SIMULATING CORRELATED DEFAULT PROCESSES USING COPULAS: A CRITERION-BASED APPROACH

Sanjiv R. Das and Gary Geng

Date: Versions: First - October 2001, Current - February 2003.

Key words and phrases. Correlated default, copulas, hazard rates.

Sanjiv Das is at Santa Clara University, and Gary Geng is at Amaranth LLC, Connecticut, though this work was completed when he was at Gifford Fong Associates. We are extremely thankful for many constructive suggestions and illuminating discussions with Darrell Duffie, Gifford Fong, Nikunj Kapadia, John Knight, Robert Merton and Ken Singleton, and participants at the Q-conference (Arizona 2001), the Risk Conference (Boston 2002), and at Barclays Global Investors Seminar (San Francisco 2002). The first author gratefully acknowledges support from the Dean Witter Foundation, and a Breetwor Fellowship. We are also grateful to Gifford Fong Associates, and Moody's Investors Services for data and research support for this paper. The second author is supported by the Natural Sciences and Engineering Reserves Council of Canada. Please address all correspondence to Professor Sanjiv Das, Santa Clara University, Leavey School of Business, Dept of Finance, 208 Kenna Hall, Santa Clara, CA 95053-0388. Email: srdas@scu.edu.

Simulating Correlated Default Processes Using Copulas: A Criterion-Based Approach

Abstract

Modeling correlated default risk is a new phenomenon currently sweeping through the credit markets. Little is known about the drivers of default risk at the portfolio level. This paper develops a methodology to assess alternative specifications of the joint distribution of default risk. Specifications are based on three criteria: level, asymmetry, and tail-dependence in the joint default distribution. The study is based on a data set of default probabilities supplied by Moody's Risk Management Services. We undertake an empirical examination of the joint stochastic process of default risk over the period 1987-2000. Using copula functions, we separate the estimation of the marginal distributions from the joint distribution. We determine the appropriate choice of multivariate distribution based on a new metric for the assessment of joint distributions. This metric accounts for different aspects of default correlation, namely level, asymmetry and tail-dependence or extreme behavior. Our model, based on estimating a joint system of over 600 issuers, is designed to replicate the fat-tailed empirical joint distribution of defaults. A comparison of a jump model and a regime-switching model shows that the latter provides a better representation of the properties of correlated default. We also find that the skewed double-exponential distribution is the best choice for the marginal distribution of each issuer's hazard rate process, and combines well with the normal, Gumbel, Clayton and student's t copulas in the joint dependency relationship amongst issuers. As a complement to the methodological innovation, we show that (a) appropriate choices of marginal distributions and copulas are essential in modeling correlated default, (b) accounting for regimes is an important aspect of joint specifications of default risk, and (c) misspecification of credit portfolio risk can occur easily if joint distributions are inappropriately chosen. We believe this is one of the first papers to empirically compare joint default distributions using copulas for the U.S. market.

1. Introduction

Default risk at the level of an individual security has been extensively modeled using both structural and reduced form models[1] However, little is known about default risk at the portfolio level. Default dependencies among many issuers in a large portfolio play an important role in the quantification of a portfolio's credit risk exposure for the many reasons. Growing linkages in the financial markets have led to a greater degree of joint default. While the actual portfolio loss due to the default of an individual obligor may be small unless the risk exposure is extremely large, the effects of simultaneous defaults of several issuers in a well diversified portfolio could be catastrophic. In order to efficiently manage and hedge the risk exposure to joint defaults, a model for default dependencies is called for. Further, innovations in the credit market have been growing at an unprecedented pace in recent years, and will likely persist in the near future. Many newly developed financial securities, such as collateralized debt obligations (CDOs), have payoffs depending on the joint default behavior of the underlying securities[2]. In order to accurately measure and price the risk exposure of these securities, an understanding and accurate measurement of the default dependencies among the underlying securities is essential. Another area in which accurate specification of the stochastic process of default correlations is required is in the pricing of basket default contracts. In this setting, correlations specifications are critical, as the baskets are not large enough to ensure diversification. Finally, the Basle Committee on Banking Supervision has identified poor portfolio risk management as an important source of risk for banking institutions. As a result, banks and other financial institutions have been required to measure their overall risk exposures on a routine basis. Default correlation is again an important ingredient to integrate multiple credit risk exposures.

[1] See the structural models of Merton [1974], Geske [1977], Leland [1994], Longstaff and Schwartz [1995], and reduced from models of Duffie and Singleton [1999], Jarrow and Turnbull [1995], Das and Tufano [1996], Jarrow, Lando and Turnbull [1997], Madan and Unal [1999], and Das and Sundaram [2000].

[2] A CDO securitization comprises a pool of bonds (the "collateral") against which tranches of debt are issued with varying cashflow priority. Tranches vary in credit quality, from AAA to B, depending on subordination level. Seller's interest is typically maintained as a final equity tranche, carrying the highest risk. CDO collateral usually comprises from a hundred to over a thousand bond issues. Tranche cashflows critically depend on credit events during the life of the CDO, requiring Monte Carlo simulation (see for example, Duffie and Singleton [1999]) of the joint default process for all issuers in the collateral. An excellent discussion of the motivation for CDOs and the analysis of CDO value is provided in Duffie and Garleanu [2001]. For a parsimonious model of bond portfolio allocations with default risk, see Wise and Bhansali [2001].

Recently, growing effort has been devoted to the modeling of the joint default process at a portfolio level. In the framework of a structural model, dependence between default events is often derived from the dependence between latent variables, such as a firm's assets. By relating the latent variables to a small number of economic factors, the correlation matrix of latent variables can be computed. It is common practice to assume that the latent variables follow a multi-normal distribution, and consequently the default dependencies are completely captured by Pearson's correlations. Frey, McNeil and Nyfeler [2001] demonstrate that joint default events generated from latent variables are sensitive to distributional assumptions. In particular, for a given correlation matrix, models using distributions with fat tails, such as student's t, can generate more extreme risky scenarios than models with multi-normal distributions, implying that models built on the assumption of multi-normal distributions may not adequately capture the potential risk exposure in a portfolio. We find that jumps and regime-shifting behavior represent the joint dependence of defaults well, implying that extreme risky scenarios are embedded in the empirical behavior of defaults.

In the approach of a reduced form model, default dependencies are introduced either through the co-movement of the default intensities of the obligors, or through a joint default event triggered by a separate point process[3]. Theoretically speaking, reduced form models provide a more flexible framework for modeling the dynamics and the term structure of credit risk based on market variables, such as credit spreads. However, calibration of the model to market variables is not trivial because of the sparsity of default data and the need to model a large number of variables simultaneously. Furthermore, in most cases, analytical formulae do not exist, or are very difficult to derive. Consequently, numerical methods, such as Monte Carlo Simulation, are often used to compute desired risk measures, such as credit VaR. This is particularly true if the portfolio contains many positions in credit-related derivatives. Since individual default probabilities in a given time period are typically small, it usually needs a substantial number of simulations to reach the required distributions with an acceptable accuracy. Consequently, simulating the joint default events for a large and complex portfolio is computationally expensive, and may not even be achievable in some cases. In order to reduce the computation time to a reasonable amount, unrealistic assumptions about individual default behavior and the associated default correlations have to be made. However, as demonstrated by Das, Fong and

[3]For details on the reduced-form models of default correlation and multiple defaults, see, for example, Davis and Lo [1999a], [1999b], Duffie and Singleton [1999], Duffie and Garleanu [2001], Jarrow and Yu [2000] and Kijima [2000].

Geng [2001], Gersbach and Lipponer [2000], both the individual default probability and default correlations have significant impact on the value of a credit portfolio. Ignoring the stochastic properties of default probabilities and default correlations will likely produce distorted risk measurements.

In this paper, we use Moody's unique and comprehensive default database to develop a parsimonious numerical method of modeling and simulating correlated default processes for hundred of issuers which allows us to best capture the stylized facts of individual default probabilities as well as the dynamics of default correlations documented in Das, Freed, Geng and Kapadia (DFGK [2001]). Our methodology is "criterion-based", i.e. it uses a metric that compares alternative specifications of the joint default distribution using three criteria: (a) the *level* of default risk, (b) the *asymmetry* in default correlations, and (c) the *tail-dependence* of joint defaults. We elaborate on these in the sequel.

The base unit of joint credit risk is the probability of default (PD) of each issuer. Many popular approaches exist for computing PDs in the market place, developed by firms such as KMV Corporation, Moodys Risk Management Systems (MRMS), RiskMetrics, etc. In the spirit of reduced form models, default probabilities are usually expressed as hazard rates, which we denote as $\lambda_i(t),\ i = 1..N$. The hazard rates for all N issuers vary over time. The survival probabilities over a horizon T are denoted $s_i(T) = E\left\{\exp[-\int_0^T \lambda_i(t)dt]\right\}$, and the probability of default is therefore $p_i(t) = 1 - s_i(t)$. This paper empirically examines the *joint* stochastic process for hazard rates, $\lambda_i(t),\ i = 1...N$.

Intuitively, in order to obtain the joint statistical distribution for all issuers in a portfolio that fits the correlated default process well, we need not only to find the best fitted marginal distribution for each issuer which captures the observed skewness and fat tails of each individual default probability, but also to model the joint stochastic process with the observed asymmetric correlation structure. Our proposed numerical method includes the following elements. Firstly, we model the dynamics of the mean default process of each rating group using a normal jump process as well as a regime switching model. Both models are found to be apt in accommodating the fat-tails of the distributions of default probabilities. The individual default probability is modeled as a CIR process with the long-term mean being a linear function of the mean default rate of the rating group it belongs to. This two-step modeling approach not only captures the dependence of default probabilities on economy-wide factors, but also greatly simplifies the simulation procedure as demonstrated in the rest of this paper.

Secondly, we allow the individual default probability to have distributions with different tail properties. Specifically, we consider three distributional assumptions, a normal distribution, a student's t distribution and a skewed double exponential distribution. We find the best distribution for each issuer using four goodness-of-fit criteria.

Thirdly, We use a *copula* approach to merge the marginal distributions for each issuer into a single multivariate distribution with the desired correlation between issuers. Four types of copulas are considered to embed the tail-dependence with a view to modeling correlation as emanating from the joint occurrence of outliers[4].

Finally, we develop a metric to determine how well different multivariate distributions fit the observed covariation in PDs. The metric accounts for the asymmetry of default correlations between issuers in the Moody's database of probabilities of default, and enables us to determine the best model from 56 joint distributions considered.

The methodology developed in this paper is numerically intensive, yet easy to implement. One of our primary goals is to uncover a satisfactory joint distribution of hazard rates so as to foster accurate credit portfolio choice and subsequent risk-management. Pursuant to this goal, our methodology and the accompanying metric is aimed at capturing three aspects of the dependence of defaults amongst issuers. First, we wish to correctly capture the level of default correlations. Second, it is clear from default history that there are periods in which defaults cluster, while in other periods, this aspect is absent. This results in an asymmetry in correlation which we capture, i.e. defaults are more dependent when hazard rates are high. Third, we capture default tail-dependence, whereby correlation is higher when conditioned on being in the tails of the multivariate default distribution. Our metric assesses all candidate stochastic processes for the best fit to all 3 aspects.

There are many important benefits from our analyses. The econometric specifications developed here may be used for structuring and pricing collateralized debt obligations (CDOs). The estimated model may be used to compute risk-management

[4]Extreme value distributions allow for the fact that processes tend to evidence higher correlations when tail values are experienced. This leads to a choice of fatter-tailed distributions, with a reasonable degree of "tail dependence". The support of the distribution should also be such that PD intensity $\lambda_i(t)$ lies in the range $[0, \infty)$.

characteristics of credit portfolios, for example, the credit Value-at-Risk of a portfolio (the CVaR). Trading houses may use the model to assess joint exposure to myriad counterparties, especially when performance risk becomes a critical issue in derivative books.

To the best of our knowledge, this is the first *empirical* examination using extreme-value distributions in modeling joint default risk for U.S. corporate data. The main issue we look at here is "tail-dependence". Intuitively speaking (we define this formally later), tail dependence is the feature of the joint distribution that determines how much of the correlation between the hazard rate processes comes from extreme observations than central observations. For example, imagine two different joint distributions, one multivariate normal, the other multivariate student's t with low degrees of freedom, both with the same correlation matrix. The former has lower tail-dependence than the latter, which generates more joint tail observations. Hence, from a risk point of view, the multivariate student's t distribution would be riskier for a large class of risk-averse investors. In order to flexibly model tail-dependence, we employ the copula function technique.

We find that the best choice of copula depends on the marginal distributions. The student's t copula is best when the jump-normal model is chosen for the mean hazard rate of each rating class. However, the normal or Clayton copula is more appropriate when the regime-switching model is chosen for mean hazard rates. In general, the metric prefers the regime-switching model to the jump-normal one irrespective of the choice of marginal distributions, and a perusal of the time series plot of the data (see Figure 1) provides intuitive support for this result. While the regime-switching model for mean hazard rates is best, no clear winner emerges amongst the competing copulas. Copula performance is found to be dependent on the choice of marginal distributions. We also find that choosing the best fit marginal distribution for each issuer does not necessarily ensure the best fit for the joint distribution! This interdependence between copula and marginal distribution justifies our wide-ranging search over 56 different specifications.

We also assess the impact of different copulas on risk-management measures. An examination of the tail loss distributions shows substantial differences are possible amongst the 56 econometric specifications. Arbitrarily assigning fat-tailed distributions to a fat-tailed copula may even result in the unnecessary overstatement of joint default risk. Since different copulas inject varied levels of tail dependence, the metric developed in this paper allows fine-tuning of the specification, which enhances the accuracy of credit VaR calculations.

In Section 3 of the paper, we briefly review the literature on copulas. There is now a vast collection of theoretical papers justifying the use of copulas in modeling extreme-value distributions. This paper provides a much-needed empirical complement to this literature.

The rest of the paper proceeds as follows. In section 2 we briefly describe the data on hazard rates. As mentioned, section 3 provides the finance reader with a brief introduction to copulas, and section 4 contains the estimation procedure and results. Section 5 presents the simulation model and the metric used to compare different copulas. Section 6 concludes.

2. Description of the Data

Our data set comprises issuers tracked by Moodys from 1987-2000. For each issuer, we have probabilities of default (PDs) based on Moodys econometric models for every month that the issuer was tracked by Moodys. Issuers are divided into seven rating classes. Rating classes 1 through 6 reflect progressively declining credit quality. Many more issuers fall into rating class 7, which comprises unrated issuers, and the PDs within this class range from high to low, resulting in an average PD that is close to the median PD of all the other rating classes. Table 1 describes Moody's rating categories. In our analysis, we do not consider PDs from rating class 7. Since DFGK [2001] demonstrated that cross-sectional variation by rating class is a feature of the data, class 7, a mixture of unrated firms from all classes, was dropped to avoid confounding results. Thus, we obtained data on a total of 620 issuers classified into 6 rating classes.

The time series of average default probabilities is presented in Figure 1. Table 2 presents the descriptive statistics of our data from rating classes 1 through 6. In Panel A, the statistical properties in levels are presented. As expected, the mean hazard rate increases from rating class 1 to 6, as does the standard deviation. PD changes tend to be higher for lower grade debt. However, the skewness and kurtosis of changes are high for the better rating classes 1 and 2. In this grade of debt, PDs change less frequently and when they do, it is material, resulting in more outliers. Panel C depicts the correlations of PDs (upper right) and PD changes (lower left). Notice that the correlations are far higher for high grade debt than for low grade debt. This highlights the fact that high grade debt is more systematically related, and low grade debt experiences greater idiosyncratic risk. Correlations in levels are higher than that of changes, which is as expected.

Table 3 contains the dependency between rating classes measured by Kendall's τ statistic. We will explain in the next section why this is a better measure of dependency than the standard correlation measure (Pearson's correlation) for modeling joint default. We obtain similar results for dependency as before with Kendall's τ. Higher grade debt evidences more rank correlation in PDs than low grade debt.

In the following section, we summarize a few theoretical results pertaining to copulas, that are useful in the empirical fitting of joint default distributions.

3. A Brief Introduction to Copulas

3.1. **Definition.** We are interested in modeling an n-variate distribution. A random draw from this distribution comprises a vector $X \in R^n = \{X_1, ..., X_n\}$. Each of the n variates has its own marginal distribution, $F_i(X_i),\ i = 1, ..., n$. The joint distribution is denoted as $F(X)$. The copula associated with $F(X)$ is a multivariate distribution function defined on the unit cube $[0, 1]^n$ with uniformly distributed marginals. When all the marginals, $F_i(X_i),\ i = 1, ..., n$, are continuous, the associated copula is unique and can be written as:

$$C(u_1, ..., u_n) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)). \tag{1}$$

For an extensive discussion of copulas, see Nelsen [1999][5].

Copulas allow the modeling of the marginal distributions separately from their dependence structure. This greatly simplifies the estimation problem of a joint stochastic process for a portfolio with many issuers. Instead of estimating all the distributional parameters simultaneously, we can estimate the marginal distributions separately from the joint distribution. Given the estimated marginal distribution for each issuer, we then use appropriate copulas to construct the joint distribution with a desired correlation structure. The best copula can be found by examining the statistical fit of the different copulas to the data[6].

[5] There is a growing literature on copulas. The following references contain many papers from which this summary section was developed: Embrechts, Kluppelberg, and Mikosch [1997], Embrechts, McNeil and Straumann [1999a], [1999b], Frees and Valdez [1998], Frey and McNeil [2001], Li [1999], Nelsen [1999], Lindskog [2000], Schonbucher and Schubert [2001], Wang [2000].

[6] Durrleman , Nikeghbali and Roncalli [2000] discuss several parametric and non-parametric estimation method of copulas

Copula techniques lend themselves to two types of credit risk analysis. First, given a copula, we can choose different marginal distribution for each individual issuer. By changing the types of marginal distributions and their parameters, we can examine how the individual default affects the joint default behavior of many issuers in a credit portfolio. Second, given marginal distributions, we can vary the correlation structures by choosing different copulas, or the same copula with different parameter values. It therefore enables us to quantify the effects of default correlations on a credit portfolio.

3.2. **Properties.** Let $U_i = F_i(X_i)$ for all $i = 1, ..., n$. Then $U = (U_1, ..., U_n)^T$ is a vector of random variables with uniform marginal distributions. Copulas possess the following properties:

(1) Every joint distribution may be written as a copula. This is Sklar's Theorem first introduced in Sklar [1959] and [1973].
(2) If all the marginals, $F_1(X_1), ..., F_n(X_n)$, are continuous, then the associated copula C is unique.
(3) $C(u_1, ..., u_n)$ is increasing in each u_i and $C[1, ..., 1, u_i, ..., 1] = u_i, i = 1, ..., n$.
(4) If $\alpha_1, ..., \alpha_n$ are strictly increasing on $X_1, ..., X_n$ respectively, then $(\alpha_1(X_1), ..., \alpha_n(X_n))^T$ also has a copula C.
(5) Let $\bar{C}(u_1, ..., u_n) = C(U_1 > u_1, ..., U_n > u_n)$ with $u_1 = F_1(x_1), ..., u_n = F_n(x_n)$, and define the survival copula $\hat{C}(1 - F_1(x_1), ..., 1 - F_n(x_n)) = F(X_1 > x_1, ..., X_n > x_n)$. Then $\bar{C}(u_1, ..., u_n) = \hat{C}(1 - u_1, ..., 1 - u_n)$.

There are other properties too, which are not mentioned here, as they are not directly required for the empirical objectives of this paper. The interested reader may review the more detailed expositions in the bibliography.

3.3. **Examples of Copulas.** In this paper, in order to capture the observed properties of joint default process, we consider the following four types of copulas:

- **Normal Copula**: The normal copula of the $n - variate$ normal distribution with correlation matrix $\boldsymbol{\rho}$ is defined as

$$C_{\boldsymbol{\rho}}(u_1, ..., u_n) = \Phi^n_{\boldsymbol{\rho}}(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)),$$

where $\Phi^n_{\boldsymbol{\rho}}$ denotes the joint distribution function of the $n - variate$ normal distribution with correlation matrix $\boldsymbol{\rho}$. In the bivariate case, the normal copula can be written as:

$$C_{\boldsymbol{\rho}}(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \left\{ -\frac{x_1^2 - 2\rho_{12}x_1x_2 + x_2^2}{2(1-\rho_{12}^2)} \right\} dx_1 dx_2, \tag{2}$$

where ρ_{12} is the linear correlation coefficient between the two normal variables.

- **Student's t Copula** Let $T_{\boldsymbol{\rho},\nu}$ be the standardized multivariate Student's t distribution with ν degrees of freedom and correlation matrix $\boldsymbol{\rho}$. The multivariate Student's t copula is then defined as follows:

$$C(u_1, ..., u_n; \boldsymbol{\rho}, \nu) = T_{\boldsymbol{\rho},\nu}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_n)),$$

where t_ν^{-1} is the inverse of the univariate student's t distribution.

- **Gumbel Copula** Gumbel copula was first introduced by Gumbel [1960] and can be expressed as follows:

$$C(u_1, ..., u_n) = \exp\left[-\left(\sum_{i=1}^{n}(-\ln u_i)^\alpha\right)^{1/\alpha}\right] \tag{3}$$

where α is the parameter determining the tail of the distribution.

- **Clayton Copula** The Clayton copula introduced in Clayton [1978] is as follows:

$$C(u_1, ..., u_n) = \left[\sum_{i=1}^{n} u_i^{-\alpha} - 1\right]^{-1/\alpha} \tag{4}$$

Again, $\alpha > 1$ is a tail-dependence parameter.

Simulating a n-dimensional random vector from the above copulas can be found in Wang [2000], Bouye, Durrleman, Nikeghbali, Riboulet and Roncalli [2000] and Embrechts, Lindskog and McNeil [2001].

Example: Figure 2 illustrates the use of a normal copula. The figure presents plots of simulated random variables from the bivariate normal distribution, and from a bivariate student's t distribution with degrees of freedom equal to 5. This distribution is reasonably fat-tailed and is embedded in a normal copula for the simulation. The difference in the fatness of the tails of the student's t distribution in comparison to the normal is apparent. The student's t distribution results in much higher leptokurtosis – there are more outliers, and the center of the distribution is also tightly packed. (Note that the axes are much wider for the plots of the student's t distribution.)

3.4. **Embedding correlation in copulas.** Two important concepts which are closely related to the copula method are the *correlations* and the *tail-dependence*. In this section, we will discuss the concept of correlations and will look at the tail-dependence in the next section.

Traditionally, Pearson's correlation, ρ, is widely used in the measurement of associations between random variables. However, as pointed out by Embrechts, McNeil and Straumann [1999b], ρ tends to have several deficiencies. First, it is not a measure of general dependency, but only a measure of linear dependence. Second, all correlation levels in the range $[-1, 1]$ are not necessarily attainable in the joint distribution. Third, a correlation of zero does not imply that risks are independent. Fourth, ρ is not invariant under transforms. In contrast to ρ, rank correlation, such as Kendall's τ and Spearman's ρ^s, do not suffer from most of these deficiencies, and therefore are more useful in the copula framework.

In the following, we will use Kendall's τ as an example and look at some of the most important properties of rank correlations. Let us first define Kendall's τ in the bivariate framework. If (X_1, Y_1) and (X_2, Y_2) are two independent draws from a joint distribution, then

$$\tau = Pr[(X_2 - X_1)(Y_2 - Y_1) > 0] - Pr[(X_2 - X_1)(Y_2 - Y_1) < 0]. \tag{5}$$

If $(X_2 - X_1)(Y_2 - Y_1) > 0$, then the pair of random draws is *concordant*, else it is *discordant*. If we define c as the number of concordant pairs, and d as the number of discordant pairs, then we can also define the rank correlation as:

$$\begin{aligned} \tau &= \frac{c-d}{c+d} = \frac{c-d}{\frac{n(n-1)}{2}} \\ &= \frac{2}{n(n-1)} \sum_{i<j} \text{sign}[(X_i - X_j)(Y_i - Y_j)] \end{aligned}$$

If τ is computable analytically, then we get the parametric expression:

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 F_{X,Y}(x, y) d^2 F_{X,Y}(x, y) - 1. \tag{6}$$

else the non-parametric version as stated before is used. Rank correlation has the following properties:

- $\tau \in [-1, 1]$.
- If X, Y are *comonotonic*, then $\tau = 1$. Two random variables X, Y are comonotonic if there exists a random variable z such that $X = f(z), Y = g(z)$ with probability 1, where $f(.), g(.)$ are non-decreasing.

- If $X, -Y$ are *comonotonic*, then $\tau = -1$.
- If X, Y are independent, then $\tau = 0$.
- τ is invariant under strictly monotonic transformations, that is if $f(X), g(Y)$ are strictly increasing or decreasing functions. Then, $\tau[f(X), g(Y)] = \tau(X, Y)$.

The Kendall's τ is closely associated with the parameters of copulas. For example, for a normal copula, we have

$$\tau[\Phi(X), \Phi(Y)] = \tau[X, Y] = \frac{2}{\pi} \arcsin(\rho(X, Y)) \tag{7}$$

which connects the τ with the correlation coefficient ρ. Similar relationships hold for other copulas. For a Gumbel copula, $\alpha = \frac{1}{1-\tau}$. With a Clayton copula, we have $\alpha = \frac{2\tau}{1-\tau}$. In our work using the student's t copula, we choose the degree of freedom to be 6.

3.5. **Tail dependence.** An important feature of the use of copulas is that it permits varying degrees of *tail-dependence*. Tail dependence refers to the extent to which the dependence (or correlation) between random variables arises from extreme observations. There is plenty of evidence that large market moves are characterized by higher levels of correlation than occurs during quiet markets (see Das and Uppal [2000]). A similar phenomenon appears to exist for correlated default. Das, Freed, Geng and Kapadia [2001] find that correlation levels amongst PDs are higher when the PDs are in the high regime, as compared to that in the low regime. The use of copulas that allow for varying tail dependence is an important aspect of the flexibility of this modeling approach.

Tail dependence is low in the normal distribution. Most of the dependency in multivariate normal distributions comes from central observations, i.e. phenomena such as contagion, herding, etc., are difficult to model when random variables are specified as joint normal. It seems more appropriate to model tail dependency with other marginal distributions. Moreover, tail dependency, which is a copula property, may be specified for each tail. Hence, we could have upper tail dependency, lower tail dependency, or both.

Suppose (X_1, X_2) is a continuous random vector, with marginal distributions F_1 and F_2. The coefficient of upper tail dependence is:

$$\lambda_U = \lim_{z \to 1} Pr[X_2 > F_2^{-1}(z) | X_1 > F_1^{-1}(z)] \tag{8}$$

If $\lambda_U > 0$, then upper tail dependence exists. Intuitively, upper tail dependence exists when there is a positive probability of positive outliers occurring jointly. Lower tail dependence is symmetrically defined. The coefficient of lower tail dependence is:

$$\lambda_L = \lim_{z \to 0} Pr[X_2 < F_2^{-1}(z) | X_1 < F_1^{-1}(z)] \tag{9}$$

If $\lambda_L > 0$, then lower tail dependence exists.

For example, the Gumbel copula has upper tail dependence with $\lambda_U = 2 - 2^{\frac{1}{\delta}}$. The Clayton copula has lower tail dependence with $\lambda_L = 2^{-\frac{1}{\delta}}$. The student's t has equal upper and lower tail dependence with $\lambda_U = 2\bar{t}_{v+1}\left(\frac{\sqrt{v+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)$.

Table 4 provides a snapshot of the tail dependence between rating classes. The results in the table present the correlations between rating pairs over observations in the tails of the bivariate distribution. Results are provided both, in levels and changes. There is strong positive correlation in the upper tail, evidencing upper tail dependence. Correlations in the lower tail are often low and negative, and hence, there is not much evidence of lower tail dependence.

Figure 3 presents bivariate scatter plots of Monte Carlo samples from the four copulas we use in this study. The fatter tails of the student's t versus the normal copula are evident. The upper tail dependence of the Gumbel copula, and the lower tail dependence of the Clayton copula are also seen. Figures 4 and 5 depict a closer look at the tails of these copula distributions by zooming in on the respective regions.

3.6. **Empirical features of dependency in the joint distribution.** In this section we explore three features of the dependency structure in the processes of joint default. We also explore possible ways of determining the empirical estimates that provide the best representation of the correlations amongst hazard rates. Traditional maximum-likelihood provides a means of assessing the fit of various marginal distributions, and the likelihood ratio test (suitably corrected for degrees of freedom by the Akaike criterion) enables the choice of the best marginal for each time series of default rates, indexed by issuer i.

However, it is far harder to use maximum-likelihood to compare the quality of fit across various joint distributions. There are many reasons for this. First, estimating joint distributions for N processes, where $N > 2$ is usually very difficult, as the number of parameters explodes rapidly, and numerical convergence is hard to attain. Second, multivariate distributions are not available when the marginals are drawn from different univariate distributions. Third, fixing the type of joint distribution constrains the dependency structure across the variates, making calibration difficult. Hence, we use copula functions to "stitch" various marginal distributions together.

Different copulas result in different joint distributions. In order to compare the statistical fitting of joint distributions associated with different copulas, we develop a metric. Before presenting the metric, we discuss three features of the dependency relationship we wish to capture in the joint distribution. These correlation properties are as follows:

- *Correlation levels*: We wish to ensure that our copula permits the empirically observed correlation levels in conjunction with the other moments. This depends on the copula and the attributes of the data.
- *Correlation asymmetry*: Hazard rate correlations are level dependent, and are higher when PD levels are high. This means that correlations are higher when PD levels jump up, rather than down.
- *Tail-dependence*: It is important that we capture the correct degree of tail dependence in the data, i.e. the extent to which extreme values drive correlations.

Following Ang and Chen [2002], and Longin and Solnik [2001], we can present the correlations for different rating classes in a correlation diagram (see Figure 6). This plot is created as follows. Using all issuers i, we compute the total hazard rate (THR) at each point in time (t), i.e. $\text{THR}_t = \sum_{i=1}^{N} \lambda_i(t), \ \forall t$. We normalize the THR by subtracting the mean from each observation and dividing by the standard deviation. We then segment our data set based on exceedance levels (ξ) determined by the normalized THR. Our exceedance levels are drawn from the set $\xi = \{-1.5, -1.0, -0.5, 0, 0, 0.5, 1.0, 1.5\}$. Ang and Chen [2002] developed a version of this for bivariate processes. Our procedure here is a modification for the multivariate case. For example, the exceedance correlation at level $-\xi$ is determined by extracting the time series of PDs for which the normalized THR is less than $-\xi$, and computing the correlation matrix therefrom. We then find the average value of all entries in this correlation matrix, to obtain a summary number for the correlation in the rating class at the given exceedance level. Figure 6 presents the results for our data set, where i indexes the average of hazard rates for the rating class.

The three features of the dependency relationship are immediately evident from the graph. First, the height of the exceedance line indicates the level of correlation, and we can see that high grade debt has greater correlation. Second, there is clear evidence of correlation asymmetry, as correlation levels are much higher on the right side of the graph, i.e. when hazard rate changes are positive. Third, the amount of tail-dependence is inferred from the slope of the correlation line as exceedance level increases. The flatter the line the greater the amount of tail-dependence. As absolute exceedance levels increase, we can see that correlation levels drop. However, they fall less slowly if there is greater tail dependence. We can see that lower grade debt appears to have more tail-dependence than higher grade debt.

Having reviewed literature on copulas in this section, we will present an efficient numerical method of simulating correlated default processes in the next section. We will demonstrate through Monte Carlo simulation how to use copulas to obtain a joint default process with desired correlation structure.

4. Determining the Joint Default Process

4.1. **Overview.** Determining the best model for correlated default comprises three steps:

(1) *Estimation of the Model*: Estimation is done in two steps. First, we estimate the stochastic processes for the mean default rates of each rating class. Second, we estimate the default process for each issuer within each rating class.
(2) *Simulation of the Model*: This is undertaken in two stages. In the first step we generate samples from the joint distribution of the mean default intensity for each rating class. In the second step we draw joint samples *within* each rating class conditional on the previously generated mean hazard rates. We choose the best distribution of PDs for each issuer, and then stitch these together into a joint distribution using copulas.
(3) *Model Evaluation*: This step consists of evaluating the different candidate joint distributions for best fit. We develop a new metric to determine the best statistical model. Our metric derives from the objective of best matching the correlation plot discussed in the previous section with the correlation plot emanating from the Monte Carlo implementation. Our goal in this step is to examine if the simulation achieves the three features of interest regarding correlations, i.e. level, asymmetry and tail-dependence.

The following subsections discuss the estimation step. We also delve into the technical details of the implementation. Subsequent sections will cover the simulation and evaluation phases.

4.2. **The estimation phase.** Estimation is undertaken in a two-stage manner: (a) the estimation of the stochastic process for the mean hazard rate of each rating class, and (b) the estimation of the best candidate distribution for the stochastic process of each individual issuer's hazard rate. We chose two processes for step (a). Both choices were made to inject fat-tails into the hazard rate distribution. One, we used a normal-jump model. Two, we used a regime-switching model.

4.2.1. *Estimation of the mean of each rating class using a jump model.* For our six rating classes (indexed by k), and all issuers (indexed by j), we compute hazard rates from the probabilities of default ($P_{kj}, j = 1..N_k, k = 1..M$) in our data set. The hazard rates are computed as: $\lambda_{kj} = -\ln(1 - P_{kj}) \geq 0$. We denote M as the total number of rating classes and N_k as the total number of issuers within the rating class for which data is available.

Let $\lambda_k(t)$ be the average hazard rate across issuers within rating class k. Therefore, $\lambda_k(t) = \frac{1}{N_k}\sum_{j=1}^{N_k} \lambda_{kj}(t), \;\; \forall t$. We assume that $\lambda_k(t)$ follows the stochastic process:

$$
\begin{aligned}
\Delta\lambda_k(t) &= \kappa_k[\theta_k - \lambda_k(t)]\Delta t + x_k(t)\sqrt{\lambda_k(t)\Delta t} \\
x_k(t) &= \epsilon_k(t) + J_k(t)L_k(q_k, t) \\
\epsilon_k(t) &\sim N[0, \sigma_k^2] \\
J_k(t) &\sim U[a_k, b_k] \\
L_k(q_k, t) &= \begin{cases} 1 & \text{w/prob} \quad q_k \\ 0 & \text{w/prob} \quad 1 - q_k \end{cases}
\end{aligned}
\tag{10}
$$

We assume that the jump size J follows a uniform distribution over $[a, b]$[7]. This regression accommodates a mean-reverting version of the stochastic process, and κ_k calibrates the persistence of the process. This process is estimated for each rating class k, and the parameters are used for subsequent simulation. Copulas are used to obtain the joint distribution of correlated default. The correlation matrix used in the copula comes from the residuals $x_k(t)$ computed in the regression above.

We use maximum-likelihood estimation to obtain the parameters of this process. Since there is a mixture of a normal and a jump component, we can decompose the conditional density function for the residual term $x_k(t)$ into the following:

$$
f[x_k(t)] = q_k f[x_k(t)|L_k = 1] + (1 - q_k) f[x_k(t)|L_k = 0] \tag{11}
$$

[7] This specification does permit the hazard rate to populate negative values. While this is not an issue during the estimation phase, the simulation phase is adjusted to truncate the shock if the negative support is accessed. However, we remark that this occurs in very rare cases, since $\lambda_k(t)$ is the average across all issuers within the rating class, and the averaging drives the probability of negative hazard rates to minuscule levels.

The latter density, $f[x_k(t)|L_k = 0]$ is conditionally normal. The marginal distribution of $x_k(t)$ conditional on $L_k = 1$ can be derived as follows:

$$f[x_k(t)|L_k = 1] = \int_a^b \frac{1}{(b-a)\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_k(t)-J)^2}{2\sigma^2}\right\} dJ \tag{12}$$

Let $z = \frac{J - x_k(t)}{\sigma}$. Then

$$\begin{aligned} f[x_k(t)|L_k = 1] &= \int_{\frac{a-x_k(t)}{\sigma}}^{\frac{b-x_k(t)}{\sigma}} \frac{1}{b-a} \exp\{-\frac{z^2}{2}\} \frac{1}{\sqrt{2\pi}} dz \\ &= \frac{1}{b-a}\left\{\Phi\left[\frac{b-x_k(t)}{\sigma}\right] - \Phi\left[\frac{a-x_k(t)}{\sigma}\right]\right\} \end{aligned}$$

In the case when $a = b$, the process has a constant jump size with probability 1 and the marginal distribution of $x_k(t)$ degenerates to a normal distribution with mean a and standard deviation σ. This can be seen by letting $b = a + c$. We then have

$$\begin{aligned} f[x_k(t)|L_k = 1] &= \lim_{c \to 0} \frac{1}{c}\left\{\Phi\left(\frac{a+c-x_k(t)}{\sigma}\right) - \Phi\left(\frac{a-x_k(t)}{\sigma}\right)\right\} \\ &= \lim_{c \to 0} \frac{1}{\sigma}\phi\left(\frac{a+c-x_k(t)}{\sigma}\right) \\ &= \frac{1}{\sigma}\phi\left(\frac{x_k(t)-a}{\sigma}\right) \end{aligned}$$

where $\phi(.)$ and $\Phi(.)$ are the standard normal density and distribution functions respectively. Since jumps are permitted to be of any sign, we may expect that $b \geq 0$ and $a \leq 0$. Also, if $|b| > |a|$, it implies a greater probability of positive jumps. When a is close to b, or when they have the same sign, the process has a constant or one-direction jump respectively.

The estimation results are presented in Table 5. The mean of the hazard rate process (θ_k) increases with rating class k as expected. The variance (σ_k) also increases with declining credit quality. Generally speaking, the probability of a jump (q_k) in the hazard rate is higher for lower quality issuers. We can see that rating classes 2-4 have a constant jump and rating 6 has only positive jumps. While the mean jump for all rating classes appears to be close to zero, that for the poorest rating class is much higher than zero. Hazard rates jump drastically when a firm approaches default.

After obtaining the parameters, we compute residuals for each rating class. The residuals include randomness from both the normal and jump terms. The covariance matrix of these residuals is stored for later use in Monte Carlo simulation.

4.2.2. *Estimating the mean processes in a regime-switching environment.* From Figure 1, we notice that there are periods in which PDs are low, interjected by smaller, sporadic regimes of spikes in the hazard rates. A natural approach to capturing this behavior is to use a regime-switching model[8].

In order to determine regimes, we first computed the average hazard rate $\bar{\lambda} = \sum_{i=1}^{N} \lambda_i$ across all issuers in our database. Within each regime the hazard rate is assumed to follow a square-root volatility model represented in discrete-time as follows:

$$\Delta\bar{\lambda}_r(t) = \kappa_r[\theta_r - \bar{\lambda}_r(t)]\Delta t + \sigma_r\sqrt{\bar{\lambda}_r(t)\Delta t}\,\epsilon(t), \quad r = \{HI, LO\}. \tag{13}$$

The two regimes are indexed by r, which is either HI or LO. κ_r is the rate of mean-reversion. The mean value within the regime is θ_r and σ_r is the volatility parameter.

The probability of switching between regimes comes from a logit model based on a transition matrix: $\begin{bmatrix} p_{LO} & 1-p_{LO} \\ 1-p_{HI} & p_{HI} \end{bmatrix}$, where $p_r = \frac{\exp(\alpha_r)}{1+\exp(\alpha_r)}, r \in \{LO, HI\}$. Estimation is undertaken using maximum-likelihood. We fixed the values of θ_r based on the historical PDs. θ_{LO} was found to be 1.68% and θ_{HI} to be 2.40%[9]. We then estimated the rest of the parameters in the above model. The estimation results are presented in Table 6. In the high PD regime, the higher level of hazard rates is matched by a higher level of volatility.

Our next step is to use the determined regimes to estimate the parameters of the mean process for each rating class, i.e. $\bar{\lambda}_k, k = 1..6$. We fit regime-shifting models to each of the rating classes, using a stochastic process similar to that in equation (13):

$$\Delta\bar{\lambda}_{k,r}(t) = \kappa_{k,r}[\theta_{k,r} - \bar{\lambda}_{k,r}(t)]\Delta t + \sigma_{k,r}\sqrt{\bar{\lambda}_{k,r}(t)\Delta t}\,\epsilon(k,t), \quad r = \{HI, LO\}, k = 1..6. \tag{14}$$

The results of the estimation are presented in Table 7. The parameters are as expected. The mean of the processes is higher in the high regime, as is the volatility. We now proceed to look at the choice of individual issuer marginal distributions.

4.2.3. *Estimation of the stochastic process for each individual issuer.* Assume that the hazard rate for each individual issuer follows a square-root process:

$$\Delta\lambda_{kj}(t) = \kappa_{kj}\{\theta_{kj} + \gamma_{kj}\lambda_k - \lambda_{kj}(t)\} + \sigma_{kj}\sqrt{\lambda_{kj}(t)}\,\varepsilon_{kj}(t) \tag{15}$$

[8] Ang and Chen [2002] found this type of model to be good at capturing the three stated features of asymmetric correlation.

[9] We found that the estimations of the θ_r were very sensitive to the initial values used in the optimization. By fixing these two values, the estimation turned out to be more stable.

We normalize the residuals by dividing both sides by $\sqrt{\lambda_{kj}(t)dt}$. We then use OLS to estimate parameters, which are stored for use in simulations. The resulting residuals are used to form the covariance matrix for each class.

The long-term mean is defined as the sum of a constant number θ_k, and the group mean $\lambda_k(t)$ multiplied by a constant, γ_{kj}. This indexes the mean hazard rate for each issuer to the current value of the mean for the rating class. Consequently, there are four parameters for each regression, κ_{kj}, θ_{kj}, γ_{kj} and σ_{kj}. Notice that κ_{kj} should be positive, θ_{kj} and γ_{kj} can be positive or negative as long as they are not both negative at the same time. For the 620 issuers, none of the issuers have negative θ_{kj} and γ_{kj} at the same time, although 20 issuers have negative κ_{kj}. We therefore deleted the 20 issuers with negative κ from our dataset for further analysis [10].

4.2.4. *Estimation of the marginal distributions.* We fitted the residuals from the previous section to the normal, student's t and skewed double-exponential distributions. While tail-dependence is low for the normal distribution, it is present in the student's t and skewed double-exponential distributions. We used prepackaged functions in Matlab for maximum-likelihood estimation for the normal and student's t distributions. The likelihood function for the skewed double exponential distribution is as follows. Assume that the residual ε_{kj} has a normal distribution with mean γV and variance V. Further, the variance V is assumed to have an exponential distribution with the following density function:

$$pdf(V) = \frac{1}{V_0}\exp(-\frac{V}{V_0}) \tag{16}$$

Then it can be shown that ε_{kj} has a skewed double exponential distribution with the following density function:

$$pdf(\varepsilon) = \frac{\lambda}{V_0}\exp\left(-\frac{|\varepsilon_{kj}|}{\lambda} - \gamma\varepsilon_{kj}\right), \tag{17}$$

where

$$\lambda = \sqrt{\frac{V_0}{2+\gamma^2 V_0}}$$

[10] The estimation of equation (15) relies on the assumption that mean around which the individual hazard rate oscillates depends on the initial rating class of the issuer, even though this may change over time. Extending the model to map default probabilities to rating classes is a non-trivial problem, and would complicate the estimation exercise here beyond the scope of this paper. Indeed, this problem in isolation from other estimation issues is complicated enough to warrant separate treatment and has been addressed in a paper by Das, Fan and Geng [2002].

See the Appendix for a derivation of these results. The log-likelihood function L is given as follows:

$$L = -\frac{n}{2}\left\{\log(V_0) + \log(2+\gamma^2 V_0)\right\} - \sum_{i=1}^{n} |\varepsilon_i| \sqrt{\frac{2+\gamma^2 V_0}{V_0}} + \gamma \sum_{i=1}^{n} \epsilon_i.$$

4.2.5. *Goodness of fit of marginal distributions.* Four criteria are used in selecting the best marginal distribution for each issuer. These are as follows:

- The Kolmogorov distance: this is defined as the supremum over the absolute differences between two cumulative density functions, the empirical one, $F_{emp}(x)$ and the estimated, fitted one, $F_{est}(x)$.
- The Anderson & Darling Statistic: which is given by

$$AD = max_{x\in R} \frac{|F_{emp}(x) - F_{est}(x)|}{\sqrt{F_{est}(x)[1-F_{est}(x)]}}$$

  The AD statistic puts more weight on the tails compared to Kolmogorov statistic.
- The L^1 distance: equal to the average of the absolute differences between the empirical and statistical distributions.
- The L^2 distance: which is the root mean-squared difference between the two distributions.

We used these criteria to choose between the normal, student's t and skewed double exponential distributions for each issuer. For each criterion, we counted the number of times each distribution provided the best fit. The results are reported in Table 8, which provides interesting features. Based on the Kolmogorov, Anderson-Darling and L^2 statistics, the skewed double-exponential distribution is the most likely to fit the marginals best. However, the L^1 statistic finds that the normal distribution fits most issuers better. The better fit of the skewed double-exponential as marginal distribution comes from its better capturing the fatness in the tails of the distribution.

5. Simulating Correlated Defaults & Model Comparisons

5.1. **Overview.** In this section, we discuss the simulation procedure. This step uses the estimated parameters from the previous section to generate correlated samples of hazard rates. The goal is for the simulation approach to deliver a model with the three properties described in the previous section. Consistent with this objective, the asymmetric correlation plot from the simulated data should appear similar to that in

Figure 6. Since there is a high degree of correlation between the six rating classes, we should expect to see that simulated series also appear to be correlated. This is seen in Figure 7 where we plot the times series of $\lambda_k,\ k = 1..6$ for a random simulation of the sample path of hazard rates.

We implemented the simulation model with an additional constraint, whereby we ensure that the hazard rates $(\lambda_k(t))$ are monotonically increasing in k. This is to prevent the average PD for a rating class from being lower than the average PD for the next better rating category. The average hazard rates $[\lambda_k(t)]$ should be such that for all time periods t, it must be that if $i < j$, then $\lambda_i(t) < \lambda_j(t)$. This check was instituted during the simulation as follows. During the Monte Carlo step, if $\lambda_i(t) > \lambda_j(t)$ when $i < j$, then we set $\lambda_j(t) = \lambda_i(t)$.

Finally, the simulation is driven in two steps. We use a copula system with jumps or regime-switching to generate the rating class levels of hazard rates. Next we exploit the regression in equation (15) to obtain the individual issuer hazard rates.

5.2. **Illustrative Monte Carlo Experiment.** As a simple illustration that our approach arrives at a correlation plot fairly similar to that seen in the data, we ran a naive Monte Carlo experiment. In this exercise, we assumed that all error terms were normal, except under some conditions, when we assumed the student's t distribution. We ran the Monte Carlo model 25 times and computed the average exceedance correlations for all rating classes. The plot of the simulated exceedance correlations is provided in Figure 8. From Figure 6 we see that there is appreciable asymmetry in the correlations. In order to generate this, we need to have higher correlations when hazard rates are high. To achieve this, the simulation uses the student's t distribution when the average level of hazard rates in the previous period in the simulation is above the empirical average of hazard rates. The chosen student's t distribution is fat-tailed with degrees of freedom equal to 6. These features provided the results in Figure 8. The similarity between Figures 6 and 8 shows that the two-stage Monte Carlo model is able to achieve the three correlation properties of interest.

5.3. **Determining Goodness of Fit.** Our objective is to correctly capture the dependency structure amongst the hazard rates, as depicted in the asymmetric correlation plots. We develop a measure to compare different specifications of the joint distribution. Our metric comprises the average squared point-wise difference between the empirical exceedance correlation plot and the simulated one. This is a natural distance metric. The points in each plot that are used are for the combination of rating class (from 1 to 6) and exceedance levels (from −1.5 to +1.5). Hence there are a total of 48 points in each plot which are used for computing the metric. Define the points in the empirical plot as $h_{k,x}$, where k indexes the rating class and x indexes the exceedance levels. The corresponding points in the simulated plot are denoted $h'_{k,x}$. The metric q (a RMSE statistic) is as follows:

$$q = \sqrt{\frac{1}{48} \sum_{k=1}^{6} \sum_{x=-1.5}^{1.5} \left(h_{k,x} - h'_{k,x}\right)^2} \tag{18}$$

A smaller value of q implies a better fit of the joint dependency relationship. In the following section, we use the simulation approach and this metric to compare various models of correlated default.

5.4. **Monte Carlo procedure.** Parameter estimation leads to the choice of appropriate process for the mean hazard rate, marginal distributions and copulas. Based on these choices, the step-wise procedure to generate hazard rates with the correct dependency structure is as follows:

- Step 1: Compute the rank correlation for each rating group from the residuals obtained from equation (15) for each individual issuer. As discussed earlier, rank correlation does not suffer from many of the deficiencies of Pearson's correlation, and is more appropriate for this purpose. We store this for use in each step of the Monte Carlo procedure.
- Step 2: Begin the simulation at the mean values of the hazard rates for each rating group. From that point on, simulate the mean for each rating group from the estimated parameters of the normal-jump model or the regime-switching model (as the case may be) for rating-level hazard rates. We generate correlated innovations based on the correlation matrix of residuals from the estimation in equation (10) or (13).
- Step 3: Simulate residuals for each issuer using the rank correlation from Step 1 and the chosen copula based on earlier estimation.
- Step 4: Calculate the PDs for each issuer using the group mean $[\lambda_k(t)]$ from Step 2, the statistical relationship of the individual issuer to the average rating class (as in equation 15), and residuals from Step 3.
- Step 5: Repeat Step 2 to Step 4 for T periods and calculate the asymmetric correlation plot.
- Step 6: The final plot is the average of 100 individual correlation plots. Hence, we run steps 1-5 a hundred times, and average the plots from each iteration.

We ran our Monte Carlo procedure for various choices of the marginal distribution for the individual issuers. Asymmetric correlation data (and the metric q) are computed for the following seven cases:

(1) All marginal distributions were chosen to be normal;
(2) All marginal distributions are student's t;
(3) All marginals are from the skewed double-exponential family;
(4) For each issuer, the best distribution is chosen based on the Kolmogorov criterion;
(5) For each issuer, the best distribution is chosen based on the Anderson & Darling statistic.
(6) Each marginal is chosen based on the best L^1 distance.
(7) Each marginal is chosen based on the best L^2 distance.

A sample correlation plot from the regime-switching model is presented in Figure 9. The results are presented in Table 9, which is best read with the results in Table 8. The inferences from the fitting exercise are as follows.

We note that the joint dependency system across all issuer hazard rates depends on three aspects, (a) the marginal distributions used for individual issuers, (b) the marginal distributions used for each rating class, and (c) the copula used to implement the correlations between rating classes. Hence, it is not necessary that optimizing the choice of individual issuer marginal distributions will achieve the best dependency relationship, since the copula chosen must also be compatible. In this paper, we focus on four copulas (normal, Gumbel, Clayton and student's t), and allow the rating class hazard rate stochastic process to be either a jump-normal model or a regime-shifting model. Hence, the results depend on the interaction of the marginal distributions with the chosen copula.

Tables 9 and 10 provide a comparison of the jump and the regime-switching models across copulas. We see that the values of the q metric are consistently smaller for the regime-switching model using all test statistics. This implies that the latter model provides a better representation of the stochastic properties amongst the issuer hazard rates. In particular, it fits the asymmetric correlations better[11]. We conclude that

[11]This result is consistent with that of Ang and Chen [2002], who undertake a similar exercise with equity returns.

injecting tail-fatness through a regime-switching model performs better than using a jump-normal model[12].

We can use Table 9 to compare the fit provided by the four copulas. We find that when the jump-normal model is used, the student's t copula performs best, no matter which marginal distribution is used. However, when the regime-switching model is used, sometimes the normal copula works best and at other times, the Clayton copula is better. The normal copula works best when the marginals are all normal or student's t, and when the marginals are chosen using the Anderson-Darling criterion. The Clayton copula is best when the marginals are all skewed double exponential or when the marginal criteria are Kolmogorov or L^2. The student's t copula is best when the L^1 criterion is chosen. We can conclude that while the regime-switching model dominates the jump-normal one when choosing the stochastic process for rating class hazard rates, the choice of copula depends critically on the choice of marginal distributions for the individual issuer hazard rates.

A comparison of the correlations from the skewed double exponential distribution with those from the raw data reveals that the model overestimates the up-tail dependencies a lot. The Clayton copula decreases the up-tail dependencies and increases the low-tail dependencies. As a result, it corrects some of the over-estimation from the double-exponential marginals, resulting in better fitting. Hence, the Clayton copula seems to combine well with skewed double-exponential marginals. On the other hand, the q-metric favors the normal copula if the criterion chooses normal marginals (as with the normal, Student's t and L^1 criteria), since less balancing of tail-dependencies is required.

5.5. **Tail distributions for different copulas.** As discussed earlier, the degree of tail-dependence varies with the choice of copula. It is ultimately an empirical question as to whether the parameterized joint distribution does result in differing tail risk in credit portfolios. To explore this question, we simulated defaults using a portfolio comprising all the 620 issuers in this study, under the regime-switching model. Remember that we fitted 4 copulas, and 7 different choices of marginal distributions, resulting in 28 different models, each with its attendant parameter set. To compare copulas, we fix the marginal distribution, and then vary the copula.

[12]The fact that the student's t marginal distribution works best for the jump models may indicate that jump-model does not capture the dependencies as well as the regime-switching model does, since the student's t distribution is good at injecting fat tails into the conditional distribution of the marginals.

As an illustration, we present Figure 10. This plots the tail loss distributions from the four copulas when all marginals are assumed to be normal. The line most to the right (the Gumbel copula) has the most tail dependence. There is roughly a 90% chance that the number of defaults will be less than 75, i.e. a 10% probability that the number of defaults will exceed 75. For the same level of 75 defaults, the leftmost line (from the normal copula), there is only a 7% probability of the number of losses being greater than 75. By examining all four panels of the figure, we see that the ranking of copulas by tail dependence is unaffected by the choice of marginal distribution, i.e. the normal copula is the leftmost plot, followed by the student's t, Clayton, and Gumbel copulas. It is intuitive that the Clayton and Gumbel copulas have the most tail dependence, since they are popular forms of "extreme-value" copulas.

In Figure 11 we plot the tail loss distributions for two models, the best fitting one and the worst. The best fit model combines the Clayton copula, and marginal distributions based on the Kolmogorov criterion. The worst fit copula combines the student's t copula with student's t marginals. A comparison of the two models shows that the worst fitted copula in fact grossly overstates the extent of tail loss. Therefore, while there is tail-dependence in the data, careful choice of copula and marginals is needed to avoid either under or over-estimation of tail-dependence.

6. Conclusions and Extensions

This paper develops a criterion-based methodology to assess alternative specifications of the joint distribution of default risk across hundreds of issuers. We believe this is the first paper to empirically fit joint default distributions using copulas in the U.S. corporate market. The study is based on a data set of default probabilities supplied by Moody's Risk Management Services. We undertake an empirical examination of the joint stochastic process of default risk over the period 1987-2000. Using copula functions, we separate the estimation of the marginal distributions from the estimation of the joint distribution. Using a two-step Monte Carlo model, we determine the appropriate choice of multivariate distribution based on a new metric for the assessment of joint distributions.

We explored fifty-six different specifications for the joint distribution of hazard rates. Our methodology uses two alternative specifications (jumps and regimes) for the means of default-rates in rating classes. We consider three marginal distributions for individual issuer hazard rates, combined using four different copulas. Other than the myriad specifications, there are many useful features of the analysis for modelers of portfolio credit risk. First, we developed a simple metric to measure best fit of the joint default process. This metric accounts for different aspects of default correlation, namely level, asymmetry and tail-dependence or extreme behavior. Second, the simulation model, based on estimating the joint system of 620 issuers, is able to replicate the empirical joint distribution of default. Third, a comparison of the jump model and the regime-switching model shows that the latter provides a better representation of the properties of correlated default. Fourth, the skewed double exponential distribution is a suitable choice for the marginal distribution of each issuer hazard rate process, and combines well with the Clayton copula in the joint dependency relationship amongst issuers. Our simulation approach is fast and robust, allowing for rapid generation of scenarios to assess risk in credit portfolios. Finally, The results show that it is important to correctly capture the interdependency of marginal distributions and copula to achieve the best joint distribution depicting correlated default. Thus, this paper delivers the empirical counterpart to the body of theoretical papers advocating the usage of copulas in modeling correlated default.

Appendix A. The Skewed Double Exponential Distribution

Assume that a random variable X has a normal distribution with mean $\mu + \gamma V$, and variance V, where V is has an exponential distribution with the following density function:

$$pdf(V) = \frac{1}{V_0}\exp(-\frac{V}{V_0}), \tag{19}$$

then X is has a skewed double exponential distribution.

A.1. **Density function.**

The density function is derived as follows:

$$\begin{aligned}
pdf(x) &= \int_0^\infty \frac{1}{\sqrt{2\pi V}}\exp\left\{-\frac{(x-\mu-\gamma V)^2}{2V}\right\}\frac{1}{V_0}\exp\left(-\frac{V}{V_0}\right)dV \\
&= \frac{1}{\sqrt{2\pi}V_0}\exp\{\gamma(x-\mu)\}\int_0^\infty \frac{1}{\sqrt{V}}\exp\left\{-\frac{(x-\mu)^2}{2V} - \frac{\gamma^2 V_0 + 2}{2V_0}V\right\}dV \\
&= \frac{1}{\sqrt{2\pi}V_0}\exp\{\gamma(x-\mu)\}\sqrt{\frac{2\pi V_0}{2+\gamma^2 V_0}}\exp\left\{-|x-\mu|\sqrt{\frac{2+\gamma^2 V_0}{V_0}}\right\} \\
&= \frac{1}{V_0}\sqrt{\frac{V_0}{2+\gamma^2 V_0}}\exp\left\{-|x-\mu|\sqrt{\frac{2+\gamma^2 V_0}{V_0}} + \gamma(x-\mu)\right\} \\
&= \frac{\lambda}{V_0}\exp\left\{-\frac{|x-\mu|}{\lambda} + \gamma(x-\mu)\right\}
\end{aligned}$$

A.2. **Maximizing the log-likelihood function.** This subsection deals with the technical details of the maximization of log-likelihoods for the skewed double exponential model. The likelihood function is:

$$L = -\frac{n}{2}\left\{\log(V_0) + \log(2+\gamma^2 V_0)\right\} - \sum_{i=1}^{n} |x_i - \mu| \sqrt{\frac{2+\gamma^2 V_0}{V_0}} + \gamma \sum_{i=1}^{n} (x_i - \mu)$$

where m is the number of observations where x_i is greater than μ, l is the number of observations otherwise, and $n = m + l$.

The first order condition can be derived as follows:

$$\begin{aligned}
\frac{dL}{d\mu} &= \sum_{i=1}^{n} \frac{d|x_i-\mu|}{d\mu}\left(\sqrt{\frac{2+\gamma^2 V_0}{V_0}}\right) + \gamma \sum_{i=1}^{n}(-1) \\
&= -n\gamma - \sqrt{\frac{2+\gamma^2 V_0}{V_0}} \sum_{i=1}^{n} \frac{d|x_i-\mu|}{d\mu} = 0 \\
\frac{dL}{d\gamma} &= -\frac{n}{2}\left(\frac{2\gamma V_0}{2+\gamma^2 V_0}\right) - \sum_{i=1}^{n} |x_i-\mu| \left(\frac{\gamma}{\sqrt{\frac{2+\gamma^2 V_0}{V_0}}}\right) + \sum_{i=1}^{n}(x_i-\mu) \\
&= -\frac{n\gamma V_0}{2+\gamma^2 V_0} - \sum_{i=1}^{n} \gamma |x_i-\mu| \sqrt{\frac{V_0}{2+\gamma^2 V_0}} + \sum_{i=1}^{n}(x_i-\mu) = 0 \\
\frac{dL}{dV_0} &= -\frac{n}{2}\left(\frac{1}{V_0} + \frac{\gamma^2}{2+\gamma^2 V_0}\right) - \sum_{i=1}^{n} |x_i-\mu| \left(\frac{\frac{\gamma^2 V_0 - 2 - \gamma^2 V_0}{V_0^2}}{2\sqrt{\frac{2+\gamma^2 V_0}{V_0}}}\right) \\
&= -\frac{n(1+\gamma^2 V_0)}{V_0(2+\gamma^2 V_0)} + \sum_{i=1}^{n} \frac{|x_i-\mu|}{V_0\sqrt{V_0(2+\gamma^2 V_0)}} = 0
\end{aligned}$$

Solving the above first order conditions gives:

$$\begin{aligned}
\mu &= \frac{m^2 \sum_{x\in A1} x_i + l^2 \sum_{x \in A2} x_i}{nml} \\
V_0 &= \frac{2ml(\frac{1}{n}\sum_{i=1}^{n} x_i - \mu)^2}{(l-m)^2} \\
\gamma &= \frac{(l-m)^2}{2ml(\frac{1}{n}\sum_{i=1}^{n} x_i - \mu)}
\end{aligned}$$

where $A1 = \{x : x_i - \mu >= 0\}$ and $A2 = \{x : x_i - \mu < 0\}$.

Given a dataset, we may not be able to find a value of μ to satisfy the above equations. However, by choosing μ to be as close as given by the first order conditions, we can fit the data well with the skewed double exponential distribution.

References

[1972] Abramowitz, M., and I. A. Stegun (1972). "The Handbook of Mathematical Functions," Dover, New York.

[2002] Ang, A., and J. Chen (2002). "Asymmetric Correlation of Equity Portfolios," *Journal of Financial Economics*, v63(3), 443-494.

[2000] Bouye, E., V. Durrelman, A. Nikeghbali, G. Riboulet, and T. Roncalli (2000). "Copulas for Finance: A Reading Guide and some Applications," working paper, Credit Lyonnais, Paris.

[1978] Clayton, D.G. (1978). "A Model for Association in Bivariate Life Tables and its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence," *Biometrika*, v65, 141-151.

[2000] Crouhy, Michel., Dan Galai, and Robert Mark (2000). "A Comparative Analysis of Current Credit Risk Models," *Journal of Banking and Finance*, v24, 59-117.

[2002] Das, Sanjiv., Rong Fan, and Gary Geng (2002). "Bayesian Migration in Credit Ratings based on Probabilities of Default," *Journal of Fixed Income*, v12(3), 17-23

[2001] Das, Sanjiv., Laurence Freed, Gary Geng, and Nikunj Kapadia (2001). "Correlated Default Processes," working paper, Santa Clara University and Gifford Fong Associates.

[2001] Das, S., G. Fong., and G. Geng, (2001), "The Impact of Correlated Default Risk on Credit Portfolios", *Journal of Fixed Income*, December, v11(3), 9-19.

[2000] Das, Sanjiv., and R. Sundaram (2000). "A Discrete-Time Approach to Arbitrage-Free Pricing of Credit Derivatives," *Management Science*, v46(1), 46-62.

[1996] Das, Sanjiv., and P. Tufano (1996). "Pricing Credit Sensitive Debt when Interest Rates, Credit Ratings and Credit Spreads are Stochastic," *The Journal of Financial Engineering*, v5(2), 161-198.

[2000] Das, Sanjiv., and Raman Uppal (2000). "Systemic Risk and International Portfolio Choice," working paper, London Business School.

[1999] Davis, Mark., and Violet Lo (1999). " Modeling Default Correlation in Bond Portfolios," working paper, Vienna University of Technology.

[1999a] Davis, M., and L. Violet (1999): "Infectious Default", Working Paper, Imperial College, London.

[1999b] Davis, M., and L. Violet (1999): "Modeling Default Correlation in Bond Portfolios", In Carol Alexander, editor,*ICBI Report on Credit Risk*.

[1999] Dowd, Kevin (1999). "The Extreme Value Approach to VaR: An Introduction," *Financial Engineering News*, Issue 11.

[2001] Duffie, Darrel, J., and Nicole Garleanu (2001). "Risk and Valuation of Collateralized Debt Obligations," *Financial Analysts Journal*, v57(1), (January-February), 41-59.

[1999] Duffie, Darrel, J., and Kenneth J. Singleton (1999). " Simulating Correlated Defaults," working paper, Stanford University, Graduate School of Business.

[2000] Durrleman, V., A. Nikeghbali and T. Roncalli, (2000), "Which Copula is the Right One?", Working paper, Groupe de Recherche Operationnelle, Credit Lyonnais, France.

[1997] Embrechts, Paul., Claudia Kluppelberg, and T. Mikosch (1997). "Modeling Extremal Events for Insurance and Finance," Springer-Verlag, Berlin.

[1999a] Embrechts, Paul., Alexander McNeil, and Daniel Straumann (1999). "Correlation and Dependence in Risk Management: Properties and Pitfalls," working paper, University of Zurich.

[1999b] Embrechts, Paul., Alexander McNeil, and Daniel Straumann (1999). "Correlation: Pitfalls and Alternatives," working paper, Department Mathematik, ETH Zentrum, Zurich.

[2001] Embrechts, Paul., Filip Lindskog, and Alexander McNeil (2001). "Modeling Dependence with Copulas and Applications to Risk Management," working paper CH-8092, Dept. of Mathematics, ETZH Zurich.

[1998] Frees, Edward, W., and Emiliano A. Valdez (1998). "Understanding Relationships using Copulas," *North American Actuarial Journal*, v2(1), 1-25.

[2001] Frey, Rudiger., and Alexander J. McNeil (2001). "Modeling Dependent Defaults," working paper, University of Zurich.

[1987] Genest, C. (1987). "Frank's Family of Bivariate Distributions," Biometrika, v74, 549-555.

[1993] Genest, C., and L. Rivest (1993). "Statistical Inference Procedures for Bivariate Archimedean Copulas,' *Journal of the American Statistical Association*, v88, 1034-1043.

[1977] Geske, R. (1977): "The Valuation of Corporate Liabilities as Compound Options", *Journal of Financial and Quantitative Analysis*, 12(4), 541-552.

[2000] Gersbach, H., and A. Lipponer, (2000), "The Correlation Effect", Working paper, University of Heidelberg.

[1960] Gumbel, E.J. (1960). "Distributions des valeurs extremes en plusiers dimensions," *Publ. Inst. Statist. Univ. Paris*, v9, 171-173.

[1995] Jarrow, R.A., and S.M. Turnbull (1995): "Pricing Derivatives on Financial Securities Subject to Credit Risk", *Journal of Finance*, 50(1), 53-85.

[1997] Jarrow, R.A., D. Lando, and S.M. Turnbull (1997): "A Markov Model for the Term Structure of Credit Spreads", *Review of Financial Studies*,v10, 481-523.

[2000] Jarrow, R.A., and F. Yu (2000): "Counterparty Risk and the Pricing of Defaultable Securities", Working Paper, Johnson GSM, Cornell University.

[2000] Kijima, M. (2000): "Credit Events and the Valuation of Credit Derivatives of Basket Type", *Review of Derivatives Research*, 4, 55-79.

[1993] Lee, A.J. (1993). "Generating Random Binary Deviates having fixed Marginal Distributions and Specified Degree of Association," *The American Statistician*, v47, 209-215.

[1994] Leland, H.E. (1994): "Corporate Debt Value, Bond Covenants and Optimal Capital Structure", *Journal of Finance*,49(4), 1213-1252.

[1999] Li, David, X. (1999). "On Default Correlation: A Copula Function Approach," working paper 99-07, The RiskMetrics Group, New York.

[2000] Lindskog, Filip (2000). "Modeling Dependence with Copulas and Applications to Risk Management," working paper, Risklab, ETH Zurich.

[2001] Longin, F., and B. Solnik (2001). "Extreme Correlation of International Equity Markets," *Journal of Finance*, v56, 649-676.

[1995] Longstaff, F.A., and E.S. Schwartz (1995):"A Simple Approach to Valuing Risky Fixed and Floating Rate Debt", *Journal of Finance*, 50(3), 789 - 819.

[1999] Madan, D., and H. Unal (1999): "Pricing the Risks of Default", *Review of Derivatives Research*, v2(2/3), 121-160.

[1974] Merton, R. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, v29, 449-470.

[1999] Nelsen, Roger, B (1999). "An Introduction to Copulas," Springer-Verlag, New York.

[1959] Sklar, A. (1959). "Functions de repartition a n dimensions et leurs marges," *Publ. Inst. Statist. Univ. Paris*, v8, 229-231.

[1973] Sklar, A (1973). "Random Variables, Joint Distributions, and Copulas," *Kybernetica*, v9, 449-460.

[2001] Schonbucher, P., and D. Schubert (2001). "Copula-Dependent Default Risk in Intensity Models," working paper, Dept. of Statistics, Bonn University.

[2000] Wang, Shaun, S., (2000). "Aggregation of Correlated Risk Portfolios: Models and Algorithms," working paper, CAS.

[2001] Wise, M., and V. Bhansali (2001). "Portfolio Allocation to Corporate Bonds with Correlated Defaults," working paper No. CALT-68-2365, California Institute of Technology.

[2001] Zhou, C., (2001). "An Analysis of Default Correlations and Multiple Default," *Review of Financial Studies*, v14, 555-576.

TABLE 1. **Rating classes**

The database comprises 7 rating categories, in descending order of credit quality. Our data set does not contain data from firms in the financial sector.

Rating Classes

1 Aaa/Aa: Bonds rated Aaa/Aa carry the smallest degree of investment risk and are generally known as high-grade bonds.

2 A: Bonds rated A possess many favorable investment attributes and are to be considered as upper-medium-grade obligations

3 Baa: Bonds rated Baa are considered as medium-grade obligations. Such bonds lack outstanding investment characteristics and in fact have speculative characteristics as well.

4 Ba: Bonds rated Ba are judged to have speculative elements. Often the protection of interest and principal payments may be very moderate and thereby not well safeguarded during both good and bad times over the future

5 B: Bonds with B rating generally lack characteristics of the desirable investment. Assurance of interest and principal payments over any long period of time may be small

6 Caa/Ca/C: These are the bonds with high degree of speculation. Such issues may either be in default or there may be present elements of danger with respect to principal or interest payments.

7 Not Rated.

Table 2. **Descriptive statistics for probabilities of default**

This table presents the moments of the time series of average PDs (probabilities of default) for each rating class. Each time series represents a diversified portfolio of issuers within each rating class, equally weighted. Statistics are presented for PDs in levels (Panel A) and changes (Panel B). Panel C contains the correlations of PDs in the upper right triangle and the correlations of PD changes in the lower left triangle. The same data is used in DFGK [2001].

Panel A: D-stats for PD levels

Rating Category	Mean	Std Dev.	Skewness	Kurtosis
1	0.0728	0.1051	7.5681	92.8751
2	0.1701	0.4884	13.2578	227.5997
3	0.6113	1.6808	6.2708	49.0043
4	1.4855	2.6415	3.6180	17.9157
5	3.8494	4.6401	1.5985	4.6140
6	6.5389	4.7731	0.6683	2.4210

Panel B: D-stats for PD changes

Rating Category	Mean	Std Dev.	Skewness	Kurtosis
1	0.0012	0.0441	0.1996	353.0732
2	0.0036	0.1468	8.1615	405.5719
3	0.0044	0.7317	1.2879	231.7470
4	0.0066	1.2694	-0.0524	58.8742
5	0.0138	2.3576	-0.1901	18.1022
6	0.0165	3.2075	-0.1150	9.8268

Panel C: Correlations Matrices

Rating	1	2	3	4	5	6
1	1.0000	0.9608	0.7859	0.5367	0.6765	0.2939
2	0.5537	1.0000	0.8540	0.6051	0.7416	0.3120
3	0.2099	0.3452	1.0000	0.7603	0.8394	0.3544
4	0.1766	0.2780	0.2530	1.0000	0.6751	0.3464
5	0.0438	0.1776	0.3494	0.3291	1.0000	0.3730
6	-0.0379	0.0106	0.0331	0.1627	0.0760	1.0000

TABLE 3. **Kendall's τ for probabilities of default**

In this table we report Kendall's τ, which is a measure of dependency between any two time series. It is analogous to measuring the rank correlation. The upper right triangle below reports the measure for PD levels. The lower left triangle reports the correlation for PD changes.

Rank Correlations Matrices

Rating	1	2	3	4	5	6
1	1.0000	0.6508	0.2724	0.2868	0.2987	0.2762
2	0.4254	1.0000	0.4671	0.4569	0.4052	0.1794
3	0.1849	0.2312	1.0000	0.6416	0.4483	0.1080
4	0.1447	0.1393	0.0543	1.0000	0.5382	0.1708
5	0.0378	0.0877	0.0352	0.1245	1.0000	0.2911
6	0.0362	-0.0429	0.0238	-0.0276	0.1032	1.0000

Table 4. **Measures of tail dependence between the different rating classes**

This table presents correlations between hazard rates amongst the rating classes based on different tail percentiles. We examine correlation amongst observations in the bottom 10, 20 and 30th percentiles. We also present the correlations in the top 10, 20 and 30th percentiles, i.e. the cutoffs are the 70th, 80th and 90th percentiles. The values in the table are the correlations when observations from the second rating in the pair of rating classes lies in the designated portion of the tail of the distribution. Results are presented for levels of hazard rates and for changes in hazard rates.

Panel A: Tail dependence in Levels

Rating Pair		Lower tail dependency			Upper tail dependency		
Rating	Rating	10%	20%	30%	70%	80%	90%
1	2	0.4853	0.2879	0.2212	0.7518	0.7386	0.7206
1	3	0.0147	-0.3371	-0.1673	0.5837	0.5795	0.5441
1	4	-0.1471	0.0379	-0.0008	0.6376	0.4545	0.1471
1	5	-0.0882	-0.1439	-0.1314	0.6865	0.6061	0.2059
1	6	-0.2206	-0.1705	-0.0727	0.5298	0.4583	-0.0735
2	3	-0.1912	-0.2197	0.1004	0.5967	0.5985	0.5588
2	4	-0.2353	-0.1326	-0.0857	0.6000	0.4015	0.1324
2	5	-0.0147	-0.2045	-0.2065	0.6718	0.5189	0.0882
2	6	-0.1912	-0.3447	-0.1559	0.4171	0.3826	-0.1765
3	4	0.2500	0.2424	0.2539	0.5559	0.3902	0.1765
3	5	-0.0294	-0.1402	0.0694	0.5576	0.4205	0.2647
3	6	-0.1765	-0.1591	-0.0955	0.3535	0.4583	0.0147
4	5	0.0588	-0.1023	0.0155	0.7257	0.6439	0.4412
4	6	0.1471	-0.0833	-0.0318	0.4922	0.6326	0.4853
5	6	0.2206	0.0455	0.1020	0.4563	0.5152	0.2059

Panel B: Tail dependence in Changes

Rating Pair		Lower tail dependency			Upper tail dependency		
Rating	Rating	10%	20%	30%	70%	80%	90%
1	2	0.5000	0.4394	0.4106	0.4253	0.4015	0.3971
1	3	0.0441	0.0189	0.0367	0.0776	0.0038	0.2206
1	4	0.1029	0.1439	0.0188	0.1429	0.0644	0.1618
1	5	0.1029	-0.0455	-0.0269	0.1069	0.1439	0.2794
1	6	0.0588	-0.0455	-0.0580	-0.2033	-0.1856	0.0735
2	3	0.1618	0.0606	0.0873	0.0204	-0.0606	0.1324
2	4	-0.2647	-0.0076	-0.0580	0.1984	0.0606	0.2647
2	5	-0.0147	0.0189	0.0645	0.0302	0.2045	0.3382
2	6	0.1618	0.0455	0.0482	-0.1559	-0.1667	-0.2500
3	4	0.2941	0.1364	-0.0433	0.0122	-0.0720	-0.0441
3	5	-0.0441	-0.1212	-0.0367	0.2082	-0.0152	0.0294
3	6	0.0882	0.0076	0.1216	-0.0694	-0.1136	-0.3382
4	5	-0.0588	-0.0227	0.1265	0.2555	0.3864	0.3824
4	6	-0.0735	0.1515	0.0971	0.0041	0.1098	-0.1324
5	6	-0.0294	0.1553	0.0171	0.0547	0.0341	-0.0147

TABLE 5. **Estimation of the Average Rating Process**

In this table we report the estimation results of the stochastic process for the mean of the hazard rates of each rating class. Numbers below the estimates are the t-statistics from the estimation, which is undertaken by maximum-likelihood.

Parameter	Rating Class					
	1	2	3	4	5	6
θ	0.0524	0.1336	0.3499	1.1550	2.3812	5.4848
	7.73	16.74	2.75	4.01	1.99	7.12
κ	0.4813	0.0018	0.6680	0.4811	0.4305	2.8350
	2.84	0.01	1.88	1.87	2.01	4.19
σ	0.0487	0.1050	0.2518	0.2633	0.2872	0.5743
	14.17	13.39	12.48	11.44	8.51	5.85
b	0.0839	0.1316	0.1241	0.2070	0.2684	1.1231
	12.39	5.64	1.94	1.58	1.52	5.50
a	-0.0707	-0.0644	0.1241	0.2069	0.2684	0.0771
	-6.33	-3.46	1.96	1.61	2.37	0.21
q	0.1016	0.0777	0.1514	0.0950	0.2461	0.4374
	3.87	3.18	2.94	2.07	2.08	2.07
Log-likelihood	-622.16	-462.95	-205.48	-133.56	-32.82	-35.03

Table 6. **Estimation of the regime-switching model across all issuers**

This table provides estimation results for regimes estimated on the average hazard rate process ($\bar{\lambda} = \sum_{i=1}^{N} \lambda_i$) across all issuers in the data set. We designated periods where $\bar{\lambda} <= 2\%$ as the "low-PD" regime, and periods in which $\bar{\lambda} > 2\%$ as the "high-PD" regime. The two regimes are indexed by r, which is either HI or LO. κ_r is the rate of mean-reversion. The mean value within the regime is θ_r and σ_r is the volatility parameter. The regimes were bifurcated exogenously, as described, and we estimated the values of θ_r by simply averaging the hazard rate within each regime. The probability of switching between regimes comes from a logit model based on a transition matrix: $\begin{bmatrix} p_{LO} & 1-p_{LO} \\ 1-p_{HI} & p_{HI} \end{bmatrix}$, where $p_r = \frac{\exp(\alpha_r)}{1+\exp(\alpha_r)}, r \in \{LO, HI\}$. Estimation is undertaken using maximum-likelihood.

Parameter	Low-PD regime	High-PD regime
θ	1.68%	2.40%
κ	1.5043	1.7746
t-stat	2.4053	2.0558
σ	0.1681	0.2397
t-stat	14.2093	7.9849
α	4.1286	3.3644
t-stat	4.6649	2.5428
Loglik	-195.5948	

Table 7. Estimation of the regime-switching model for the mean of each rating class.

This table presents parameters for the regime-switching model applied to the mean process for each rating class, i.e. $\bar{\lambda}_k, k = 1..6$.

	Rating Class					
Parameter	1	2	3	4	5	6
θ_{LO}	0.0559	0.1625	0.5678	1.3588	3.6367	6.3138
t-stat	0.40	7.06	11.90	14.22	45.24	74.32
θ_{HI}	0.0991	0.3757	0.9818	2.2358	5.1216	7.8295
t-stat	0.36	6.55	4.91	19.29	18.74	8.07
κ_{LO}	1.3535	0.5094	1.5855	1.1317	2.2365	6.2456
t-stat	0.23	1.31	2.86	2.25	3.49	6.45
κ_{HI}	0.6152	0.1775	1.1717	2.8567	2.0833	1.8862
t-stat	0.35	1.74	0.95	2.52	1.67	1.81
σ_{LO}	0.0568	0.1224	0.3214	0.2986	0.3338	0.6798
t-stat	17.63	17.27	16.19	16.57	16.92	16.02
σ_{HI}	0.1103	0.2214	0.3275	0.3305	0.3681	0.8336
t-stat	8.41	7.83	7.84	7.93	8.15	8.24

Table 8. **Fitting marginal distributions to individual PDs**

In this table we report the estimation results of the fit of individual PD residuals to various distributions. The 3 distributions chosen were: double exponential, normal and the student's t. We used four distance metrics to compare the empirical residuals to standardized distributions: the Kolmogorov distance, the Anderson-Darling statistic, and distances in the L^1 and L^2 norms. The table below presents the number of individual issuers that best fit each of the distributions under the different metrics. A total of 619 issuers was classified in this way.

Statistic	Double Exponential	Normal	Student-T
Kolmogorov	522	55	42
Anderson-Darling	409	7	203
L^2	529	53	37
L^1	118	500	1

Table 9. Metric for best correlated default model (normal and Gumbel copulas)

This table presents the summary statistic for the asymmetric correlation metric to determine the best simulation model. For each model we generate the asymmetric correlation plot and then compute the distance metric. The asymmetric correlations (the metric q) are computed for the following seven cases: normal distribution, student-t distribution, skewed double exponential distribution, the combination of the best distributions based on Kolmogorov criterion, the combination of the best distributions based on Anderson & Darling statistic, the combination based on the L^1 and L^2 norms. Our metric comprises the average squared point-wise difference between the empirical exceedance correlation plot and the simulated one. This a natural distance metric. The points in each plot that are used are for the combination of rating class (from 1 to 6) and exceedance levels (from -1.5 to $+1.5$). Hence there are a total of 48 points in each plot which are used for computing the metric. Define the points in the empirical plot as $h_{k,x}$, where k indexes the rating class and x indexes the exceedance levels. The corresponding points in the simulated plot are denoted $h'_{k,x}$. The metric q (a RMSE statistic) is as follows: $q = \sqrt{\frac{1}{48}\sum_{k=1}^{6}\sum_{x=-1.5}^{1.5}[h_{k,x} - h'_{k,x}]^2}$ We report the results for both models, the jump-diffusion set up and the regime-switching one, and two copulas, the Gaussian and Gumbel copulas.

	Normal Copula		Gumbel Copula	
Model	Jump-diffusion	Regime-shifting	Jump-diffusion	Regime-shifting
Normal	0.0508	0.0276	0.0486	0.0462
Student-T	0.0375	0.0327	0.0416	0.0515
Double-exponential	0.0752	0.0355	0.0700	0.0381
Kolmogorov	0.0802	0.0387	0.0743	0.0364
Anderson-Darling	0.0447	0.0315	0.0452	0.0450
L^2	0.0764	0.0367	0.0710	0.0376
L^1	0.0667	0.0318	0.0625	0.0401

Table 10. **Metric for best correlated default model (Clayton and student-T copulas)**

This table presents the summary statistic for the asymmetric correlation metric to determine the best simulation model. For each model we generate the asymmetric correlation plot and then compute the distance metric. The asymmetric correlations (the metric q) are computed for the following seven cases: normal distribution, student-t distribution, skewed double exponential distribution, the combination of the best distributions based on Kolmogorov criterion, the combination of the best distributions based on Anderson & Darling statistic, the combination based on the L^1 and L^2 norms. Our metric comprises the average squared point-wise difference between the empirical exceedance correlation plot and the simulated one. This a natural distance metric. The points in each plot that are used are for the combination of rating class (from 1 to 6) and exceedance levels (from -1.5 to $+1.5$). Hence there are a total of 48 points in each plot which are used for computing the metric. Define the points in the empirical plot as $h_{k,x}$, where k indexes the rating class and x indexes the exceedance levels. The corresponding points in the simulated plot are denoted $h'_{k,x}$. The metric q (a RMSE statistic) is as follows: $q = \sqrt{\frac{1}{48}\sum_{k=1}^{6}\sum_{x=-1.5}^{1.5}[h_{k,x} - h'_{k,x}]^2}$ We report the results for both models, the jump-diffusion set up and the regime-switching one, and two copulas, the Clayton and T copulas.

	Clayton Copula		student-T Copula	
Model	Jump-diffusion	Regime-shifting	Jump-diffusion	Regime-shifting
Normal	0.0484	0.0395	0.0446	0.0506
Student-T	0.0415	0.0473	0.0370	0.0564
Double-exponential	0.0694	0.0304	0.0672	0.0433
Kolmogorov	0.0738	0.0293	0.0714	0.0404
Anderson-Darling	0.0445	0.0397	0.0418	0.0514
L^2	0.0703	0.0296	0.0681	0.0423
L^1	0.0622	0.0324	0.0596	0.0452

FIGURE 1. **Time series of average PDs**

This figure depicts the average level of default probabilities in the data set. The data shows the presence of two regimes, one in which PDs were high, as in the early and latter periods of the data. In the other regime, PDs were much lower, less than half of those seen in the high PD regime. Complementary analysis of correlated default on the same data set by subperiod in this same time frame is presented in DFGK [2001] from where this figure is taken.

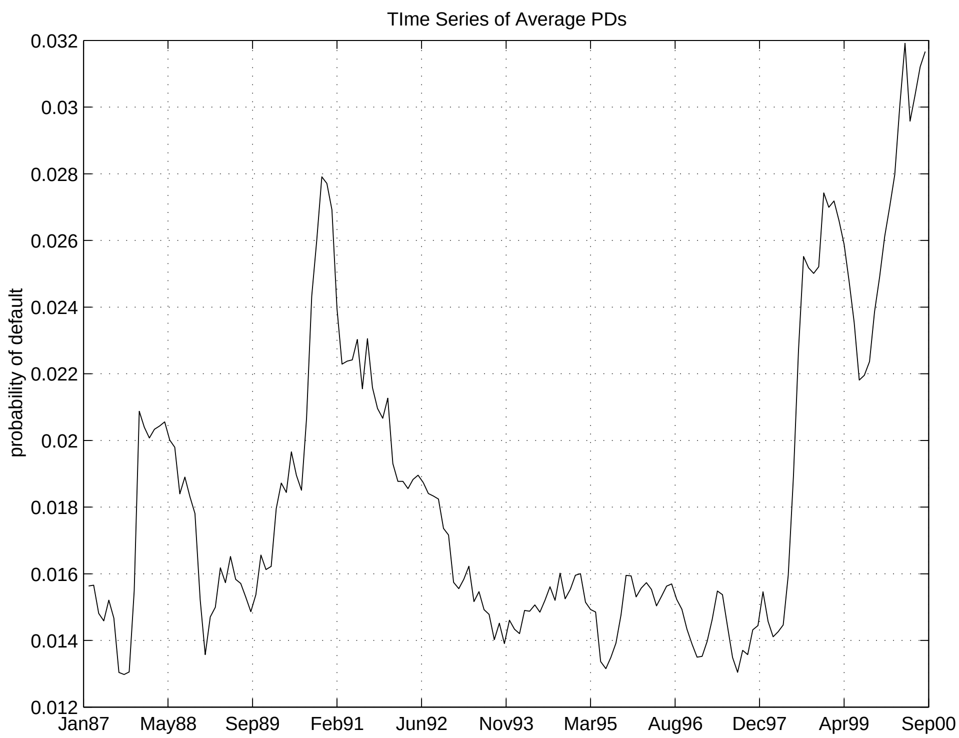

FIGURE 2. **Using the normal copula with a student's t distribution**

This figure presents 6 plots of 1000 bivariate samples based on the normal and student's t distributions. The 3 plots on the left side are based on the bivariate standard normal distribution with correlation 0,0.5 and -0.5 respectively. The next 3 plots on the right are based on the student's t distribution with correlation 0,0.5 and -0.5 respectively. The simulation of the student's t distribution is done using a normal copula.

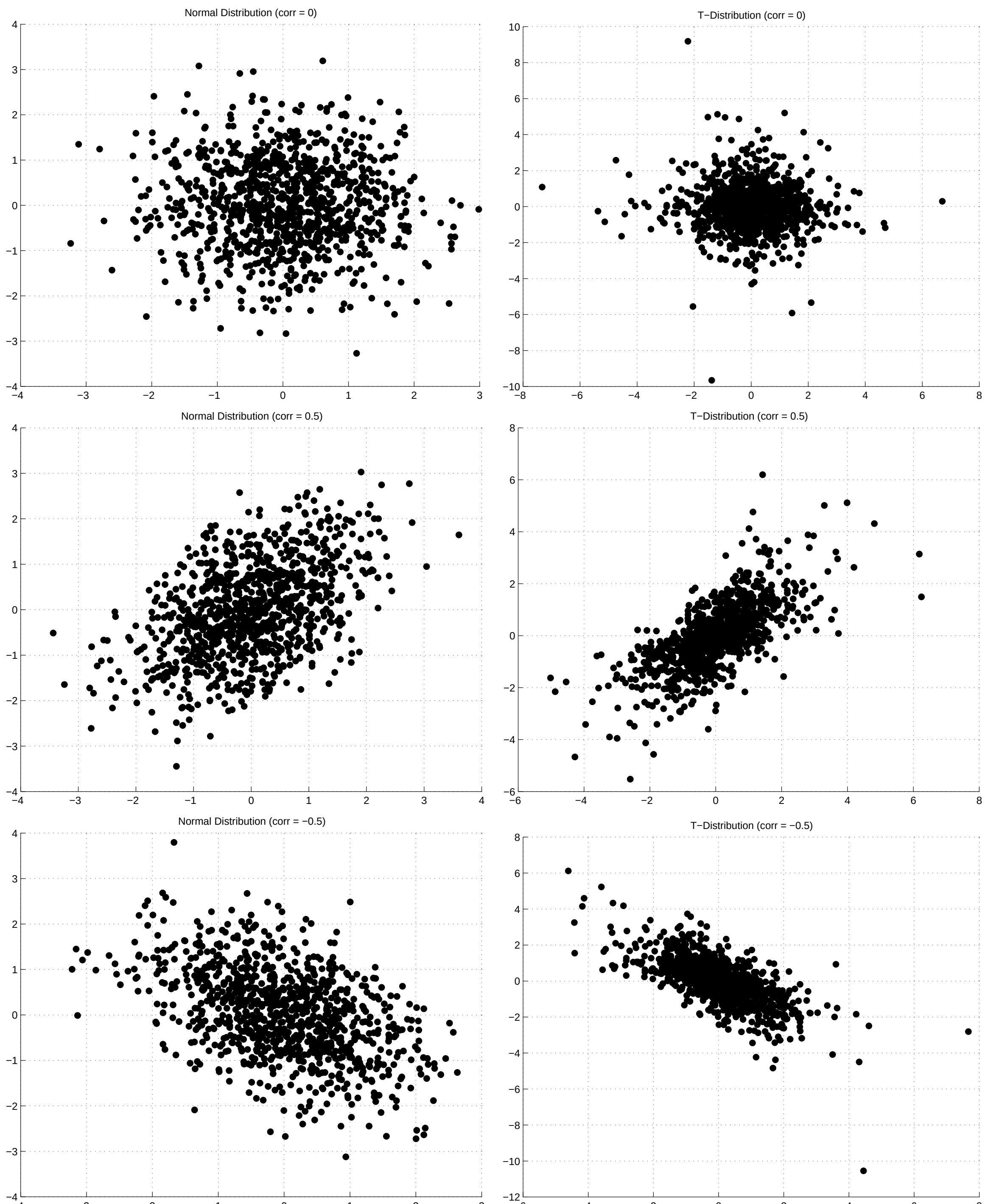

FIGURE 3. **Scatter Plots of Tail Dependence of Four Copulas**

The 4 plots present scatter plots of 2,500 samples drawn from each of the 4 copulas used in this paper. The plots offer a comparison of the tail-dependence of the various copulas.

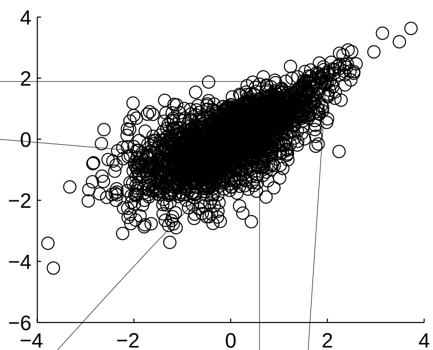

FIGURE 4. **Scatter Plots of Upper Tail Dependence of Four Copulas**

The 4 plots present scatter plots of 2,500 samples drawn from each of the 4 copulas used in this paper. The plots offer a comparison of the upper tail-dependence of the various copulas.

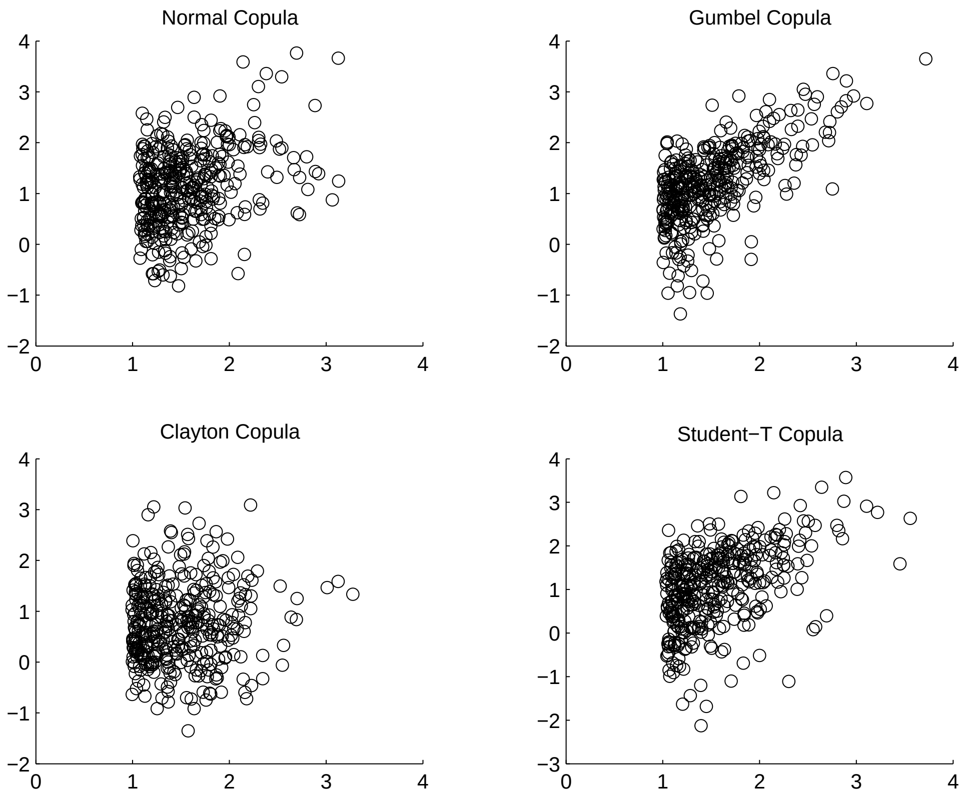

Figure 5. **Scatter Plots of Lower Tail Dependence of Four Copulas**

The 4 plots present scatter plots of 2,500 samples drawn from each of the 4 copulas used in this paper. The plots offer a comparison of the lower tail-dependence of the various copulas.

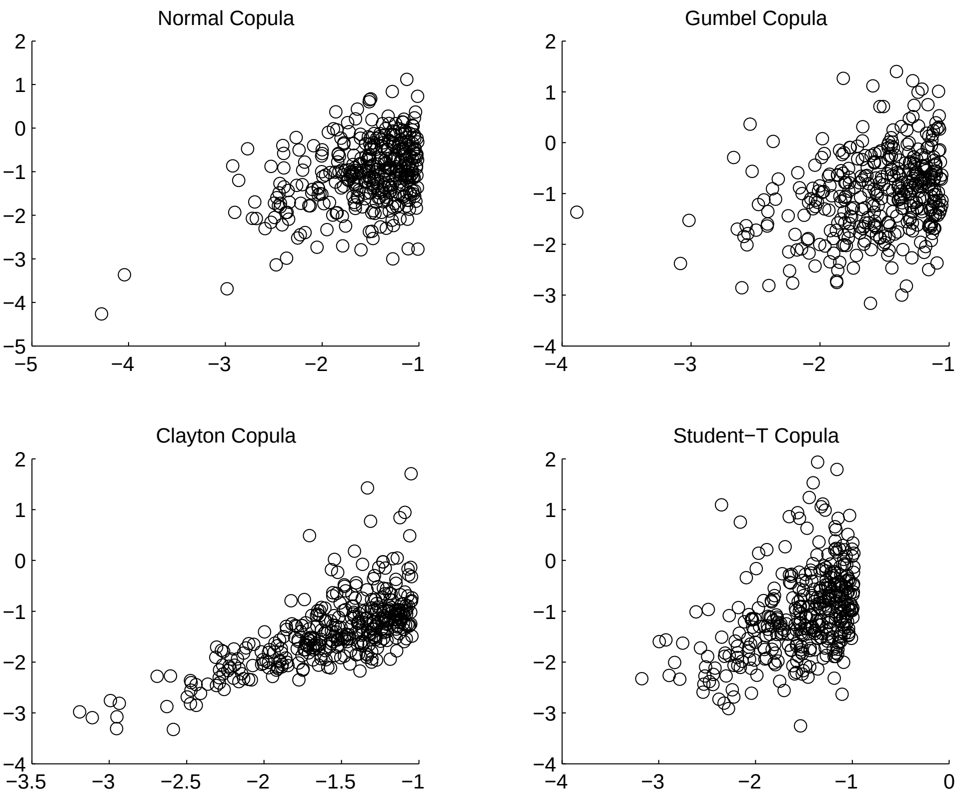

Figure 6. **Asymmetric Correlation of PDs**

This figure graphs the exceedance correlations for the different rating classes in our study. The figure breaks this out into the individual ratings. The plots are consistent with the presence of asymmetric correlation. The degree of asymmetry appears to be higher for the better rating categories. Complementary analysis of correlated default by subperiod in this same time frame is presented in DFGK [2001] where different analyses are conducted on the same data set. This figure corresponds to their paper.

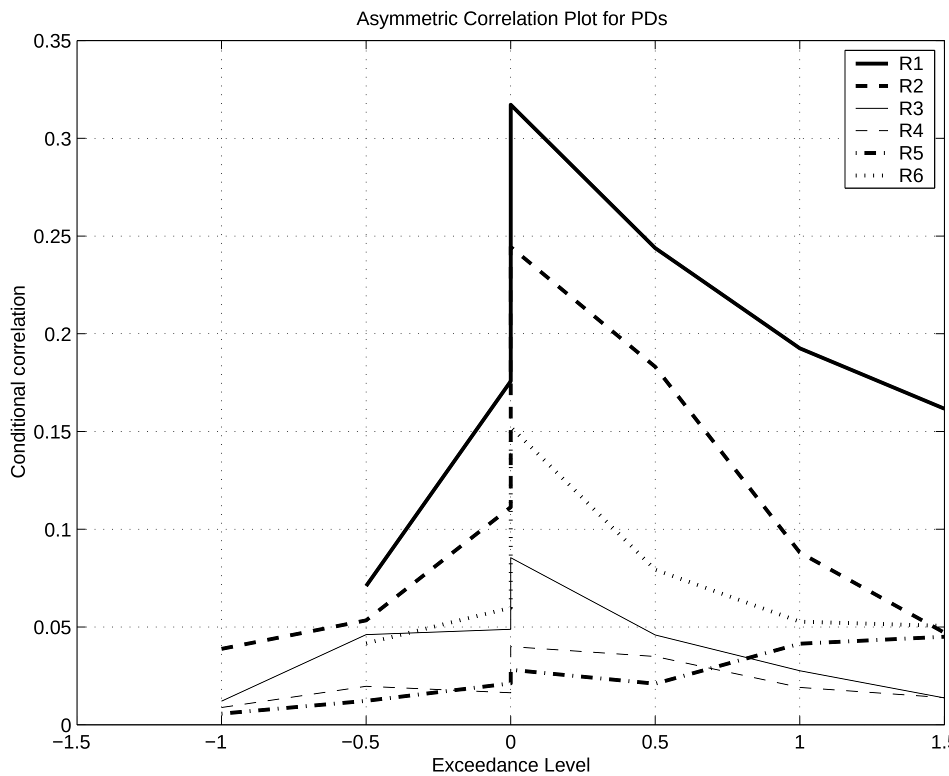

FIGURE 7. **Simulated series of average PDs by Rating Class**

This figure depicts the average level of default probabilities in the data set by rating class. The hazard rates are (from top to bottom of the graph) from rating class six to rating class one. The simulation ensures that $\lambda_i(t) < \lambda_j(t)$ if $i < j$.

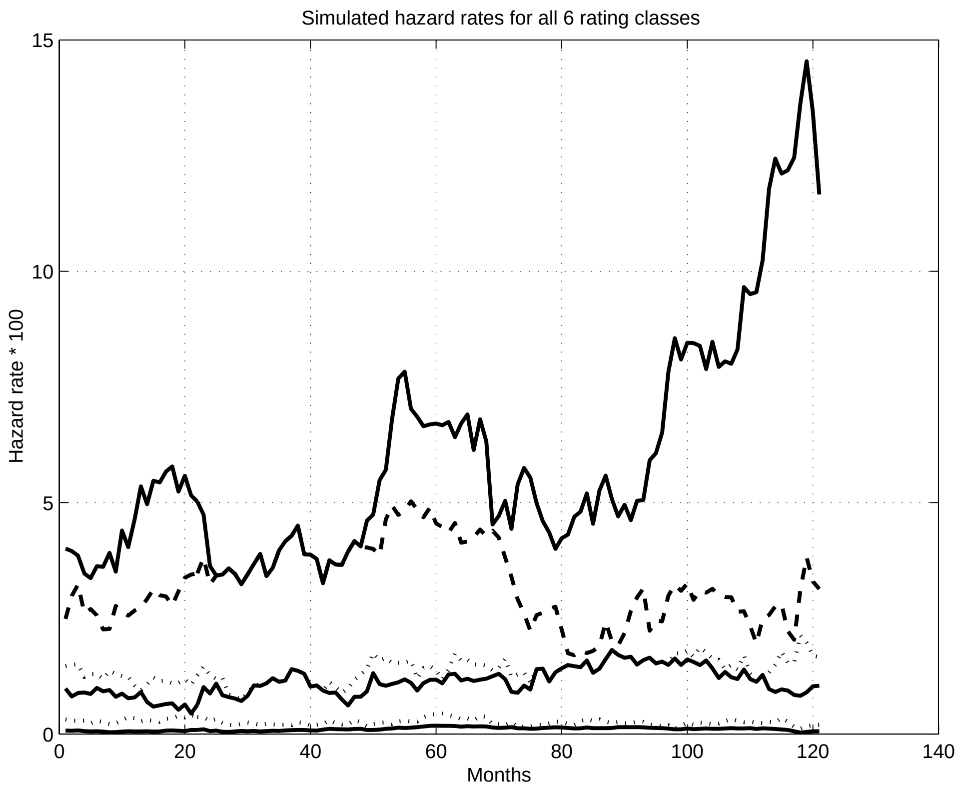

Figure 8. **Asymmetric Correlation of Simulated PDs**

This figure graphs the exceedance correlations for the different rating classes in our study, based on simulated data. The figure breaks this out into the individual ratings. The plots are consistent with the presence of asymmetric correlation. The degree of asymmetry is higher for the better rating categories.

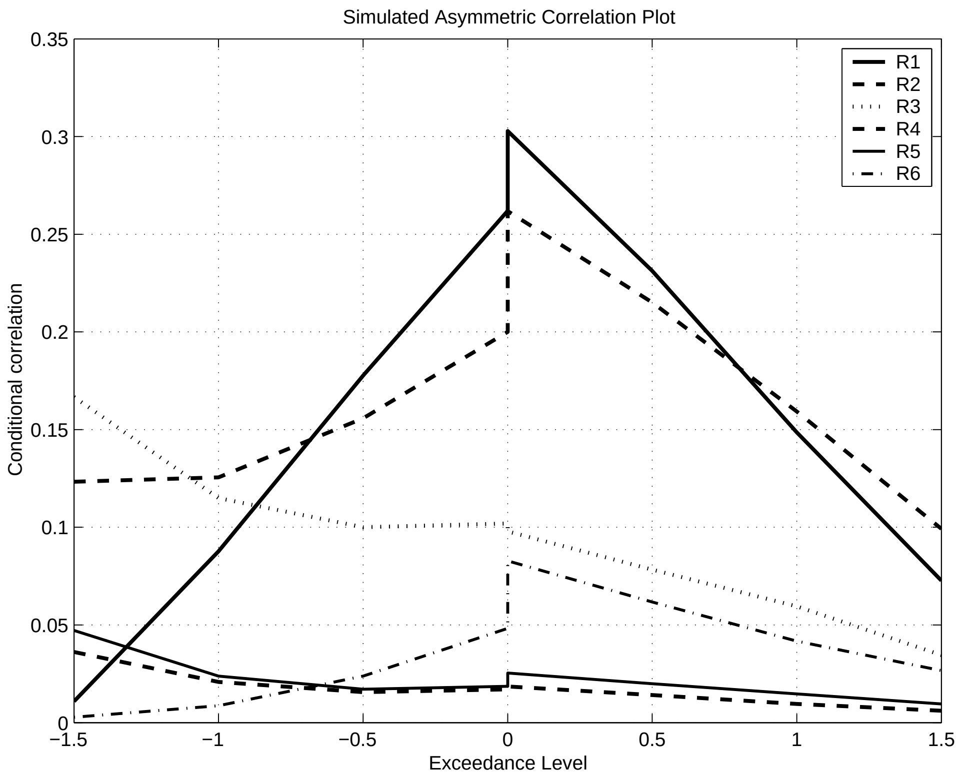

FIGURE 9. **Asymmetric Correlation of Simulated PDs from the Regime-Shifting model**

This figure graphs the exceedance correlations for the different rating classes in our study, based on the regime-switching model. The figure breaks this out into the individual ratings. The plots are consistent with the presence of asymmetric correlation. The degree of asymmetry is higher for the better rating categories.

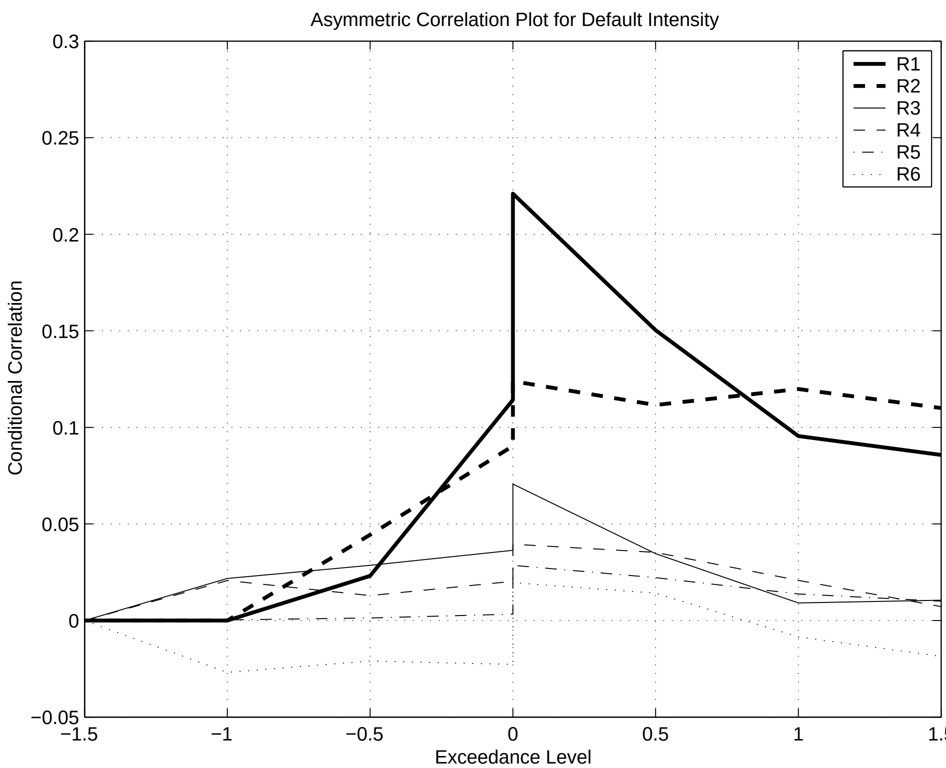

FIGURE 10. **Comparing copula tail loss distributions**

This figure presents plots of the tail loss distributions for 4 copulas, when the marginal distribution is normal. The x-axis shows the number of losses out of 620 issuers, and the y-axis depicts the percentiles of the loss distribution. The simulation runs over a horizon of 5 years, and account for regime shifts as well. The copulas used are: normal, Gumbel, Clayton, student's t.

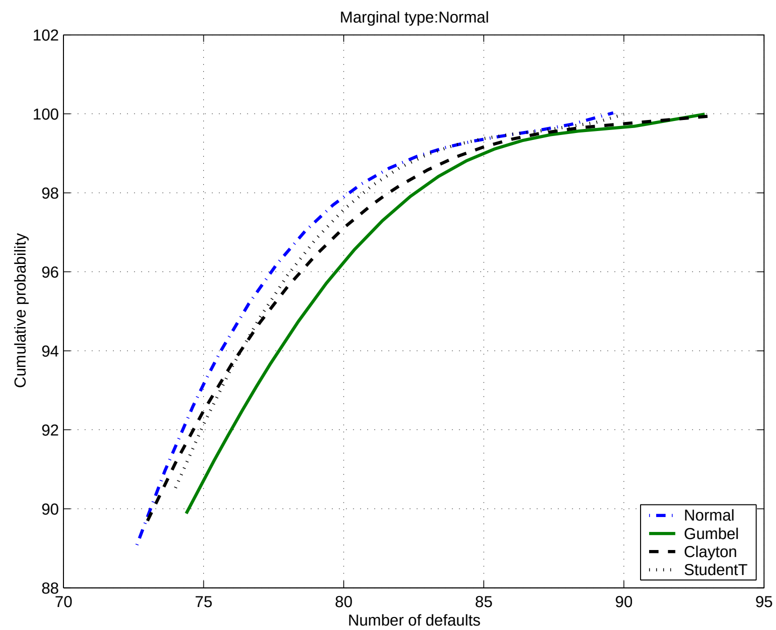

FIGURE 11. **Comparing copula tail loss distributions**

In this figure we plot the tail loss distributions for two models, the best fitting one and the worst. The best fit model combines the Clayton copula, and marginal distributions based on the Kolmogorov criterion. The worst fit copula combines the student's t copula with student's t marginals. The simulation runs over a horizon of 5 years, and account for regime shifts as well.