td>1.33</td>
</tr>
</tbody>
</table>

Table 5.5: The average, minimum, maximum and standard deviation of the SSE for the standard model, the double-\(t\) factor copula (Double-\(t\)), the normal inverse Gaussian factor copula (NIG), the mixture factor copula (Mixture) and the random factor loadings copula (RFL). The values are obtained by calibrating the models to CDO tranches on the CDX index with five year maturities. The weights in the SSE (5.13) are given by \(w_i = Q_i^{-1}\).

<table>
<thead>
<tr>
<th></th>
<th>standard</th>
<th>Double-(t)</th>
<th>NIG</th>
<th>mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg</td>
<td>57.0</td>
<td>2.28</td>
<td>1.79</td>
<td>6.13</td>
<td>2.04</td>
</tr>
<tr>
<td>min</td>
<td>21.0</td>
<td>0.023</td>
<td>0.012</td>
<td>0.006</td>
<td>0.016</td>
</tr>
<tr>
<td>max</td>
<td>337.6</td>
<td>17.8</td>
<td>20.8</td>
<td>97.1</td>
<td>12.0</td>
</tr>
<tr>
<td>stdev</td>
<td>44.0</td>
<td>2.89</td>
<td>2.56</td>
<td>13.9</td>
<td>2.34</td>
</tr>
</tbody>
</table>

A model can provide a good fit to market data, except for a small number of dates on which a poor fit is found, as is the case for the mixture factor copula. In this case the average and maximum might not be a good measure of the performance of the model. We therefore have added figures where the SSE are plotted for each date in the data set. The graphs of the SSEs for the four different cases can be found in Appendix 5.B in Figures 5.1 - 5.4. In these figures we see that in general the SSE are quite close for the four models, although for a small number of dates in the first half of the data set we can observe that the mixture factor copula performs worse than the other models, when calibrating the CDO quotes on the CDX index. In the second half of the data set however, we see that the mixture copula performs better than the other models. From the figures, it does not become clear whether the models calibrate better before or after May 2005.

5.6.2 Fair Premia vs. Market Quotes

In the previous section we have investigated the SSE, where we have looked at the average SSE for each of the alternative models, and compared these values with the standard model and with the other alternative models. Further, we have looked at some figures showing the SSE for each day in the data set. The SSE, even for a single day, only gives a global view on the calibration performance of the models. From this statistic it does not become clear how well, or how poor, the model fits to each
CDO tranche. In this section we look at the fair spreads, that are obtained from the model with the optimal parameters for each CDO tranche, and we compare these to the corresponding market quotes. This makes it easier to compare the performance of the models and it becomes easier to compare the calibration with the different sets of weights which is not possible by only looking at the SSE.

We first consider the weights $Q_i^{-2}$, where we calculate the average absolute differences as well as the relative absolute difference with respect to the corresponding market quote. In Tables 5.6 and 5.7 the averages are presented.

<table>
<thead>
<tr>
<th>CDO tranche</th>
<th>differences</th>
<th>Standard</th>
<th>Double-$t$</th>
<th>NIG</th>
<th>Mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0%,3%)</td>
<td>absolute</td>
<td>4.26</td>
<td>4.08</td>
<td>3.74</td>
<td>3.65</td>
<td>4.16</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>16.79%</td>
<td>16.52%</td>
<td>15.60%</td>
<td>14.44%</td>
<td>16.86%</td>
</tr>
<tr>
<td>(3%,6%)</td>
<td>absolute</td>
<td>67.97</td>
<td>7.74</td>
<td>5.02</td>
<td>5.93</td>
<td>4.45</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>76.73%</td>
<td>9.02%</td>
<td>5.61%</td>
<td>6.47%</td>
<td>4.90%</td>
</tr>
<tr>
<td>(6%,9%)</td>
<td>absolute</td>
<td>9.43</td>
<td>2.11</td>
<td>1.57</td>
<td>1.42</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>33.91%</td>
<td>8.28%</td>
<td>5.87%</td>
<td>5.01%</td>
<td>6.52%</td>
</tr>
<tr>
<td>(9%,12%)</td>
<td>absolute</td>
<td>8.62</td>
<td>0.81</td>
<td>1.07</td>
<td>0.46</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>67.13%</td>
<td>5.69%</td>
<td>7.47%</td>
<td>3.18%</td>
<td>6.39%</td>
</tr>
<tr>
<td>(12%,22%)</td>
<td>absolute</td>
<td>7.11</td>
<td>0.38</td>
<td>0.27</td>
<td>0.25</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>93.55%</td>
<td>4.08%</td>
<td>3.85%</td>
<td>2.57%</td>
<td>2.52%</td>
</tr>
</tbody>
</table>

Table 5.6: The averages of the absolute and absolute relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the iTraxx index as underlying and the weights $Q_i^{-2}$ are used.

<table>
<thead>
<tr>
<th>CDO tranche</th>
<th>differences</th>
<th>Standard</th>
<th>Double-$t$</th>
<th>NIG</th>
<th>Mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0%,3%)</td>
<td>absolute</td>
<td>4.82</td>
<td>4.32</td>
<td>4.09</td>
<td>3.98</td>
<td>4.80</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>12.69%</td>
<td>11.48%</td>
<td>11.04%</td>
<td>10.56%</td>
<td>12.71%</td>
</tr>
<tr>
<td>(3%,7%)</td>
<td>absolute</td>
<td>100.08</td>
<td>15.22</td>
<td>13.12</td>
<td>20.82</td>
<td>11.61</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>78.42%</td>
<td>10.04%</td>
<td>10.12%</td>
<td>13.33%</td>
<td>7.25%</td>
</tr>
<tr>
<td>(7%,10%)</td>
<td>absolute</td>
<td>9.23</td>
<td>3.14</td>
<td>2.13</td>
<td>2.56</td>
<td>2.92</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>22.72%</td>
<td>8.73%</td>
<td>5.89%</td>
<td>6.86%</td>
<td>8.64%</td>
</tr>
<tr>
<td>(10%,15%)</td>
<td>absolute</td>
<td>9.28</td>
<td>1.03</td>
<td>1.23</td>
<td>0.62</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>62.94%</td>
<td>7.78%</td>
<td>9.23%</td>
<td>3.82%</td>
<td>8.54%</td>
</tr>
<tr>
<td>(15%,30%)</td>
<td>absolute</td>
<td>7.29</td>
<td>0.56</td>
<td>0.46</td>
<td>0.38</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>94.19%</td>
<td>6.29%</td>
<td>6.74%</td>
<td>3.48%</td>
<td>3.48%</td>
</tr>
</tbody>
</table>

Table 5.7: The averages of the absolute and absolute relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the CDX index as underlying and the weights $Q_i^{-2}$ are used.

The results given in these tables give more insights into the performance of the models, since we have broken down the calibration error over the five different tranches. We again observe the superior behavior of the alternative models with respect to the
standard model. Further we can see that the absolute as well as the relative errors are reasonably small for the most senior tranches. The errors on the equity and junior mezzanine tranches are quite large. This is what we expected from using these types of weights, as smaller quotes are given larger weights. Comparing the alternative models, we can observe that none of the models performs better than the other, as one model performs better on one tranche and another model performs better on a different tranche.

Next we consider the same statistics, only in this case we use the weights \( Q^{-1}_i \). In Section 5.5.1 we have mentioned that the two sets of weights should provide a different fit to the market data, as the first set of weights, \( w_i = Q^{-2}_i \), focusses more on the more senior tranches and the second set, \( w_i = Q^{-1}_i \) focusses less on the senior tranches and more on the junior mezzanine tranches. We can thus expect that the errors on the senior tranches become larger and the errors on the junior mezzanine tranches become smaller.

Tables 5.8 and 5.9 present the results from calibrating with weights \( w_i = Q^{-1}_i \).

<table>
<thead>
<tr>
<th>CDO tranche</th>
<th>differences</th>
<th>Standard</th>
<th>Double-t</th>
<th>NIG</th>
<th>Mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0%,3%)</td>
<td>absolute</td>
<td>6.95</td>
<td>3.99</td>
<td>3.82</td>
<td>3.62</td>
<td>4.27</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>27.95%</td>
<td>16.18%</td>
<td>15.93%</td>
<td>14.31%</td>
<td>17.30%</td>
</tr>
<tr>
<td>(3%,6%)</td>
<td>absolute</td>
<td>18.29</td>
<td>2.33</td>
<td>2.18</td>
<td>2.75</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>17.17%</td>
<td>2.77%</td>
<td>2.36%</td>
<td>2.84%</td>
<td>1.54%</td>
</tr>
<tr>
<td>(6%,9%)</td>
<td>absolute</td>
<td>18.40</td>
<td>2.29</td>
<td>1.48</td>
<td>1.61</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>74.16%</td>
<td>9.26%</td>
<td>5.57%</td>
<td>5.86%</td>
<td>4.94%</td>
</tr>
<tr>
<td>(9%,12%)</td>
<td>absolute</td>
<td>12.77</td>
<td>1.10</td>
<td>1.17</td>
<td>0.76</td>
<td>1.19</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>92.63%</td>
<td>8.40%</td>
<td>8.24%</td>
<td>5.62%</td>
<td>9.26%</td>
</tr>
<tr>
<td>(12%,22%)</td>
<td>absolute</td>
<td>7.63</td>
<td>1.03</td>
<td>0.79</td>
<td>0.68</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>98.85%</td>
<td>14.78%</td>
<td>10.70%</td>
<td>9.51%</td>
<td>6.05%</td>
</tr>
</tbody>
</table>

**Table 5.8:** The averages of the absolute and absolute relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the iTraxx index as underlying and the weights \( Q^{-1}_i \) are used.

Looking at the tables, we can indeed observe that the calibration errors on the senior tranches have increased, and that the errors on the junior mezzanine tranches have decreased. Again, we cannot clearly point out a model that performs better than the other models.

The Tables 5.6 - 5.9 again provide a global view on the performance, since we consider averages over all valuation dates at the same time. In order to analyze the day-to-day performance we could look at figures where the market quotes are plotted together with the fair spreads calculated with the alternative models. We do not show such figures here, as this would require twenty different figures, which do not add much additional information. Instead, we confine ourselves with a few remarks about the results. First, it was observed that the fair premia from the alternative models, tend
Section 5.6 · Results

<table>
<thead>
<tr>
<th>CDO tranche</th>
<th>differences</th>
<th>Standard</th>
<th>Double-t</th>
<th>NIG</th>
<th>Mixture</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0%,3%)</td>
<td>absolute</td>
<td>6.96</td>
<td>4.28</td>
<td>3.47</td>
<td>3.91</td>
<td>4.85</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>18.00%</td>
<td>11.53%</td>
<td>9.59%</td>
<td>10.39%</td>
<td>13.03%</td>
</tr>
<tr>
<td>(3%,7%)</td>
<td>absolute</td>
<td>34.80</td>
<td>4.30</td>
<td>3.95</td>
<td>13.08</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>21.85%</td>
<td>2.79%</td>
<td>2.76%</td>
<td>8.42%</td>
<td>2.10%</td>
</tr>
<tr>
<td>(7%,10%)</td>
<td>absolute</td>
<td>22.40</td>
<td>3.12</td>
<td>2.31</td>
<td>3.81</td>
<td>2.32</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>76.73%</td>
<td>8.11%</td>
<td>6.65%</td>
<td>9.88%</td>
<td>6.75%</td>
</tr>
<tr>
<td>(10%,15%)</td>
<td>absolute</td>
<td>13.97</td>
<td>1.75</td>
<td>1.56</td>
<td>1.46</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>91.76%</td>
<td>11.33%</td>
<td>10.12%</td>
<td>7.63%</td>
<td>12.15%</td>
</tr>
<tr>
<td>(15%,30%)</td>
<td>absolute</td>
<td>7.68</td>
<td>1.58</td>
<td>1.36</td>
<td>1.74</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>relative</td>
<td>98.89%</td>
<td>17.29%</td>
<td>18.17%</td>
<td>17.31%</td>
<td>12.03%</td>
</tr>
</tbody>
</table>

Table 5.9: The averages of the absolute and relative differences between the fair spread obtained with the alternative models and the market quotes. The CDO tranches have the CDX index as underlying and the weights $Q_i^{-1}$ are used.

to closely follow each other and to a lesser extend the market quotes. Second, in the second half of the data set, the fair premia for the alternative models are very close to the market quotes on the most senior tranches when using the relative weights. The fair premia were very close to the market quotes on the junior mezzanine tranche when using the weights $w_i = Q_i^{-1}$. Further similar conclusions to the ones from Tables 5.6 - 5.9 can be drawn.

5.6.3 Parameter Stability

In the previous two sections we have investigated the calibration errors of the alternative one-factor copula models. We did this first by directly looking at these errors, and secondly by looking at the differences between model and market prices. In this section we focus on a third important issue, namely the parameter stability. A model might be able to match market prices on a daily basis, but with different parameter values on each day. This could indicate that we are overfitting the market data, which is an undesirable feature of a model.

To test the parameter stability, we compute a number of statistics for the optimal parameters, such as averages, standard deviations and correlations between the different model parameters. As each model has different kinds of parameters, we present these statistics separately for each model. For the one-factor Gaussian copula we only have to deal with a single parameter. In Table 5.10 one can find the average correlations and its standard deviations.

From this table we can observe that the optimal correlation is in general small, and it is quite stable, as the standard deviation is not very large. However, the results for the one-factor Gaussian copula are not very interesting, as it cannot match the market quotes.
Table 5.10: The average and standard deviations of the optimal values for the correlation parameter $\rho$ in the one-factor Gaussian copula.

<table>
<thead>
<tr>
<th></th>
<th>$w_i = Q_i^{-2}$</th>
<th>$w_i = Q_i^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iTraxx</td>
<td>CDX</td>
</tr>
<tr>
<td>avg. $\rho$</td>
<td>0.107</td>
<td>0.125</td>
</tr>
<tr>
<td>std.dev. $\rho$</td>
<td>0.048</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Next we consider the double-$t$ factor copula, which matches the market quotes considerably better. There are three parameters that we need to consider in this case, which are the correlation $\rho$ and the degrees of freedom for the common and idiosyncratic factor, $\nu_Y$ and $\nu_{\epsilon}$, respectively. In Table 5.11 one finds the average and standard deviation of the optimal choices of these parameters.

Table 5.11: The average and standard deviations of the optimal parameter values in the double-$t$ factor copula.

<table>
<thead>
<tr>
<th></th>
<th>$w_i = Q_i^{-2}$</th>
<th>$w_i = Q_i^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iTraxx</td>
<td>CDX</td>
</tr>
<tr>
<td>avg. $\rho$</td>
<td>0.550</td>
<td>0.388</td>
</tr>
<tr>
<td>avg. $\nu_Y$</td>
<td>4.89</td>
<td>2.53</td>
</tr>
<tr>
<td>avg. $\nu_{\epsilon}$</td>
<td>50.26</td>
<td>21.74</td>
</tr>
<tr>
<td>std.dev. $\rho$</td>
<td>0.312</td>
<td>0.265</td>
</tr>
<tr>
<td>std.dev. $\nu_Y$</td>
<td>4.23</td>
<td>0.82</td>
</tr>
<tr>
<td>std.dev. $\nu_{\epsilon}$</td>
<td>120.72</td>
<td>62.86</td>
</tr>
</tbody>
</table>

From this table we can observe that the standard deviation of $\nu_{\epsilon}$ is very large, just as for $\nu_Y$ in case we look at the market quotes for CDO tranches on the CDX index, with the weights $w_i = Q_i^{-1}$. Furthermore, the corresponding averages of the degrees of freedom are also quite high. These high values are caused by very high optimal values for the idiosyncratic degrees of freedom on about 30% of the dates considered. In these cases the degrees of freedom are such that the corresponding factor is in fact very close to being normally distributed. Ignoring these dates yields much lower and more stable degrees of freedom, although the fraction of 30% does not justify ignoring this procedure. We therefore conclude that the optimal parameters are not very stable.

For the NIG factor copula, we have to consider the correlation $\rho$ and the two parameters related to the NIG distribution, namely $\alpha$ and $\beta$, as introduced in Section 5.4. For the optimal parameter values we have again calculated the averages and standard deviations, which can be found in Table 5.12.

Considering the standard deviations, we observe these are lower than for the double-$t$ factor copula, indicating more stable optimal parameters. Furthermore we see that the average correlation is about the same for the four different cases that are consid-
The average and standard deviations of the optimal parameter values in the NIG factor copula.

<table>
<thead>
<tr>
<th></th>
<th>( w_i = Q_i^{-2} )</th>
<th>( w_i = Q_i^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iTraxx</td>
<td>CDX</td>
</tr>
<tr>
<td>avg. ( \alpha )</td>
<td>0.828</td>
<td>0.711</td>
</tr>
<tr>
<td>avg ( \beta )</td>
<td>-0.552</td>
<td>-0.427</td>
</tr>
<tr>
<td>avg. ( \rho )</td>
<td>0.335</td>
<td>0.366</td>
</tr>
<tr>
<td>std.dev. ( \alpha )</td>
<td>0.544</td>
<td>0.591</td>
</tr>
<tr>
<td>std.dev. ( \beta )</td>
<td>0.729</td>
<td>0.705</td>
</tr>
<tr>
<td>std.dev. ( \rho )</td>
<td>0.070</td>
<td>0.065</td>
</tr>
</tbody>
</table>

Table 5.12: The average and standard deviations of the optimal parameter values in the NIG factor copula.

A point of concern for this model, or at least for the current implementation of the model and the calibration, is that there is a high negative correlation. For each of the cases this is less than -94%. On a large number of dates, this means that approximately \( \alpha = -\beta \). As one divides by \( \sqrt{\alpha^2 - \beta^2} \) in the specification of the model, this could lead to numerical difficulties.

Next, we consider the mixture factor copula, where we have to consider three correlation parameters, namely \( \rho_1 \), \( \rho_2 \), and \( \rho_3 \) and the corresponding probabilities \( p_1 \), \( p_2 \) and \( p_3 = 1 - p_1 - p_2 \). As the calibration algorithm does not necessarily keep the correlations ordered, we first order the optimal correlation parameters, and thereafter we calculate the averages and standard deviations, which are given in Table 5.13.

<table>
<thead>
<tr>
<th></th>
<th>( w_i = Q_i^{-2} )</th>
<th>( w_i = Q_i^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iTraxx</td>
<td>CDX</td>
</tr>
<tr>
<td>avg. ( \rho_1 )</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>avg. ( \rho_2 )</td>
<td>0.222</td>
<td>0.253</td>
</tr>
<tr>
<td>avg. ( \rho_3 )</td>
<td>0.947</td>
<td>0.948</td>
</tr>
<tr>
<td>avg. ( p_1 )</td>
<td>0.637</td>
<td>0.628</td>
</tr>
<tr>
<td>avg. ( p_2 )</td>
<td>0.265</td>
<td>0.263</td>
</tr>
<tr>
<td>std.dev. ( \rho_1 )</td>
<td>0.009</td>
<td>0.002</td>
</tr>
<tr>
<td>std.dev. ( \rho_2 )</td>
<td>0.092</td>
<td>0.135</td>
</tr>
<tr>
<td>std.dev. ( \rho_3 )</td>
<td>0.117</td>
<td>0.101</td>
</tr>
<tr>
<td>std.dev. ( p_1 )</td>
<td>0.189</td>
<td>0.155</td>
</tr>
<tr>
<td>std.dev. ( p_2 )</td>
<td>0.153</td>
<td>0.160</td>
</tr>
</tbody>
</table>

Table 5.13: The average and standard deviations of the optimal parameter values in the mixture factor copula.

From this table we observe that the lowest correlation value is almost (always) equal to zero, as the average as well as the standard deviation are close to zero. Furthermore, the highest correlation value is close to one, again with a low standard deviation. The remaining correlation parameter, \( \rho_2 \), takes values around 0.25, where the standard deviation is largest when we calibrate the CDO tranches on the CDX basket, with the weights \( w_i = Q_i^{-1} \). Similar to the correlations, also the probabilities show
a quite stable behavior, as the average probabilities are roughly equal across the test cases, and the standard deviations are quite low. Comparing these results with those of the double-\( t \) and NIG factor copulas, we observe a higher parameter stability for the mixture factor copula.

We conclude this section with an investigation of the optimal parameter values of the random factor loadings factor copula, where we have to consider four parameters, namely \( \alpha, \beta, \gamma \) and \( \delta \), where \( \alpha \) and \( \beta \) determine the minimum and maximum correlation value, \( \delta \) determines the speed at which the correlation decreases and \( \gamma \) determines the point where the correlations are exactly halfway between \( \alpha \) and \( \beta \). In Table 5.14 the averages and standard deviations of the optimal parameters are shown.

<table>
<thead>
<tr>
<th></th>
<th>( w_i = Q_i^{-2} )</th>
<th></th>
<th>( w_i = Q_i^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iTraxx</td>
<td>CDX</td>
<td>iTraxx</td>
</tr>
<tr>
<td>avg. ( \alpha )</td>
<td>0.131</td>
<td>0.151</td>
<td>0.111</td>
</tr>
<tr>
<td>avg. ( \beta )</td>
<td>0.794</td>
<td>0.825</td>
<td>0.817</td>
</tr>
<tr>
<td>avg. ( \gamma )</td>
<td>-2.757</td>
<td>-2.690</td>
<td>-2.730</td>
</tr>
<tr>
<td>avg. ( \delta )</td>
<td>17.609</td>
<td>3.583</td>
<td>14.214</td>
</tr>
<tr>
<td>std.dev. ( \alpha )</td>
<td>0.106</td>
<td>0.107</td>
<td>0.106</td>
</tr>
<tr>
<td>std.dev. ( \beta )</td>
<td>0.202</td>
<td>0.141</td>
<td>0.186</td>
</tr>
<tr>
<td>std.dev. ( \gamma )</td>
<td>0.344</td>
<td>0.255</td>
<td>0.335</td>
</tr>
<tr>
<td>std.dev. ( \delta )</td>
<td>98.224</td>
<td>5.703</td>
<td>78.435</td>
</tr>
</tbody>
</table>

Table 5.14: The average and standard deviations of the optimal parameter values in the random factor loadings factor copula.

From this table we observe that the first three averages are roughly equal across the four test cases, the \( \delta \) however can be quite large, and its standard deviation is very high, especially when we use the weights \( w_i = Q_i^{-2} \). These high values are a direct result of a small number of very high values of \( \delta \). When we would ignore these results, which occur on less than 5% of the dates, we find much lower average deltas, between 2.1 and 2.7, with standard deviations between 1.1 and 2.1, which shows that, except for the ‘outliers’, the parameters for the random factor loadings factor copula are quite stable.

5.7 Concluding Remarks

In this chapter we have compared four different one-factor copula models. First of all we have looked at the ability of the models to fit to market data for CDO tranches on the iTraxx and CDX index. We have compared the calibration errors of these models, with those of the one-factor Gaussian copula, and with errors of the other models. To investigate if we overfit to the market data, we have carried out a stability analysis of the optimal parameters. Based on the test results that we have presented in this chapter, we summarize a number of general conclusions with respect to the models.
The first, and most straightforward, conclusion that can be drawn from the results is that each of the four models performs much better than the one-factor Gaussian copula. Comparing the calibration errors for the factor copula models, we can see that none of the models has the lowest error in each of the four test cases. On the other hand we have seen that the errors for the models, do not differ much.

With respect to the comparison between the fair premia and the market quotes we have seen that the two different types of weights used in the calibration function favor different tranches. Furthermore, it was observed that the fair premia of the models are quite close to each other. Based on the test results we cannot select a model that overall significantly outperforms the other models.

When looking at the optimal parameters of the models, a number of interesting conclusions has been drawn. As we have seen in Section 5.6.3, the optimal parameters for the double-$t$ factor copula are not very stable, and for the NIG factor copula, the parameters $\alpha$ and $\beta$ showed an undesirable correlation. The two random correlation models however, showed more stable parameters, although for a small number of cases in the random factor loadings factor copula the optimal values of $\delta$ can be very large.

## 5.A The Levenberg-Marquardt Algorithm

The Levenberg-Marquardt algorithm is designed for optimization problems where the parameters are all real valued and the objective function has the form

$$S(p) := \sum_{i=1}^{N} w_i (f(x_i, p) - y_i)^2,$$

where the parameters are given by $p \in \mathbb{R}^M$, and the goal is to fit the function $f(x, p)$ to the data $\{(x_i, y_i) | i = 1, \ldots, N\}$. In this section we briefly describe the Levenberg-Marquardt algorithm. For a more detailed description of the algorithm we refer to [PTVF02].

The Levenberg-Marquardt algorithm combines two techniques that can be used to find the minimum of a problem of the form (5.14). The first is the steepest descent method, where one determines the gradient in a given point and change the parameters by a constant times the gradient, such that the value of the objective function decreases. This method works well if we are far away from the minimum, but if we get closer to the minimum we might wander around the minimum, since the derivatives are close to zero.

The second approach uses a quadratic approximation of (5.14). It is easy to find the minimum of this quadratic approximation. This results in an iterative procedure to find the minimum of the nonlinear problem. This approach works well if the quadratic approximation is accurate, which is typically the case when we are reasonably close to the minimum. By combining both methods it is possible to construct
an algorithm that should work well in both cases. In the remainder of this section we provide the Levenberg-Marquardt algorithm.

Start with an initial choice \( p^{(0)} \in \mathbb{R}^M \) and set \( \lambda = \lambda_0 \), where \( \lambda \) is used to smoothly switch between the two methods. The initial value \( \lambda_0 \) is chosen to be a modest number, e.g. 0.001.

At step \( m \) we want to find \( p^{(m+1)} \in \mathbb{R}^M \) based on \( p^{(m)} \). Therefore we need to compute the following quantities

- **the gradient**
  \[
  \beta_k = 2 \sum_{i=1}^{N} w_i (f(x_i, p) - y_i) \frac{\partial f(x_i, p^{(m)})}{\partial p_k},
  \]

- **an approximation to the Hessian**
  \[
  \alpha_{kl} = 2 \sum_{i=1}^{N} w_i \frac{\partial f(x_i, p^{(m)})}{\partial p_k} \frac{\partial f(x_i, p^{(m)})}{\partial p_l}.
  \]

Terms involving second derivatives are ignored, as these are typically small. For additional motivation of this approximation we refer to [PTVF02].

- **the matrix** \( A_{kl} = \left\{ \begin{array}{ll} \alpha_{kl} & k \neq l \\ \alpha_{kl} (1 + \lambda) & k = l \end{array} \right. \).

We obtain a candidate for \( p^{(m+1)} \) by solving (5.15) for \( \delta p = p^{(m+1)} - p^{(m)} \)

\[
A \cdot \delta p = \beta. \tag{5.15}
\]

When \( \delta p \) is found, we set \( \tilde{p} = p^{(m)} + \delta p \) and we calculate \( S(\tilde{p}) \). Then we check if \( \tilde{p} \) results in an improvement.

- **\( S(\tilde{p}) < S(p^{(m)}) \):** Set \( p^{(m+1)} = \tilde{p} \) and divide \( \lambda \) by \( \lambda_{\text{scale}} \), where \( \lambda_{\text{scale}} \) is a substantial number, e.g. 10. By this division the algorithm behaves more like the quadratic approximation method.

- **\( S(\tilde{p}) \geq S(p^{(m)}) \):** Set \( p^{(m+1)} = p^{(m)} \) as there was no improvement. Further we multiply \( \lambda \) by \( \lambda_{\text{scale}} \) such that the algorithm behaves more like the steepest descent method.

We stop the algorithm when the change in \( S(p) \) is very small. The algorithm should not be stopped when there was no improvement, as \( \lambda \) might now have reached to proper level. In case there was an improvement we determine the absolute improvement \( S(p^{(m)}) - S(p^{(m+1)}) \) together with the relative improvement \( \frac{S(p^{(m)}) - S(p^{(m+1)})}{S(p^{(m)})} \).

We stop the algorithm if one or both of these errors are below some small level, e.g. 0.0001.
5.B Figures

In this section one can find the figures that are referred to in this paper.

Figure 5.1: SSE’s for the alternative models calibrated to CDO tranches on the iTraxx index with a five year maturity and using the weights $Q_i^{-2}$. 
Figure 5.2: SSE’s for the alternative models calibrated to CDO tranches on the CDX index with a five year maturity and using the weights $Q_i^{-2}$.
Figure 5.3: SSE’s for the alternative models calibrated to CDO tranches on the iTraxx index with a five year maturity and using the weights $Q_i^{-1}$. The results are plotted on a logarithmic scale due to the wide range.
Figure 5.4: SSE’s for the alternative models calibrated to CDO tranches on the CDX index with a five year maturity and using the weights $Q_i^{-1}$. The results are plotted on a logarithmic scale due to the wide range.