In the Core of Correlation

Jon Gregory and Jean-Paul Laurent

April 2004

Introduction

The modelling of dependence between defaults is a key issue for the valuation and risk management of multi-name credit derivatives. The Gaussian copula model seems to have become an industry standard for pricing. It's appeal is partly due to its ease of implementation via Monte Carlo simulation and the fact that the underlying dependence structure has for a long time been linked to equity returns correlation. Furthermore, a big driving force behind the adoption of this approach has been the tractability in reduced dimension, with fast (semi-)analytical calculations of prices and deltas of CDO tranches and basket default swaps. The simplest form of the model is the so-called one factor Gaussian copula.

There do, however, remain some drawbacks in the applicability of the one factor Gaussian model:

- There is a reported "correlation smile" in the CDO market[1] similar to the well-known Black-Scholes implied volatility smile. Whether this is due to liquidity effects or a more theoretical issue such as the choice of copula (e.g. fat tails in the loss distribution) remains an open question.
- The one-factor correlation structure is rather limited, for example, we cannot have separated "regions" such as a highly correlated domain amongst a background of low correlation.
- There are some practical issues with the representation and aggregation of correlation risk, for example calibration done at the transaction level means that a name can exist in different portfolios with different associated correlation parameters. It is more realistic to view risk to changes in underlying correlations rather than the factor(s) themselves.

This paper aims to provide practical results in light of such issues. We show that in a Gaussian copula framework we can keep the appeal of analytical tractability and:

- Provide a more intuitive correlation structure, leading itself readily to correlation risk analysis.
- Compute correlation sensitivities either via the above structure or analytically in an extension of the one-factor model.
- Introduce some dependence between recovery rates and between recovery rates and defaults.

While this paper is dedicated to CDO tranches, the results can be directly applied to k^{th} to default swaps and to portfolio credit risk analysis.

Dependence in default times and Gaussian copula

Default dates or default events usually exhibit positive dependence. Due to economic cycles or firm interactions, defaults tend to cluster together. There is a variety of approaches to tackle this phenomenon. For instance, one may use dependent default intensities as in Duffie and Garleanu (2001). A related approach is that of Jarrow and Yu (2000), where the inputs are the jumps in credit spreads at default times. Arvanitis and Gregory (2001), Finger (2000) and Hull and White (2001) have proposed a discrete time firm value approach to the valuation of multiname credit derivatives.

The most widely used models in the industry are based on the copula approach initiated by Li (2000) and further developed by Schönbucher and Schubert (2001). The Gaussian copula also corresponds to CreditMetrics[TM] (see Gupton, Finger and Bahia (1997)). Other copulas such as Clayton (Rogge and Schönbucher (2003)) or Student t copula (Mashal and Naldi (2001)) have been proposed to capture tail dependence effects. Nevertheless, Schlögl and O'Kane (2003) show that Student t copula

---

[1] The relevant implied correlation is the flat correlation implied from the market price of a particular tranche.

provides a worse characterisation of the correlation smile than the Gaussian approach. Further work thus seems to be required to assess the importance of a dependence structure.

Let us introduce some notations. We consider $i = 1,...,n$ names with default times $\tau_1,...,\tau_n$. We denote by $S_i(t) = Q(\tau_i > t)$ the marginal survival functions[2], $F_i(t) = Q(\tau_i \leq t)$ the marginal cdf and by $S(t_1,...,t_n) = Q(\tau_i > t_1,...,\tau_n > t_n)$ the joint survival function. For notational simplicity, we omit the dependence on the pricing date. We will assume that the marginal survival functions are calibrated from credit curves on the different names and thus that one needs only to specify the dependence function (or copula) for a full characterisation of the joint distribution of default times. In the Gaussian copula model, the default times are obtained as:

$$\tau_i = F_i^{-1}\big(\Phi(V_i)\big), \quad i = 1,\ldots,n$$

where $(V_1,\ldots,V_n)$ is a Gaussian vector, F_i^{-1} denotes the (generalised) inverse of F_i and Φ is the Gaussian cdf.

The conditional independence approach is important in portfolio credit risk modelling (see Finger (1999), Crouhy, Galai and Mark (2000), Merino and Nyfeler (2002), Pykhtin and Dev (2002), Gordy (2003) and Frey and McNeil (2004)). This approach is often coupled with large sample approximation techniques in the case of homogeneous portfolios and leads to simple computations of loss distributions over a given time horizon. In that framework, Gordy and Jones (2003) analyse the risks within CDO tranches.

In order to deal with numerical issues, Gregory and Laurent (2003) and Laurent and Gregory (2003) have described a semi-analytical approach, based on conditional independence, for the pricing and hedging of basket credit derivatives and CDOs. This topic is also discussed by, among others, Andersen, Sidenius and Basu (2003), Galiani (2003), Hull and White (2003), Mina and Stern (2003) and Friend and Rogge (2004).

In these approaches, we deal with a low dimensional factor V such that default times are independent conditionally on V. This standard framework for the modelling of loss distributions can be extended to consistently account for various time horizons (see Laurent and Gregory (2003) for more details). The factor approach makes it simple to deal with a large number of names and leads to very tractable pricing results. We will denote by $p_t^{i|V} = Q(\tau_i \leq t | V)$ and $q_t^{i|V} = Q(\tau_i > t | V)$ the conditional default and survival probabilities. Conditionally on V, the joint survival function is:

$$S(t_1,\cdots,t_n | V) = \prod_{1 \leq i \leq n} q_{t_i}^{i|V}$$

As an example, the one factor Gaussian copula has been introduced by Vasicek (1987). In this setting, $V_i = \rho_i V + \sqrt{1 - \rho_i^2}\, \overline{V}_i$ and $\tau_i = F_i^{-1}(\Phi(V_i))$ for $i = 1,\ldots,n$ where $V, \overline{V}_i$ are independent Gaussian random variables. Then:

$$p_t^{i|V} = \Phi\left(\frac{-\rho_i V + \Phi^{-1}\big(F_i(t)\big)}{\sqrt{1 - \rho_i^2}} \right).$$

Correlation sensitivities

We focus on sensitivities with respect to correlation parameters in the one factor Gaussian model. The correlation terms between names i and j are of the form $\rho_i \rho_j$ for $i, j = 1,\ldots,n$, $i \neq j$. We want to bump a specific correlation parameter say, between names k and l, around the previous one factor correlation structure. Of course, once bumped, the new correlation matrix does not anymore correspond to a one factor correlation structure. However, it is still possible to treat such a local bump

---

[2] From now on, Q denotes a pricing or risk-neutral measure.

in an analytical framework. Let us consider the following Gaussian structure: $V_i = \rho_i V + \sqrt{1-\rho_i^2}\,\overline{V}_i$

for $i \neq k$ and $V_k = \rho_k V + \sqrt{1-\rho_k^2}\left(\lambda\overline{V}_l + \sqrt{1-\lambda^2}\,\overline{V}_k\right)$, where $\lambda \in [0,1]$ and $V, \overline{V}_i$, $i = 1,\ldots,n$ are independent standard Gaussian variables. Thus, the correlation between names i and j remains unchanged when $(i,j) \neq (k,l)$. The correlation between V_k and V_l is equal to $\rho_k \rho_l + \lambda\sqrt{1-\rho_k^2}\sqrt{1-\rho_l^2}$. That correlation matrix is the initial one bumped for the couple of names k,l where the parameter λ controls the magnitude of the correlation shift. This approach can readily be extended to bumping all pairs within a group of names, such as those in a particular sector.

Let us now consider the probability generating function of the accumulated losses at time t, $L(t) = \sum_{i=1}^{n} M_i N_i(t)$ where M_i is the loss given default on name i and $N_i(t) = 1_{\{\tau_i \leq t\}}$ is the corresponding default indicator. We assume here that the losses given default are not stochastic. By conditioning on V, the probability generating function of $L(t)$, $\psi_{L(t)}(u) = E\left[u^{L(t)}\right]$ is given by:

$$\psi_{L(t)}(u) = E\left[E\left[u^{L(t)}|V\right]\right].$$

Using conditional independence upon V, we have:

$$E\left[u^{L(t)}|V\right] = E\left[u^{M_k N_k(t) + M_l N_l(t)}|V\right] \times \prod_{i \neq k,l} E\left[u^{M_i N_i(t)}|V\right].$$

We recall that $E\left[u^{M_i N_i(t)}|V\right] = q_t^{i|V} + p_t^{i|V} u^{M_i}$. Let us now compute $E\left[u^{M_k N_k(t) + M_l N_l(t)}|V\right]$. This can also be written as:

$$E\left[u^{M_k N_k(t) + M_l N_l(t)}|V\right] = Q\left(N_k(t) = 1, N_l(t) = 1|V\right)u^{M_k + M_l} + Q\left(N_k(t) = 1, N_l(t) = 0|V\right)u^{M_k}$$
$$+ Q\left(N_k(t) = 0, N_l(t) = 1|V\right)u^{M_l} + Q\left(N_k(t) = 0, N_l(t) = 0|V\right)$$

This requires the computation of the joint default probabilities conditionally on the factor V. For instance, we can write the joint survival probability $Q\left(N_k(t) = 0, N_l(t) = 0|V\right)$ as:

$$Q\left(N_k(t) = 0, N_l(t) = 0|V\right) = Q\left(\tau_k > t, \tau_l > t|V\right) = Q\left(V_k > \Phi^{-1}(F_k(t)), V_l > \Phi^{-1}(F_l(t))|V\right)$$

Let us denote by $x_k = \Phi^{-1}(F_k(t))$, $x_l = \Phi^{-1}(F_l(t))$. The joint survival probability is equal to:

$$Q\left(V_k > x_k, V_l > x_l|V = v\right) = Q\left(\sqrt{1-\rho_k^2}\left(\lambda\overline{V}_l + \sqrt{1-\lambda^2}\,\overline{V}_k\right) > x_k - \rho_k v, \sqrt{1-\rho_l^2}\,\overline{V}_l > x_l - \rho_l v\right).$$

Using Fubini's theorem and integrating firstly with respect to $\overline{V}_l$, we can eventually write the joint survival probability $Q\left(N_k(t) = 0, N_l(t) = 0|v\right)$ as:

$$\int \overline{\Phi}\left(\max\left(\frac{1}{\lambda}\left(\frac{\Phi^{-1}(F_k(t)) - \rho_k v}{\sqrt{1-\rho_k^2}} - \sqrt{1-\lambda^2}\,u\right), \frac{\Phi^{-1}(F_l(t)) - \rho_l v}{\sqrt{1-\rho_l^2}}\right)\right)\varphi(u)du \, ,$$

where $\varphi(x)$ denotes the Gaussian pdf and $\overline{\Phi}(x) = 1 - \Phi(x) = \Phi(-x)$. Similarly, the joint default probability $Q(N_k(t) = 1, N_l(t) = 1|v)$ is provided by:

$$\int \Phi\left(\min\left(\frac{1}{\lambda}\left(\frac{\Phi^{-1}(F_k(t)) - \rho_k v}{\sqrt{1 - \rho_k^2}} - \sqrt{1 - \lambda^2} u\right), \frac{\Phi^{-1}(F_l(t)) - \rho_l v}{\sqrt{1 - \rho_l^2}}\right)\right)\varphi(u)du$$

The conditional on V probabilities of one name being in default and not the other are given by:

$$Q(N_k(t) = 1, N_l(t) = 0|v) = \int \left(\Phi\left(\frac{\Phi^{-1}(F_k(t)) - \rho_k v}{\lambda\sqrt{1 - \rho_k^2}} - \frac{\sqrt{1 - \lambda^2} u}{\lambda}\right) - \Phi\left(\frac{\Phi^{-1}(F_l(t)) - \rho_l v}{\sqrt{1 - \rho_l^2}}\right)\right)^+ \varphi(u)du$$

and:

$$Q(N_k(t) = 1, N_l(t) = 0|v) = \int \left(\Phi\left(\frac{\Phi^{-1}(F_l(t)) - \rho_l v}{\sqrt{1 - \rho_l^2}}\right) - \Phi\left(\frac{\Phi^{-1}(F_k(t)) - \rho_k v}{\lambda\sqrt{1 - \rho_k^2}} - \frac{\sqrt{1 - \lambda^2} u}{\lambda}\right)\right)^+ \varphi(u)du$$

As a consequence, $E\left[u^{M_k N_k(t) + M_l N_l(t)}|V\right]$ can be easily computed through one dimensional integration. The probability generating function of the accumulated losses is then provided by integration over the distribution of V:

$$\psi_{L(t)}(u) = \int E\left[u^{M_k N_k(t) + M_l N_l(t)}|v\right] \times \prod_{i \neq k, l}\left(q_t^{i|v} + p_t^{i|v} u^{M_i}\right)\varphi(v)dv$$

Eventually, the distribution of $L(t)$ is obtained from the coefficients of the polynomial $\psi_{L(t)}(u)$. We refer to Laurent and Gregory (2003) for further details on the computation of the CDO tranche premiums from the loss distributions over different time horizons.

To illustrate the above approach, we consider three tranches of a euro denominated CDO structure with 50 names and five years maturity. The attachment points for the tranches are $A = 4\%$ and $B = 15\%$. The credit spreads for the names are equal to 25, 30, 35,... up to 270 basis points. Recovery rates are assumed to be deterministic and equal to 40%. We assume that correlation coefficients are constant and equal to 25%. The premiums of the equity, mezzanine and senior tranches are such that the present value of the default leg equals the present value of the corresponding premium leg. We illustrate the sensitivities from the point of view of the protection buyer.

We consider in Figure 1 the PV impact of a bump from 25% to 35% of each single correlation coefficient. Calculating pair-wise correlation sensitivities emphasises the different contributions of the names depending on the level of credit spreads. The equity tranche correlation sensitivities are negative as an increase in a given pair-wise correlation reduces the present value of the tranche since poorer diversification means that the expected loss on the tranche is smaller (the equity tranche has a "negative vega"). That effect is more pronounced for bigger spreads, which are associated to the names that are more likely to contribute to default first. The senior tranche has a positive sensitivity with respect to an increase in a single correlation coefficient due to the fact that protection buyer is effectively long a call option on the aggregated losses and thus has a positive vega. For the mezzanine tranche the effects are blurred since the protection buyer has a call spread, i.e. a long position in a call option with strike A and a short position in call option with strike B. The majority of correlation sensitivities are close to zero, except for the high credit spread names where there are positive and similar in magnitude to the senior tranche sensitivities.

Figure 1. Pairwise tranche correlation sensitivities as a function of the spread of the names being perturbed, equity (top), mezzanine (middle) and senior (bottom).
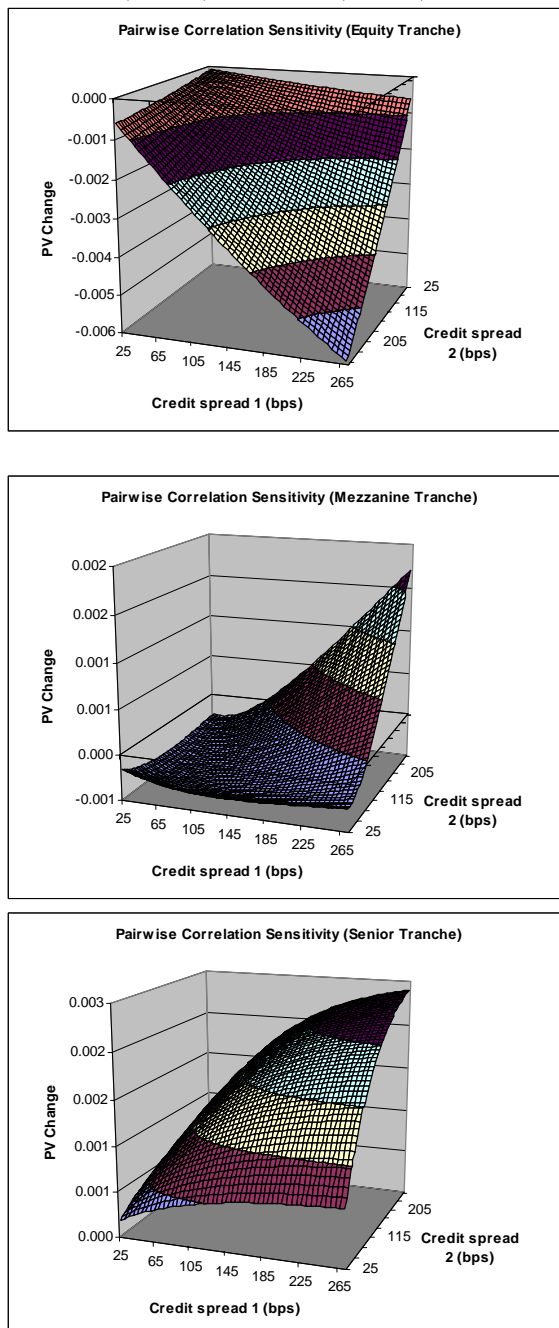
Dealing with more general correlation structures

The one factor correlation structure may seem too restrictive. Of course one may add additional factors albeit at exponentially increasing computational cost. Andersen, Sidenius and Basu (2003) propose a principal components analysis (PCA) to build a low dimensional correlation structure from a given correlation matrix[3]. However, perhaps a more obvious approach in trying to balance

---

[3] Such an approach obviously may lead to some extremely time consuming calculations as many factors are required to provide a suitable fit to the correlation matrix. Furthermore, it does not lend itself readily to calculation and aggregation of various correlation risks as each sequential perturbation of the matrix would require a new PCA.

flexibility and computational burden is to consider a structure built from groups specifying intra and inter-group correlation coefficients. The grouping can be done in any arbitrary way, obvious choices being sector[4] (which we refer to from now on) or geography. We will describe an analytical framework to deal with such a correlation structure, the main advantage being that we can have many different sectors with different intra-sector correlations.

We assume that the underlying Gaussian variables V_i have the following correlation structure:

$$V_i = \rho_{k(i)} W_{k(i)} + \sqrt{1 - \rho_{k(i)}^2}\, \overline{V}_i \,,$$

where the $\overline{V}_i, W_{k(i)}$ are independent standard Gaussian variables and $k(i)$ denotes the sector to which name i is related. Thus, we have an homogeneous one factor structure within a given sector. For simplicity, we have considered identical exposures to factor W for all names within a sector. We then relate the sector risk factors $W_{k(i)}$ through a second single factor structure:

$$W_j = \lambda_j W + \sqrt{1 - \lambda_j^2}\, \overline{W}_j \,,$$

where the $W, \overline{W}_j, \overline{V}_i$ are independent standard Gaussian variables. Our purpose is to aggregate credit risks arising from different sectors while keeping the one factor Gaussian framework for homogeneous portfolios. We can write:

$$V_i = \rho_{k(i)} \lambda_{k(i)} W + \rho_{k(i)} \sqrt{1 - \lambda_{k(i)}^2}\, \overline{W}_{k(i)} + \sqrt{1 - \rho_{k(i)}^2}\, \overline{V}_i \,.$$

For two names in the same sector, the correlation is equal to $\rho_{k(i)}^2$. For two names in different sectors $k(i), k(j)$, $k(i) \neq k(j)$, the correlation is equal to $\rho_{k(i)} \rho_{k(j)} \lambda_{k(i)} \lambda_{k(j)}$. Normally, the parameters λ_j would be taken between 0 and 1 and inter-sector correlation coefficients will be smaller than intra-sector correlation coefficients which seems a reasonable feature. The most obvious correlation structure we can deal with this way is where all inter-sector correlations are equal to γ (say) but with no limitations on the number of sectors:

$$\begin{pmatrix}
1 & \beta_1 & \beta_1 & & & & & & \\
\beta_1 & 1 & \beta_1 & & & & \gamma & & \\
\beta_1 & \beta_1 & 1 & & & & & & \\
& & & 1 & & & & & \\
& & & & . & & & & \\
& & & & & . & & & \\
& & & & & & 1 & & \\
& & & & & & 1 & \beta_m & \beta_m \\
& \gamma & & & & & \beta_m & 1 & \beta_m \\
& & & & & & \beta_m & \beta_m & 1
\end{pmatrix}$$

This is calibrated through $\rho_i = \sqrt{\beta_i}$ and $\lambda_i = \sqrt{\dfrac{\gamma}{\beta_i}}$ for $i = 1, \ldots, m$. Let us now compute the probability generating function of the cumulated losses at time t, $L(t) = \sum_{i=1}^{n} M_i N_i(t)$. We assume here that the losses given default are not stochastic. By conditioning on W, we get:

$$\psi_{L(t)}(u) = E\left[u^{L(t)} \right] = E\left[E\left[u^{L(t)} \middle| W \right] \right]$$

We now separate the losses over the different sectors, $L_j(t) = \sum_{i, k(i)=j} M_i N_i(t)$ and $L(t) = \sum_j L_j(t)$. Given W and $L_j(t)$ are independent,

---

[4] A similar approach is also taken by the rating agencies in CDO tranching models (e.g. S&P). See also Li and Skarabot (2003).

$$\psi_{L(t)}(u) = E\left[\prod_j E\left[u^{L_j(t)} \Big| W \right] \right].$$

Let us now evaluate $E\left[u^{L_j(t)} \Big| W \right]$. We have:

$$E\left[u^{L_j(t)} \Big| W \right] = E\left[E\left[u^{L_j(t)} \Big| W, \overline{W}_j \right] \Big| W \right].$$

For i, $k(i) = j$, $N_i(t)$ are independent conditionally on $W, \overline{W}_j$. Thus,

$$E\left[u^{L_j(t)} \Big| W, \overline{W}_j \right] = \prod_{i,k(i)=j} E\left[u^{M_i N_i(t)} \Big| W, \overline{W}_j \right] = \prod_{i,k(i)=j} \left(q_t^{i|W_j} + p_t^{i|W_j} u^{M_i} \right),$$

with $p_t^{i|W_j} = \Phi\left(\dfrac{-\rho_i W_j + \Phi^{-1}\left(F_i(t) \right)}{\sqrt{1-\rho_i^2}} \right)$ and $q_t^{i|W_j} = 1 - p_t^{i|W_j}$, since W_j is known whenever $W, \overline{W}_j$ are

known. We then get:

$$E\left[u^{L_j(t)} \Big| W \right] = E\left[\prod_{i,k(i)=j} \left(q_t^{i|W_j} + p_t^{i|W_j} u^{M_i} \right) \Big| W \right]$$

This results in a one dimensional integral (with respect to the distribution of $\overline{W}_j$). Eventually,

$$\psi_{L(t)}(u) = E\left[\prod_j E\left[\prod_{i,k(i)=j} \left(q_t^{i|W_j} + p_t^{i|W_j} u^{M_i} \right) \Big| W \right] \right],$$

where the global expectation is computed as a one dimensional integral with respect to the distribution of W. It can be seen that the burden of computing the probability generating function is similar to a two-factor Gaussian structure. The distribution of $L(t)$ is obtained by some inversion technique such as FFT (Gregory and Laurent (2003)).

We now take a practical example based on tranches of the five-year TRAC-X Europe index with names arbitrarily grouped into 5 sectors. We assume also that the intra-sector correlations are all equal and that the inter-sector correlation is $\gamma = 20\%$. We show in Table 1 the effect of varying the intra-sector correlation on the tranche pricing. As expected, an increase in intra-sector correlation means less diversification of credit risk and lower equity tranche premiums, while we observe an increase of the premiums associated with the senior and mezzanine tranches, the latter surprisingly showing monotonic. This illustrates that a tranche that is insensitive to parallel moves in the correlation structure may be rather sensitive to some local correlation risk.

Table 1. Pricing (bp pa) of TRAC-X europe tranches as a function of the intra-sector correlation using the correlation structure described in the text.

	0-3%	3-6%	6-9%	9-12%	12-22%
20%	1273.9	287.5	93.4	33.3	6.0
30%	1226.6	294.4	102.7	39.9	7.9
40%	1168.9	303.5	114.0	47.3	10.3
50%	1100.5	314.2	127.6	56.3	13.3
60%	1020.9	325.8	143.8	67.2	17.0
70%	929.1	337.5	163.6	80.8	21.6
80%	821.9	349.3	188.0	98.8	27.2

We then calculated an implied flat correlation from each tranche premium using the one-factor model (with constant correlations) with the results shown in Table 2. The constant implied correlation in the one factor model increases with the intra-sector correlation. For a 30% level of the intra-sector correlation, the implied correlation structure remains flat, while it exhibits a bump for a 40% level. This is not surprising since the mezzanine tranches have smaller correlation sensitivities. Even more significantly, for higher intra-sector correlations, it is not possible to match some prices, in such a situation a tranche could falsely appear rather correlation neutral.

Table 2. Implied flat correlation from the pricing of the TRAC-X tranches with different values of the intra-sector correlation parameter. An asterisk denotes that the premium could not be matched with the single-parameter model, due primarily to the small correlation sensitivity of the lower mezzanine tranches.

	0-3%	3-6%	6-9%	9-12%	12-22%
20%	20.0%	20.0%	20.0%	20.0%	20.0%
30%	22.2%	22.6%	22.1%	22.2%	22.0%
40%	25.0%	27.6%	25.2%	24.6%	24.2%
50%	28.5%	*	29.7%	27.3%	26.8%
60%	32.8%	*	40.5%	30.6%	29.8%
70%	44.9%	*	*	34.8%	33.1%
80%	44.8%	*	*	41.3%	37.1%

Default times, recovery rates and the correlation smile

Up to now recovery rates were deterministic or at least independent from default dates. We propose here an extension to the analytical Gaussian framework where default dates and recovery rates are dependent. We refer to Chabaane et al (2004) and the references therein for related models in portfolio credit risk analysis. We consider the two-factor Gaussian structure:

$$\begin{cases} V_i = \rho_i V + \sqrt{1-\rho_i^2}\,\overline{V}_i \\ \xi_i = \beta_i \xi + \sqrt{1-\beta_i^2}\,\overline{\xi}_i \end{cases}$$

where V, ξ are the Gaussian factors and $\overline{V}_i, \overline{\xi}_i$ the Gaussian specific risks. We postulate that the specific risks are independent from the factors. We denote by η the correlation between V and ξ and by γ the correlation between $\overline{V}_i$ and $\overline{\xi}_i$. We then define default dates and recovery rates from:

$$\begin{cases} \tau_i = F_i^{-1}\big(\Phi(V_i)\big) \\ M_i = N_i \times \sum_{k=0}^{K} (1-\delta_k)\mathbb{1}_{b_{i,k} \le \xi_i < b_{i,k+1}} \end{cases}$$

where N_i corresponds to the nominal associated with name i. δ_k, $k = 0,\ldots,K$ are possible values for the recovery rates, the $b_{i,k}$ are some thresholds related to name i. The default dates are as in the usual one factor Gaussian copula model. Dependence between recovery rates and default dates comes both from factor correlation when names are different and from factor correlation and specific correlation for a given name.

Let us now consider the distribution of the aggregated loss at time t, $L(t) = \sum_{i=1}^{n} M_i N_i(t)$. We consider the probability generating function of the losses which can be written as:

$$\Psi_{L(t)}(u) = E\left[u^{\sum_{i=1}^{n} M_i N_i(t)} \right] = E\left[E\left[u^{\sum_{i=1}^{n} M_i N_i(t)} \Big| V, \xi \right] \right] = E\left[\prod_{i=1}^{n} E\big[u^{M_i N_i(t)} | V, \xi \big] \right],$$

by using the conditional independence framework. We can notice that:

$$M_i N_i(t) = 1_{V_i < \Phi^{-1}(F_i(t))} \times \sum_{k=1}^{K} N_i (1-\delta_k) 1_{b_{i,k} \leq \xi_i < b_{i,k+1}},$$

The terms $E\left[u^{M_i N_i(t)} \middle| V, \xi\right]$ can be written as:

$$\sum_{k=1}^{K} \left(\Phi_2 \left(\gamma, \frac{\Phi^{-1}(F_i(t)) - \rho_i V}{\sqrt{1-\rho_i^2}}, \frac{b_{i,k+1} - \beta_i \xi}{\sqrt{1-\beta_i^2}} \right) - \Phi_2 \left(\gamma, \frac{\Phi^{-1}(F_i(t)) - \rho_i V}{\sqrt{1-\rho_i^2}}, \frac{b_{i,k} - \beta_i \xi}{\sqrt{1-\beta_i^2}} \right) \right) u^{N_i(1-\delta_k)},$$

where Φ_2 denotes the bivariate Gaussian distribution. To obtain the probability generating function of the accumulated losses at time t, we need to integrate:

$$\prod_{i=1}^{n} \sum_{k=1}^{K} \left(\Phi_2 \left(\gamma, \frac{\Phi^{-1}(F_i(t)) - \rho_i V}{\sqrt{1-\rho_i^2}}, \frac{b_{i,k+1} - \beta_i \xi}{\sqrt{1-\beta_i^2}} \right) - \Phi_2 \left(\gamma, \frac{\Phi^{-1}(F_i(t)) - \rho_i V}{\sqrt{1-\rho_i^2}}, \frac{b_{i,k} - \beta_i \xi}{\sqrt{1-\beta_i^2}} \right) \right) u^{N_i(1-\delta_k)},$$

with respect to the joint distribution of the factors V, ξ. From the probability generating function, we can obtain the distribution of $L(t)$ and eventually the CDO tranche premiums.

As a simple example, consider that $\eta = \xi = 0$ and so we have independence between default events and recovery rates but correlation amongst the recovery rates themselves. The possible recovery rates are $\delta_k = k\%$, $k = 0,\ldots,100\%$ and we assume that their distribution has some symmetry, e.g. the recovery rate on name i is given by $\tilde{\delta}_i = \frac{1}{100} \sum_{k=1}^{100} k 1_{b_k \leq \xi_i < b_{k+1}}$ and $\xi_i = \beta \xi + \sqrt{1-\beta^2} \overline{V}_i$. The probability that the recovery rate is smaller than $k/100$ is $\Phi(b_k)$ [5] and b_k is calibrated from $b_k = \Phi^{-1}(\text{Beta}(k/100))$, where $\text{Beta}(.)$ corresponds to a Beta distribution function with mean 40% and standard deviation 25%. Under these recovery rate assumptions, we then priced TRAC-X tranches with a flat default correlation of 25% and calculated the implied correlation corresponding to these prices. Some results are shown in Figure 2 for recovery rate correlations of $\beta^2 = 50\%$ and 70%. Rather interestingly, we observe a correlation smile which is qualitatively similar to that seen in the market.

Figure 2. Correlation smile implied from the correlated recovery rates example.
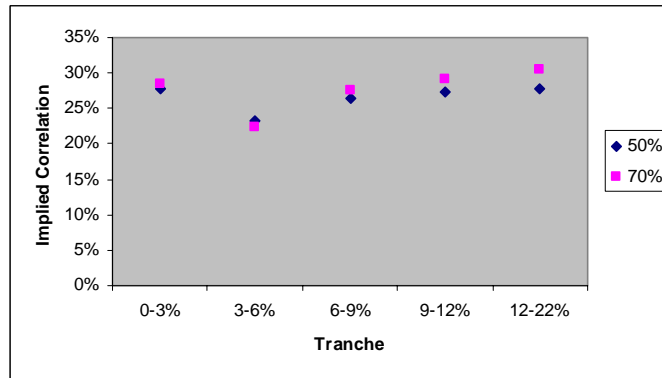
The increase in the implied correlation (decrease of premium) of the equity tranche can be understood as arising from a decrease in its expected loss. The expected loss can be written as $\sum_l Q(N(t) = l) E\left[(3\% - L(t))^+ \middle| N(t) = l\right]$ where $N(t)$ is the number of names in default at time t. Since

---

[5] We set $b_0 = -\infty$ and $b_{101} = +\infty$.

we have equal nominals, and using the independence between default events and recovery rates, $E\left[(3\% - L(t))^+ \mid N(t) = l\right] = E\left[(3\% - M_1 - \ldots - M_l)^+\right]$. Thus increasing the correlation between the M_i will lead to an increase in the volatility of $M_1 + \ldots + M_l$ and to a decrease in the expected loss of the equity tranche. Since the reverse is true for the most senior tranche, the appearance of a correlation smile is not surprising.

The previous example illustrates a potential consequence of the correlation between only recovery rates but there are obviously many ways in which to parameterise the above model. As a simple example of the impact of correlation between default events and recovery rates consider the extreme case of perfect correlation between the latent variables driving default and recoveries, that is $V_i = \xi_i$ for $i = 1, \ldots, n$. Here, $V_i = \xi_i$ may be seen as the asset value in a Merton type firm-value model since that single variable will drive both default and recoveries.

We assume that the recovery rates can only take values equal to zero and one (which corresponds to a very large volatility) and we let the recovery rate distribution be time-varying. This corresponds to the following model:

$$
\begin{cases}
\tau_i = F_i^{-1}\left(\Phi(V_i)\right) \\
M_i(t) = N_i \times \left(1 - 1_{s_i(t) \leq V_i \leq t_i(t)}\right)
\end{cases}
$$

where $s_i(t)$ and $t_i(t)$ are some thresholds to be determined. The time t loss associated with name i is given by:

$$
M_i(t)N_i(t) = N_i \times \left(1 - 1_{s_i(t) \leq V_i \leq t_i(t)}\right) \times 1_{V_i \leq \Phi^{-1}(F_i(t))}
$$

We assume that the expected losses $E\left[M_i(t)N_i(t)\right]$ are given and calibrated on individual credit curves. We take $t_i(t) = \Phi^{-1}\left(F_i(t)\right)$. This provides $M_i(t)N_i(t) = N_i \times 1_{V_i \leq s_i(t)}$ where $s_i(t)$ is provided by the calibrating equation $E\left[M_i(t)N_i(t)\right] = N_i \Phi\left(s_i(t)\right)$. We can notice that this model is equivalent to:

$$
\begin{cases}
\tau_i = s_i^{-1}(V_i) \\
M_i = N_i
\end{cases}
$$

that is a simple one factor Gaussian model with deterministic (equal to zero) recovery rates. A perfect correlation between recovery rates and default events thus leads to a one factor Gaussian model with zero recovery rates. One can then easily assess the magnitude of the correlation effect between recovery rates and default events.

Conclusion
This article was dedicated to treatment of correlation for credit derivatives within a Gaussian copula framework. We show that we can address analytically the computation of correlation risk and propose a more flexible two-factor model to incorporate a realistic correlation structure. We also showed how to incorporate dependence between recovery rates and default times in a tractable way and illustrated that such an approach may match better the observed "correlation smile" across portfolio tranches. Despite the popularity of the Gaussian copula model, there are clear and valid questions over its theoretical foundations, for example, the implied dynamics of credit spreads seems unrealistic. However, we believe it will probably remain an industry standard for some time which is why we have tried to further the depth and analytical tractability around the model. Clearly, there is room for further analyses and comparisons between different dependence structures and investigations of tractable ways for incorporating reasonable credit spread dynamics.

References

Andersen, L., J. Sidenius and S. Basu, 2003, *All your Hedges in one Basket*, 2003, RISK, November, 67-72.

Arvanitis, A., and J. Gregory, 2001, *Credit: The Complete Guide to Pricing, Hedging and Risk Management*, Risk Publications.

Chabaane, A., J-P. Laurent and J. Salomon, 2004, *Double Impact: Credit Risk Assessment and Collateral Value*, working paper, ISFA Actuarial School, University of Lyon & BNP-Paribas, available on *www.defaultrisk.com*.

Crouhy, M., D. Galai and R. Mark, 2000, *A Comparative Analysis of Current Credit Risk Models*, Journal of Banking and Finance, 24, 59-117.

Duffie, D., and N. Garleanu, 2001, *Risk and the Valuation of Collateralized Debt Obligations*, Financial Analysts Journal, 57, 41-59.

Finger, C. C., 1999, *Conditional Approaches for CreditMetrics Portfolio Distributions*, CreditMetrics Monitor, April, 14-33.

Finger, C. C., 2000, *A Comparison of Stochastic Default Rate Models*, RiskMetrics Journal, Volume 1, November, 49-73.

Frey, R., and A. McNeil, 2004, *Dependent Defaults in Models of Portfolio Credit Risk*, To appear in Journal of Risk.

Friend, A., and E. Rogge, 2004, *Correlation at First Sight*, working paper, ABN AMRO.

Galiani, S., 2003, *Copula Functions and their Application in Pricing and Risk Managing Multiname Credit Derivatives Products*, working paper, King's College.

Gordy, M., 2003, *A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules*, Journal of Financial Intermediation, 12(3), 199-232.

Gordy, M., and D. Jones, 2003, *Random Tranches*, RISK, March, 78-83.

Gregory, J., and J-P. Laurent, 2003, *I Will Survive*, RISK, June, 103-107.

Gupton, G., C. Finger and M. Bahia, 1997, *CreditMetrics: The Benchmark for Understanding Credit Risk*, Technical Document, JP Morgan.

Hull, J., and A White, 2001, *Valuing Credit Default Swaps II: Modeling Default Correlations*, Spring, Volume 8(3), 12-22.

Hull, J., and A. White, 2003, *Valuation of a CDO and an nth–to-default CDS without Monte Carlo Simulation*, working paper, university of Toronto.

Jarrow, R., and F. Yu, 2000, *Counterparty Risk and the Pricing of Defaultable Securities*, Journal of Finance, 56, 1765-1799.

Laurent, J-P., and J. Gregory, 2003, *Basket Default Swaps, CDO's and Factor Copulas*, working paper ISFA Actuarial School, Univerity of Lyon & BNP-Paribas available on *www.defaultrisk.com*.

Li, D., 2000, *On Default Correlation: A Copula Function Approach*, Journal of Fixed Income, Volume 9(4), 43-54.

Li, D., and J. Skarabot, 2003, *Valuation and Risk Analysis of Synthetic Collateralised Debt Obligations: a Copula Function Approach*, in *Credit Derivatives: The Definitive Guide*, RISK Books, J. Gregory (ed.).

Mashal, R., and M. Naldi, 2001, *Pricing Multiname Credit Derivatives: Heavy Tailed Hybrid Approach*, working paper, Columbia University and Lehman Brothers.

Merino, S., and M. Nyfeler, 2002, *Calculating Portfolio Loss*, RISK, August, 82-86.

Mina, J., and E. Stern, 2003, Examples and Applications of Closed-Form CDO Pricing, Riskmetrics Journal, 4(1), 5-34.

Pykhtin, M., and A. Dev, 2002, *Analytical Approach to Credit Risk Modelling*, RISK, March, S26-S32.

Rogge, E., and P. Schönbucher, 2003, *Modelling Dynamic Portfolio Credit Risk*, working paper.

Schlögl, L., and D. O'Kane, 2003, *Assessing Correlation Skew Versus Tail Dependence in CDO Tranche Valuation*, Lehman Brothers, Presentation at ICBI conference, Geneva, December.

Schönbucher, P., and D. Schubert, 2001, *Copula Dependent Default Risk in Intensity Models*, working paper, Department of Statistics, Bonn University.

Vasicek, O., 1987, *Probability of Loss on Loan Portfolio*, KMV Corporation, *www.kmv.com*