# ENHANCING BASEL METHOD VIA CONDITIONAL DISTRIBUTIONS THAT CAPTURE STRONGER CONNECTION AMONG CREDIT LOSSES IN DOWNTURNS

**Fernando F. Moreira**
**Credit Research Centre, University of Edinburgh**
**Email: F.F.Moreira@sms.ed.ac.uk**

## ABSTRACT

This paper suggests a formula able to capture potential stronger 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 (stronger dependence across higher losses) to estimate the probability of credit losses in extreme scenarios (crises). Next, we present an example based on a specific copula that indicates upper-tail dependence among credit losses. Simulations show that, for both credit classes tested (retail and corporate), the alternative method outperforms the Basel formula which, in turn, is prone to result in insufficient capital when the losses have positively-skewed distributions (long tail in the right side) and are upper-tail dependent (a realistic representation of credit portfolios, according to the literature). The method proposed is extendable to any differentiable copula family which gives flexibility to future practical applications of the model.

JEL codes: G28, G21, G32, C46, C49
Keywords: Financial Regulation, Basel Accords, credit risk, conditional distributions, copulas.

# 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 across losses. However, these assumptions do not allow the identification of possible stronger dependence across losses in extreme scenarios (which seems to be the case for several financial assets, loans included) and, therefore, the Basel method may underestimate 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 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 an example, we present a formula originated from the Clayton Copula that is able to detect stronger connection (tail dependence) among low values of latent variables (which is equivalent to identify higher dependence among high credit losses). The resultant formula to calculate unexpected (extreme) credit losses does not assume any kind of distribution for the variables considered and therefore overcomes the limitations of the existing method with regard to the assumption of normality and the use of the linear correlation.

Simulations corroborate our hypothesis that the Basel formula tends to underestimate the likelihood of joint extreme defaults when losses present positively-skewed distributions and right-tail dependence (which, according to the pertinent literature[1], characterize loan portfolios held by financial institutions) and show that the alternative approach yields better estimates of unexpected losses when compared to Basel.

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.

This paper is organized as follows. In Section 2, we introduce copula functions and explain how conditional distributions can be derived from them. Then, we present two derivations of the Basel formula used to estimate extreme credit losses. Section 4 contains an example of conditional distributions that can capture potential asymmetric dependence across losses. Simulations are used in Section 5 to compare the performance of the proposed model to the performance of the Basel model. Section 6 concludes.

## 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:

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

[1] Mentioned ahead.

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

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

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), and Bouyé and Salmon (2009). Two families of particular interest here are the Gaussian (Normal) and the Clayton that respectively generate the conditional distributions stated in [1] and [2]:

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

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

where $\Phi$ and $\Phi^{-1}$ 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 $\theta_{12}$ is the Gaussian Copula parameter[3] between $X_1$ and $X_2$.

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

where $F(.|.)$ is the distribution of $X_1$ conditional on $X_2$, $F(.)$ is an unconditional distribution and $\theta$ is the Clayton Copula parameter between $X_1$ and $X_2$.

Note that this same concept of conditional distributions is treated in Bouyé and Salmon (2009) as *nonlinear quantile regressions*. The models proposed here are equivalent to quantile regressions but we do not use this terminology to keep the discussion as simple as possible.

The strength of the dependence (copula) is expressed by a parameter $\theta$ which is closely related to rank correlations Kendall's tau ($\tau$). 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 $\theta$ of their representative copula can be inferred from[4]:

$$\tau = 4\iint_{[0,1]^2} C(u_1,u_2)dC(u_1,u_2) - 1 \qquad [3]$$

[3] The parameter of the Gaussian Copula is usually represented by $\rho$. We adopt the notation $\theta$ 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.

[4] The proof is given in Nelsen (2006, chapter 5).

## 3. BASEL METHOD: TWO DERIVATIONS

### 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:

$$[LGD * K_V - LGD * PD] * Maturity = [LGD * (K_V - PD)] * Maturity \qquad [\ 4\ ]$$

where $LGD$ is the "loss given default", i.e. the percentage of exposure the lender will lose if borrowers default and $PD$ stands for 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 [4], $K_V$, is the expected default rate at the 99.9% percentile of the $PD$ distribution ("Vasicek Formula") - see Vasicek (1991, 2002) - and is calculated as:

$$K_V = \Phi\left(\frac{\Phi^{-1}(PD) + \sqrt{\rho}\,\Phi^{-1}(0.999)}{\sqrt{1-\rho}}\right) \qquad [\ 5\ ]$$

where:

$\Phi$ and $\Phi^{-1}$ represent the standard normal cumulative distribution and its inverse, respectively;

$PD$ is the probability of default of the loan portfolio (average);

$\Phi^{-1}(0.999)$, which is equal to $-\Phi^{-1}(0.001)$, is the level of the economy (confidence) chosen to represent an extreme scenario when unexpected losses may occur. Therefore, the systematic factor is assumed to be normally distributed; and

Rho ($\rho$) 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 function of $PD$ and (in the case of corporate debt) the size of debtors (measured in terms of annual sales). Thus, for retail loans, the correlation is given by[5]:

$$\rho \; = \; 0.03 * \left(\frac{1 - e^{-35*PD}}{1 - e^{-35}}\right) + 0.16\left(1 - \frac{1 - e^{-35*PD}}{1 - e^{-35}}\right) \qquad [\ 6\ ]$$

And, for corporate debts, the correlation is calculated as:

$$\rho \; = \; 0.12 * \left(\frac{1 - e^{-50*PD}}{1 - e^{-50}}\right) + 0.24\left(1 - \frac{1 - e^{-50*PD}}{1 - e^{-50}}\right) - 0.04 * \left(1 - \left(\frac{Size - 5}{45}\right)\right) \qquad [\ 7\ ]$$

where “Size” (in € million) refers to the obligors’ size and is applied for annual sales between €5 million and €50 million.

### 3.2 Derivation from factor models

Some leading industry credit risk models, such as CreditMetrics® and KMV®, rely on the presumptions of structural models (initially proposed by Merton, 1974) according to which an obligor defaults when a latent variable associated to it (typically interpreted as the log-returns of its assets) falls below a threshold (the amount needed to pay the outstanding debt).

The dependence across defaults of different obligors is estimated in line with factor models which assume that the correlation among defaults is driven by the debtors’ latent variables (see, for instance, Crouhy et al., 2000 and Bluhm et al., 2002). Such underlying variables are impacted by common (systematic) factors that affect all obligors and specific (idiosyncratic) factors that have effect only on the respective borrowers.

The idiosyncratic factors are assumed to be independent from one another and therefore do not contribute to asset return correlations which are exclusively determined by the systematic factors.

To illustrate this idea, consider a case based on an example given by Bluhm et al. (2002). If two automotive companies A and B operating in country C are debtors, the ability of those firms to

[5] This formula does not apply to revolving and mortgage credits, for which the correlations were specified in Basel as 0.04 and 0.15, respectively.

pay their obligations is likely affected *in the same direction* by the underlying factor *automotive industry*. That is, if the activity in that sector falls, the default probability of A and B increases simultaneously. Another aspect that certainly have influence on the performance of those companies is the country C's economic level. So this is another systematic factor that may change the default probability of A and B in the same way. In contrast, if the firm A's CEO steps down or one of its factories is flooded, this event will, in principle, impact only the default likelihood of A (not B's). Hence, this would be an idiosyncratic risk of A.

Naturally, there are many common factors that act together and influence debtors' situation. However this model may be simplified if we consider that the asset returns of all borrowers are driven by only one common factor (the "economic status"). The latent variable ($Y$), the single systematic factor ($E$), and the specific factor ($\varepsilon$) are assumed to be standardized normally distributed. Also, each idiosyncratic risk is uncorrelated with the systematic risk and the specific risks of all other obligors. For simplicity, all pairs of asset returns (*i* and *j*) are considered to present the same correlation ($\rho_{ij}$). The correlation between the systematic factor and the asset return of each debtor is denoted $\rho_{YE}$.

Owen and Steck (1962) show that equally correlated and jointly standard normal variables may be expressed as a function of their correlation coefficient and two other standard normal variables.

Thus, considering all assumptions of factor models, in the case of a single common risk, the latent variable $Y$ for a debtor $i$ may be expressed as a function of $E$, $\varepsilon$, and $\rho_{YE}$, namely:

$$Y_i = E\rho_{YE} + \varepsilon_i\sqrt{1 - \rho_{YE}^2} \qquad [\ 8\ ]$$

Due to some properties of jointly standard normal variables, we have:

$$\rho_{YE} = \sqrt{\rho_{ij}} \qquad [\ 9\ ]$$

This equality is essential to the subsequent calculations since there is usually no adequate proxy for $E$ (which is not observable) and, consequently, $\rho_{YE}$ cannot be directly estimated from empirical data. On the other hand, we can infer the correlation between asset returns, $\rho_{ij}$, from historical losses (default rates). Expression [9] is often mentioned in the literature but its derivation is rarely presented. In order to fill this gap, we present one possible derivation in Appendix A. By replacing $\rho_{YE}$ with $\sqrt{\rho_{ij}}$ in [8], we get:

$$Y_i = E\sqrt{\rho_{ij}} + \varepsilon_i\sqrt{1-\rho_{ij}} \qquad [\ 10\ ]$$

where $\sqrt{\rho_{ij}}$ and $\sqrt{1-\rho_{ij}}$ indicate how much of the variability of $Y_i$ is explained by $E$ and $\varepsilon_i$, respectively.

Apart from the doubtful presumption of normal behavior for some of these variables, the use of the linear correlation coefficient is a limitation given that it does not capture asymmetric dependence which could indicate more or less intense association across some variables in certain scenarios (see Embrechts et al., 2002).

In general, $K_V$ (defined in [5]) follows the main presumptions of factor models (see, e.g., Gordy, 2003) where each latent variable ($Y_i$) is a linear function of an unobserved single factor (systematic risk, $E$) and specific characteristics of the respective obligor (idiosyncratic risk, $\varepsilon_i$). The single factor is assumed to be standard normally distributed and impacts all obligors equally (same correlation $\rho_{YE}$). The latent variables are considered equicorrelated (same $\rho_{ij}$ for all pairs) and also follow the standard normal distribution. This leads to expression [10] mentioned above.

For each loan $i$, the probability of default is the likelihood that the latent variable $Y_i$ becomes smaller than the cutoff $y_c$, that is, $PD = \Pr[Y_i < y_c]$. Extreme credit losses happen when the economy $E$ reaches an extremely unfavourable level $e*$. In other words, these high losses are the probability of default conditional on a poor economic status. Representing this probability as $PD*$, we have $PD* = \Pr[Y_i < y_c \mid E = e*]$ and using [10] with the simplified notation for $\rho_{ij} = \rho$:

$$PD* = \Pr[E\sqrt{\rho} + \varepsilon_i\sqrt{1-\rho} < y_c \mid E = e*]$$

Solving for $\varepsilon_i$ and replacing $E$ with $e*$:

$$PD* = \Pr\left[\varepsilon_i < \frac{y_c - e*\sqrt{\rho}}{\sqrt{1-\rho}}\right]$$

As mentioned above, $\varepsilon_i$ is presumed to be normally distributed with mean 0 and variance 1. Thus, the previous equation turns into:

$$PD* = \Phi\left(\frac{y_c - e*\sqrt{\rho}}{\sqrt{1-\rho}}\right)$$

where $\Phi$ indicates the cdf of the standard normal distribution.

Since $Y_i$ is also normally distributed, $PD = \Phi(y_c)$ which implies that $y_c = \Phi^{-1}(PD)$, i.e. the cutoff of the latent variable below which default occurs is the inverse of the normal distribution, $\Phi^{-1}$, evaluated at $PD$. Basel demands confidence of 99.9% which means that the capital is supposed to be sufficient to cover the losses whenever the economy is above (better than) the 0.01 percentile of its distribution (also assumed to be normal). Hence the extreme adverse scenario $e*$ is given by $\Phi^{-1}(0.001)$. Due to two properties of the standard normal distribution (symmetry and mean 0), $\Phi^{-1}(0.001) = -\Phi^{-1}(0.999)$. Using this fact and replacing $e*$ with $-\Phi^{-1}(0.999)$ and $y_c$ with $\Phi^{-1}(PD)$ in the prior equation, we get the formula presented in Basel Accord (here the extreme loss, $PD*$, is denoted as $K_V$):

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

Readers interested in more details about this derivation of Vasicek formula ( $K_V$ ) should consult, for instance, Schönbucher (2000), Perli and Nayda (2004), and Crook and Bellotti (2010).

Some models have been proposed to transform [5] into another expression that does not have the limitation regarding the assumption of normality. Starting from [10], Hull and White (2004) relax the distributions[6] of $Y_i$, $E$ and $\varepsilon_i$, such that they can, for example, present heavy tails (which tends to increase the joint occurrences of extreme realizations of the latent variables). Representing the distributions of those three variables respectively by $F$, $G$ and $H$ and following the same steps that derived [10] from [5], the expression to estimate the probability of default conditional on an unfavorable economic status (the worst 0.1% scenario, i.e. with confidence of 99.9%) turns into:

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

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 probability of default and $\rho$ is the linear correlation between returns of obligors' assets. Obviously, the expression above cannot be solved unless the shapes of the three distributions $F$, $G$ and $H$ are known.

Some studies, such as Bluhm et al. (2002), Kostadinov (2005) and Kang (2005), have suggested the Student t distribution for $E$ and $\varepsilon_i$ to characterize the existence of more events (than the normal distribution) in the tails. In this case, it is not possible to define the distribution of the latent variable in [10] and the probability of default in downturns (at the 0.1% worst scenario) is:

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

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. In view of the impossibility of the estimation of the probability of default in adverse economic scenarios when

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

one (or more) of the variables in [10] are not normally distributed, we propose a different setup to incorporate Copula Theory into this analysis and capture potential tail dependence even if we do not know *any* of the distributions concerning the latent variable, the economic factor and the idiosyncratic factors (which is the reality in financial institutions).

### 3.3 Derivation from the Gaussian Copula

Departing 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 defaults happen and $X_2$ is the economic status (single factor). So, that formula gives the likelihood of the latent variable $X_1$ being below a specific value $x_1$ conditional on $X_2 = x_2$. Assume that both variables follow the standard normal distribution.

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

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 the latent variable cutoff[7]. $F_2(x_2) = \Phi(x_2)$ 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%; as said before, $\Phi^{-1}(0.999) = -\Phi^{-1}(0.001)$. The parameter $\theta_{12}$ 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, $\theta_{12}$ will be equal to the linear correlation between the variables (denoted here as $\rho_{12}$) which cannot be estimated given that there is no sufficient information on the economic status. Assume we can assess the linear correlation between the latent variables (based on the observed probabilities of default). Under the conditions specified (i.e. the latent variables and the economic factor follow the standard

[7] $\Phi$ represents the standard normal distribution and $\Phi^{-1}$ indicates its inverse.

normal distribution) and according to [9], $\theta_{12}$ can be associated to the linear correlation $\rho$ 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:

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

## 4. EMPLOYING ALTERNATIVE CONDITIONAL DISTRIBUTIONS TO CAPTURE TAIL DEPENDENCE

### 4.1 An example to detect higher dependence across losses in downturns

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. Recalling 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 a scatterplot like the one in Figure 1.

*[Insert Figure 1 here]*

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 this context. The correspondence between Figure 1 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). 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.

*[Insert Figure 2 here]*

Such dependence structure can be represented by, for example, the Clayton Copula and, in this case, 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]:

$$F(y_c \mid E = e^*) = \{F_E(e^*)^{\theta_{YE}}[F_Y(y_c)^{-\theta_{YE}} - 1] + 1\}^{(-1-\theta_{YE})/\theta_{YE}} \qquad [\ 11\ ]$$

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) probability of default and $\theta_{YE}$ is the copula parameter between $Y$ and $E$.

Among the three variables necessary to compute extreme losses by applying [11], two, $F_Y(y_c)$ and $F_E(e^*)$, are readily available; the former is the expected probability of default of the homogeneous portfolio and the latter is to be set according to the confidence demanded for the economic scenario[8]. 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 particular case of the Clayton Copula, this monotonically increasing behavior of $F(y_c|E)$ with respect to $\theta_{YE}$ 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 $\theta_{YE}$ and starts falling after rising up to a specific value. Therefore the calculation of the regulatory capital will yield more consistent results if the extreme economic level is restricted to percentiles smaller than or equal

[8] 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.

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 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 [11], $\theta_{YE}$, is to derive it from the rank correlation between $Y$ and $E$ (Kendall's tau, $\tau_{YE}$). As shown in [3], the Kendall's tau between two variables is associated with the parameter of the copula that represents their dependence. For some families this association has closed form (see some examples in Nelsen, 2006, chapter 5) and the Clayton Copula is one of them such that its parameter can be calculated as:

$$\theta_{YE} = \frac{2\tau_{YE}}{1 - \tau_{YE}} \qquad [\ 12\ ]$$

However we do not have enough information on $E$ to estimate $\tau_{YE}$. 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 [9]). Thus, assuming the rank correlation between the latent variables, $\tau_{ij}$, can be inferred from datasets 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 $\tau_{YE}$ and $\tau_{ij}$ so that $\theta_{YE}$ can be calculated and plugged into [11].

### 4.2 Relationship between rank correlations

Kendall's tau ($\tau$) 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., $\tau$ = *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:

$$\tau = \frac{c - d}{c + d} \qquad [13]$$

Table 1 illustrates the co-movements of the three variables considered here: $Y_i$, $Y_j$ and $E$. The arrows "↑" and "↓" indicate the direction in which the variables move. 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". Denote the number of concordant pairs as $c_{ij}$, $c_{Ei}$, and $c_{Ej}$ for the pairs $(Y_i,Y_j)$, $(E,Y_i,)$ and $(E,Y_j)$ respectively. The related number of discordant pairs will be represented by $d_{ij}$, $d_{Ei}$, and $d_{Ej}$. Let $N$ be the total number of observations (which will be obviously the same for all variables). So, *for any pair*, $c + d = N$ and, from [13], $c - d = \tau N$. Combine these two expressions, we get:

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

*[Insert Table 1 here]*

Since we are assuming that the latent variable of each debtor has equal dependence in terms of the economic factor, $c_{Ei} = c_{Ej}$ (and also $d_{Ei} = d_{Ej}$). 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 all observations of $Y_i$ and $Y_j$ in Panel B of Table 1 and the two first observations of those two variables in Panels C and D). 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 are discordant, $E$ must be concordant with $Y_j$ ($Y_i$) in another period when the latent variables are discordant.

Panel A represents the only case in which $\tau_{ij}$ (rank correlation for each pair of latent variables $Y_i$ and $Y_j$) implies a single value of $\tau_{YE}$ (rank correlation related to each $Y$ and $E$), i.e. when $\tau_{ij} = -1$. 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 $\tau_{ij} = -1$, we know for sure that, given the assumption of equal dependence between each latent variable and the single economic factor, $\tau_{YE} = 0$.

Apart from this special case, there is no mapping from $\tau_{ij}$ to a unique value of $\tau_{YE}$. Panel B shows the highest rank correlation between the latent variables ($\tau_{ij} = 1$) where all *pairs* of arrows in the two first 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 B is an example). This means that if the latent variables present the strongest possible connection, any value for $\tau_{YE}$ is possible.

Fortunately, credit losses tend not to be perfectly correlated and this reduce the range of feasible values of $\tau_{YE}$ when $\tau_{ij}$ can be estimated (or assumed based on some reasonable presumptions).

Whenever $\tau_{ij}$ is different from -1 and 1, there will be concordant and discordant pairs of $Y_i$ and $Y_j$. Panels C and D in Table 1 help us to identify the minimum and maximum possible $\tau_{YE}$ (i.e. its bounds) for a given $\tau_{ij}$ in that interval. Both panels symbolize pairs $(Y_i,Y_j)$ with identical observations: the two first lines are concordant and the others are discordant (the directions of the arrows are just illustrative).

From [13], it is clear that the minimum $\tau_{YE}$ in this scenario will happen when $c_{Ei}$ $(=c_{Ej})$ is minimum and this happens when $E$ is discordant with the concordant pairs $(Y_i,Y_j)$; see the first two lines of $Y_i$ and $Y_j$ in Panel C where the arrows of $E$ 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 C). From this, we deduce that the minimum number of concordant pairs between $E$ 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 C, $c_{Ei}^{\min}$, 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 $\tau_{YE}$ 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 two first lines of Panel D in Table 1. As before, $E$ must be concordant with each $Y$ half of the discordant observations (see the last four lines in Panel D). 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 D, $c_{Ei}^{\max} = 2 + 0.5d_{ij} = 2 + 0.5(4) = 4$ (the first four lines in Panel D) 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:

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

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 between $E$ and each latent variable, $\tau_{YE}$ (which is the same for both loans *i* and *j* due to the assumption of homogeneous dependence) 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 $\tau_{ij}$ is, the shorter the range of $\tau_{YE}$ is[9]. Note that [17] and [18] are also compatible with the extreme cases mentioned earlier ($\tau_{ij}$ = -1 and $\tau_{ij}$ = 1) and the maximum $\tau_{YE}$, for example, is respectively equal to 0 (the only admissible value for $\tau_{YE}$ when $\tau_{ij}$ = -1) and 1 (the maximum theoretical $\tau_{YE}$, which reflects the fact that $\tau_{ij}$ = 1 allows any rank correlation between $E$ and the latent variables). Another interesting example is the possible range of $\tau_{YE}$ when the loan defaults (i.e. the latent variables) are independent. When $\tau_{ij}$ = 0, $\tau_{YE}$ 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.

In principle, any value in the interval $[(-(\tau_{ij}+1))/2, (\tau_{ij}+1)/2]$ can 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. However, in the particular case of the Clayton Copula, the parameter $\theta_{YE}$ is in the interval (0,∞). Thus [19] becomes:

---

[9] In accordance to what was said before, the shortest range is associated with $\tau_{ij}$ = -1 (the smallest possible rank correlation between the latent variables) which results in a single value for $\tau_{YE}$ (= 0).

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

In a prudential regulatory context, a reasonable choice for $\tau_{YE}$ 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 worse 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 $\tau_{YE}$ can be employed at the discretion of regulators and practitioners. In the simulations ahead we will test three levels of the rank correlation between each latent variable and the economic factor: ⅓ of the maximum $\tau_{YE}$, the average $\tau_{YE}$ and the highest $\tau_{YE}$. In the instance of the Clayton Copula, considering [20] these three levels are respectively given by:

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

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 estimated by means of [11], restated below:

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

where the parameter $\theta_{YE}$ will be defined according to the level of the rank correlation between credit losses and the economic factor. Three options are ⅓ of the maximum value, the average value, and the highest value, calculated respectively as (by combining [12] and [21]):

$$\theta_{YE} = \frac{2(\tau_{ij} + 1)}{5 - \tau_{ij}}, \ \theta_{YE} = \frac{2(\tau_{ij} + 1)}{3 - \tau_{ij}}, \text{ and } \theta_{YE} = \frac{2(\tau_{ij} + 1)}{1 - \tau_{ij}} \qquad [22]$$

Recall that $\tau_{ij}$ is the observable (computable) rank correlation (Kendall's tau) across probabilities of default and can be determined in the same way the linear correlation in [5] was defined by several credit classes in Basel Accord.

This model is flexible and the copula family can be changed according to different dependence shapes empirically found. One example could be the case of credit losses that present tail dependence in a symmetric joint distribution. An option to express this situation would be the Student t Copula. Thus, [11] would be replaced with the first derivative of the Student t Copula (displayed, for example, in Aas et al., 2009) where the distribution of the latent variable evaluated at the cutoff point (= $PD$) and the chosen percentile of the economic factor would be the conditioned and the conditioning variables, respectively.

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

### 5.1 Initial comparisons

Simulations were run to check whether the estimation of unexpected losses based on formulas derived from a left-tail-dependent copula outperforms the formula used in the Basel Accord. Following the evidence presented in the literature according to which high asset losses are prone to be more connected than low losses[10], credit portfolios with right-tail-dependent losses were simulated with 520 observations each[11]. Two classes of loans were considered, retail (Table 2) and corporate (Table 3), because the Basel Accord uses different ranges of correlation for those two groups: between 0.03 and 0.16 for the former and between 0.12 and 0.24 for the latter[12]. In Basel, expressions [6] and [7] refer to correlations across *asset returns of obligors* but, given that these correlations are assumed to drive the relationship among defaults, we take the values calculated from [6] and [7] as proxies for correlations among *PDs*.

In both classes (retail and corporate), correlations were taken as decreasing functions of $PD$ (see [6] and [7]) following the idea that riskier obligors (higher $PD$) present more idiosyncratic risk

[10] See for example, Ang and Bekaert (2002), Di Clemente and Romano (2004), Das and Geng (2006), Ning (2006), Patton (2006) and Rosenberg and Schuermann (2006).

[11] This is equivalent to around 10 years of weekly data or 43 years of monthly data.

[12] The retail class in these simulations excludes revolving credit and mortgages given that they have correlation indices fixed respectively at 0.04 and 0.15 in Basel Accords. The correlation across corporate loans is also a function of the size of the obligors. In these simulations, all corporate debtors were set at the maximum size (annual sales = €50 million) stipulated in Basel.

and therefore the systematic portion of their risk is smaller which results in lower correlation (see, for instance, BCBS, 2005 and Das and Geng, 2006). So, to avoid poor performance of the Basel approach due to mispecified correlations, only portfolios with the respective levels of association determined in the Basel Accord were considered in the comparison (see second column of Tables 2 and 3 where the correlation coefficients displayed are consistent with the formulas adopted in Basel [6] and [7][13]). In order to simplify the calculations for corporate portfolios, the maturity term (see [4]) was set equal to one year.

As in the Basel approach, the portfolios are assumed to be homogeneous, i.e. all pairs have the same dependence (shape and intensity). Whilst Basel implicitly presumes the Gaussian dependence (copula), which represents a symmetric relationship without tail dependence, the simulations run in this section follow empirical evidence from the literature according to which high credit losses are typically more connected than low losses. Thus the losses in all simulated portfolios are upper-tail dependent and the Gumbel Copula was used to generate such relationship.

Following Kalyvas et al. (2006), according to whom the distributions of credit losses are skewed to the right (positive skewness), we used two distributions (beta and gamma) to represent the loss distributions[14]. The shape of such distributions is generally like the one shown in Figure 3.

[*Insert Figure 3 here*]

Ten $PD$ levels (expected losses) were tested (from 0.01 to 0.10). So, for each loan class (retail and corporate), the capital was estimated in 20 scenarios (ten $PD$ levels times two distributions). The simulation of scenario was repeated 1,000 times to eliminate potential randomness effects.

The alternative approach is illustrated by a conditional distribution derived from the Clayton Copula (as in Section 4). Since there is not a unique possible parameter for the copula between each latent variable and the economic factor, three parameters were tested. They were based on the average rank correlation ($\tau_{YE}$) between $Y_i$ and $E$, ⅓ of the maximum possible $\tau_{YE}$, and the maximum $\tau_{YE}$. The rank correlation between the latent variables of two obligors $(\tau_{ij})$ and the

[13] Difference no greater than 0.001.

[14] Both distributions were simulated such that they were positively skewed (longer tail in the right side), i.e., extremely high credit losses were farther from the mean than the extremely small losses were.

three copula parameters $(\theta_{YE})$ resulted from the three rank correlations $\tau_{YE}$ used are presented in Appendices B (retail credit) and C (corporate credit).

*[Insert Tables 2 and 3 here]*

According to what was said in Section 4.1, the calculation of the regulatory capital via the first derivative of the Clayton Copula yields more consistent results (capital increasing with the dependence across defaults) if the extreme economic level is restricted to percentiles smaller than or equal to the percentiles of the latent variables (which represents the average default probability, $PD$, of the portfolio), i.e. if $F_E(e^*) \leq PD$. As an example, we set the confidence (percentile of the economic level) equal to 99% (i.e. $F_E(e^*) = 0.01$), which complies with all $PD$ values considered.

In general, the results indicate that the formula currently used in Basel Accord tends to underestimate extreme (unexpected) losses for both retail and corporate portfolios when their losses present positively-skewed distributions and are right-tail dependent (higher losses more associated).

As for retail credit, Table 2 details the estimations for ten *PD* levels whose correlations are consistent with those determined in Basel (see the second column). The maximum *unexpected* losses (i.e. the maximum losses minus the average losses) are displayed in the third column. The next four columns present the estimates resultant from Basel approach and from the alternative method based on three possible values of the rank correlation, $\tau_{YE}$, between each obligor and the economic factor. The last four columns show the absolute difference between the maximum simulated unexpected losses and the four estimates (the values in each of these four column is added and the total is shown to indicate which approach gave the *overall* best approximation to the maximum losses). In all scenarios of both distributions tested (beta in Panel A and gamma in Panel B), at least one alternative estimate was better than the Basel evaluation. The best approximation (smallest difference between estimates and the unexpected losses) in each scenario is displayed in boldface.

Since we are following the same presumption adopted in Basel (namely, that high *PDs* are less connected than low *PDs*), the rank correlation, $\tau_{ij}$, between loans' losses is lower for higher

*PDs*. As a consequence, the maximum rank correlation, $\tau_{YE}$, between each obligor and the economic factor gets smaller as *PD* increases[15] and, logically, the same applies to the range of potential values of $\tau_{YE}$ (see [19] for all copulas and [20] for the Clayton Copula).

In complement to this fact, our results suggest that, in general, the best value of $\tau_{YE}$ in a given interval follows a similar relationship: low levels of $\tau_{YE}$ yield better results for high *PDs* and vice versa. So, in Panel A of Table 2 (beta-distributed losses), the highest level of $\tau_{YE}$ tested resulted in the best estimate of unexpected losses for the lowest *PD* (0.01). The extreme losses for the next three values of *PDs* were better captured by the intermediary $\tau_{YE}$. The lowest considered $\tau_{YE}$ gave the best outcomes for the six highest *PDs*. In Panel B (gamma-distributed losses), the three lowest *PDs* had the better results by means of the average $\tau_{YE}$. As *PD* increases, the best rank correlation becomes the smallest one (⅓ of the maximum $\tau_{YE}$).

Table 3 (pertaining to corporate loans) confirmed the pattern showed in Table 2 for both distributions.

### 5.2 Calibration of the dependence between the latent variables and the economic factor

Consonant with the reasoning that supports the Basel model (i.e. high *PDs* are less subject to systematic risk and therefore less connected than low *PDs*), Tables 2 and 3 showed that the rank correlations related to the economic factor and the latent variable of each loan ($\tau_{YE}$) that generated the best estimates decreased with the portfolios' *PDs*. Hence, based on the function used in Basel to calculate the correlation across asset returns (see [6] and [7]), we assume that $\tau_{YE}$ is an exponential function of *PD*. Nevertheless, we adopt a different expression to keep the consistency with the interval specified in [20]:

$$\tau_{YE} = \frac{(1 - PD)}{e^{PD*K}} * \tau_{YE\_MAX} \qquad [\ 23\ ]$$

[15] See [21] where $\tau_{YE}$ increases monotonically with $\tau_{ij}$ for any of the levels chosen (⅓ of the maximum $\tau_{YE}$, average $\tau_{YE}$, and maximum $\tau_{YE}$).

where K is a constant to be set according to the characteristics of the portfolios and $\tau_{YE\_MAX}$ is the maximum possible $\tau_{YE}$ (see [20]). This results in the shapes shown in Figure 4 (for three values of K) where $\tau_{YE}$ is in the range presented in [20], $\tau_{YE} = \tau_{YE\_MAX}$ when *PD* = 0 and $\tau_{YE} = 0$ when *PD* = 1.

*[Insert Figure 4 here]*

However, additional simulations (not displayed here) revealed that the estimation of $\tau_{YE}$ is improved when K is defined as a decreasing function of *PD*, such that, K = (K1 – K2**PD*). So, [23] becomes:

$$\tau_{YE} = \frac{(1 - PD)}{e^{PD*(K1-K2*PD)}} * \tau_{YE\_MAX} \qquad [ 24 ]$$

and, for relatively "small" *PDs*, this new function has a similar shape to the function represented in Figure 4. Concerning the simulated portfolios (Gumbel-dependent with beta and gamma distributions), the best approximations were obtained when K1= 30 and K2 = 200. Tables 4 (retail credit) and 5 (corporate credit) compare the capital estimated via the alternative approach (where the copula parameter, $\theta_{YE}$, is derived from the rank correlation computed according to [24][16]) to the capital calculated in line with the Basel method.

[*Insert Tables 4 and 5 here*]

The results show that the alternative model outperforms Basel method in all scenarios given that, for all *PD* levels tested, the absolute difference between the "true" (simulated) unexpected losses and the alternative estimate is smaller than the absolute difference between the true unexpected losses and the Basel estimate (compare the seventh column to the last one).

[16] We used [12] to estimate $\theta_{YE}$.

## 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 underestimate the capital necessary to face credit 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 a model that keeps the basic structure of the current method but uses different conditional distributions able to detect possible tail dependence among losses. The suggested method is flexible and can capture several dependence shapes since it can be adapted to any differentiable copula family. Its implementation is as simple as the implementation of the existent model and has the advantage of identifying potential higher association between losses in downturns. As an example, we set a formula derived from the Clayton Copula that can capture the supposed stronger dependence across defaults in downturns.

There are typically several possible rank correlations between the economic factor and the latent variable of each loan (called $\tau_{YE}$ in this paper) for each rank correlation across loans (named here as $\tau_{ij}$) and following the assumption present in the Basel method (according to which, high *PDs* are less connected than low *PDs*), we proposed a closed-form expression to estimate $\tau_{YE}$ as a decreasing function of the portfolios' *PD*. We are currently working on simplified formulas to calculate $\tau_{YE}$.

If the losses have small rank correlation, the model proposed gets 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 solution.

By simulating credit losses that potentially represent defaults observed in real loan portfolios (i.e. with positively skewed distributions and upper-tail dependence[17]) and comply with the dependence levels specified in the Basel Accord, we confirm that the current model tends to underestimate joint extreme losses. We also demonstrate that the alternative formula outperforms Basel in all scenarios tested.

[17] See, for example, Kalyvas et al. (2006), Rosenberg and Schuermann (2006), Di Clemente and Romano (2004), and Das and Geng (2006).

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 ( $\tau_{YE}$ ) and the dependence we can infer from empirical data ( $\tau_{ij}$ ). 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.

**REFERENCES**

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

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

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 are a dynamic dependence in Forex markets. *European Journal of Finance*, Vol. 15, Issue 7, pp. 721-750.

Crook, J., T. Bellotti, (2010). Time Varying and Dynamic Models for Default Risk in Consumer Loans. *Journal of Statistical Society Series A (Statistics in Society)*, Vol. 173, Issue 2, pp. 283-305.

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.

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.

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.

Gordy, Michael B. (2003). A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation*, 12, pp. 199-232.

Hull, J., A. White (2004). Valuation of a CDO and an $n^{th}$ 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.

Kalyvas, L., I. Akkizdis, I. Zourka, V. Bouchereau (2006). *Integrating Market, Credit and Operational Risk. A complete guide for bankers and risk professionals*. Pamplona, Spain: Risk Books.

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.

Merton, Robert C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 28, pp. 449-470.

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

Ning, Cathy (2006). Dependence Structure between the Equity Market and the Foreign Exchange Market – a Copula Approach. *Working Paper*.

Owen, D., G. Steck (1962). Moments of Order Statistics from the Equicorrelated Multivariate Normal Distribution. *The Annals of Mathematical Statistics*, Vol. 33, n. 4, pp. 1286-1291.

Patton, Andrew J. (2006). Modelling Asymmetric Exchange Rate Dependence. *International Economic Review*. Vol. 47, no 2, pp. 527-556.

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

Rosenberg, J. V., T. Schuermann (2006). A general approach to integrated risk management with skewed, fat-tailed risks. *Journal of Financial Economics*, 79, pp. 569-614.

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

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

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

## APPENDIX A

**Derivation of equality [9],** $\rho_{YE} = \sqrt{\rho_{ij}}$

The latent variables $Y$ that explain defaults are assumed to follow the standard normal distribution. Therefore those variables can be decomposed into a linear combination of two other standardized normal variables. This results in expressions similar to [10] in which the latent variables pertaining to two obligors $i$ and $j$ are guided by a single factor, $E$, and their respective idiosyncratic risks, $\varepsilon_i$:

$$Y_i = aE + b\varepsilon_i \quad \text{[A.1]}$$

$$Y_j = aE + b\varepsilon_j \quad \text{[A.2]}$$

where $Y_i$ and $Y_j$ are the latent variables of the obligors $i$ and $j$ respectively; $a$ and $b$ indicate how much of the $Y$'s oscillation is explained by the single factor, $E$, and the their idiosyncratic risk, $\varepsilon_i$ and $\varepsilon_j$, respectively; $E$, $\varepsilon_i$, $\varepsilon_j$, $Y_i$ and $Y_j$ are standard normal variables; and the pairs ($E, \varepsilon_i$), ($E$, $\varepsilon_j$) and ($\varepsilon_i, \varepsilon_j$) are assumed to be three pairs of independent variables.

In principle, the coefficients $a$ and $b$ are unknown but owing to some properties of standard normal variables and other additional assumptions made above, they happen to be straightly related to the linear correlations between $Y_i$ and $Y_j$ (= $\rho_{ij}$) and between any $Y$ and $E$ (= $\rho_{YE}$).

At this point, we know that:

(i) $Var(Y_i) = Var(Y_j) = 1$ (since these variables are standard normal);

(ii) $Var(E) = Var(\varepsilon_i) = Var(\varepsilon_j) = 1$ (since these variables are standard normal);

(iii) $Cov(Y_i, E) = \rho_{YE}$ where $\rho_{YE}$ is the linear correlation between $Y_i$ and $E$. Recall that $\rho_{YE} = \dfrac{Cov(Y_i, E)}{\sigma(Y_i)\sigma(E)}$ and, due to (i) and (ii), the denominator has no effect on this calculation.

(iv) $Cov(Y_i, Y_j) = \rho_{ij}$ where $\rho_{ij}$ is the linear correlation between $Y_i$ and $Y_j$. Similarly to the prior item, $\rho_{ij} = \dfrac{Cov(Y_i, Y_j)}{\sigma(Y_i)\sigma(Y_j)}$ and, due to (i), the denominator disappears.

(v) $Cov(E,\varepsilon_i) = Cov(E,\varepsilon_j) = 0$ because the single factor $E$ is assumed to be independent of the idiosyncratic risk of each obligor; and

(vi) $Cov(\varepsilon_i, \varepsilon_j) = 0$ given that the idiosyncratic risks are taken as independent of each other.

The covariance and the variance pertaining to variables composed of the summation of two other variables can be computed as:

$$\begin{aligned} Cov(mX_1 + nX_2,\ pX_3 + qX_4) = mpCov(X_1,\ X_3) + mqCov(X_1,\ X_4) + \\ + npCov(X_2,\ X_3) + nqCov(X_2,\ X_4) \end{aligned} \qquad [A.3]$$

and

$$Var(mX_1 + nX_2) = m^2Var(X_1) + n^2Var(X_2) + 2mnCov(X_1,\ X_2) \qquad [A.4]$$

If we rewrite $Cov(Y_i, E)$ as $Cov(aE + b\varepsilon_i, rE + sE)$ for $\forall$ $r$ and $s$ such that $r + s = 1$ (and therefore $rE + sE = E$) we can use [A.3] to calculate this covariance. For example, doing $r = s = 0.5$, using (v) and considering that the covariance between two identical variables is equal to 1:

$$\begin{aligned} Cov(aE + b\varepsilon_i, 0.5E + 0.5E) = a(0.5)Cov(E,\ E) + a(0.5)Cov(E,\ E) + b(0.5)Cov(E,\ \varepsilon_i) + \\ + b(0.5)Cov(E,\ \varepsilon_i) = a(0.5)(1) + a(0.5)(1) + b(0.5)(0) + b(0.5)(0) = a \end{aligned} \qquad [A.5]$$

So, combining (iii) and [A.5]:

$$\rho_{YE} = a \qquad [A.6]$$

This means that the coefficient that gives the relationship between each latent variable and the single factor in expressions [A.1] and [A.2] is equal to the correlation between this factor and each of the underlying variables. Nonetheless we do not have enough information to calculate such correlation and we still need to find the second coefficient ($b$). As an alternative, we will estimate $a$ and $b$ as a function of $\rho_{ij}$ (the correlation between the latent variables) and, consequently, find the relationship between $\rho_{YE}$ and $\rho_{ij}$.

Since $Y_i$ is the weighted sum of two random variables (the same applies to $Y_j$), we can calculate its variance by using [A.4] and items (ii) and (v) above:

$$Var(Y_i) = Var(aE + b\varepsilon_i) = a^2 Var(E) + b^2 Var(\varepsilon_i) + 2abCov(E, \varepsilon_i) =$$
$$= a^2(1) + b^2(1) + 2ab(0) = a^2 + b^2$$

According to (i), $Var(Y_i) = 1$, so :

$$a^2 + b^2 = 1 \qquad [A.7]$$

Using [A.1], [A.2], [A.3], (v), (vi) and the fact the covariance between two equal variables is 1:

$$Cov(Y_i, Y_j) = Cov(aE + b\varepsilon_i, aE + b\varepsilon_j) = a^2 Cov(E, E) + abCov(E, \varepsilon_j) +$$
$$+ abCov(\varepsilon_i, E) + b^2 Cov(\varepsilon_i, \varepsilon_j) = a^2(1) + ab(0) + ab(0) + b^2(0) = a^2 \qquad [A.8]$$

As stated in (iv), $Cov(Y_i, Y_j) = \rho_{ij}$, and from [A.8] we have:

$$\rho_{ij} = a^2 \Leftrightarrow a = \sqrt{\rho_{ij}} \qquad [A.9]$$

Plugging [A.9] into [A.7]:

$$b = \sqrt{1 - \rho_{ij}}$$

Hence, comparing [A.6] and [A.9]:

$$\rho_{YE} = \sqrt{\rho_{ij}}$$

## APPENDIX B

## Complementary information concerning simulated data for retail credit

| Average $PD$ | $\rho_{ij}$* | $\tau_{ij}$ ** | $\theta_{YE}$ derived from ⅓ of maximum $\tau_{YE}$ *** | $\theta_{YE}$ derived from average $\tau_{YE}$ *** | $\theta_{YE}$ derived from maximum $\tau_{YE}$ *** |
|---|---|---|---|---|---|
| **Panel A: Beta-distributed losses** | | | | | |
| 0.01 | 0.12 | 0.0346 | 0.4168 | 0.6980 | 2.1462 |
| 0.02 | 0.09 | 0.0369 | 0.4179 | 0.7001 | 2.1566 |
| 0.03 | 0.08 | 0.0365 | 0.4177 | 0.6997 | 2.1545 |
| 0.04 | 0.06 | 0.0334 | 0.4162 | 0.6969 | 2.1413 |
| 0.05 | 0.05 | 0.0337 | 0.4164 | 0.6972 | 2.1427 |
| 0.06 | 0.05 | 0.0266 | 0.4129 | 0.6907 | 2.1117 |
| 0.07 | 0.04 | 0.0257 | 0.4125 | 0.6899 | 2.1078 |
| 0.08 | 0.04 | 0.0250 | 0.4121 | 0.6892 | 2.1045 |
| 0.09 | 0.04 | 0.0201 | 0.4097 | 0.6848 | 2.0841 |
| 0.10 | 0.03 | 0.0187 | 0.4091 | 0.6835 | 2.0782 |
| **Panel B: Gamma-distributed losses** | | | | | |
| 0.01 | 0.12 | 0.0408 | 0.4198 | 0.7037 | 2.1735 |
| 0.02 | 0.09 | 0.0344 | 0.4167 | 0.6978 | 2.1455 |
| 0.03 | 0.08 | 0.0327 | 0.4159 | 0.6964 | 2.1396 |
| 0.04 | 0.06 | 0.0275 | 0.4133 | 0.6915 | 2.1156 |
| 0.05 | 0.05 | 0.0376 | 0.4182 | 0.7007 | 2.1592 |
| 0.06 | 0.05 | 0.0248 | 0.4120 | 0.6890 | 2.1038 |
| 0.07 | 0.04 | 0.0277 | 0.4134 | 0.6917 | 2.1169 |
| 0.08 | 0.04 | 0.0246 | 0.4119 | 0.6889 | 2.1027 |
| 0.09 | 0.04 | 0.0240 | 0.4116 | 0.6883 | 2.1004 |
| 0.10 | 0.03 | 0.0206 | 0.4100 | 0.6853 | 2.0864 |

Copula parameters ($\theta_{YE}$) calculated according to [22]

* $\rho_{ij}$ is the linear correlation between the default probability of two loans $i$ and $j$ (which represents the correlation of all pairs in the portfolio).

** $\tau_{ij}$ is the rank correlation between the default probability of two loans $i$ and $j$.

*** $\tau_{Ei}$ is the rank correlation (Kendall's tau) between the latent variable of each obligor and the economic factor.

## APPENDIX C

## Complementary information concerning simulated data for corporate credit

| Average $PD$ | $\rho_{ij}$* | $\tau_{ij}$ ** | $\theta_{YE}$ derived from ⅓ of maximum $\tau_{YE}$ *** | $\theta_{YE}$ derived from average $\tau_{YE}$ *** | $\theta_{YE}$ derived from maximum $\tau_{YE}$ *** |
|---|---|---|---|---|---|
| **Panel A: Beta-distributed losses** | | | | | |
| 0.01 | 0.19 | 0.0907 | 0.4444 | 0.7501 | 2.4040 |
| 0.02 | 0.16 | 0.0920 | 0.4451 | 0.7514 | 2.4110 |
| 0.03 | 0.15 | 0.0882 | 0.4432 | 0.7477 | 2.3904 |
| 0.04 | 0.14 | 0.0882 | 0.4432 | 0.7477 | 2.3910 |
| 0.05 | 0.13 | 0.0841 | 0.4411 | 0.7438 | 2.3702 |
| 0.06 | 0.13 | 0.0785 | 0.4384 | 0.7385 | 2.3440 |
| 0.07 | 0.12 | 0.0765 | 0.4373 | 0.7366 | 2.3342 |
| 0.08 | 0.12 | 0.0760 | 0.4371 | 0.7361 | 2.3316 |
| 0.09 | 0.12 | 0.0764 | 0.4373 | 0.7365 | 2.3331 |
| 0.10 | 0.12 | 0.0686 | 0.4334 | 0.7292 | 2.2972 |
| **Panel B: Gamma-distributed losses** | | | | | |
| 0.01 | 0.19 | 0.0929 | 0.4455 | 0.7522 | 2.4148 |
| 0.02 | 0.16 | 0.0890 | 0.4436 | 0.7485 | 2.3950 |
| 0.03 | 0.15 | 0.0921 | 0.4451 | 0.7513 | 2.4095 |
| 0.04 | 0.14 | 0.0863 | 0.4422 | 0.7459 | 2.3814 |
| 0.05 | 0.13 | 0.0783 | 0.4382 | 0.7383 | 2.3428 |
| 0.06 | 0.13 | 0.0781 | 0.4381 | 0.7381 | 2.3415 |
| 0.07 | 0.12 | 0.0733 | 0.4358 | 0.7337 | 2.3195 |
| 0.08 | 0.12 | 0.0725 | 0.4353 | 0.7328 | 2.3149 |
| 0.09 | 0.12 | 0.0719 | 0.4351 | 0.7324 | 2.3128 |
| 0.10 | 0.12 | 0.0734 | 0.4358 | 0.7338 | 2.3203 |

Copula parameters ($\theta_{YE}$) calculated according to [22]

* $\rho_{ij}$ is the linear correlation between the default probability of two loans $i$ and $j$ (which represents the correlation of all pairs in the portfolio).

** $\tau_{ij}$ is the rank correlation between the default probability of two loans $i$ and $j$.

*** $\tau_{YE}$ is the rank correlation (Kendall's tau) between the latent variable of each obligor and the economic factor.

Table 1: Representation of concordant and discordant variables for several values of $\tau_{ij}$

| Panel A<br>$\tau_{ij}$ = -1 | | |
|---|---|---|
| $Y_i$ | $Y_j$ | $E$ |
| ↑ | ↓ | ↑ |
| ↓ | ↑ | ↓ |
| ↑ | ↓ | ↑ |
| … | … | … |
| ↑ | ↓ | ↓ |

| Panel B<br>$\tau_{ij}$ = 1 | | |
|---|---|---|
| $Y_i$ | $Y_j$ | $E$ |
| ↑ | ↑ | ↓ |
| ↓ | ↓ | ↑ |
| ↑ | ↑ | ↑ |
| … | … | … |
| ↑ | ↑ | ↓ |

| Panel C<br>(minimum when $-1 < \tau_{ij} < 1$) | | |
|---|---|---|
| $Y_i$ | $Y_j$ | $E$ |
| ↑ | ↑ | ↓ |
| ↓ | ↓ | ↑ |
| ↑ | ↓ | ↑ |
| ↑ | ↓ | ↑ |
| ↑ | ↓ | ↓ |
| ↑ | ↓ | ↓ |

| Panel D<br>(maximum when $-1 < \tau_{ij} < 1$) | | |
|---|---|---|
| $Y_i$ | $Y_j$ | $E$ |
| ↑ | ↑ | ↑ |
| ↓ | ↓ | ↓ |
| ↑ | ↓ | ↑ |
| ↑ | ↓ | ↑ |
| ↑ | ↓ | ↓ |
| ↑ | ↓ | ↓ |

TABLE 2: Comparison between unexpected retail credit losses estimated via Basel and alternative methods

| Average *PD* | $\rho_{ij}$* | Observed unexpected losses (OUL) | Basel estimation (BE) | Estimation with ⅓ of maximum $\tau_{YE}$ ** (AE1) | Estimation with average $\tau_{YE}$ ** (AE2) | Estimation with maximum $\tau_{YE}$ ** (AE3) | Absolute difference (BE) - (OUL) | Absolute difference (AE1) - (OUL) | Absolute difference (AE2) - (OUL) | Absolute difference (AE3) - (OUL) |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A: Beta-distributed losses** | | | | | | | | | | |
| 0.01 | 0.12 | 0.3040 | 0.0825 | 0.1144 | 0.1876 | **0.3626** | 0.2215 | 0.1896 | 0.1164 | **0.0586** |
| 0.02 | 0.09 | 0.3353 | 0.1039 | 0.1831 | **0.3134** | 0.7248 | 0.2315 | 0.1522 | **0.0219** | 0.3895 |
| 0.03 | 0.08 | 0.3508 | 0.1107 | 0.2291 | **0.3910** | 0.8419 | 0.2401 | 0.1217 | **0.0402** | 0.4911 |
| 0.04 | 0.06 | 0.3745 | 0.1164 | 0.2687 | **0.4520** | 0.8885 | 0.2580 | 0.1058 | **0.0775** | 0.5140 |
| 0.05 | 0.05 | 0.3756 | 0.1187 | **0.2982** | 0.4938 | 0.9043 | 0.2570 | **0.0774** | 0.1182 | 0.5287 |
| 0.06 | 0.05 | 0.3810 | 0.1205 | **0.3190** | 0.5206 | 0.9067 | 0.2605 | **0.0620** | 0.1396 | 0.5257 |
| 0.07 | 0.04 | 0.3843 | 0.1226 | **0.3378** | 0.5437 | 0.9057 | 0.2617 | **0.0464** | 0.1594 | 0.5214 |
| 0.08 | 0.04 | 0.4011 | 0.1260 | **0.3548** | 0.5630 | 0.9015 | 0.2751 | **0.0463** | 0.1619 | 0.5004 |
| 0.09 | 0.04 | 0.4199 | 0.1302 | **0.3678** | 0.5761 | 0.8946 | 0.2898 | **0.0522** | 0.1561 | 0.4747 |
| 0.10 | 0.03 | 0.4185 | 0.1341 | **0.3791** | 0.5870 | 0.8876 | 0.2844 | **0.0394** | 0.1685 | 0.4691 |
| Total …… | | | | | | | 2.5794 | **0.8930** | 1.1598 | 4.4733 |
| **Panel B: Gamma-distributed losses** | | | | | | | | | | |
| Average *PD* | $\rho_{ij}$* | Observed unexpected losses (OUL) | Basel estimation (BE) | Estimation with ⅓ of maximum $\tau_{YE}$ ** (AE1) | Estimation with average $\tau_{YE}$ ** (AE2) | Estimation with maximum $\tau_{YE}$ ** (AE3) | Absolute difference (BE) - (OUL) | Absolute difference (AE1) - (OUL) | Absolute difference (AE2) - (OUL) | Absolute difference (AE3) - (OUL) |
| 0.01 | 0.12 | 0.2536 | 0.0826 | 0.1155 | **0.1892** | 0.3650 | 0.1711 | 0.1382 | **0.0645** | 0.1113 |
| 0.02 | 0.09 | 0.3021 | 0.1038 | 0.1823 | **0.3121** | 0.7226 | 0.1984 | 0.1199 | **0.0099** | 0.4204 |
| 0.03 | 0.08 | 0.3347 | 0.1110 | 0.2286 | **0.3901** | 0.8409 | 0.2237 | 0.1062 | **0.0554** | 0.5061 |
| 0.04 | 0.06 | 0.3487 | 0.1148 | **0.2638** | 0.4445 | 0.8840 | 0.2340 | **0.0850** | 0.0958 | 0.5352 |
| 0.05 | 0.05 | 0.3616 | 0.1182 | **0.2988** | 0.4950 | 0.9055 | 0.2434 | **0.0628** | 0.1333 | 0.5439 |
| 0.06 | 0.05 | 0.3692 | 0.1202 | **0.3177** | 0.5188 | 0.9062 | 0.2490 | **0.0515** | 0.1497 | 0.5371 |
| 0.07 | 0.04 | 0.3983 | 0.1233 | **0.3399** | 0.5463 | 0.9058 | 0.2750 | **0.0584** | 0.1480 | 0.5075 |
| 0.08 | 0.04 | 0.3991 | 0.1258 | **0.3544** | 0.5625 | 0.9016 | 0.2732 | **0.0447** | 0.1634 | 0.5025 |
| 0.09 | 0.04 | 0.4088 | 0.1296 | **0.3685** | 0.5775 | 0.8956 | 0.2792 | **0.0403** | 0.1687 | 0.4869 |
| 0.10 | 0.03 | 0.4257 | 0.1343 | **0.3802** | 0.5883 | 0.8877 | 0.2914 | **0.0455** | 0.1626 | 0.4620 |
| Total …… | | | | | | | 2.4384 | **0.7524** | 1.1512 | 4.6128 |

* $\rho_{ij}$ is the linear correlation between the default probability of two loans *i* and *j* (which represents the correlation for all pairs in the portfolio).

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

The best estimate and the smallest difference between the unexpected losses and the estimates for each *PD* level are highlighted in boldface.

TABLE 3: Comparison between unexpected corporate credit losses estimated via Basel and alternative methods

**Panel A: Beta-distributed losses**

| Average *PD* | $\rho_{ij}$* | Observed unexpected losses (OUL) | Basel estimation (BE) | Estimation with ⅓ of maximum $\tau_{YE}$ ** (AE1) | Estimation with average $\tau_{YE}$ ** (AE2) | Estimation with maximum $\tau_{YE}$ ** (AE3) | Absolute difference (BE) - (OUL) | Absolute difference (AE1) - (OUL) | Absolute difference (AE2) - (OUL) | Absolute difference (AE3) - (OUL) |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.01 | 0.19 | 0.3214 | 0.1308 | 0.1212 | 0.1971 | **0.3691** | 0.1905 | 0.2002 | 0.1243 | **0.0478** |
| 0.02 | 0.16 | 0.3528 | 0.1697 | 0.1943 | **0.3320** | 0.7577 | 0.1830 | 0.1585 | **0.0208** | 0.4049 |
| 0.03 | 0.15 | 0.3611 | 0.1939 | 0.2444 | **0.4165** | 0.8711 | 0.1671 | 0.1166 | **0.0555** | 0.5100 |
| 0.04 | 0.14 | 0.3901 | 0.2146 | 0.2844 | **0.4777** | 0.9085 | 0.1755 | 0.1057 | **0.0876** | 0.5184 |
| 0.05 | 0.13 | 0.3986 | 0.2329 | **0.3139** | 0.5188 | 0.9183 | 0.1657 | **0.0847** | 0.1202 | 0.5197 |
| 0.06 | 0.13 | 0.4160 | 0.2508 | **0.3376** | 0.5488 | 0.9184 | 0.1651 | **0.0783** | 0.1329 | 0.5025 |
| 0.07 | 0.12 | 0.4142 | 0.2677 | **0.3580** | 0.5729 | 0.9147 | 0.1465 | **0.0562** | 0.1586 | 0.5005 |
| 0.08 | 0.12 | 0.4346 | 0.2825 | **0.3745** | 0.5909 | 0.9092 | 0.1521 | **0.0601** | 0.1562 | 0.4745 |
| 0.09 | 0.12 | 0.4532 | 0.2979 | **0.3904** | 0.6069 | 0.9017 | 0.1553 | **0.0628** | 0.1538 | 0.4486 |
| 0.10 | 0.12 | 0.4650 | 0.3104 | **0.3988** | 0.6135 | 0.8937 | 0.1546 | **0.0662** | 0.1485 | 0.4288 |
| Total ………………………………………………………………………………………… | | | | | | | 1.6555 | **0.9894** | 1.1583 | 4.3556 |

**Panel B: Gamma-distributed losses**

| Average *PD* | $\rho_{ij}$* | Observed unexpected losses (OUL) | Basel estimation (BE) | Estimation with ⅓ of maximum $\tau_{YE}$ ** (AE1) | Estimation with average $\tau_{YE}$ ** (AE2) | Estimation with maximum $\tau_{YE}$ ** (AE3) | Absolute difference (BE) - (OUL) | Absolute difference (AE1) - (OUL) | Absolute difference (AE2) - (OUL) | Absolute difference (AE3) - (OUL) |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.01 | 0.19 | 0.2669 | 0.1302 | 0.1209 | **0.1963** | 0.3648 | 0.1367 | 0.1461 | **0.0707** | 0.0979 |
| 0.02 | 0.16 | 0.3108 | 0.1699 | 0.1938 | **0.3312** | 0.7561 | 0.1409 | 0.1170 | **0.0204** | 0.4453 |
| 0.03 | 0.15 | 0.3383 | 0.1951 | 0.2470 | **0.4207** | 0.8749 | 0.1432 | 0.0913 | **0.0824** | 0.5366 |
| 0.04 | 0.14 | 0.3560 | 0.2157 | **0.2850** | 0.4785 | 0.9086 | 0.1403 | **0.0710** | 0.1225 | 0.5526 |
| 0.05 | 0.13 | 0.3714 | 0.2320 | **0.3108** | 0.5141 | 0.9166 | 0.1394 | **0.0606** | 0.1427 | 0.5452 |
| 0.06 | 0.13 | 0.3986 | 0.2505 | **0.3370** | 0.5481 | 0.9184 | 0.1481 | **0.0616** | 0.1495 | 0.5198 |
| 0.07 | 0.12 | 0.4082 | 0.2674 | **0.3565** | 0.5708 | 0.9143 | 0.1408 | **0.0518** | 0.1625 | 0.5061 |
| 0.08 | 0.12 | 0.4326 | 0.2831 | **0.3737** | 0.5896 | 0.9086 | 0.1495 | **0.0589** | 0.1570 | 0.4760 |
| 0.09 | 0.12 | 0.4454 | 0.2982 | **0.3888** | 0.6048 | 0.9012 | 0.1472 | **0.0567** | 0.1593 | 0.4558 |
| 0.10 | 0.12 | 0.4528 | 0.3116 | **0.4018** | 0.6170 | 0.8934 | 0.1412 | **0.0510** | 0.1642 | 0.4405 |
| Total ………………………………………………………………………………………… | | | | | | | 1.4274 | **0.7659** | 1.2312 | 4.5758 |

* $\rho_{ij}$ is the linear correlation between the default probability of two loans *i* and *j* (which represents the correlation for all pairs in the portfolio).

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

The best estimate and the smallest difference between the unexpected losses and the estimates for each *PD* level are highlighted in boldface.

TABLE 4: Additional comparison between unexpected retail credit losses estimated via Basel and alternative methods

| **Panel A: Beta-distributed losses** | | | | | | | |
|---|---|---|---|---|---|---|---|
| Average *PD* | $\tau_{YE}$ * | $\theta_{YE}$ ** | Observed unexpected losses (OUL) | Basel estimation (BE) | Alternative estimation (AE) | Absolute difference (BE) - (OUL) | Absolute difference (AE) -( OUL) |
| 0.01 | 0.3836 | 1.2444 | 0.2953 | 0.0818 | 0.2752 | 0.2135 | 0.0201 |
| 0.02 | 0.3016 | 0.8636 | 0.3378 | 0.1032 | 0.3771 | 0.2346 | 0.0393 |
| 0.03 | 0.2428 | 0.6414 | 0.3582 | 0.1124 | 0.3630 | 0.2458 | 0.0048 |
| 0.04 | 0.2055 | 0.5172 | 0.3545 | 0.1154 | 0.3350 | 0.2391 | 0.0195 |
| 0.05 | 0.1798 | 0.4385 | 0.3728 | 0.1183 | 0.3138 | 0.2545 | 0.0590 |
| 0.06 | 0.1648 | 0.3946 | 0.3846 | 0.1200 | 0.3034 | 0.2646 | 0.0812 |
| 0.07 | 0.1562 | 0.3701 | 0.3844 | 0.1226 | 0.3025 | 0.2619 | 0.0819 |
| 0.08 | 0.1538 | 0.3635 | 0.4058 | 0.1262 | 0.3133 | 0.2796 | 0.0925 |
| 0.09 | 0.1575 | 0.3738 | 0.4223 | 0.1301 | 0.3362 | 0.2923 | 0.0862 |
| 0.10 | 0.1698 | 0.4089 | 0.4215 | 0.1342 | 0.3791 | 0.2872 | 0.0424 |
| Total ……… | | | | | | 2.5731 | 0.5268 |
| **Panel B: Gamma-distributed losses** | | | | | | | |
| Average *PD* | $\tau_{YE}$ * | $\theta_{YE}$ ** | Observed unexpected losses (OUL) | Basel estimation (BE) | Alternative estimation (AE) | Absolute difference (BE) - (OUL) | Absolute difference (AE) - (OUL) |
| 0.01 | 0.3863 | 1.2588 | 0.2580 | 0.0817 | 0.2781 | 0.1763 | 0.0201 |
| 0.02 | 0.3004 | 0.8589 | 0.3167 | 0.1040 | 0.3790 | 0.2127 | 0.0623 |
| 0.03 | 0.2449 | 0.6485 | 0.3236 | 0.1117 | 0.3651 | 0.2119 | 0.0415 |
| 0.04 | 0.2048 | 0.5151 | 0.3352 | 0.1158 | 0.3345 | 0.2194 | 0.0007 |
| 0.05 | 0.1804 | 0.4402 | 0.3588 | 0.1180 | 0.3146 | 0.2408 | 0.0442 |
| 0.06 | 0.1646 | 0.3941 | 0.3692 | 0.1202 | 0.3034 | 0.2490 | 0.0658 |
| 0.07 | 0.1562 | 0.3702 | 0.3910 | 0.1229 | 0.3031 | 0.2681 | 0.0879 |
| 0.08 | 0.1537 | 0.3632 | 0.4049 | 0.1261 | 0.3130 | 0.2788 | 0.0919 |
| 0.09 | 0.1568 | 0.3720 | 0.4188 | 0.1290 | 0.3330 | 0.2898 | 0.0857 |
| 0.10 | 0.1690 | 0.4066 | 0.4253 | 0.1339 | 0.3765 | 0.2915 | 0.0488 |
| Total ……… | | | | | | 2.4383 | 0.5488 |

* $\tau_{YE}$, the rank correlation (Kendall's tau) between the latent variable of each obligor and the economic factor, was calculated according to [24].

** $\theta_{YE}$ is the parameter of the copula that expresses the dependence between the latent variable of each obligor and the economic factor.

TABLE 5: Additional comparison between unexpected corporate credit losses estimated via Basel and alternative methods

| Panel A: Beta-distributed losses | | | | | | | |
|---|---|---|---|---|---|---|---|
| Average *PD* | $\tau_{YE}$ * | $\theta_{YE}$ ** | Observed unexpected losses (OUL) | Basel estimation (BE) | Alternative estimation (AE) | Absolute difference (BE) - (OUL) | Absolute difference (AE) -( OUL) |
| 0.01 | 0.4082 | 1.3793 | 0.3080 | 0.1328 | 0.3001 | 0.1752 | 0.0079 |
| 0.02 | 0.3214 | 0.9471 | 0.3502 | 0.1686 | 0.4044 | 0.1816 | 0.0542 |
| 0.03 | 0.2563 | 0.6893 | 0.3746 | 0.1961 | 0.3887 | 0.1784 | 0.0141 |
| 0.04 | 0.2163 | 0.5519 | 0.3885 | 0.2150 | 0.3572 | 0.1735 | 0.0313 |
| 0.05 | 0.1915 | 0.4737 | 0.4047 | 0.2323 | 0.3369 | 0.1724 | 0.0678 |
| 0.06 | 0.1731 | 0.4187 | 0.4101 | 0.2515 | 0.3226 | 0.1586 | 0.0875 |
| 0.07 | 0.1643 | 0.3932 | 0.4248 | 0.2672 | 0.3212 | 0.1576 | 0.1036 |
| 0.08 | 0.1615 | 0.3851 | 0.4361 | 0.2835 | 0.3315 | 0.1526 | 0.1046 |
| 0.09 | 0.1657 | 0.3972 | 0.4378 | 0.2964 | 0.3547 | 0.1414 | 0.0831 |
| 0.10 | 0.1775 | 0.4316 | 0.4505 | 0.3117 | 0.3981 | 0.1388 | 0.0523 |
| Total ………………………………………………………………………………………………………………. | | | | | | 1.6302 | 0.6064 |

| Panel B: Gamma-distributed losses | | | | | | | |
|---|---|---|---|---|---|---|---|
| Average *PD* | $\tau_{YE}$ * | $\theta_{YE}$ ** | Observed unexpected losses (OUL) | Basel estimation (BE) | Alternative estimation (AE) | Absolute difference (BE) - (OUL) | Absolute difference (AE) - (OUL) |
| 0.01 | 0.4082 | 1.3793 | 0.2621 | 0.1311 | 0.2945 | 0.1310 | 0.0324 |
| 0.02 | 0.3199 | 0.9407 | 0.2946 | 0.1692 | 0.4037 | 0.1254 | 0.1091 |
| 0.03 | 0.2617 | 0.7090 | 0.3175 | 0.1928 | 0.3933 | 0.1247 | 0.0758 |
| 0.04 | 0.2171 | 0.5548 | 0.3640 | 0.2146 | 0.3585 | 0.1494 | 0.0055 |
| 0.05 | 0.1918 | 0.4745 | 0.3601 | 0.2323 | 0.3374 | 0.1278 | 0.0227 |
| 0.06 | 0.1732 | 0.4191 | 0.3863 | 0.2518 | 0.3234 | 0.1345 | 0.0629 |
| 0.07 | 0.1642 | 0.3930 | 0.4066 | 0.2665 | 0.3204 | 0.1400 | 0.0862 |
| 0.08 | 0.1616 | 0.3854 | 0.4328 | 0.2825 | 0.3310 | 0.1503 | 0.1018 |
| 0.09 | 0.1658 | 0.3975 | 0.4302 | 0.2983 | 0.3566 | 0.1319 | 0.0737 |
| 0.10 | 0.1769 | 0.4298 | 0.4459 | 0.3109 | 0.3961 | 0.1350 | 0.0498 |
| Total ………………………………………………………………………………………………………………. | | | | | | 1.3500 | 0.6199 |

* $\tau_{YE}$, the rank correlation (Kendall's tau) between the latent variable of each obligor and the economic factor, was calculated according to [24].

** $\theta_{YE}$ is the parameter of the copula that expresses the dependence between the latent variable of each obligor and the economic factor.

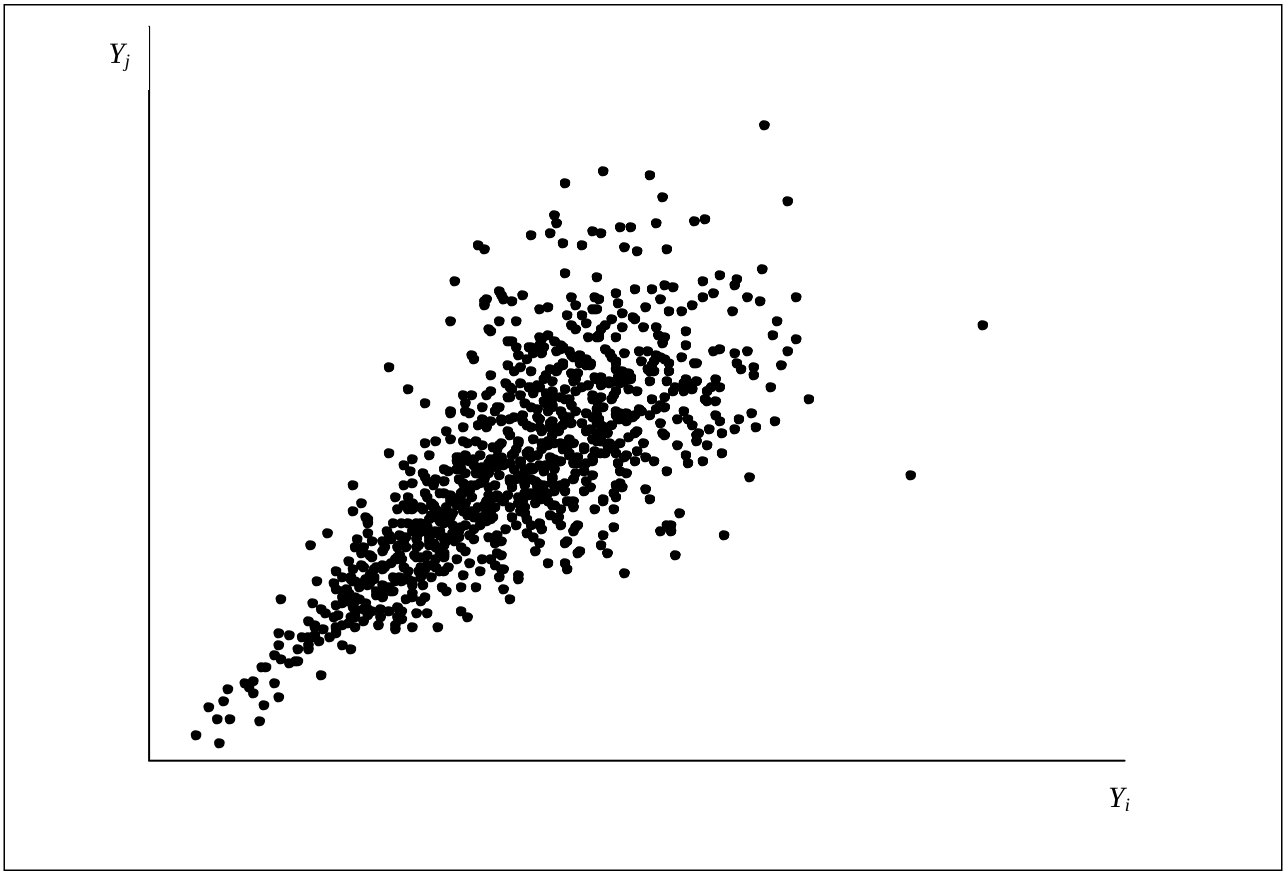


FIGURE 1: A representation of the stronger connection across small values of the latent variables ($Y_i$ and $Y_j$) in downturns.

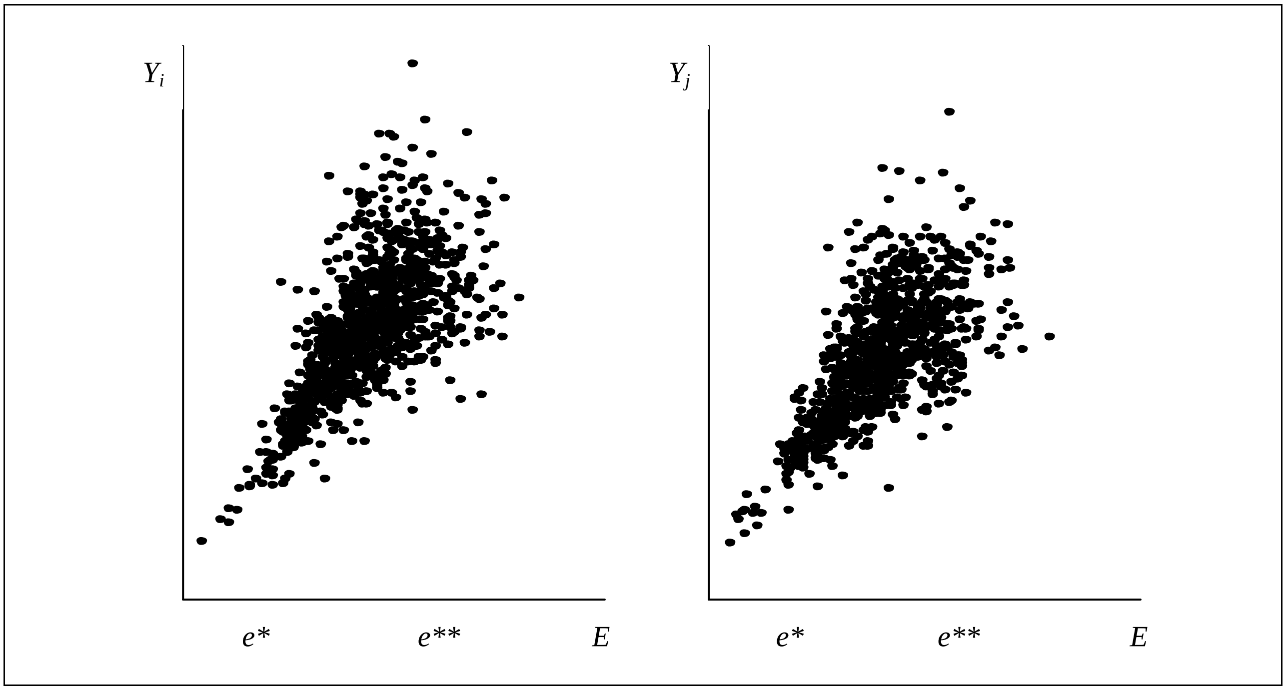


FIGURE 2: A representation of the stronger connection across small values of each latent variable ($Y_i$ on the left and $Y_j$ on the right) and the economic status. When $E$ is reduced ($e^*$), indicating an unfavorable scenario, both $Y_i$ and $Y_j$ tend to be small. When $E$ increases ($e^{**}$),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

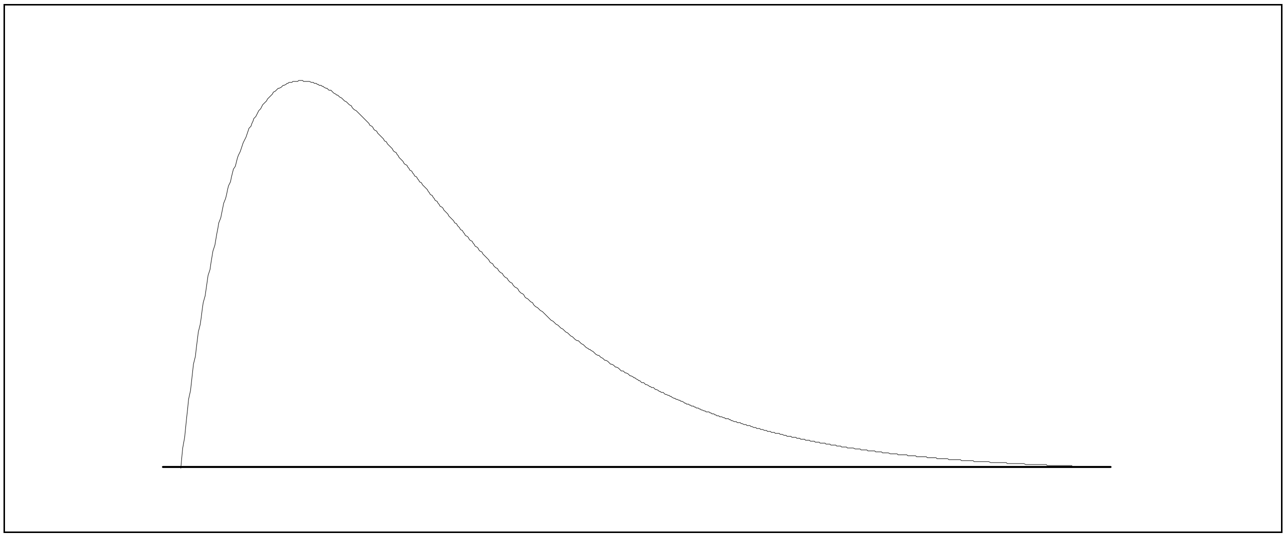

FIGURE 3 – Representation of a positively-skewed distribution.

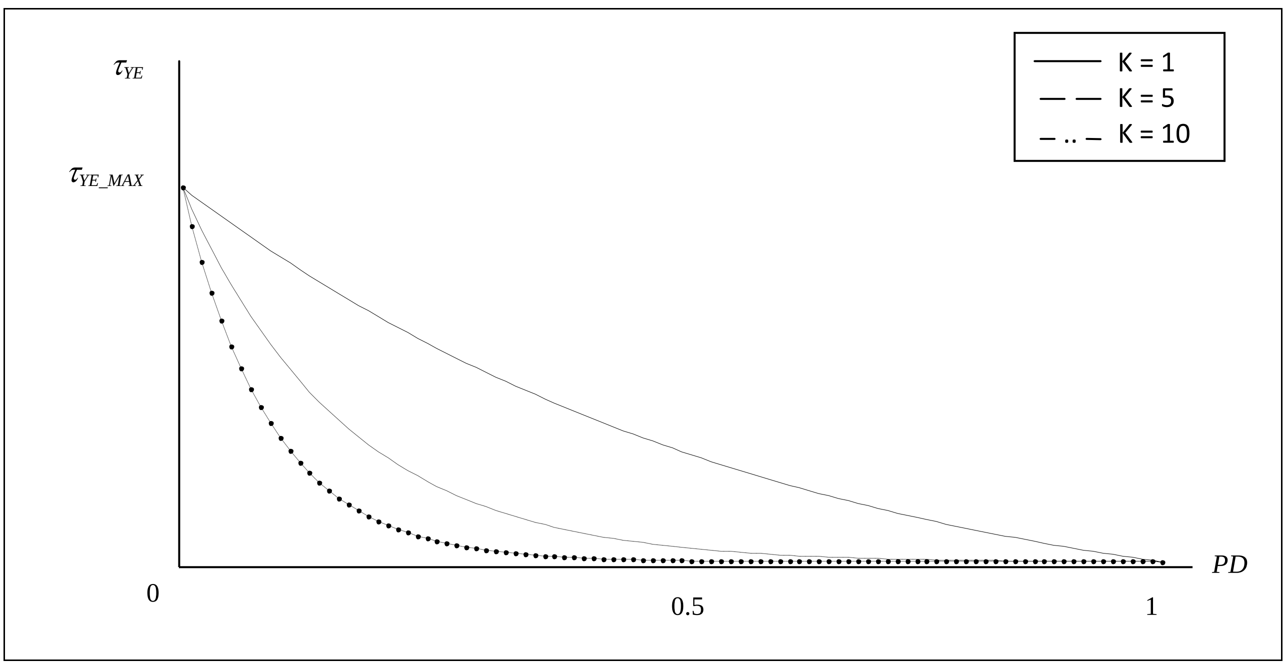


FIGURE 4 – Rank correlation ($\tau_{YE}$) between the latent variable of each obligor and the economic factor expressed as an exponential function of the probability of default (*PD*).