The Skewed t Distribution for Portfolio Credit Risk*

Wenbo Hu[†]

Bell Trading

Alec N. Kercheval[‡]

Florida State University

Abstract

Portfolio credit derivatives, such as basket credit default swaps (basket CDS), require for their pricing an estimation of the dependence structure of defaults, which is known to exhibit tail dependence as reflected in observed default contagion. A popular model with this property is the (Student's) t copula; unfortunately there is no fast method to calibrate the degree of freedom parameter.

In this paper, within the framework of Schönbucher's copula-based trigger-variable model for basket CDS pricing, we propose instead to calibrate the full multivariate t distribution. We describe a version of the EM algorithm that provides very fast calibration speeds compared to the current copula-based alternatives.

The algorithm generalizes easily to the more flexible skewed t distributions. To our knowledge, we are the first to use the skewed t distribution in this context.

*Journal of Economic Literature Classification Codes: C16, G12. Key words: Portfolio Credit Risk, Basket Credit Default Swaps, Skewed t Distribution, t Distribution, t Copula.

[†]Bell Trading. `whu@belltrading.us`

[‡]Corresponding author, `kercheval@math.fsu.edu`, Department of Mathematics, Florida State University, Tallahassee, FL 32306-4510

1 Introduction

For portfolio risk modeling and basket derivative pricing, it is essential to understand the *dependence structure* of prices, default times, or other asset-related variables. This structure is completely described by the second moments (the covariance matrix) for jointly Normal variables, so practitioners often use the covariance matrix as a simple proxy for multivariate dependence.

However, it is widely acknowledged that prices, returns, and other financial variables are not Normally distributed. They have fat tails, and exhibit "tail dependence" (see Section 4), in which correlations are observed to rise during extreme events.

Therefore there is a need for practical uses of more general multivariate distributions to model joint price behavior. This raises the question of how to choose these distributions, and, once chosen, how to efficiently calibrate them to data. In this paper we look at the multivariate (Student's) t distribution, which has become a popular choice because of its heavy tails and non-zero tail dependence, and it's generalization, the skewed t distribution, described, for example, by Demarta and McNeil (2005) – see Section 2 below.

It has become popular and useful to isolate the dependence structure of a distribution from the individual marginal distributions by looking at its *copula* (see Section 3). Copulas that come from known distributions inherit their names – e.g. we have the Gaussian copulas, the t copulas, etc.

There are now many financial applications of copulas. For example, Di Clemente and Romano (2003b) used copulas to minimize expected shortfall (ES) in modeling operational risk. Di Clemente and Romano (2004) applied the same framework in the portfolio optimization of credit default swaps. Masala et al. (working paper) used the t copula and a transition matrix with a gamma-distributed hazard rate and a beta-distributed recovery rate to compute the efficient frontier for credit portfolios by minimizing ES.

The success of copulas greatly depends both on good algorithms for calibrating the copula itself, and on the availability of a fast algorithm to calculate the cumulative distribution functions (CDF) and quantiles of the corresponding one dimensional marginal distributions.

The calibration of a t copula is very fast if we fix the degree of freedom parameter ν, which in turn is optimized by maximizing a log likelihood. However, the latter is slow. Detailed algorithms for calibrating t copulas can be found in the work of many researchers, such as Di Clemente and Romano (2003a), Demarta and McNeil (2005), Mashal and Naldi (2002), and Galiani (2003).

The calibration of a t copula is (by definition) separate from the calibration of marginal distributions. It is generally suggested to use the empirical distributions to fit the margins, but empirical distributions tend to have poor performance in the tails. A hybrid of the parametric and non parametric method considers the use of the empirical distribution in the center and a generalized Pareto distribution (GPD) in the tails. Some use a Gaussian distribution in the center. To model multivariate losses, Di Clemente and Romano (2003a) used a t copula and Gaussian distribution in the center and left tail and a GPD in the right tail for the margins. We will be able to avoid these issues because we can effectively calibrate the full distribution directly by using t or skewed t distributions.

In this paper, the primary application we have in mind is portfolio credit risk – specifically, the pricing of multiname credit derivatives such as kth-to-default basket credit default swaps (basket CDS).

For this problem, the most important issue is the correlation structure among the default obligors as described by the copula of their default times. Unfortunately, defaults are rarely observed, so it is difficult to calibrate their correlations directly. In this paper we follow Cherubini et al. (2004) and use the distribution of daily equity prices to proxy the dependence structure of default times. See Section 6.2 below.

Several groups have discussed the pricing of basket CDS and CDO via copulas, such as Galiani (2003), Mashal and Naldi (2002), and Meneguzzo and Vecchiato (2002), among others. However, in this paper, we find that calibrating the full joint distribution is much faster than calibrating the copula separately, because of the availabilty of the EM algorithm discussed below.

In Hu (2005), we looked at the large family of generalized hyperbolic distributions to model multivariate equity returns by using the EM algorithm (see Section 2). We showed that the skewed t has better performance and faster convergence than other generalized hyperbolic distributions. Furthermore, for the t distribution, we have greatly simplified formulas and an even faster algorithm. For the t copula, there is still no good method to calibrate the degree of freedom ν except to find it by direct search. The calibration of a t copula takes *days* while the calibration of a skewed t or t distribution via the EM algorithm takes *minutes*. To our knowledge, we are the first to directly calibrate the skewed t or t distributions to price basket credit default swaps.

This paper is organized as follows. In section 2, we introduce the skewed t distribution from the normal mean variance mixture family and provide a version of the EM algorithm to calibrate it, including the limiting t distribution. We give an introduction to copulas in section 3, and review rank correlation and tail dependence in section 4.

In section 5, we follow Rutkowski (1999) to review the reduced form approach to single name credit risk. In section 6, we follow Schönbucher (2003) to provide our model setup for calculating default probabilities for the k-th to default using a copula-based trigger variable method. There we also discuss the calibration problem.

In section 7 we apply all the previous ideas to describe a method for pricing basket credit default swaps. We illustrate how selecting model copulas with different tail dependence coefficients influences the relative probabilities of first and last to default. We then argue that calibrating the skewed t distribution is the best and fastest approach, among the common alternatives.

2 Skewed t distributions and the EM algorithm

2.1 Skewed t and t distributions

Definition 2.1 Inverse Gamma Distribution. *The random variable X has an* inverse gamma distribution*, written $X \sim InverseGamma(\alpha, \beta)$, if its probability density function is*

$$f(x) = \beta^{\alpha} x^{-\alpha-1} e^{-\beta/x} / \Gamma(\alpha), \quad x > 0, \quad \alpha > 0, \quad \beta > 0, \tag{2.1}$$

where Γ is the usual gamma function.

We have the following standard formulas:

$$E(X) = \frac{\beta}{\alpha - 1}, \quad \text{if} \ \ \alpha > 1 \tag{2.2}$$

$$Var(X) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \quad \text{if} \ \ \alpha > 2 \tag{2.3}$$

$$E(\log(X)) = \log(\beta) - \psi(\alpha), \tag{2.4}$$

where

$$\psi(x) = d \log(\Gamma(x))/dx \tag{2.5}$$

is the digamma function.

The skewed t distribution is a subfamily of the generalized hyperbolic distributions – see McNeil et al. (2005), who suggested the name "skewed t". It can be represented as a normal mean-variance mixture, where the mixture variable is inverse gamma distributed.

Definition 2.2 Normal Mean-Variance Mixture Representation of Skewed t Distribution. *Let $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ be parameter vectors in $\mathbb{R}^d$, Σ a $d \times d$ real positive semidefinite matrix, and $\nu > 2$. The d* dimensional skewed t distributed random vector $\mathbf{X}$*, which is denoted by*

$$\mathbf{X} \sim SkewedT_d(\nu, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma}),$$

is a multivariate normal mean-variance mixture variable with distribution given by

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + W\boldsymbol{\gamma} + \sqrt{W}\mathbf{Z}, \tag{2.6}$$

where

1. $\mathbf{Z} \sim N_d(\mathbf{0}, \Sigma)$*, the multivariate normal with mean* $\mathbf{0}$ *and covariance* Σ*,*
2. $W \sim InverseGamma(\nu/2, \nu/2)$*, and*
3. W *is independent of* $\mathbf{Z}$*.*

Here, $\boldsymbol{\mu}$ are location parameters, $\boldsymbol{\gamma}$ are skewness parameters and ν is the degree of freedom.

From the definition, we can see that

$$\mathbf{X} \mid W \sim N_d(\boldsymbol{\mu} + W\boldsymbol{\gamma}, W\Sigma). \tag{2.7}$$

This is also why it is called a normal mean-variance mixture distribution. We can get the following moment formulas easily from the mixture definition:

$$E(\mathbf{X}) = \boldsymbol{\mu} + E(W)\boldsymbol{\gamma}, \tag{2.8}$$

$$(2.9) \qquad COV(\mathbf{X}) = E(W)\Sigma + var(W)\boldsymbol{\gamma}\boldsymbol{\gamma}',$$

when the mixture variable W has finite variance $var(W)$.

Definition 2.3 *Setting $\boldsymbol{\gamma}$ equal to zero in Definition 2.2 defines the* multivariate t distribution*,*

$$(2.10) \qquad \mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \sqrt{W}\mathbf{Z}.$$

For convenience we next give the density functions of these distributions. Denote by $K_\lambda(x)$, $x > 0$, the modified Bessel function of the third kind, with index λ:

$$K_\lambda(x) = \frac{1}{2}\int_0^\infty y^{\lambda-1}e^{-\frac{x}{2}(y+y^{-1})}\,dy.$$

The following formula may be computed using (2.7), and is given in McNeil et al. (2005).

Proposition 2.4 Skewed t Distribution. *Let $\mathbf{X}$ be skewed t distributed, and define*

$$(2.11) \qquad \rho(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu}).$$

Then the joint density function of $\boldsymbol{X}$ is given by

$$(2.12) \qquad f(\boldsymbol{x}) = c\frac{K_{\frac{\nu+d}{2}}\left(\sqrt{(\nu+\rho(\boldsymbol{x}))\,(\boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})}\right)e^{(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}\boldsymbol{\gamma}}}{\left(\sqrt{(\nu+\rho(\boldsymbol{x}))\,(\boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})}\right)^{-\frac{\nu+d}{2}}(1+\frac{\rho(\boldsymbol{x})}{\nu})^{\frac{\nu+d}{2}}},$$

where the normalizing constant is

$$c = \frac{2^{1-\frac{\nu+d}{2}}}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}.$$

The mean and covariance of a skewed t distributed random vector $\mathbf{X}$ are

$$E(\mathbf{X}) = \boldsymbol{\mu} + \boldsymbol{\gamma}\frac{\nu}{\nu-2}, \tag{2.13}$$

$$COV(\mathbf{X}) = \frac{\nu}{\nu-2}\Sigma + \boldsymbol{\gamma}\boldsymbol{\gamma}'\frac{2\nu^2}{(\nu-2)^2(\nu-4)}, \tag{2.14}$$

where the covariance matrix is only defined when $\nu > 4$, and the expectation only when $\nu > 2$.

Furthermore, in the limit as $\boldsymbol{\gamma} \to \mathbf{0}$, we get the joint density function of the t distribution:

$$f(\boldsymbol{x}) = \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}(1 + \frac{\boldsymbol{\rho(x)}}{\nu})^{-\frac{\nu+d}{2}}. \tag{2.15}$$

The mean and covariance of a t distributed random vector $\mathbf{X}$ are

$$E(\mathbf{X}) = \boldsymbol{\mu}, \tag{2.16}$$

$$COV(\mathbf{X}) = \frac{\nu}{\nu-2}\Sigma. \tag{2.17}$$

2.2 Calibration of t and Skewed t Distributions Using the EM Algorithm

The mean-variance representation of the skewed t distribution has a great advantage: the so-called EM algorithm can be applied to such a representation. See McNeil et al. (2005) for a general discussion of this algorithm for calibrating generalized hyperbolic distributions.

The EM (Expectation-Maximization) algorithm is a two-step iterative process in which (the E-step) an expected log likelihood function is calculated using current parameter values, and then (the M-step) this function is maximized to produce updated parameter values. After each E and M step, the log likelihood is increased, and the method converges to a maximum log likelihood estimate of the distribution parameters.

What helps this along is that the skewed t distribution can be represented as a conditional normal distribution, so most of the parameters $(\Sigma, \boldsymbol{\mu}, \boldsymbol{\gamma})$ can be calibrated, conditional on W, like a Gaussian distribution. We give a brief summary of our version of the EM algorithms for skewed t and t distributions here. Detailed derivations, along with comparisons to other versions, can be found in Hu (2005).

To explain the idea, suppose we have i.i.d. data $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$ that we want to fit to a skewed t distribution.

We seek parameters $\boldsymbol{\theta} = (\nu, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$ to maximize the log likelihood

$$\log L(\boldsymbol{\theta}; \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \sum_{i=1}^{n} \log f(\boldsymbol{x}_i; \boldsymbol{\theta}),$$

where $f(\cdot; \boldsymbol{\theta})$ denotes the skewed t density function.

The method is motivated by the observation that if the latent variables $w_1, \ldots, w_n$ were observable, our optimization would be straightforward. We define the augmented log-likelihood function

$$\log \tilde{L}(\boldsymbol{\theta}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, w_1, \ldots, w_n) = \sum_{i=1}^{n} \log f_{\boldsymbol{X},W}(\boldsymbol{x}_i, w_i; \boldsymbol{\theta}),$$

$$= \sum_{i=1}^{n} \log f_{\boldsymbol{X}|W}(\boldsymbol{x}_i|w_i; \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma}) + \sum_{i=1}^{n} \log h_W(w_i; \nu)$$

where $f_{\boldsymbol{X}|W}(\cdot|w; \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$ is the conditional normal $N(\boldsymbol{\mu} + w\boldsymbol{\gamma}, w\Sigma)$ and $h_W(\cdot; \nu)$ is the the density of $InverseGamma(\nu/2, \nu/2)$.

These two terms could be maximized separately if the latent variables were observable. Since they are not, the method is instead to maximize the *expected value* of the augmented log-likelihood $\tilde{L}$ conditional on the data *and* on a guess for the parameters $\boldsymbol{\theta}$. We must condition on the parameters because the distribution of the latent variables depends on the parameters. This produces an *updated guess* for the parameters, which we then use to repeat the process until convergence.

To be more explicit, suppose we have a step k parameter estimate $\boldsymbol{\theta}^{[k]}$. We carry out the following steps.

- E-step: compute an objective function

$$Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{[k]}) = E(\log \tilde{L}(\boldsymbol{\theta}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, W_1, \ldots, W_n)|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n; \boldsymbol{\theta}^{[k]})$$

  This can be done analytically and requires formulas for quantities like $E(W_i|\boldsymbol{x}_i, \boldsymbol{\theta}^{[k]})$, $E(1/W_i|\boldsymbol{x}_i, \boldsymbol{\theta}^{[k]})$, and $E(\log W_i|\boldsymbol{x}_i, \boldsymbol{\theta}^{[k]})$, which can all be explicitly derived from the definitions.

- M-step: Maximize Q to find $\boldsymbol{\theta}^{[k+1]}$.

Using our explicit formulas for the skewed t distribution, we can compute the expectation and the subsequent maximum $\boldsymbol{\theta}$ explicitly. Below we summarize the resulting formulas needed for directly implementing this algorithm.

We will use a superscript in square brackets to denote the iteration counter. Given, at the kth step, parameter estimates $\nu^{[k]}$, $\Sigma^{[k]}$, $\boldsymbol{\mu}^{[k]}$, and $\boldsymbol{\gamma}^{[k]}$, let, for $i = 1, \ldots, n$,

$$\rho_i^{[k]} = (\boldsymbol{x}_i - \boldsymbol{\mu}^{[k]})'(\Sigma^{[k]})^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}^{[k]}).$$

Define the auxiliary variables $\theta_i^{[k]}$, $\eta_i^{[k]}$, and $\xi_i^{[k]}$ by

$$\theta_i^{[k]} = \left(\frac{\rho_i^{[k]} + \nu^{[k]}}{\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]}}\right)^{-\frac{1}{2}} \frac{K_{\frac{\nu+d+2}{2}}\left(\sqrt{(\rho_i^{[k]} + \nu^{[k]})(\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]})}\right)}{K_{\frac{\nu+d}{2}}\left(\sqrt{(\rho_i^{[k]} + \nu^{[k]})(\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]})}\right)} \tag{2.18}$$

$$\eta_i^{[k]} = \left(\frac{\rho_i^{[k]} + \nu^{[k]}}{\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]}}\right)^{\frac{1}{2}} \frac{K_{\frac{\nu+d-2}{2}}\left(\sqrt{(\rho_i^{[k]} + \nu^{[k]})(\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]})}\right)}{K_{\frac{\nu+d}{2}}\left(\sqrt{(\rho_i^{[k]} + \nu^{[k]})(\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]})}\right)} \tag{2.19}$$

$$\xi_i^{[k]} = \frac{1}{2}\log\left(\frac{\rho_i^{[k]} + \nu^{[k]}}{\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]}}\right) + \frac{\frac{\partial K_{-\frac{\nu+d}{2}+\alpha}\left(\sqrt{(\rho_i^{[k]}+\nu^{[k]})(\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]})}\right)}{\partial\alpha}\Big|_{\alpha=0}}{K_{\frac{\nu+d}{2}}\left(\sqrt{(\rho_i^{[k]} + \nu^{[k]})(\boldsymbol{\gamma}^{[k]\prime}\Sigma^{[k]-1}\boldsymbol{\gamma}^{[k]})}\right)}. \tag{2.20}$$

In the special case of the multivariate t distributions, we have simpler forms for above formulas:

$$\theta_i^{[k]} = \frac{\nu^{[k]} + d}{\rho_i^{[k]} + \nu^{[k]}} \tag{2.21}$$

$$\eta_i^{[k]} = \frac{\rho_i^{[k]} + \nu^{[k]}}{\nu^{[k]} + d - 2} \tag{2.22}$$

$$\xi_i^{[k]} = \log(\frac{\rho_i^{[k]} + \nu^{[k]}}{2}) - \psi(\frac{d + \nu^{[k]}}{2}). \tag{2.23}$$

Let us denote

$$\bar{\theta} = \frac{1}{n}\sum_1^n \theta_i,\ \bar{\eta} = \frac{1}{n}\sum_1^n \eta_i,\ \bar{\xi} = \frac{1}{n}\sum_1^n \xi_i. \tag{2.24}$$

Algorithm 2.5 *EM algorithm for calibrating the t and skewed t distributions*

1. *Set the iteration counter k=1. Select staring values for $\nu^{[1]}$, $\boldsymbol{\gamma}^{[1]}$, $\boldsymbol{\mu}^{[1]}$ and $\Sigma^{[1]}$. Reasonable starting value for mean and dispersion matrix are the sample mean and sample covariance matrix.*

2. *Calculate $\theta_i^{[k]}$, $\eta_i^{[k]}$, and $\xi_i^{[k]}$ and their averages $\bar{\theta}$, $\bar{\eta}$ and $\bar{\xi}$.*

3. *Update $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$ and Σ according to*

$$\boldsymbol{\gamma}^{[k+1]} = \frac{n^{-1}\sum_{i=1}^n \theta_i^{[k]}(\bar{\boldsymbol{x}} - \boldsymbol{x}_i)}{\bar{\theta}^{[k]}\bar{\eta}^{[k]} - 1} \tag{2.25}$$

$$\boldsymbol{\mu}^{[k+1]} = \frac{n^{-1}\sum_{i=1}^n \theta_i^{[k]}\boldsymbol{x}_i - \boldsymbol{\gamma}^{[k+1]}}{\bar{\theta}^{[k]}} \tag{2.26}$$

$$\Sigma^{[k+1]} = \frac{1}{n}\sum_{i=1}^n \theta_i^{[k]}(\boldsymbol{x}_i - \boldsymbol{\mu}^{[k+1]})(\boldsymbol{x}_i - \boldsymbol{\mu}^{[k+1]})' - \bar{\eta}^{[k]}\boldsymbol{\gamma}^{[k+1]}\boldsymbol{\gamma}^{[k+1]\prime}. \tag{2.27}$$

4. *Compute $\nu^{[k+1]}$ by numerically solving the equation*

$$-\psi(\frac{\nu}{2}) + \log(\nu/2) + 1 - \bar{\xi}^{[k]} - \bar{\theta}^{[k]} = 0. \tag{2.28}$$

5. *Set counter $k := k+1$ and go back to step 2 unless the relative increment of log likelihood is small and in this case, we terminate the iteration.*

The result of this algorithm is an estimate of the maximum likelihood parameter values for the given data.

3 Copulas

Copulas are used to describe the dependence structure of a multivariate distribution. A good general reference is Nelsen (1999). One of the definitions can be found in Li (1999), the first one to use copulas to price portfolio credit risk.

Definition 3.1 Copula Functions. *U is a* uniform random variable *if it has a uniform distribution on the interval* $[0, 1]$.

For d uniform random variables $U_1, U_2, \cdots, U_d$, the joint distribution function C,defined as

$$C(u_1, u_2, \cdots, u_d) = P[U_1 \leq u_1, U_2 \leq u_2, \cdots, U_d \leq u_d],$$

is called a copula *function.*

Proposition 3.2 ***Sklar's Theorem.*** *Let F be a joint distribution function with margins $F_1, F_2, \cdots, F_d$, then there exists a copula C such that for all $(x_1, x_2, \cdots, x_d) \in \mathbb{R}^d$,*

$$F(x_1, x_2, \cdots, x_d) = C(F_1(x_1), F_2(x_2), \cdots, F_d(x_d)). \tag{3.1}$$

If $F_1, F_2, \cdots, F_d$ are continuous, then C is unique. Conversely, if C is a copula and $F_1, F_2, \cdots, F_d$ are distribution functions, then the function F defined by equation 3.1 is a joint distribution function with margins $F_1, F_2, \cdots, F_d$.

Corollary 3.3 *If $F_1, F_2, \cdots, F_m$ are continuous, then, for any $(u_1, \cdots, u_m) \in [0,1]^m$, we have*

$$C(u_1, u_2, \cdots, u_m) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \cdots, F_m^{-1}(u_m)), \tag{3.2}$$

where $F_i^{-1}(u_i)$ denotes the inverse of the cumulative distribution function, namely, for $u_i \in [0,1]$, $F_i^{-1}(u_i) = \inf\{x : F_i(x) \geq u_i\}$.

The name copula means a function that couples a joint distribution function to its marginal distributions. If $X_1, X_2, \cdots, X_d$ are random variables with distributions $F_1, F_2, \cdots, F_d$, respectively, and a joint distribution F, then the corresponding copula C is also called the copula of $X_1, X_2, \cdots, X_d$, and $(U_1, U_2, \cdots, U_d) = (F_1(X_1), F_2(X_2), \cdots, F_d(X_d))$ also has copula C. We will use this property to price basket credit default swaps later.

We often assume the marginal distributions to be *empirical distributions*. Suppose that the sample data is $\mathbf{x_i} = (x_{i,1}, \cdots, x_{i,d})$, where $i = 1, \cdots, n$, then we may take the empirical estimator of jth marginal distribution function to be

$$\hat{F}_j(x) = \frac{\sum_{i=1}^{n} I_{\{x_{i,j} \leq x\}}}{n+1}. \tag{3.3}$$

(Demarta and McNeil (2005) suggested dividing by $n+1$ to keep the estimation away from the boundary 1.) By using different copulas and empirical or other margins, we can create a rich family of multivariate distributions. It is not required that the margins and joint distribution be the same type of distribution.

Two types of copulas are widely used: Archimedean copulas and elliptical copulas. Archimedean copulas form a rich family of examples of bivariate copulas, including the well-known Frank, Gumbel and Clayton copulas. These have only one parameter and are easy to calibrate. However, the usefulness of Archimedean copulas of more than two variables is quite limited: they have only one or two parameters, and enforce a lot of symmetry in the dependence structure, such as bivariate exchangeability, that is unrealistic for a portfolio of heterogeneous firms.

Therefore we now restrict attention to the elliptical copulas, which are created from multivariate elliptical distributions, such as the Gaussian and t distributions, and their immediate generalizations, such as the skewed t copula.

Definition 3.4 Multivariate Gaussian Copula. *Let R be a positive semidefinite matrix with $diag(R) = \mathbf{1}$ and let Φ_R be the standardized multivariate normal distribution function with correlation matrix R. Then the* multivariate Gaussian copula *is defined as*

$$C(u_1, u_2, \cdots, u_m; R) = \Phi_R(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \cdots, \Phi^{-1}(u_m)), \tag{3.4}$$

where $\Phi^{-1}(u)$ denotes the inverse of the standard univariate normal cumulative distribution function.

Definition 3.5 Multivariate t Copula. *Let R be a positive semidefinite matrix with $diag(R) = \mathbf{1}$ and let $T_{R,\nu}$ be the standardized multivariate t*

distribution function with correlation matrix R and ν degrees of freedom. Then the multivariate t copula *is defined as*

$$C(u_1, u_2, \cdots, u_m; R, \nu) = T_{R,\nu}(T_\nu^{-1}(u_1), T_\nu^{-1}(u_2), \cdots, T_\nu^{-1}(u_m)), \tag{3.5}$$

where $T_\nu^{-1}(u)$ denotes the inverse of standard univariate t cumulative distribution function.

4 Measures of Dependence

All dependence information is contained in the copula of a distribution. However, it is helpful to have real-valued measures of the dependence of two variables. The most familiar example of this is Pearson's linear correlation coefficient; however, this does not have the nice properties we will see below.

4.1 Rank Correlation

Definition 4.1 Kendall's Tau. *Kendall's tau rank correlation for the bivariate random vector (X, Y) is defined as*

$$\tau(X, Y) = P((X - \hat{X})(Y - \hat{Y}) > 0) - P((X - \hat{X})(Y - \hat{Y}) < 0), \tag{4.1}$$

where $(\hat{X}, \hat{Y})$ is an independent copy of (X, Y).

As suggested by Meneguzzo and Vecciato (2002), the sample consistent estimator of Kendall's tau is given by

$$\hat{\tau} = \frac{\sum_{i,j=1,i<j}^{n} sign[(x_i - x_j)(y_i - y_j)]}{n(n-1)/2}, \tag{4.2}$$

where $sign(x) = 1$ if $x \geq 0$, otherwise $sign(x) = 0$, and n is the number of observations.

In the case of elliptical distributions, Lindskog et al. (2003) showed that

$$\tau(X, Y) = \frac{2}{\pi} \arcsin(\rho), \tag{4.3}$$

where ρ is Pearson's linear correlation coefficient between random variable X and Y. However, Kendall's tau is more useful in discussions of dependence structure because it depends in general only on the copula of (X, Y) (Nelsen, 1999):

$$\tau(X, Y) = 4 \int\int_{[0,1]^2} C(u, v) dC(u, v) - 1. \tag{4.4}$$

It has nothing to do with the marginal distributions. Sometimes, we may need the following formula

$$\tau(X, Y) = 1 - 4 \int\int_{[0,1]^2} C_u(u, v) C_v(u, v) du dv, \tag{4.5}$$

where C_u denotes the partial derivative of $C(u, v)$ with respect to u and C_v denotes the partial derivative of $C(u, v)$ with respect to v.

Proposition 4.2 Copula of Transformations (Nelsen, 1999)**.** *Let X and Y be continuous random variables with copula C_{XY}. If both $\alpha(X)$ and $\beta(Y)$ are strictly increasing on $RanX$ and $RanY$ respectively, then $C_{\alpha(X)\beta(Y)} = C_{XY}$. If both $\alpha(X)$ and $\beta(Y)$ are strictly decreasing on $RanX$ and $RanY$ respectively, then $C_{\alpha(X)\beta(Y)}(u, v) = u + v - 1 + C_{XY}(1 - u, 1 - v)$.*

Corollary 4.3 Invariance of Kendall's Tau Under Monotone Transformation . *Let X and Y be continuous random variables with copula C_{XY}. If both $\alpha(X)$ and $\beta(Y)$ are strictly increasing or strictly decreasing on $RanX$ and $RanY$ respectively, then $\tau_{\alpha(X)\beta(Y)} = \tau_{XY}$.*

Proof: we just need to show the second part. If both $\alpha(X)$ and $\beta(Y)$ are strictly decreasing, then $C_{\alpha(X)\beta(Y)}(u,v) = u+v-1+C_{XY}(1-u,1-v)$. From equation 4.5, we have

$$\tau_{\alpha(X)\beta(Y)} = 1-4\int\int_{[0,1]^2}(1-C_1(1-u,1-v))(1-C_2(1-u,1-v))dudv$$

where C_i denotes the partial derivative with respect to ith variable to avoid confusion. By replacing $1-u$ by x and $1-v$ by y, we have

$$\tau_{\alpha(X)\beta(Y)} = 1-4\int\int_{[0,1]^2}(1-C_1(x,y))(1-C_2(x,y))dxdy.$$

Since

$$\int\int_{[0,1]^2} C_1(x,y)dxdy = \int_{[0,1]} ydy = 0.5,$$

we have $\tau_{\alpha(X)\beta(Y)} = \tau_{XY}$. ∎

These results are the foundation of modeling of default correlation in the pricing of portfolio credit risk. From now on, when we talk about correlation, we will mean Kendall's tau rank correlation.

4.2 Tail Dependence

Corresponding to the heavy tail property in univariate distributions, tail dependence is used to model the co-occurrence of extreme events. For credit risk, this is the phenomenon of default contagion. Realistic portfolio credit risk models should exhibit positive tail dependence, as defined next.

Definition 4.4 Tail Dependence Coefficient(TDC). *Let (X_1, X_2) be a bivariate vector of continuous random variables with marginal distribution functions F_1 and F_2. The level of* upper tail dependence λ_U *and* lower tail dependence λ_L *are given, respectively, by*

$$\lambda_U = \lim_{u \uparrow 1} P[X_2 > F_2^{-1}(u) | X_1 > F_1^{-1}(u)], \tag{4.6}$$

$$\lambda_L = \lim_{u \downarrow 0} P[X_2 \leq F_2^{-1}(u) | X_1 \leq F_1^{-1}(u)]. \tag{4.7}$$

If $\lambda_U > 0$, then the two random variables (X_1, X_2) are said to be *asymptotically dependent in the upper tail.* If $\lambda_U = 0$, then (X_1, X_2) are *asymptotically independent in the upper tail.* Similarly for λ_L and the lower tail.

Joe (1997) gave the copula version of TDC,

$$\lambda_U = \lim_{u \uparrow 1} \frac{[1 - 2u + C(u,u)]}{1-u}, \tag{4.8}$$

$$\lambda_L = \lim_{u \downarrow 0} \frac{C(u,u)}{u}. \tag{4.9}$$

For elliptical copulas, $\lambda_U = \lambda_L$, denoted simply by λ. Embrechts et al. (2001) showed that for a Gaussian copula, $\lambda = 0$, and for a t copula,

$$\lambda = 2 - 2t_{\nu+1}\left(\sqrt{\nu+1}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right), \tag{4.10}$$

where ρ is the Pearson correlation coefficient. We can see that λ is an increasing function of ρ and a decreasing function of the degree of freedom ν.

The t copula is a tail dependent copula. We can see the difference of the tail dependence between Gaussian copulas and t copulas from figure 1.

FIGURE 1 ABOUT HERE

5 Single Name Credit Risk

Before looking at the dependence structure of defaults for a portfolio, we first review the so-called reduced form approach to single firm credit risk, sometimes called stochastic intensity modeling. We follow the approach of Rutkowski (1999).

5.1 Defaultable Bond Pricing

Suppose that τ is the default time of a firm. Let $H_t = I_{\tau \leq t}$, and $\mathcal{H}_t = \sigma(H_s : s \leq t)$ denote the default time information filtration. We denote by F the right-continuous cumulative distribution function of τ, i.e., $F(t) = P(\tau \leq t)$.

Definition 5.1 Hazard Function. *The function* $\Gamma : \mathbb{R}^+ \to \mathbb{R}^+$ *given by*

$$\Gamma(t) = -\log(1 - F(t)), \forall t \in \mathbb{R}^+ \tag{5.1}$$

is called the hazard function. If F *is absolutely continuous, i.e.,* $F(t) = \int_0^t f(u)du$*, where* f *is the probability density function of* τ*, then so is* $\Gamma(t)$*, and we define the* intensity function

$$\lambda(t) = \Gamma'(t).$$

It is easy to check that

$$F(t) = 1 - e^{-\int_0^t \lambda(u)du}, \tag{5.2}$$

and

$$f(t) = \lambda(t)S(t), \tag{5.3}$$

where $S(t) = 1 - F(t) = e^{-\int_0^t \lambda(u)du}$ is called the *survival function.*

For simplicity, we suppose the risk free short interest rate $r(t)$ is a non-negative deterministic function, so that the price at time t of a unit of default free zero coupon bond with maturity T equals $B(t,T) = e^{-\int_t^T r(u)du}$.

Suppose now we have a defaultable zero-coupon bond that pays c at maturity T if there is no default, or pays a recovery amount $h(\tau)$ if there is a default at time $\tau < T$. The time-t present value of the bond's payoff is therefore

$$Y_t = I_{\{t<\tau\leq T\}} h(\tau) e^{-\int_t^\tau r(u)du} + I_{\{\tau>T\}} c e^{-\int_t^T r(u)du}.$$

When the only information is contained in the default filtration $\mathcal{H}_t$, we have the following pricing formula.

Proposition 5.2 *Rutkowski (1999). Assume that $t \leq T$, and Y_t is defined as above. If $\Gamma(t)$ is absolutely continuous, then*

$$E(Y_t|\mathcal{H}_t) = I_{\{\tau>t\}} \left(\int_t^T h(u)\lambda(u) e^{-\int_t^u \hat{r}(v)dv} du + c e^{-\int_t^T \hat{r}(u)du} \right), \tag{5.4}$$

where $\hat{r}(v) = r(v) + \lambda(v)$.

The first term is the price of the default payment, the second is the price of the survival payment. Note that in the first term we have used equation 5.3 to express the probability density function of τ. In the case of zero recovery, the formula tells us that a defaultable bond can be valued as if it were default free by replacing the interest rate by the sum of the interest rate and a default intensity, which can be interpreted as a credit spread. We use this proposition to price basket credit default swaps.

5.2 Credit Default Swaps

A credit default swap (CDS) is a contract that provides insurance against the risk of default of a particular company. The buyer of a CDS contract obtains the right to sell a particular bond issued by the company for its par value once a default occurs. The buyer pays to the seller a periodic payment, at time $t_1, \cdots, t_n$, as a fraction q of the nominal value M, until the maturity of the contract $T = t_n$ or until a default at time $\tau < T$ occurs. If a default occurs, the buyer still needs to pay the accrued payment from the last payment time to the default time. There are $1/\theta$ payments a year (for semiannual payments, $\theta = 1/2$), and every payment is $\theta q M$.

5.3 Valuation of Credit Default Swaps

Set the current time $t_0 = 0$. Let us suppose the only information available is the default information, interest rates are deterministic, the recovery rate R is a constant, and the expectation operator $E(\cdot)$ is relative to a risk neutral measure. We use Proposition 5.2 to get the premium leg PL, accrued payment AC, and default leg DL. PL is the present value of periodic payments and AP is the present value of the accumulated amount from last payment to default time. The default leg DL is the present value of the net gain to the buyer in case of default. We have

$$
\begin{aligned}
PL &= M\theta q \sum_{i=1}^{n} E(B(0,t_i)I\{\tau > t_i\}) \\
&= M\theta q \sum_{i=1}^{n} B(0,t_i)e^{-\int_0^{t_i}\lambda(u)du},
\end{aligned} \tag{5.5}
$$

$$
\begin{aligned}
AP &= M\theta q \sum_{i=1}^{n} E\left(\frac{\tau - t_{i-1}}{t_i - t_{i-1}} B(0,\tau) I\{t_{i-1} < \tau \le t_i\}\right) \\
&= M\theta q \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \frac{u - t_{i-1}}{t_i - t_{i-1}} B(0,u)\lambda(u) e^{-\int_0^u \lambda(s)ds} du,
\end{aligned} \tag{5.6}
$$

$$
\begin{aligned}
DL &= M(1-R)E\left(B(0,\tau)I_{\{\tau \le T\}}\right) \\
&= M(1-R)\int_0^T B(0,u)\lambda(u)e^{-\int_0^u \lambda(s)ds} du.
\end{aligned} \tag{5.7}
$$

The spread price q^* is the value of q such that the value of the credit default swap is zero,

$$
PL(q^*) + AP(q^*) = DL. \tag{5.8}
$$

5.4 Calibration of Default Intensity: Illustration

As Hull (2002) points out, the credit default swap market is so liquid that we can use credit default swap spread data to calibrate the default intensity using equation (5.8).

In table 1, we have credit default spread prices for five companies on 07/02/2004 from GFI (http://www.gfigroup.com). The spread price is quoted in basis points. It is the annualized payment made by the buyer of the CDS per dollar of nominal value. The mid price is the average of bid price and ask price.

TABLE 1 ABOUT HERE

We denote the maturities of the CDS contracts as $(T_1, \cdots, T_5) = (1, 2, 3, 4, 5)$. It is usually assumed that the default intensity is a step function, with step size of 1 year, expressed in the following form (where $T_0 = 0$),

$$\lambda(t) = \sum_{i=1}^{5} c_i I_{(T_{i-1}, T_i]}(t). \tag{5.9}$$

We can get c_1 by using the 1 year CDS spread price first. Knowing c_1, we can estimate c_2 using the 2 year CDS spread price. Following this procedure, we can estimate all the constants c_i for the default intensity.

In our calibration, we assume a recovery rate R of 0.4, a constant risk free interest rate of 0.045, and semiannual payments ($\theta = 1/2$). In this setting, we can get PL, AP and DL explicitly. The calibrated default intensity is shown in table 2.

TABLE 2 ABOUT HERE

6 Portfolio Credit Risk

6.1 Setup

Our setup for portfolio credit risk is to use default trigger variables for the survival functions (Schönbucher and Schubert, 2001), as a means of introducing default dependencies through a specified copula.

Suppose we are standing at time t=0.

Model Setup and Assumptions. *Suppose there are d firms. For each obligor $1 \le i \le d$, we define*

1. The default intensity $\lambda^i(t)$: a deterministic function. We usually assume it to be a step function.

2. The survival function $S^i(t)$:

$$S^i(t) := \exp\left(-\int_0^t \lambda^i(u)du\right). \tag{6.1}$$

3. The default trigger variables U_i: uniform random variables on $[0,1]$. The d-dimensional vector $\mathbf{U} = (U_1, U_2, \cdots, U_d)$ is distributed according to the d-dimensional copula C (see Definition 3.1).

4. The time of default τ_i of obligor i, where $i = 1, \cdots, d$,

$$\tau_i := \inf\{t : S^i(t) \le 1 - U_i\}. \tag{6.2}$$

The copula C of $\mathbf{U}$ is also called the survival copula of $\mathbf{1} - \mathbf{U}$. (See Georges et al. (2001) for more details about survival copulas.)

From equation (6.2), we can see that the default time τ_i is a increasing function of the uniform random variable U_i, so the rank correlation Kendall's tau between default times is the same as the Kendall's tau between the uniform random variables, and the copula of $\boldsymbol{\tau}$ equals the copula of $\mathbf{U}$. Equivalently, the copula of $\mathbf{1} - \mathbf{U}$, is the survival copula of $\boldsymbol{\tau}$.

Define the *default function*, $F^i(t) = 1 - S^i(t)$.

Theorem 6.1 Joint Default Probabilities. *The joint default probabilities of $(\tau_1, \tau_2, \cdots, \tau_d)$ are given by*

$$P[\tau_1 \leq T_1, \tau_2 \leq T_2, \cdots, \tau_d \leq T_d] = C(F^1(T_1), \cdots, F^d(T_d)). \tag{6.3}$$

Proof: From the definition of default in equation (6.2), we have

$$P[\tau_1 \leq T_1, \cdots, \tau_d \leq T_d] = P[1 - U_1 \geq S^1(T_1), \cdots, 1 - U_d \geq S^d(T_d)].$$

By the definition of the copula C of $\mathbf{U}$ we have

$$P[\tau_1 \leq T_1, \cdots, \tau_d \leq T_d] = C\left(F^1(T_1), \cdots, F^d(T_d)\right).$$

$\blacksquare$

6.2 Calibration

In the preceding setup, two kinds of quantities need to be calibrated: the default intensities $\lambda^i(t)$, and the default time copula C. Calibration of the default intensities can be accomplished using the single name credit default spreads visible in the market, as described below in Section 7.

However, calibration of the default time copula C is difficult. Indeed, it is a central and fundamental problem for portfolio credit risk modeling to properly calibrate correlations of default times. The trouble is that data is scarce – for example, a given basket of blue chips may not have any defaults at all in recent history. On the other hand, calibration using market prices of basket CDS is hampered by the lack of a liquid market with observable prices. Even if frequently traded basket CDS prices were observable, we would need many different basket combinations in order to extract full correlation information among all the names.

Therefore in the modeling process we need to choose some way of proxying the required data. McNeil et al. (2005) report that asset price correlations are commonly used as a proxy for default time correlations. This is also the approach taken by Cherubini et al. (2004), who remark that it is consistent with most market practice.

From the perspective of Merton-style value threshold models of default, it makes sense to use firm value correlations, since downward value co-movements will be associated with co-defaults. However, firm values are frequently not available, so asset prices can be used instead — even if, as Schönbucher (2003) points out, liquidity effects may lead to higher correlations for asset prices than for firm values.

Another way to simplify this calibration problem is to restrict to a family of copulas with only a small number of parameters, such as Archimedean copulas. Because this introduces too much symmetry among the assets, we choose instead to use asset price correlations as a proxy for default time correlations in this paper. This specific choice does not affect our conclusions, which apply to calibrating the copula of any asset-specific data set chosen to represent default time dependence.

A good choice of copula family for calibration is the t-copula, because it naturally incorporates default contagion through tail dependence, which is not present in the Gaussian copula. An even better choice is the skewed t-copula, for which the upper and lower tail dependence need not be equal.

When applying this copula approach, a direct calibration of the t-copula or skewed t-copula is time-consuming because there is no fast method of finding the degree of freedom ν except by looping. Instead, we will show that it is much faster to find the copula by calibrating the full multivariate distribution and extracting the implied copula, as in equation (3.5). This may seem counterintuitive, since the full distribution also contains the marginals as well as the dependence structure. However, for calibrating the full distribution function, we have at our disposal the fast EM algorithm; we know of no corresponding algorithm for the copula alone. Moreover, we will see that the marginals are needed anyway to contruct uniform variates. If they are not provided as a by-product of calibrating the full distribution, they need to be separately estimated.

7 Pricing of Basket Credit Default Swaps: Elliptical Copulas vs the Skewed t Distribution

7.1 Basket CDS contracts

We now address the problem of basket CDS pricing. For ease of illustration we will look at a 5 year basket CDS, where the basket contains the five firms used in section 5.4; other maturities and basket sizes are treated in the same way.

All the settings are the same as the single CDS except that the default event is triggered by the k-th default in the basket, where k is the seniority level of this structure, specified in the contract. The seller of the basket CDS will face the default payment upon the k-th default, and the buyer will pay the spread price until k-th default or until maturity T. Let $(\tau^1, \cdots, \tau^5)$ denote the default order.

The premium leg, accrued payment, and default leg are

$$PL = M\theta q \sum_{i=1}^{n} E(B(0,t_i) I\{\tau^k > t_i\}), \tag{7.1}$$

$$AP = M\theta q \sum_{i=1}^{n} E\left(\frac{\tau_k - t_{i-1}}{t_i - t_{i-1}} B(0,\tau^k) I\{t_{i-1} < \tau^k \leq t_i\}\right), \tag{7.2}$$

$$DL = M(1-R)E\left(B(0,\tau) I_{\{\tau^k \leq T\}}\right). \tag{7.3}$$

The spread price q^* is the q such that the value of credit default swap is zero, i.e.,

$$PL(q^*) + AP(q^*) = DL. \tag{7.4}$$

7.2 Pricing Method

To solve this equation, we now need the the distribution of τ^k, the time of the k-th default in the basket, so that we can evaluate the foregoing expectations. To do this, we need all the preceding tools of this paper. Here is a summary of the steps.

1. Select firm-specific critical variables $\mathbf{X}$ whose dependence structure will proxy for the dependence structure of default times. (In the study below we use equity prices.)

2. Calibrate the copula C of $\mathbf{X}$ from a selected parametric family of copulas or distributions, such as the t copula or the skewed t distribution. In the distribution case, use the EM algorithm.

3. Separately, calibrate deterministic default intensities from single name CDS spread quotes, as in section 5.

4. Use the default intensities to calculate survival functions $S^i(t)$ for each of the firms, using equation (6.1).

5. Using the copula C, develop the distribution of kth-to-default times by monte carlo sampling of many scenarios, as follows. In each scenario,

choose a sample value of $\mathbf{U}$ from the copula C. Use equation (6.2) to determine the default time for each firm in this scenario. Order these times from first to last to define $\tau^1, \ldots, \tau^5$. By repeating this simulation over many scenarios, we can develop a simulated unconditional distribution of each of the kth-to-default times τ^k.

6. Use these distributions to compute the expectations in equation (7.4) in order to solve for the basket CDS spread price q^*.

7.3 The distribution of kth-to-default times

Before describing our empirical results for this basket CDS pricing method, we elaborate a little on item 5 above, and examine via some experiments how the distributions depend on the choice of copula, comparing four different commonly used bivariate copulas: Gaussian, t, Clayton, and Gumbel.

To simplify the picture, we assume there are two idealized firms, with Kendall's $\tau = 0.5$ for all copulas. We take a five year horizon and set the default intensity of the first firm to be a constant 0.05, and 0.03 for the second firm. We want to look at the first to default (FTD) and last to default (LTD) probabilities at different times before maturity.

7.3.1 Algorithm

We calculated the k-th to default probabilities using the following procedure.

1. Use Matlab(TM) copula toolbox 1.0 to simulate Gaussian, t, Clayton and Gumbel copulas uniform variables $u_{i,j}$with the same Kendall's tau

correlation, where $i = 1, 2$, $j = 1, \cdots, n$ and n is the number of samples.

2. From equation (6.2), we get $\tau_{i,j}$ and sort according to column. The k-th row is a series of k-th to default times τ_i^k.

3. Divide the interval from year 0 to year 5 into 500 small sub-intervals. Count the number of τ_i^k values that fall into each sub-interval and divide by the number of samples to get the default probabilities for each small sub interval, and hence an approximate probability density function.

In the following, we illustrate results for FTD and LTD using $n = 1,000,000$ samples.

7.3.2 Empirical Probabilities of Last to Default (LTD) and First to Default (FTD)

First, we recall that the t-copula is both upper and lower tail dependent, the Clayton copula is lower tail dependent, but upper tail independent, the Gumbel copula is the reverse, and the Gaussian copula is tail independent in both tails.

FIGURE 2 ABOUT HERE

We can see from Figure 2 that a copula function with lower tail dependence (Clayton copula) leads to the highest default probabilities for LTD, while a copula function with upper tail dependence (Gumbel copula) leads to the lowest default probabilities. The tail dependent t-copula leads to higher default probabilities than tail independent Gaussian copula.

Default events tend happen when the uniform random variables $\mathbf{U}$ are small (close to 0). Since the LTD requires that both uniform variables in the basket are small, a lower tail dependent copula will lead to higher LTD probalities than a copula without lower tail dependence.

FIGURE 3 ABOUT HERE

In Figure 3, we see that the Clayton copula with only lower tail dependence leads to the lowest FTD probabilities, while the Gumbel copula with only upper tail dependence leads to the highest FTD probabilities. These results illustrate the sometimes unexpected relationships between tail dependence and FTD probabilities.

7.4 Empirical Basket CDS Pricing Comparison

We now use the method of Section 7.2 to compare two approaches to the calibration of the copula C. The first approach, popular in the literature, is to directly calibrate a t copula. Since this copula has tail dependence, it provides a way to introduce default contagion explicitly into the model. In order to get uniform variates, we will still need to specify marginal distributions, which we will take to be the empirical distributions.

The second approach is to calibrate the skewed t distribution using the EM algorithm described earlier. Calibrating the full distribution frees us from the need to separately estimate the marginals. Also, the skewed t distribution, has heavier tails than the t distribution, and does not suffer from the bivariate exchangeability of the t copula, which some argue is an unrealistic symmetry in the dependence structure of defaults.

In this experiment we use for our critical variables the equity prices for the same five underlying stocks as used above: AT&T, Bell South, Century Tel, SBC, Sprint. We obtained the adjusted daily closing prices from finance.yahoo.com for the period 07/02/1998 to 07/02/2004.

7.4.1 Copula approach

We first use the empirical distribution to model the marginal distributions and transform the equity prices into uniform variables. Then we can calibrate the t copula using those variates. For comparison, we also calibrate a Gaussian copula.

If we fix in advance the degree of freedom ν, the calibration of the t copula is fast — see Di Clemente and Romano (2003a), Demarta and McNeil (2005), and Galiani (2003). However, we know of no good method to calibrate the degree of freedom ν. With this data, we find the degree of freedom to be 7.406, which is found by maximizing the log likelihood using direct search, looping ν from 2.001 to 20 with step size 0.001. Each step takes about 5 seconds, and the full calibration takes about 24 hours (2005 vintage laptop running Windows XP).

The maximum log likelihood for the Gaussian copula was 936.90, while for the t copula it was 1043.94, substantially better. After calibration, we follow the remaining steps of section 7.2 and report the results in the table below.

Demarta and McNeil (2005) also suggest using the skewed t copula, but we were not able to calibrate it directly for this study.

7.4.2 Distribution approach

We calibrate the multivariate t and skewed t distributions using the EM algorithm described in Section 2. The calibration is fast compared to the copula calibration: with the same data and equipment, it takes less than one minute, compared to 24 hours for the looping search of ν. The calibrated degree of freedom for both t and skewed t is 4.31. The log likelihood for skewed t and t are almost the same: 18420.58 and 18420.20, respectively.

Spread prices for the k-th to default basket CDS are reported in Table 3. We can see that lower tail dependent t copula, compared to the Gaussian, leads to higher default probability for LTD and lower probability for FTD, thus leads to higher spread price for LTD and lower spread price for FTD. The t distribution has almost the same log likelihood and almost the same spread price of k-th to default as the skewed t distribution. Both distributions lead to higher spread price for LTD and lower spread price for FTD.

The calibration of the t distribution is a superior approach, both because there is no extra requirement to assume a form for the marginals, and because the EM algorithm has tremendous speed advantages. Basket credit default swaps or collateralized debt obligations usually have a large number of securities. For example, a synthetic CDO called EuroStoxx50 issued on May 18, 2001 has 50 single name credit default swaps on 50 credits that belong to the DJ EuroStoxx50 equity index. In this case, the calibration of a t copula will be extremely slow. TABLE 3 ABOUT HERE

8 Summary and Concluding Remarks

We follow Rukowski's single name credit risk modeling and Schönbucher and Schubert's portfolio credit risk copula approach to price basket credit default swaps.

The t copula is widely used in the pricing of basket credit default swaps for its lower tail dependence. However, we need to specify the marginal distributions first and calibrate the marginal distributions and copula separately. In addition, there is no good (fast) method to calibrate the degree of freedom ν.

Instead, we suggest using the fast EM algorithm for t distribution and skewed t distribution calibration, where all the parameters are calibrated together. To our knowledge, we are the first to suggest calibrating the full multivariate distribution to price basket credit default swaps with this trigger-variable approach.

As compared to the Gaussian copula, the t copula leads to higher default probabilities and spread prices of basket LTD credit default swaps, and lower default probabilities and spread prices for FTD, because of the introduction of tail dependence to model default contagion.

Both the t distribution and the skewed t distribution lead to yet higher spread prices of basket LTD credit default swaps and lower spread prices for FTD than the t copula. This is suggestive of a higher tail dependence of default times than is reflected in the pure copula approach. Because default contagion has shown itself to be pronounced during extreme events, we suspect that this is a more useful model of real default outcomes.

We feel the skewed t distribution has potential to become a powerful tool for quantitative analysts doing rich-cheap analysis of credit derivatives.

References

Cherubini, U., E. Luciano, and W. Vecchiato. 2004. *Copula Methods in Finance*, John Wiley and Sons.

Demarta, S. and A. McNeil. 2005. *The t copula and related copulas*, International Statistical Review, to appear.

Di Clemente, A. and C. Romano. 2003a. *Beyond Markowitz: building the optimal portfolio using non-elliptical asset return distributions*, Rome, Italy, working paper.

———. 2003b. *A copula extreme value theory approach for modeling operational risk*, Rome, Italy, working paper.

———. 2004. *Measuring and optimizing portfolio credit risk: a copula-based approach*, Rome, Italy, working paper.

Embrechts, P., F. Lindskog, and A. McNeil. 2001. *Modelling dependence with copulas and applications to risk management*, Zurich.

Galiani, Stefano. 2003. *Copula functions and their applications in pricing and risk managing multiname credit derivative products*, Dept. of Mathematics, King's College, London, Master's thesis.

Georges, P., A. Lamy, E. Nicolas, G. Quibel, and T. Roncalli. 2001. *Multivariate survival modelling: a unified approach with copulas.*

Hu, Wenbo. 2005. *Calibration of Multivariate Generalized Hyperbolic Distributions using the EM Algorithm, with Applications in Risk Management, Portfolio Optimization and Portfolio Credit Risk*, Florida State University, Tallahassee, FL, PhD Dissertation.

Hull, John. 2002. *Options, Futures & Other Derivatives*, 5th ed., Prentice Hall.

Joe, H. 1997. *Multivariate Models and Dependence Concepts*, Monographs on statistics and applied probability **73**.

Li, David. 1999. *On default correlation: A copula function approach*, working paper.

Masala, G., M. Menzietti, and M. Micocci. *Optimization of conditional VaR in an actuarial model for credit risk assuming a student t copula dependence structure*, University of Cagliari, University of Rome, working paper.

Mashal, R. and M. Naldi. 2002. *Pricing multiname credit derivatives: heavy tailed hybrid approach*, working paper.

McNeil, A., R. Frey, and P. Embrechts. 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton Univ. Press.

Meneguzzo, D. and W. Vecciato. 2002. *Copula sensitivity in collateralized debt obligations and basket default swaps pricing and risk monitoring*, preprint.

Nelsen, Roger. 1999. *An Introduction to Copulas*, Springer Verlag.

Rutkowski, M. 1999. *On models of default risk by R. Elliott, M. Jeanblanc and M. Yor*, Warsaw Univ. of Technology, working paper.

Schönbucher, P.J. and D. Schubert. 2001. *Copula-dependent default risk in intensity models*, working paper.

Schönbucher, P. 2003. *Credit Derivatives and Pricing Models: Models, Pricing and Implementation*, John Wiley and Sons.

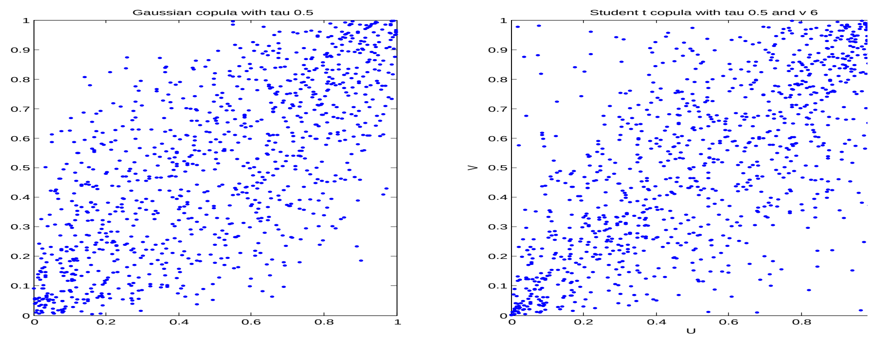

Figure 1: 1000 samples of Gaussian and t copula with Kendall's $\tau = 0.5$. There are more points in both corners for the t copula.

Company	Year 1	Year 2	Year 3	Year 4	year 5
AT&T	144	144	208	272	330
Bell South	12	18	24	33	43
Century Tel	59	76	92	108	136
SBC	15	23	31	39	47.5
Sprint	57	61	66	83	100

Table 1: Credit default swap mid price quote, where year1, $\cdots$, year5 mean maturities.

Company	Year 1	Year 2	Year 3	Year 4	year 5
AT&T	0.0237	0.0237	0.0599	0.0893	0.1198
Bell South	0.0020	0.0040	0.0061	0.0105	0.0149
Century Tel	0.0097	0.0155	0.0210	0.0271	0.0469
SBC	0.0025	0.0052	0.0080	0.0109	0.0144
Sprint	0.0094	0.0108	0.0127	0.0235	0.0304

Table 2: Calibrated default intensity

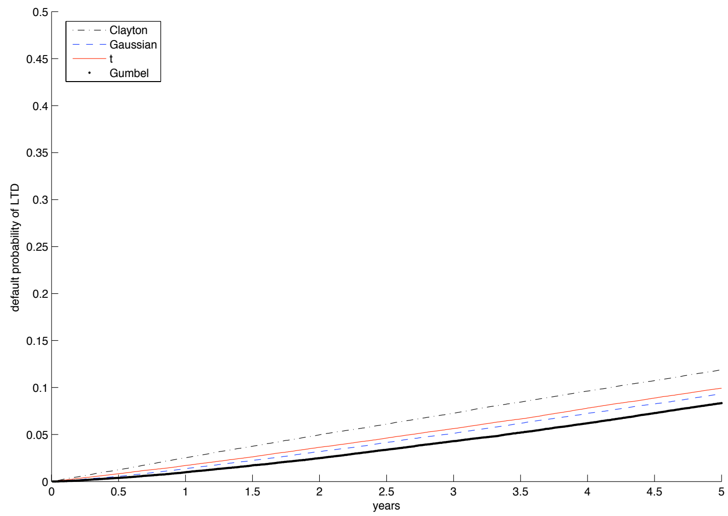

Figure 2: Default probabilities of LTD.

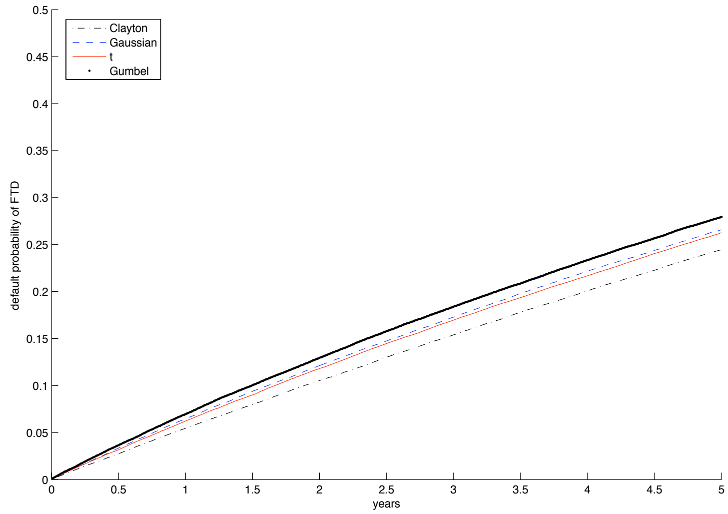

Figure 3: Default probabilities of FTD.

Model	FTD	2TD	3TD	4TD	LTD
Gaussian copula	525.6	141.7	40.4	10.9	2.2
t copula	506.1	143.2	46.9	15.1	3.9
t distribution	498.4	143.2	48.7	16.8	4.5
Skewed t distribution	499.5	143.9	49.3	16.8	4.5

Table 3: Spread price for k-th to default using different models