ESTIMATING PORTFOLIO CREDIT LOSSES IN DOWNTURNS*

Fernando F. Moreira
Keele University
Keele Management School
Email: f.moreira@mngt.keele.ac.uk
Phone: +44(0)755-123-7270

ABSTRACT

We suggest formulas able to capture potential strong connection among credit losses in downturns without assuming any specific distribution for the variables involved. We first show that the current model adopted by regulators (Basel) is equivalent to a conditional distribution derived from the Gaussian Copula (which does not identify tail dependence). We then use conditional distributions derived from copulas that express tail dependence to estimate the probability of credit losses in extreme scenarios. Next, we use data on historical credit losses incurred in American banks to compare the suggested approach to the Basel formula with respect to their performance when predicting the extreme losses observed in 2009 and 2010. Our results indicate that, in general, the copula approach outperforms the Basel method in two of the three credit segments investigated. The proposed method is extendable to other differentiable copula families which gives flexibility to future practical applications of the model.

JEL codes: G28, G21, G32, C46, C49
Keywords: Credit risk, downturns, Basel Accords, conditional distributions, copulas.

* A modified version of this paper was previously circulated under the title "Enhancing Basel method via conditional distributions that capture stronger connection among credit losses in downturns".

1. INTRODUCTION

The model (Basel Accord) adopted by regulators in many countries to calculate the capital to cover unexpected credit losses in financial institutions assumes normally-distributed variables and uses the linear correlation to measure dependence among losses. However, these assumptions do not allow the identification of possible asymmetric dependence across losses in extreme scenarios (which seems to be the case for several financial assets, loans included) and, therefore, the Basel method may misestimate joint credit losses in periods of crisis.

Albeit the formula currently used in Basel Accords has a derivation not associated to copula functions, we show that it turns out to be equivalent to the first derivative of the Gaussian Copula (which denotes symmetric association without tail dependence). Moreover, the distribution of one variable conditional on another variable can be calculated as the first derivative of the copula that represents the dependence between the considered variables with respect to the conditioning variable. In other words, the Basel formula can be interpreted as the cumulative distribution of a latent variable (asset returns of obligors, for instance) conditional on the economic status. Based on this interpretation of the Basel model, we propose the use of copulas that capture stronger dependence among high losses (stronger dependence among *low* values of debtors' asset returns) to generate alternative conditional distributions. So, we keep the basic intuition of the traditional approach but *change the dependence structure* such that we can, for example, identify higher probability of default in adverse scenarios. The alternative model is basically set as the first derivative of the copula chosen to represent the relationship between the latent variable and the economic factor with respect to the latter variable. At this point, we face a challenge pertaining to the copula parameter that measures the dependence intensity. For some copulas, this parameter can be directly deduced from the rank correlation (Kendall's tau) between the variables. Thus we need to find the rank correlation between the latent variable of each loan and the economic factor but we cannot calculate it since we do not have enough information about the second variable. To overcome this problem, we show that the rank correlation between the latent variable of each debtor and the economic factor is related to the rank correlation between two latent variables (e.g. asset returns of two obligors) which can be presumed from past losses (default rates). Once we have an estimate for the former rank correlation, we will have all necessary information to calculate the conditional probability by means of the first derivative of a copula with a given confidence (unfavorable economic level).

As examples, we present two formulas originated from the Clayton and the Student t Copulas that are able to detect stronger connection (tail dependence) among low values of latent variables (which is equivalent to identify higher dependence among high credit losses). These formulas do not assume any kind of distribution for the variables considered and therefore such approach overcomes the limitations of the existing method with regard to the assumption of normality and the use of the linear correlation.

We use aggregate data on losses in American banks to check the performance of the suggested approach and our analyses show that, for two of the three credit segments considered, the copula formulas typically outperform the Basel formula regarding the estimation of unusually high losses.

In short, our contributions are threefold: (i) we present an alternative derivation of the Basel formula and show that it corresponds to the first derivative of the Gaussian Copula; (ii) we set up a model able to capture stronger dependence among credit losses in unfavorable scenarios which results in more efficient estimations of potential extreme losses; and (iii) we propose a way to derive the dependence between a latent variable of each loan and an economic factor from the dependence observed across loans' default rates.

2. COPULAS AND CONDITIONAL DISTRIBUTIONS

Copulas are multivariate distribution functions with uniformly distributed margins in (0,1) that link marginal (individual) distributions of variables to their joint distributions:

$$F_{1...n}(x_1,...,x_n) = C(F_1(x_1),...,F_n(x_n))$$

where *F(.)* denotes a cumulative distribution function and *C* stands for a copula. Thus, C is an expression (function) with n inputs and, when evaluated at $F_1(x_1),...,F_n(x_n)$, returns the joint cumulative distribution of the n variables evaluated at $x_1,...,x_n$, i.e., the probability that all variables $X_1,...,X_n$ are concurrently below the respective values $x_1,...,x_n$.

According to Joe (1996)[2], the cumulative distribution of a random variable conditional on other variables is given by the first derivative of the copula that represents the dependence among the variables with respect to the conditioning variables (those placed after the symbol "|"):

$$F(x \mid \mathbf{v}) = \frac{\partial C_{xv_j|\mathbf{v}_{-j}}(F(x \mid \mathbf{v}_{-j}), F(v_j \mid \mathbf{v}_{-j}))}{\partial F(v_j \mid \mathbf{v}_{-j})}$$

where *F(x*|**v***)* is the distribution of *X* evaluated at *x* conditional on vector **v**, $C_{xv_j|\mathbf{v}_{-j}}$ is a copula distribution function, $\mathbf{v}_j$ is a component of vector **v** and $\mathbf{v}_{-j}$ is the vector **v** excluding this component. When **v** is univariate, the conditional distribution becomes:

[2] A detailed proof of this formula is given in Czado (2010).

$$F(x \mid v) = C_{x|v}(F(x) \mid F(v)) = \frac{\partial C_{xv}(F(x), F(v))}{\partial F(v)}$$

where x and v indicate the conditioned and the conditioning variables respectively and the remaining notation is the same used in the prior formula.

The first derivative of some bivariate copulas can be found, for example, in Joe (1997, Chapter 5), Aas et al. (2009, Appendix C) and Bouyé and Salmon (2009). Three families of particular interest here are the Gaussian (Normal), the Clayton and the Student t that respectively generate the conditional distributions stated in [1], [2] and [3]:

$$\begin{aligned} \Pr[X_1 < x_1 \mid X_2 = x_2] &= F_{1|2}(x_1 \mid X_2 = x_2) = \\ &= \Phi\left(\frac{\Phi^{-1}(F_1(x_1)) - \theta_{12}\Phi^{-1}(F_2(x_2))}{\sqrt{1-\theta_{12}^2}} \right) \end{aligned} \qquad [1]$$

where Φ and Φ^{-1} represent the standard normal distribution and its inverse respectively, *F(.|.)* is the distribution of X_1 conditional on X_2, *F(.)* is an unconditional distribution and θ_{12} is the Gaussian Copula parameter[3] that indicates the strength of the dependence between X_1 and X_2.

$$\Pr[X_1 < x_1 \mid X_2 = x_2] = F_{1|2}(x_1 \mid X_2 = x_2) = \{F_2(x_2)^{\theta_{12}}[F_1(x_1)^{-\theta_{12}} - 1] + 1\}^{(-1-\theta_{12})/\theta_{12}} \qquad [2]$$

where *F(.|.)* is the distribution of X_1 conditional on X_2, *F(.)* is an unconditional distribution and θ_{12} is the Clayton Copula parameter between X_1 and X_2.

$$\Pr[X_1 < x_1 \mid X_2 = x_2] = F_{1|2}(x_1 \mid X_2 = x_2) = T_{v+1}\left(\frac{T_v^{-1}(F_1(x_1)) - \theta_{12}T_v^{-1}(F_2(x_2))}{\sqrt{\frac{v + (T_v^{-1}(F_2(x_2)))^2(1-\theta_{12}^2)}{v+1}}} \right) \qquad [3]$$

[3] The parameter of the Gaussian Copula is usually represented by ρ. We adopt the notation θ to distinguish the Gaussian Copula parameter from the linear correlation coefficient between the variables studied. These two measures of dependence are identical only when the marginal distributions are normal.

where T_v and T_v^{-1} represent the Student t cumulative distribution function and its inverse, respectively, with v degrees of freedom ($v > 0$); the other variables are defined as in [1] and [2]. Bear in mind that as v increases, this conditional distribution approaches [1].

The copula parameter θ is closely related to rank correlations Kendall's tau (τ). For two variables X_1 and X_2 with distribution functions evaluated at x_1 and x_2, $F_1(x_1) = u_1$ and $F_2(x_2) = u_2$ correspondingly, the intensity θ of their representative copula can be inferred from[4]:

$$\tau = 4\iint_{[0,1]^2} C(u_1, u_2)\, dC(u_1, u_2) - 1 \qquad [4]$$

3. BASEL METHOD: A DERIVATION FROM THE GAUSSIAN COPULA

3.1 The calculation of extreme credit losses

For each homogeneous credit segment, the capital required to cover unexpected losses is calculated as the unexpected losses adjusted by the portfolio maturity.

In mathematical terms:

$$(LGD * K_V - LGD * PD) * Maturity \qquad [5]$$

where LGD is the "loss given default" (which is equal to 1 - recovery rate, i.e. the percentage of exposure the lender will lose if borrowers default) and PD stands for the probability of default. *Maturity* corresponds to the maturity of *corporate* loans (i.e., not applied to consumer debt) and is added to the calculation in order to give higher weight to long-term obligations which are known to be riskier. For the sake of brevity, the maturity formula is not presented here. See BCBS (2005, 2006) for more details.

The other term in [5], K_V, is the expected default rate at the 99.9th percentile of the PD distribution ("Vasicek Formula") - see Vasicek (1991, 2002) - and is calculated as:

$$K_V = \Phi\left(\frac{\Phi^{-1}(PD) + \sqrt{\rho}\,\Phi^{-1}(0.999)}{\sqrt{1-\rho}}\right) \qquad [6]$$

where:

[4] The proof is given, for example, in McNeil et al. (2005, Chapter 5) and Nelsen (2006, Chapter 5).

Φ and Φ^{-1} represent the standard normal cumulative distribution and its inverse, respectively; *PD* is the probability of default of the loan portfolio (average); $\Phi^{-1}(0.999)$ is the level of the economy (confidence) chosen to represent an extreme scenario when unexpected losses may occur and rho (ρ) is the correlation between returns of obligors' assets. $\sqrt{\rho}$ is the linear correlation between the unobserved systematic factor and those asset returns. In Basel method, the correlation between asset returns is calculated as a decreasing function of *PD* and (in the case of corporate debt) the size of debtors (measured in terms of annual sales); see formulas in BCBS (2005, 2006).

Thus, the terms $LGD * K_V$ and $LGD * PD$ in [5] represent, respectively, the extreme and the average losses net of recoveries.

3.2 Derivation of K_V from the Gaussian Copula

Expression [6] is typically derived from Factor Models which assume that the correlation among defaults is driven by the debtors' latent variables. See, e.g., Crouhy et al. (2000) and Bluhm et al. (2002) for general information about Credit Factor Models and Schönbucher (2000) and Perli and Nayda (2004) for the derivation of [6] from Factor Models.

Naturally, there are many common factors that act together and influence debtors' situation. However this model can be simplified if we consider that all latent variables (usually interpreted as asset returns of borrowers) are driven by only one common factor (the "economic status"). For simplicity, all pairs of asset returns (*i* and *j*) are considered to present the same correlation (ρ_{ij}). The correlation between the asset return (*Y*) of each debtor and the systematic factor (*E*) is denoted by ρ_{YE}. Since the variables in this approach are assumed to follow the jointly standard normal distribution, we have (see, e.g., Hull and White 2004):

$$\rho_{YE} = \sqrt{\rho_{ij}} \qquad [\ 7\ ]$$

This equality is essential to the subsequent calculations since there is usually no adequate proxy for *E* (which is not observable) and, consequently, ρ_{YE} cannot be directly estimated from empirical data. On the other hand, we can infer the correlation between asset returns, ρ_{ij}, from historical losses (default rates).

We show here that [6] is also associated with the Gaussian Copula and, to our knowledge, this has not been shown in previous studies. Starting from [1], the conditional distribution calculated from the

Gaussian Copula (restated below for convenience), consider that X_1 is a latent variable, x_1 is the level below which obligors default and X_2 is the economic status (single factor). So, this formula gives the likelihood of the latent variable X_1 being below a specific value x_1 (the probability of default) conditional on $X_2 = x_2$. Assume that both variables follow the standard normal distribution.

$$F_{1|2}(x_1 \mid X_2 = x_2) = \Phi\left(\frac{\Phi^{-1}(F_1(x_1)) - \theta_{12}\Phi^{-1}(F_2(x_2))}{\sqrt{1-\theta_{12}^2}}\right) \qquad \text{[1] restated}$$

Therefore $F_1(x_1) = \Phi(x_1) = PD$ (i.e. the probability of the latent variable X_1 being below the cutoff x_1) and $\Phi^{-1}(F_1(x_1)) = \Phi^{-1}(PD)$ returns x_1, the latent variable cutoff[5]. $F_2(x_2) = \Phi(x_2)$ is the level of the economic situation and the inverse of its distribution $\Phi^{-1}(F_2(x_2)) = \Phi^{-1}(\Phi(x_2)) = x_2$ gives the "value" of the economic variable. So, the smaller $\Phi(x_2)$ is the worse the economic status gets and to express adverse scenarios in [1] small values for $\Phi(x_2)$ should be used. Basel adopts the confidence level of 99.9% and this is expressed as $\Phi^{-1}(0.999)$, which is equal to $-\Phi^{-1}(0.001)$. The parameter θ_{12} in [1] refers to the dependence between X_1 and X_2. If we assume that X_1 and X_2 have individual normal distributions, θ_{12} will be equal to the linear correlation between the variables (denoted here as ρ_{12}) which cannot be estimated given that there is no sufficient information on the economic factor. Assume we can assess the linear correlation between the latent variables (based on the observed default rates). Under the conditions specified (i.e. the latent variables and the economic factor follow the standard normal distribution) and according to [7], θ_{12} can be associated to the linear correlation ρ between the latent variables (or the probabilities of default) such that $\rho_{12} = \sqrt{\rho}$.

In resume, setting $F_1(x_1) = PD$ and $F_2(x_2) = 0.999$, replacing $\Phi^{-1}(0.999)$ with $-\Phi^{-1}(0.001)$ and noting that $\theta_{12} = \rho_{12} = \sqrt{\rho}$, we see that the first derivative of the Gaussian Copula, [1], corresponds to the formula (restated below) used in Basel to calculate the probability of default conditional on an extremely unfavorable economic situation:

$$K_V = \Phi\left(\frac{\Phi^{-1}(PD) + \sqrt{\rho}\Phi^{-1}(0.999)}{\sqrt{1-\rho}}\right)$$

[5] As before, Φ represents the standard normal distribution and Φ^{-1} indicates its inverse.

4. EMPLOYING ALTERNATIVE CONDITIONAL DISTRIBUTIONS TO CAPTURE TAIL DEPENDENCE

4.1 Some prior suggestions

As indicated in some empirical studies (for instance, Di Clemente and Romano 2004 and Das and Geng 2006), higher credit losses tend to be more associated than low levels of losses. Some models have been proposed to transform [6] into other expressions that do not have the limitation regarding the assumption of normality and can capture skewness and heavy tails (which tends to increase the joint occurrences of extreme realizations of the latent variables). Hull and White (2004), for instance, relax the distributions[6] of the latent variable *Y*, the economic factor *E* and the idiosyncratic (specific) risk. Letting *F*, *G* and *H* denote the distributions of those three variables respectively, the probability of default conditional on an unfavorable economic status (the worst 0.1% scenario, i.e. with confidence of 99.9%) turns out to be:

$$\Pr[Y_i < y_c \mid E = e^*] = H\left(\frac{F^{-1}(PD) - \sqrt{\rho}G^{-1}(0.001)}{\sqrt{1-\rho}}\right)$$

where e^* indicates an extreme adverse economic scenario and can be calculated as the inverse distribution of *E* evaluated at 0.001 (since the critical level was set at 0.1%). *PD* is the historical (average) probability of default and ρ is the linear correlation between returns of obligors' assets. Naturally, the expression above can be solved only if the shapes of the three distributions *F*, *G* and *H* are known.

A number of studies, such as Bluhm et al. (2002) and Kang (2005), have suggested the Student t distribution to represent the economic and the idiosyncratic risks (functions *G* and *H* above). In this case, it is not possible to define the distribution *F* of the latent variable and the probability of default in downturns (at the 0.1% worst scenario) is:

$$\Pr[Y_i < y_c \mid E = e^*] = T_v\left(\frac{F^{-1}(PD) - \sqrt{\rho}T_v^{-1}(0.001)}{\sqrt{1-\rho}}\right)$$

where T_v is the Student t distribution with v degrees of freedom. Given that the latent variable's distribution *F* remains unknown, the preceding likelihood cannot be calculated. Chan-Lau (2010) reasons that this approach can be used to capture asymmetry and fat tails in the calculation of regulatory capital in financial institutions.

[6] Provided that they are scaled with mean zero and variance one.

In view of the impossibility of the estimation of the probability of default in adverse economic scenarios when one (or more) of the variables in [6] are not normally distributed, we propose a different setup to incorporate Copula Theory into this analysis and to capture potential tail dependence even if we do not know *any* of the distributions concerning the latent variable and the economic factor (which is the reality in financial institutions).

4.2 Some alternatives to detect higher dependence across losses in downturns

Recalling that credit losses imply the existence of small values of the latent variables, we can interpret the stronger connection among losses in downturns as an effect of the intensification of the dependence across small latent variables. In other words, this is evidence that small values of the latent variables tend to be more connected over adverse periods. Thus the relationship between two latent variables, Y_i and Y_j, can be represented by scatterplots like the ones in Figure 1.

[Insert Figure 1 here]

The difference between Panels A and B is that the former does not indicate right-tail dependence whilst the latter does. However this difference does not impact our analyses since we are interested in the left tail (small values of the latent variables) regardless of the variables behavior in the right tail. Therefore the representation in either of those two panels is suitable for modeling strong dependencies among losses in bear markets.

When the economic factor E is inserted in the analysis, reduced levels of this variable will present more intense association with the latent variables. Figure 2 shows the dependence between E and each latent variable in the context of Panel A in Figure 1. The correspondence between Figure 1 (Panel A) and Figure 2 can be noticed by comparing the level of Y_i and Y_j in a downturn (e^*, for example) with the level of those latent variables when the economy is booming (e^{**}, for example). In the first case, both Y_i and Y_j tend to be small whilst in the better economic scenario, e^{**}, a wider range of different values of the latent variables are associated (i.e. there is a higher likelihood that a small Y_i and an elevated Y_j, for instance, will happen at the same time when the economic is at the level e^{**}). So, this means that the lower-tail dependence characterizes not only the relationship between the underlying variables but also the link between the economic status and each latent variable. A similar reasoning applies to Panel B in Figure 1.

[Insert Figure 2 here]

The dependence structures depicted in Figure 1 can be represented by, for example, the Clayton Copula (Panel A) and the Student t Copula (Panel B). In these cases, the proportion of loans in the portfolio for which the latent variable, Y, will be smaller than the cutoff y_c (i.e. the probability of default) when the economy falls to an extremely low level (e^*) is derived from [2] and [3] respectively:

$$F(y_c \mid E = e^*) = \{F_E(e^*)^{\theta_{YE}}[F_Y(y_c)^{-\theta_{YE}} - 1] + 1\}^{(-1-\theta_{YE})/\theta_{YE}} \qquad [8]$$

and

$$F(y_c \mid E = e^*) = T_{v+1}\left(\frac{T_v^{-1}(F_Y(y_c)) - \theta_{YE}T_v^{-1}(F_E(e^*))}{\sqrt{\frac{v + (T_v^{-1}(F_E(e^*)))^2(1-\theta_{YE}^2)}{v+1}}}\right) \qquad [9]$$

where $F(.|.)$ indicates a conditional distribution, $F_E(e^*)$ is the cumulative distribution of the economic factor (which indicates adverse scenarios when it approaches 0 and booms when it gets close to 1), $F_Y(y_c)$ is the average (historically observed) default rate (a proxy for the probability of default), T_v and T_v^{-1} represent the Student t cumulative distribution function and its inverse, respectively, with v degrees of freedom ($v > 0$) and θ_{YE} is the copula parameter between Y and E.

Among the three variables necessary to compute extreme losses by applying [8] and [9], two, $F_Y(y_c)$ and $F_E(e^*)$, are readily available; the former is the expected probability of default (default rate) of the homogeneous portfolio and the latter is to be set according to the confidence demanded for the economic scenario[7]. Naturally, it is expected that the probability of the latent variable of each obligor being below a particular cutoff, given a specific economic level, increases when the dependence among the defaults becomes stronger. In the case of the Clayton Copula, this monotonically increasing behavior of $F(y_c|E)$ with respect to θ_{YE} happens only if $F_E(e^*) \leq F_Y(y_c)$. When $F_E(e^*) \geq F_Y(y_c)$, $F(y_c|E)$ is a quadratic function of θ_{YE} and starts falling after rising up to a specific value. Therefore the calculation of the regulatory capital based on the Clayton Copula will yield more consistent results if the extreme economic level is restricted to percentiles smaller than or equal to the percentiles of the latent variables, i.e. if $F_E(e^*) \leq PD$, where PD is the average default probability of the portfolio. This does not represent any

[7] Since $F_E(e^*)$ is truncated in the interval [0,1] and small values represent adverse scenarios, 0.01 indicates the confidence level of 99%, 0.05 is associated with the confidence of 95% and so on.

significant concern in this context because we are interested in small values of $F_E(e^*)$ that indicate downturns.

One way to find the other variable in [8] and [9], θ_{YE}, is to derive it from the rank correlation between Y and E (Kendall's tau, τ_{YE}). As shown in [4], the Kendall's tau between two variables is associated with the parameter of the copula that represents their dependence. For some families this association is defined in closed form (see some examples in Nelsen, 2006, Chapter 5). We present below the association between Kendall's tau and the parameters of the two copula families used here to capture high dependence across credit losses in unfavorable scenarios (Clayton and Student t, respectively). The first one is derived from a relationship presented in Nelsen (2006, Chapter 5) and the last one is derived from an association mentioned in McNeil et al. (2005, Chapter 5):

$$\theta_{YE}^{C} = \frac{2\tau_{YE}}{1-\tau_{YE}} \qquad [\ 10\ ]$$

$$\theta_{YE}^{t} = \sin\left(\frac{\tau_{YE}\pi}{2}\right) \qquad [\ 11\ ]$$

However we do not have enough information on E to estimate τ_{YE}. When the Gaussian Copula is used, this problem is resolved by replacing the correlation between Y and E with the correlation between the latent variables of debtors (expression [7]). Thus, assuming that the rank correlation between the latent variables, τ_{ij}, can be inferred from data sets pertaining to credit losses (in the same way the linear correlations across probabilities of default were estimated in Basel Accords for different loan classes), we should look for a correspondence between τ_{YE} and τ_{ij} so that θ_{YE} can be calculated and plugged into the aforementioned expressions [8] and [9].

4.3 Relationship between rank correlations

4.3.1 Rank correlation between the latent variables in homogeneous portfolios

Kendall's tau (τ) is based on the number of concordant and discordant pairs of variables. Assuming (X_1, Y_1) and (X_2, Y_2) are two independent pairs from a joint distribution, they will be *concordant* if $(X_2 - X_1)(Y_2 - Y_1) > 0$, i.e., if the two variables move in the same direction. They will be *discordant* when $(X_2 - X_1)(Y_2 - Y_1) < 0$. Kendall's tau is the difference between the proportion of concordant and

discordant pairs, i.e., τ = *Pr[concordance] – Pr[discordance]*. Defining c as the number of concordant pairs and d as the number of discordant ones, Kendall's tau is equivalently expressed as:

$$\tau = \frac{c-d}{c+d} \qquad [\ 12\ ]$$

Let N be the number of observations (which will be the same for both variables). So, *for any pair*, $c + d = N$ and, from [12], $c - d = \tau N$. By combining these two expressions, we get:

$$d = \frac{N(1-\tau)}{2} \qquad [\ 13\ ]$$

and

$$c = \frac{N(\tau+1)}{2} \qquad [\ 14\ ]$$

Table 1, Panel A, illustrates the co-movements of L loans (represented by their respective latent variables Y_i where $1 \leq i \leq L$) and the systematic factor E. The arrows "↑" and "↓" indicate the direction which the variables move in. So, if two of them have equal arrows, they move in the same direction and are therefore "concordant". Conversely, if one arrow points up while the other one points down, the pairs of variables are "discordant".

[Insert Table 1 here]

As in Basel, we assume that all pairs of loans have the same dependence (here expressed by the Kendall's tau between the latent variables, τ_{ij} for loans i and j) and that the dependence between the systematic factor E and each loan (τ_{YE}) is also the same. In a portfolio of L loans, one way to comply with the two aforementioned constraints is to assume that, in each period, the latent variables of the same number of loans move in the same direction (i.e. decreasing or increasing) – see a simple example concerning a portfolio with three loans in Panel B of Table 1 (from period 1 to period 3, each latent variable is concordant with the economic factor, E, in two periods; in periods 4 and 5, all latent variables are either concordant or discordant).

Since we are assuming that the latent variable of each debtor has equal dependence in terms of the economic factor, $c_{Ei} = c_{Ej}$ (where c_{Ei} indicates the number of concordant pairs between the economic

factor, E, and the latent variable of debtor i). The same applies to the number (d) of discordant pairs (that is, $d_{Ei} = d_{Ej}$). Hence, this condition is satisfied whenever Y_i and Y_j are concordant because the relationship between each of them and E will be always the same (this is the case of the observations of Y_1, Y_2 and Y_3 in periods 4 and 5 shown in Panel B of Table 1). On the other hand, when Y_i and Y_j are discordant, E will be necessarily concordant with one latent variable and discordant with the other one. Therefore if E is concordant with Y_i (Y_j) when the latent variables of loans i and j are discordant, E must be concordant with Y_j (Y_i) in another period when the latent variables are discordant.

4.3.2 Possible values of the rank correlation between each latent variable and the economic factor

Panel C (Table 1) represents the only case in which τ_{ij} (rank correlation for each pair of latent variables Y_i and Y_j) implies a single value of τ_{YE} (rank correlation related to each Y and E), i.e. when τ_{ij} = -1. Since Y_i and Y_j present a completely inverse behavior, all pairs in the first two columns are discordant. In this scenario, the condition $c_{Ei} = c_{Ej}$ will be met only if E is concordant with Y_i in half of the observations and concordant with Y_j in the other half such that $c_{Ei} = c_{Ej} = 0.5\,N$. Recalling that $c + d = N$, we have that $d_{Ei} = d_{Ej} = 0.5\,N$ and the Kendall's tau between E and the latent variable of each obligor (i and j) will be:

$$\tau_{YE} = \frac{c_{Ei} - d_{Ei}}{c_{Ei} + d_{Ei}} = \frac{c_{Ej} - d_{Ej}}{c_{Ej} + d_{Ej}} = \frac{0.5N - 0.5N}{N} = 0$$

So, when τ_{ij} = -1, we know for sure that, given the assumption of equal dependence between each latent variable and the single economic factor, τ_{YE} = 0.

Nonetheless, in practice, this special case (τ_{ij} = -1) is not compatible with pools of more than two assets and no other value of τ_{ij} can be mapped to a unique value of τ_{YE}. Panel D in Table 1 shows the highest rank correlation between the latent variables (τ_{ij} = 1) where all *pairs* of arrows in the first two columns point in the same direction and therefore any combination of directions in E will comply with the requirement $c_{Ei} = c_{Ej}$ (the third column of Panel D is an example of this situation). This means that if the latent variables present the strongest possible connection (τ_{ij} = 1), any value for τ_{YE} is possible.

Fortunately, credit losses tend not to be perfectly correlated and this reduce the range of feasible values of τ_{YE} when τ_{ij} can be estimated (or assumed based on some reasonable presumptions). Whenever τ_{ij} is different from -1 and 1, there will be concordant and discordant pairs of Y_i and Y_j. Panels E and F in Table 1 help us to identify the minimum and maximum possible τ_{YE} (i.e. its bounds) for a given τ_{ij} in that

interval. Both panels symbolize pairs (Y_i,Y_j) with identical observations: the first two lines are concordant and the others are discordant (the directions of the arrows are just illustrative).

From [12], it is clear that the minimum τ_{YE} in this scenario will happen when c_{Ei} (=c_{Ej}) is minimum and this happens if E is discordant with the concordant pairs (Y_i,Y_j); see the first two lines of Y_i and Y_j in Panel E where the arrows of the economic factor E have the opposite direction of the respective arrows of Y_i and Y_j. Furthermore, as explained above, when the pairs are discordant, E must be concordant with each Y half of the observations (represented in the last four lines in Panel E). From this, we deduce that the minimum number of concordant pairs between the economic factor E and a latent variable is $c_{Ei}^{\min} = c_{Ej}^{\min} = 0.5d_{ij}$, that is, half of the observations presenting discordant pairs (Y_i,Y_j). The equivalent discordant pairs will be therefore $d_{Ei}^{\min} = d_{Ej}^{\min} = c_{ij} + 0.5d_{ij}$. In Panel E, $c_{Ei}^{\min}$, for instance, is equal to $0.5d_{ij} = 0.5(4) = 2$ (which refers to the third and the fourth lines where Y_i and E are concordant) and $d_{Ei}^{\min} = c_{ij} + 0.5d_{ij} = 2 + 0.5(4) = 4$ (concerning the first, the second, the fifth and the sixth lines). From this, it follows that the minimum Kendall's tau between E and each latent variable (Y_i, for example) can be associated to the concordant and discordant pairs that generated the calculable Kendall's tau between Y_i and Y_j:

$$\tau_{YE}^{\min} = \frac{c_{Ei}^{\min} - d_{Ei}^{\min}}{c_{Ei}^{\min} + d_{Ei}^{\min}} = \frac{0.5d_{ij} - (c_{ij} + 0.5d_{ij})}{0.5d_{ij} + (c_{ij} + 0.5d_{ij})} = \frac{-c_{ij}}{c_{ij} + d_{ij}} \qquad [\ 15\ ]$$

The maximum τ_{YE} will happen when c_{Ei} (=c_{Ej}) is maximum and this occurs when E is concordant with the concordant pairs (Y_i,Y_j) as demonstrated in the first two lines of Panel F in Table 1. As before, E must be concordant with each Y half of the discordant observations (see the last four lines in Panel F). In these circumstances, the highest number of concordant pairs involving E and a latent variable is $c_{Ei}^{\max} = c_{Ej}^{\max} = c_{ij} + 0.5d_{ij}$ and the discordant pairs totalize $d_{Ei}^{\max} = d_{Ej}^{\max} = 0.5d_{ij}$. In Panel F, $c_{Ei}^{\max} = 2 + 0.5d_{ij} = 2 + 0.5(4) = 4$ (the first four lines in Panel F) and $d_{Ei}^{\max} = 0.5d_{ij} =$ $= 0.5(4) = 2$ (the last two lines). The maximum Kendall's tau relating E to each Y expressed in terms of concordant and discordant pairs between the latent variables is (taking loan i as an example):

$$\tau_{YE}^{\max} = \frac{c_{Ei}^{\max} - d_{Ei}^{\max}}{c_{Ei}^{\max} + d_{Ei}^{\max}} = \frac{(c_{ij} + 0.5d_{ij}) - 0.5d_{ij}}{(c_{ij} + 0.5d_{ij}) + 0.5d_{ij}} = \frac{c_{ij}}{c_{ij} + d_{ij}} \qquad [\ 16\ ]$$

Combining [14], where $c + d = N$, with [15] and [16], these two expressions can be rewritten respectively as[8]:

$$\tau_{YE}^{\min} = \frac{-c_{ij}}{c_{ij} + d_{ij}} = \frac{-N(\tau_{ij} + 1)/2}{N} = \frac{-(\tau_{ij} + 1)}{2} \qquad [17]$$

and

$$\tau_{YE}^{\max} = \frac{c_{ij}}{c_{ij} + d_{ij}} = \frac{N(\tau_{ij} + 1)/2}{N} = \frac{(\tau_{ij} + 1)}{2} \qquad [18]$$

This means that, when $-1 < \tau_{ij} < 1$, the rank correlation τ_{YE} between the economic factor E and each latent variable Y is always in the range whose limits are the values displayed in [17] and [18], i.e.:

$$\tau_{YE} \in [(-(\tau_{ij} + 1))/2, (\tau_{ij} + 1)/2] \qquad [19]$$

such that the smaller τ_{ij} is, the shorter the range of τ_{YE} is[9]. Note that [17] and [18] are also compatible with the extreme cases mentioned earlier (τ_{ij} = -1 and τ_{ij} = 1). Another interesting example is the possible range of τ_{YE} when the loan defaults (i.e. the latent variables) are independent. When τ_{ij} = 0, τ_{YE} may vary between -0.5 and 0.5. In other words, the independence between Y_i and Y_j does not imply that each latent variable (and consequently, the probability of default of each debtor) is free from the influence of the economy.

We emphasize that this is a crucial difference between the traditional method and our suggested approach: whilst the former has a single value for the relationship between Y and E (equality [7]), the latter is based on an interval (expression [19]).

[8] Equivalent expressions can be found if we use [13] to derive $\tau_{YE}^{\min}$ and $\tau_{YE}^{\max}$ with respect to the discordant pairs.

[9] In conformity with what was said before, the shortest range is associated with τ_{ij} = -1 (the smallest possible rank correlation between the latent variables) which results in a single value for τ_{YE} (= 0). Recall that τ_{YE} is the same for both loans i and j due to the assumption of homogeneous dependence.

4.3.3 Relationship between the latent variables' rank correlation and copulas' parameters

In principle, any value in the interval $[(-(\tau_{ij}+1))/2, (\tau_{ij}+1)/2]$ could be used to estimate the parameter of the copula that expresses the dependence between the economic factor and the latent variable at the portfolio level. This is the case of the Student t Copula. However, in the particular case of the Clayton Copula, the parameter θ_{YE} is in the interval (0,∞). Thus, according to [10], [19] becomes:

$$\tau_{YE} \in (0, (\tau_{ij} + 1) / 2] \qquad [20]$$

In a prudential regulatory context, a reasonable choice for τ_{YE} seems to be its highest value (corresponding to $(\tau_{ij}+1)/2$) since it denotes the strongest connection across the latent variables and represents the highest possible dependence among credit losses (so, the capital required will be estimated according to the worst scenario given the observed rank correlation between defaults). However this alternative may lead to the overestimation of the regulatory capital and therefore some intermediary values of τ_{YE} can be employed at the discretion of regulators and practitioners. Given the two aforementioned continuous intervals, we initially test three levels of the rank correlation between each latent variable and the economic factor: the τ_{YE} correspondent to the first tercile (tertile) in the interval[10], the average τ_{YE} and the maximum τ_{YE}. In the instance of the Clayton and the Student t Copulas, considering [20] and [19], these three levels are respectively given by:

$$\tau_{YE} = \frac{(\tau_{ij} + 1)}{6}, \ \tau_{YE} = \frac{(\tau_{ij} + 1)}{4}, \text{ and } \tau_{YE} = \frac{(\tau_{ij} + 1)}{2} \qquad [21]$$

and

$$\tau_{YE} = -\frac{(\tau_{ij} + 1)}{6}, \ \tau_{YE} = 0, \text{ and } \tau_{YE} = \frac{(\tau_{ij} + 1)}{2} \qquad [22]$$

If the Clayton Copula is adopted to represent the dependence between the economic factor and the credit losses, the capital required to cover unexpected losses with higher dependence in downturns will be

[10] That is, if we divide the continuous interval into three subintervals, this τ_{YE} will be the point between the first and the second subintervals. For example, if the range of τ_{YE} is [-0.6,0.6], the first tercile will be -0.2.

estimated by means of [8] where the parameter θ_{YE} will be defined in accordance with the level of the rank correlation between credit losses and the economic factor. The three levels presented in [21] combined with [10] give the following expressions for θ_{YE}:

$$\theta_{YE}^{C} = \frac{2(\tau_{ij}+1)}{5-\tau_{ij}},\ \theta_{YE}^{C} = \frac{2(\tau_{ij}+1)}{3-\tau_{ij}},\ \text{and}\ \theta_{YE}^{C} = \frac{2(\tau_{ij}+1)}{1-\tau_{ij}} \qquad [\ 23\ ]$$

When the Student t Copula is used, the extreme losses are estimated via [9] and the three levels of θ_{YE} are (by combining [11] and [22]):

$$\theta_{YE}^{t} = \sin\left(-\frac{(\tau_{ij}+1)\pi}{12}\right),\ \theta_{YE}^{t} = 0,\ \text{and}\ \theta_{YE}^{t} = \sin\left(\frac{(\tau_{ij}+1)\pi}{4}\right) \qquad [\ 24\ ]$$

Bear in mind that τ_{ij} is the observable (computable) rank correlation (Kendall's tau) across probabilities of default (default rates) and can be determined in the same way the linear correlation in [6] was defined by several credit classes in Basel Accords.

5. COMPARISON BETWEEN THE PERFORMANCE OF THE BASEL METHOD AND THE PERFORMANCE OF THE SUGGESTED APPROACH

5.1 Data and test setup

We use aggregate data on (non-seasonally adjusted) default rates regarding all American commercial banks to compare the performance of the traditional formula [6] with the performance of the method based on the Clayton and the Student t Copulas (formulas [8] and [9], respectively). We consider three credit classes: mortgages, credit cards and corporate loans; the first comprises the period 1991Q1-2011Q2 and the last two pertain to 1985Q1-2011Q2. The data was downloaded from the Federal Reserve Economic Data (FRED) compiled by the Federal Reserve Bank of St. Louis[11]. We assume that this aggregate data represents the default rates of "average" (typical) American banks. The evolution of the default rates is presented in Figure 3.

[Insert Figure 3 here]

[11] Available at http://research.stlouisfed.org/fred2/categories/23.

In this section, we check whether models based on copulas that express left-tail dependence could predict the high losses observed in the period 2009Q1-2010Q2 better than the Basel formula did[12]. Since our data is net of recoveries, we multiply each of the formulas related to the extreme losses ([6], [8] and [9]) by the loss given default (*LGD*) which, in turn, is calibrated according to other empirical studies. As for corporate loans, the average *LGD* according to the values found in some prior studies[13] is 30.57%. Among these papers, only Grossman et al. (1997) present specific results for mortgages (*LGD* = 42% based on the present value of the repayments and *LGD* = 29.8% based on the nominal value of the loans). There is no particular investigation for credit cards and, in this case, we use the average *LGD* (34.95%) for unsecured bank loans estimated by Carty and Lieberman (1997) and Emery (2003). So, the *LGD* used in our tests are 35%, 40% and 30% for credit cards, mortgages and corporate loans, respectively.

Given that we are using aggregate data and do not have enough information to estimate the rank correlation τ_{ij} across the latent variables (and then to calculate the parameters θ_{YE} of the copulas that link the single economic factor and each latent variable by means of formulas [23] and [24]), we use the correlation coefficients (ρ_{ij} in our notation) empirically estimated by the Basel Committee on Banking Supervision (see BCBS 2005) to find the correspondent rank correlation (Kendall's tau, τ_{ij}) between the latent variables *i* and *j* that represents the rank correlation of all pairs of latent variables in a portfolio. The only association between ρ and τ refers to bivariate normal distributions (see Kruskal 1958):

$$\tau = \frac{2}{\pi} \arcsin \rho \qquad [\ 25\ ]$$

So, we use this relationship to estimate τ_{ij} from the values of ρ_{ij} given in Basel Accords (0.04 for credit cards, 0.15 for mortgages and a function of the probability of default for corporate debt[14]). This is clearly a limitation of our study since we had previously relaxed the assumption of normality. Nonetheless this limitation is restricted to the estimation of τ_{ij} without which we would not be able to empirically compare the Basel and the Copula methods. Moreover this simplification in our approach can be easily overcome by regulators and practitioners who have enough data to estimate the rank correlation across the

[12] The last four quarters in our data set are neglected due to the considerable decrease in the default rates after 2010Q2.

[13] Asarnow and Edwards (1995), Carty and Lieberman (1997), Grossman et al. (1997), Eales and Bosworth (1998), Felsovalyi and Hurt (1998), Emery (2003), Altman (2004), Araten (2004), Franks et al. (2004), Schuermann (2004), Dermine and Neto de Carvalho (2006), and Grunert (2009).

[14] ρ is also a function of obligors' size (annual sales) and the estimates presented ahead are based on the maximum size stipulated in Basel (€50 million). We tested other sizes (results not displayed) but the relative performance of the methods compared was virtually the same.

latent variables (that is, those agents can estimate τ_{ij}, for example, in the same way the Basel Committee on Banking Supervision estimated the values of ρ_{ij} defined in the second Basel Accord).

We initially assume we are in 2008Q4 and we therefore use the default rates up to that quarter to estimate the average (historical) probability of default (*PD*). Hence, the first estimate for each credit class refers to 2009Q1. Then, for each subsequent period, we update the historical *PD* by including the periods between 2008Q4 and the period immediately before the period analyzed. For instance, when estimating the extreme losses in 2010Q1, we calculated the historical *PD* as the average of the default rates from the first period in our sample until 2009Q4.

5.2 Results

The results are shown in Table 2 where the potential extreme losses are estimated via the conventional (Basel) method (expression [6] * *LGD*) and the Clayton Copula method (expression [8] * *LGD*). Each panel pertains to a loan category (credit cards, mortgages, and corporate). The second column displays the default rates observed in the periods of higher default levels (2009Q1-2010Q2). The estimates based on the Basel formula with the usual confidence (99.9%) are in the third column. In the fourth column, we raise the confidence in the Basel formula (99.99%) to test if we can improve the performance of the Basel method when it leads to the underestimation of losses. Estimates related to the Clayton Copula with confidence of 99% are in the fifth, sixth and the seventh columns (following the three levels of θ_{YE} mentioned in [23]). The best estimate for each period and credit class is highlighted in boldface.

By calculating the *absolute* difference between the observed default rates and the estimates, we can see that, for credit cards (Panel A), the formula derived from the Clayton Copula with parameter θ_{YE} related to the first tercile in the τ_{YE} range yielded the best results for the period 2009Q1-2009Q4 (exactly when the default rates reached the highest level) besides outperforming the Basel formula (confidence 99.9%) in 2010Q1 while the Basel formula with the increased confidence 99.99% gave the best results for 2010Q1-2010Q2. Note, however, that despite resulting in the closest values to the observed default rates, the Basel approach with confidence 99.99% underestimated the extreme losses in all quarters whereas the copula method presented slightly overestimated losses.

Concerning mortgages (Panel B), the Basel method had its best performance since it gave the best estimates in four quarters (2009Q1-2009Q2 at the 99.99% confidence level and 2010Q1-2010Q2 at the 99.9% confidence level). The Clayton Copula yielded the best approximations in 2009Q3 (with the average rank correlation in the interval of all possible τ_{YE}) and 2009Q4 (with the first tercile in the rank correlation interval).

As for corporate loans (Panel C), contrary to what could be expected, the conventional Basel formula (confidence 99.9%) overestimated the credit losses in the downturn period. This does not represent a concern for regulators and practitioners since the institutions that presented losses at the average (aggregate) level and calculated the potential extreme losses by means of the Basel formula would have (more than) enough capital to cover those losses. Nonetheless the estimates based on the Clayton Copula (with the first tercile in the τ_{YE} interval) were even better as they had the smallest difference from the observed default rates for the whole period 2009Q1-2010Q2 (i.e. the overestimation from the copula method was lower than that resulted from the Basel formula).

Table 3 presents the comparison between results from the Basel (expression [6] * *LGD*) and the Student t Copula[15] (expression [9] * *LGD*) approaches. The performance of the Student t Copula is quite similar to the performance of the Clayton Copula shown in Table 2. In Panel A (credit cards), the estimates derived from the Student t Copula (at the 99% confidence level) with parameter based on the first tercile of τ_{YE} were the best approximation of the observed extreme losses in five periods (2009Q1-2010Q1). The Basel formula with confidence 99.99% gave the best result in 2010Q2. With regard to mortgages (Panel B), the copula method did not yield the best estimate in any of the periods analyzed. The results for corporate obligations (Panel C) were essentially the same as those displayed in Table 2 for this credit category: albeit the Basel formula did not result in insufficient capital to cover losses in the adverse scenarios considered in this study, the estimates from the copula approach with parameter inferred from the first tercile of τ_{YE} were closer to the observed losses and therefore avoided excessively unnecessary capital.

Thus, methods derived from the Clayton and the Student t Copulas with parameter θ_{YE} based on the first tercile of τ_{YE} would be better alternatives than the Basel formula (even with increased confidence 99.99%) to predict the unusually high losses observed in credit card and corporate loan portfolios in the period 2009Q1-2010Q2.

We tested other levels for the rank correlation τ_{YE} but, in general, none of them outperformed the first tercile (which was the level that yielded the best results for the copula approaches in Tables 2 and 3).

6. CONCLUSIONS

We show that the formula used in Basel Accord to estimate unexpected credit losses corresponds to a conditional distribution derived from the Gaussian Copula. Since this copula family does not capture tail dependence, the model largely used by regulators may misestimate the capital necessary to face credit

[15] With the minimum degree of freedom (v = 1) to assume the fattest possible tails. Other values of v can also be considered by practitioners.

losses in downturns (when the connection across defaults tends to be more intense than in periods of normal economic activity).

Based on this finding, we propose two models that keep the basic idea of the current method (namely, the first derivative of a copula) but we use different conditional distributions able to detect possible tail dependence among losses in adverse conditions. The suggested approach is flexible and can capture several dependence shapes since it can be adapted to a number of differentiable copulas. Its implementation is as simple as the implementation of the existent model and tends to identify potential higher association between losses in downturns better than the traditional approach does.

There are typically many possible rank correlations between the economic factor and the latent variable of each loan (called τ_{YE} in this paper) for each rank correlation across loans (named here as τ_{ij}). If the losses have small rank correlation, the models proposed get more accurate because the range of possible associations between the economic factor and each latent variable tends to be shorter than intervals resulted from high rank correlation between the latent variables. So, the variation of potential outcomes is reduced for low rank correlations across defaults and we move towards a unique value of τ_{YE}.

It is possible that many trials to insert copulas in this Basel framework have failed due to the lack of a link between the dependence measure we need (τ_{YE}) and the dependence we can infer from empirical data (τ_{ij}). Therefore the relationship between those two measures found in this study will certainly contribute to the application of copulas to many models dealing with dependence among variables impacted by systematic (unobservable) factors.

We test the proposed models with data pertaining to aggregate credit losses in all American commercial banks. Our results reveal that the copula methods yielded better estimates of extreme losses for credit cards (for which the Basel formula underestimates losses and the copula approaches typically present closer estimates slightly above the observed losses) and corporate loans (for which both approaches overestimate the losses but the copula one give results closer to the observed losses). On the other hand, in the case of mortgage portfolios, the estimates founded on the Basel formula are more precise than the estimates based on copulas.

The different performance of the Basel formula with regard to those three credit classes might be consequence of the distinct correlation coefficient specified in Basel Accords for each segment. If this is the case, regulators should rethink the calibration of the correlation coefficients (especially concerning credit cards for which Basel formula presented the worst results in terms of underestimation). So, this could be an alternative to keep the use of the Basel formula (based on assumptions of normality) instead of adopting copula methods. The comparison between these two possibilities is beyond the scope of this

paper and is left as a future exercise since it demands more data to estimate the linear correlation among defaults.

Another topic for further investigation is to estimate τ_{ij} from empirical data sets rather than to approximate it by using the relationship between ρ and τ in bivariate normal distributions (as we did by means of expression [25]). Moreover, our suggestions can be extended by regulators and practitioners who have access to massive data on credit losses so that the dependence across defaults can be empirically found and the first derivative of the resultant copulas (if it exists) should be used to give more realistic estimates of unexpected losses according to the properties of each portfolio.

Finally, we recognize that the approach suggested in this paper is static in the sense that a single copula family is used regardless of the economic scenario. Although most of our results are better than estimates from the Basel formula for the six quarters considered, estimations regarding the period pre-2009 (not displayed here) show that the Clayton and the Student t copulas lead to an (excessive) overestimation of losses when the economy is booming. So, the next challenge is to build a dynamic (regime-switch) model where different copula families can be chosen over time according to, for example, some macroeconomic indicators.

ACKNOWLEDGMENTS

I am grateful for helpful comments from Johan Walden, Steffen Juranek, José Feria-Dominguez and Luiz Cláudio Barcelos.

REFERENCES

Aas, K., C. Czado, A. Frigessi, H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance Mathematics & Economics* , 44, pp. 182-198.

Altman, E., A. Resti, A. Sironi (2004). Default Recovery Rates in Credit Risk Modelling: A Review of the Literature and Empirical Evidence. *Economic Notes by Banca Monte dei Paschi di Siena SpA*, Vol. 33, No. 2, pp. 183-204.

Ang, A., Bekaert, G. (2002). International Asset Allocation With Regime Shifts. *Review of Financial Studies*, 15, pp. 1137-1187.

Araten, M., M. Jacobs Jr., P. Varshney (2004). Measuring LGD on Commercial Loans: An 18-Year Internal Study. *The RMA Journal*, May, pp. 28-35.

Asarnow, E., D. Edwards (1995). Measuring loss on defaulted bank loans. A 24-year study. *Journal of Commercial Lending*, 77 (7), pp. 11-23.

Basel Committee on Banking Supervision - BCBS (2005). An Explanatory Note on the Basel II IRB Risk Weight Functions. *Bank for International Settlements*.

Basel Committee on Banking Supervision - BCBS (2006). International Convergence of Capital Measurement and Capital Standards: A Revised Framework. *Bank for International Settlements*.

Bluhm, C., L. Overbeck, C. Wagner (2002). *An Introduction to Credit Risk Modeling*. London: Chapman & Hall/CRC.

Bouyé, E., M. Salmon (2009). Dynamic copula quantile regressions and tail area dynamic dependence in Forex markets. *European Journal of Finance*, Vol. 15, Issue 7, pp. 721-750.

Carty, L., D. Lieberman (1997). Corporate bond defaults and default rates. *Moody's Investors Service*, July.

Chan-Lau, Jorge A. (2010). Regulatory Capital Charges for Too-Connected-to-Fail Institutions: A Practical Proposal. *Financial Markets, Institutions and Instruments*, 19, Issue 5, pp. 355–379.

Crouhy, M., D. Galai, R. Mark (2000). A comparative analysis of current credit risk models. *Journal of Banking & Finance*, 24, pp. 59-117.

Czado, Claudia (2010). Pair-copula constructions for multivariate copulas. In: Jaworki, P., F. Durante, W. Härdle, T. Rychlik (eds.) "*Workshop on Copula Theory and its Applications*". Dortrecht: Springer, pp. 93-110.

Das, S., G. Geng (2006). Correlated Default Processes: A Criterion-Based Copula Approach. In: Fong, H. (ed.) *"The Credit Market Handbook. Advanced Modeling Issues"*. Hoboken, New Jersey: John Wiley & Sons, pp. 186-218.

Dermine, J., C. Neto de Carvalho (2006). Bank loan losses-given-default: A case study. *Journal of Banking & Finance*, 30, pp. 1219-1243.

Di Clemente, A., Romano, C. (2004). Measuring and Optimizing Portfolio Credit Risk: A Copula-based Approach. *Economic Notes* (by Banca Monte dei Paschi di Siena SpA), Vol. 33, no 3, pp. 325-357.

Eales, R., E. Bosworth (1998). Severity of loss in the event of default in small business and larger consumer loans. *The Journal of Lending and Credit Risk Management*, pp. 58-65.

Embrechts, P., A. McNeil, D. Strautman (2002). Correlation and Dependency in Risk Management: Properties and Pitfalls. In: Dempster, M.A.H. (ed.) "*Risk Management: Value at Risk and Beyond*". Cambridge: Cambridge University Press, pp. 176-223.

Emery, Kenneth (2003). Moody's Loan Default Database as of November 2003. *Moody's Investors Service*, December.

Felsovalyi, A., L. Hurt (1998). Measuring loss on Latin American defaulted bank loans: A 27-year study of 27 countries. *Journal of Lending and Credit Risk Management*, 80, pp.41-46.

Franks, J., A. de Servigny, S. Davydenko (2004). A comparative analysis of the recovery process and recovery rates for private companies in the UK, France, and Germany. *Standard and Poor's Risk Solutions*, June.

Grossman, R.,W. Brennan, J. Vento (1997). Syndicated bank loan recovery study. *Structured Finance Credit Facilities Report, Fitch IBCA*, October.

Grunert, J., M. Weber (2009). Recovery rates of commercial lending: Empirical evidence for German companies. *Journal of Banking and Finance*, 33, pp. 505-513.

Hull, J., A. White (2004). Valuation of a CDO and an nth to Default CDS Without Monte Carlo Simulation. *Journal of Derivatives*, 12, 2, Winter.

Joe, Harry (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. *IMS Lecture Notes*, Vol. 28, pp. 120-141.

Joe, Harry (1997). *Multivariate Models and Dependence Concepts.* London, New York: Chapman & Hall/CRC.

Kang, W., P. Shahabuddin (2005). Fast simulation for multifactor portfolio credit risk in the t-copula model. *Proceedings of the 2005 Winter Simulation Conference*, pp. 1859-1868.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistics Association*, 53, pp. 814-861.

McNeil, A. J., R. Frey, P. Embrechts (2005), *Quantitative Risk Management. Concepts, Techniques and Tools*. Princeton, UK: Princeton University Press.

Nelsen, Roger B. (2006). *An Introduction to Copulas.* New York: Springer. 2nd edition.

Perli, R., W. Nayda (2004). Economic and regulatory capital allocation for revolving retail exposures. *Journal of Banking & Finance*, 28, pp. 789-809.

Schönbucher, Philipp J. (2000). Factor Models for Portfolio Credit Risk. *Discussion Paper* 16/2001, Bonn Graduate School of Economics, University of Bonn.

Schuermann, Til (2004). What Do We Know About Loss Given Default? In: Shimko, D. (Ed.). *Credit Risk Models and Management*, London, UK, Risk Books, 2nd Edition, pp. 249-274.

Vasicek, Oldrich A. (1991). Limiting Loan Loss Probability Distribution. *KMV Corporation Report*.

Vasicek, Oldrich A. (2002). The Distribution of Loan Portfolio Value. *Risk*.

TABLE 1: Representation of the movements of latent variables and one systematic factor

Panel A
A general case

Period	Y_1	Y_2	Y_3	…	Y_L	E
1	↑	↓	↑	…	↑	↓
2	↓	↑	↓	…	↓	↑
3	↑	↓	↑	…	↑	↑
…	…	…	…	…	…	…
N	↓	↓	↓	…	↑	↓

Panel B
Example of homogeneous dependence between each Y_i and E

Period	Y_1	Y_2	Y_3	E
1	↑	↑	↓	↑
2	↑	↓	↑	↑
3	↓	↑	↑	↑
4	↑	↑	↑	↑ or ↓
5	↓	↓	↓	↑ or ↓

Panel C
$\tau_{ij} = -1$

Y_i	Y_j	E
↑	↓	↑
↓	↑	↓
↑	↓	↑
…	…	…
↑	↓	↓

Panel D
$\tau_{ij} = 1$

Y_i	Y_j	E
↑	↑	↓
↓	↓	↑
↑	↑	↑
…	…	…
↑	↑	↓

Panel E
$-1 < \tau_{ij} < 1$

Y_i	Y_j	E
↑	↑	↓
↓	↓	↑
↑	↓	↑
↑	↓	↑
↑	↓	↓
↑	↓	↓

Panel F
$-1 < \tau_{ij} < 1$

Y_i	Y_j	E
↑	↑	↑
↓	↓	↓
↑	↓	↑
↑	↓	↑
↑	↓	↓
↑	↓	↓

Y_i, for $1 \leq i \leq L$, is the latent variable associated to the ith loan. E is the systematic (economic) factor. τ_{ij} is the rank correlation (Kendall's tau) between Y_i and Y_j. ↑ and ↓ indicate the direction of movements of the variables (up and down, respectively).

TABLE 2: Comparison between extreme credit losses estimated via Basel method and the Clayton Copula

Year/ Quarter	Observed extreme losses	Basel estimate confidence = 99.9%	Basel estimate confidence = 99.99%	Copula estimate with the first tercile of τ_{YE} *	Copula estimate with average τ_{YE} *	Copula estimate with maximum τ_{YE} *
Panel A: Credit Cards						
2009Q1	0.1010	0.0467	0.0571	**0.1124**	0.1779	0.3284
2009Q2	0.1012	0.0472	0.0576	**0.1133**	0.1789	0.3289
2009Q3	0.1016	0.0476	0.0582	**0.1140**	0.1800	0.3295
2009Q4	0.1097	0.0481	0.0587	**0.1148**	0.1810	0.3300
2010Q1	0.0855	0.0486	**0.0593**	0.1157	0.1821	0.3305
2010Q2	0.0770	0.0489	**0.0596**	0.1162	0.1828	0.3308
Panel B: Mortgages						
2009Q1	0.0243	0.0169	**0.0288**	0.0164	0.0192	0.0040
2009Q2	0.0285	0.0182	**0.0308**	0.0181	0.0219	0.0057
2009Q3	0.0245	0.0198	0.0331	0.0201	**0.0251**	0.0080
2009Q4	0.0214	0.0210	0.0349	**0.0217**	0.0277	0.0104
2010Q1	0.0191	**0.0220**	0.0364	0.0230	0.0299	0.0126
2010Q2	0.0199	**0.0228**	0.0376	0.0241	0.0317	0.0148
Panel C: Corporate Loans						
2009Q1	0.0254	0.0393	0.0627	**0.0364**	0.0561	0.0872
2009Q2	0.0265	0.0397	0.0632	**0.0370**	0.0571	0.0907
2009Q3	0.0189	0.0401	0.0637	**0.0375**	0.0582	0.0945
2009Q4	0.0176	0.0403	0.0639	**0.0378**	0.0587	0.0965
2010Q1	0.0172	0.0405	0.0641	**0.0381**	0.0592	0.0983
2010Q2	0.0144	0.0406	0.0643	**0.0383**	0.0597	0.0999

* τ_{YE} stands for the rank correlation (Kendall's tau) between the latent variable of each obligor and the economic factor.

The best estimate (which presents the smallest difference from the observed extreme losses) in each period are highlighted in boldface.

TABLE 3: Comparison between extreme credit losses estimated via Basel method and the Student t Copula

Year/ Quarter	Observed extreme losses	Basel estimate confidence = 99.9%	Basel estimate confidence = 99.99%	Copula estimate with the first tercile of τ_{YE} *	Copula estimate with average τ_{YE} *	Copula estimate with maximum τ_{YE} *
Panel A: Credit Cards						
2009Q1	0.1010	0.0467	0.0571	**0.0956**	0.1365	0.2767
2009Q2	0.1012	0.0472	0.0576	**0.0960**	0.1370	0.2771
2009Q3	0.1016	0.0476	0.0582	**0.0964**	0.1374	0.2775
2009Q4	0.1097	0.0481	0.0587	**0.0967**	0.1379	0.2779
2010Q1	0.0855	0.0486	0.0593	**0.0971**	0.1383	0.2783
2010Q2	0.0770	0.0489	**0.0596**	0.0973	0.1386	0.2785
Panel B: Mortgages						
2009Q1	0.0243	0.0169	**0.0288**	0.0054	0.0067	0.0045
2009Q2	0.0285	0.0182	**0.0308**	0.0065	0.0081	0.0058
2009Q3	0.0245	**0.0198**	0.0331	0.0078	0.0100	0.0076
2009Q4	0.0214	**0.0210**	0.0349	0.0090	0.0116	0.0094
2010Q1	0.0191	**0.0220**	0.0364	0.0100	0.0130	0.0110
2010Q2	0.0199	**0.0228**	0.0376	0.0108	0.0142	0.0125
Panel C: Corporate Loans						
2009Q1	0.0254	0.0393	0.0627	**0.0250**	0.0368	0.0733
2009Q2	0.0265	0.0397	0.0632	**0.0256**	0.0377	0.0767
2009Q3	0.0189	0.0401	0.0637	**0.0263**	0.0387	0.0802
2009Q4	0.0176	0.0403	0.0639	**0.0266**	0.0393	0.0822
2010Q1	0.0172	0.0405	0.0641	**0.0269**	0.0397	0.0838
2010Q2	0.0144	0.0406	0.0643	**0.0272**	0.0401	0.0854

* τ_{YE} stands for the rank correlation (Kendall's tau) between the latent variable of each obligor and the economic factor.

The best estimate (which presents the smallest difference from the observed extreme losses) in each period are highlighted in boldface.

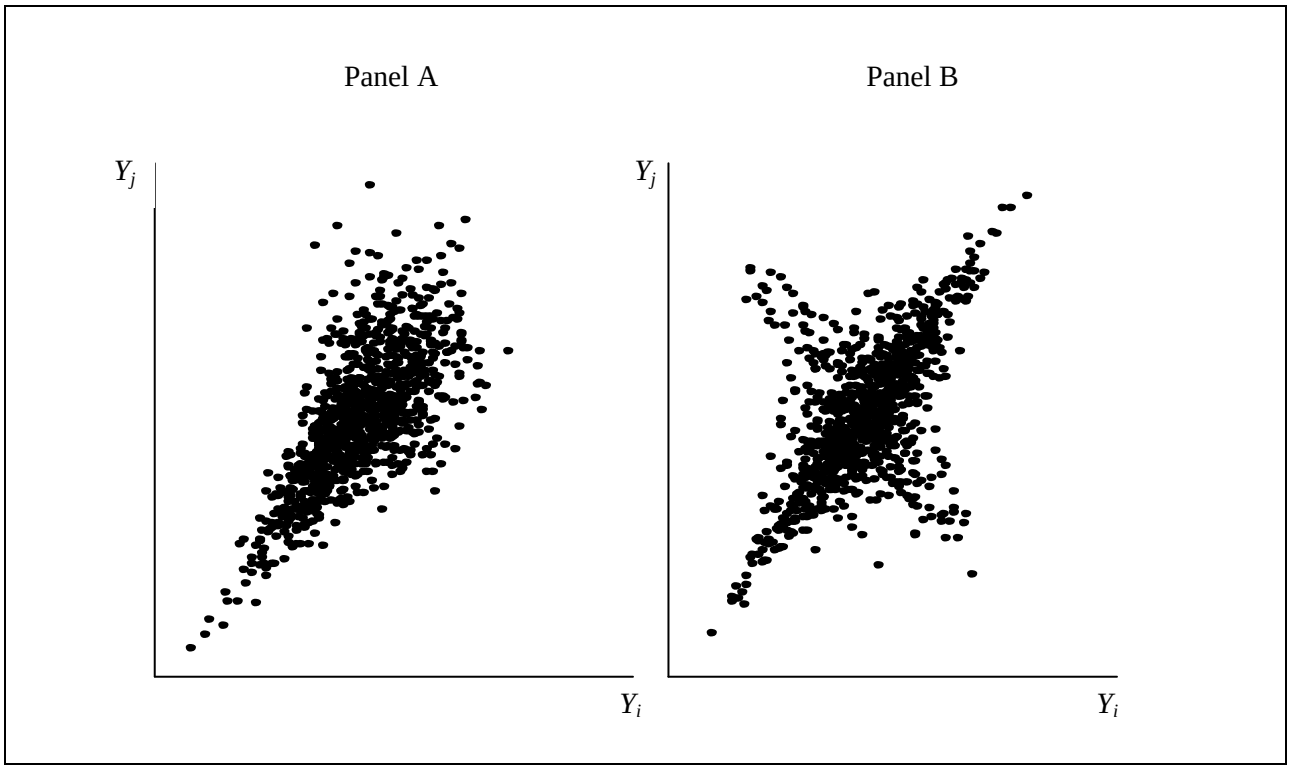

FIGURE 1: Two possible representations of the stronger connection across small values of the latent variables (Y_i and Y_j) in downturns.

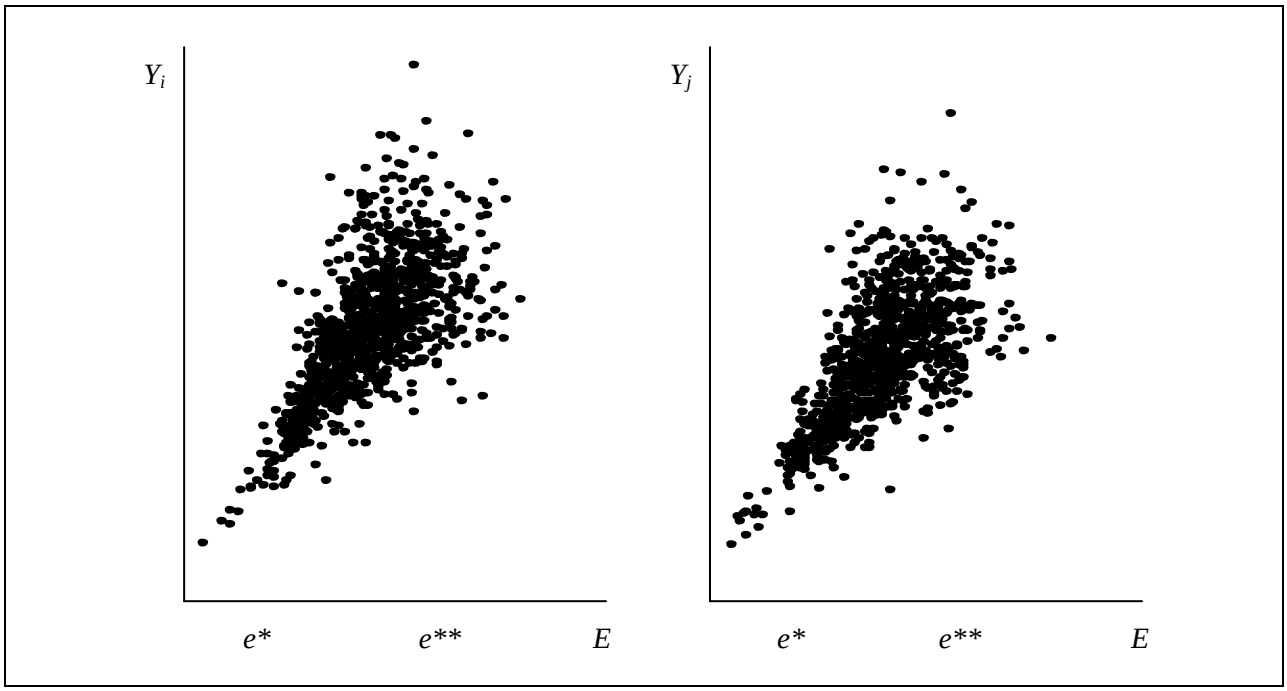

FIGURE 2: A representation of the stronger connection across small values of each latent variable (Y_i on the left and Y_j on the right) and the economic status. When E is reduced (e^*), indicating an unfavorable scenario, both Y_i and Y_j tend to be small. When E increases (e^{**}), denoting higher economic activity, different levels of the latent variables are associated.

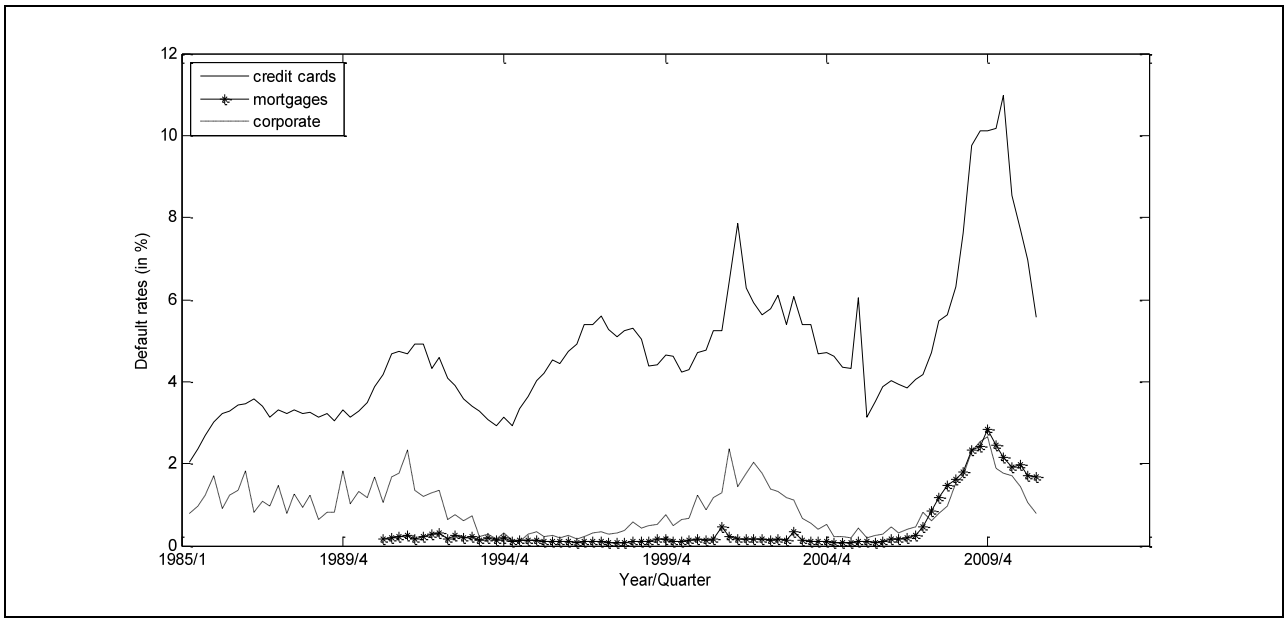

FIGURE 3: Evolution of default rates in American commercial banks (1985Q1-2011Q2).