Scientific Research Publishing

Low Default Portfolios—A Proposed Rule to Identify Differences between Imprudence, Conservatism, and Exaggeration

David J. C. Dinis*

ISEG—Lisbon School of Economics and Management, Lisbon, Portugal
Email: daviddinis2013@gmail.com

Abstract

Internal models may be used by banks to calculate their regulatory capital for credit risk. There are a variety of methodologies for estimating default probabilities, which leads to major differences in credit provisions and capital requirements. Using either a classical or a Bayesian technique, the computation of default probabilities can be ensured. Reduced form models are a choice. These models, however, might not be used to quantify economic capital because they assume independence among default events. Banks are compelled to employ structural models since defaults in the real world of banking are not solely due to exogenous causes. Because of the diversification effects between credit losses for one obligor and credit losses for other obligors in each bank's portfolio, total unexpected losses do not equal the sum of individual unexpected losses. Those two types of models—reduced form and structural—are provided in either a theoretical or a numerical format. This paper covers both the classical and Bayesian techniques, with the latter employing a broader set of prior functions that offer considerably different probabilities. Distinguishing between imprudence, conservatism, and exaggeration might be difficult in the context of low default portfolios with scarce data. A realistic rule is proposed for finding the minimum and maximum bounds and therefore assessing the required conservatism margin by comparing classical and Bayesian probabilities.

Keywords

Credit Risk, Default Probability, Correlated Distributions, Classical Approach, Bayesian Approach, Conservative Zone

---

*The opinions expressed in this work are a purely personal matter.

---

1. Introduction

Banks may use internal models to assess their credit risk exposure within an internal ratings-based framework under the Basel Accords issued by the Basel Committee on Banking Supervision, according to capital requirements rules adopted by banking institutions and transposed into the legal system across the European Union[1]. The default probability is the same for different internal rating systems, whether foundation internal rating-based or advanced internal rating-based—nomenclature used in such accords.

Loans with no defaults over a lengthy period of time are a utopian concept. Indeed, during an economic downturn, it is projected that unemployment will grow and, as a result, higher default rates will emerge, projections that are heightened by the contagion risk. It will be impossible to generate accurate and realistic estimations of default probabilities if there is no correlation among individual exposures within portfolios.

Therefore, the zero-default assumption for all risk classes will not be presented here. In any case, if the observed default number is nil, the formulas in this paper can be immediately converted to the zero-presume assumption.

A utopia would likewise imply independence of default events foreseen in reduced form models. This independence will be studied in order to compare the results to those obtained from structural models (which account for the existence of a dependence structure among borrowers in the same risk class) and to assess the significance of the asset correlation component.

The methodology used in this document is explained in Chapter 2, which is divided into four sections. Section 2.1 addresses general considerations (specifically, asset correlation) that are necessary for each axis of credit risk models applied to low default portfolios—reduced form models and structural models—and for each axis of statistical approaches used in credit risk assessment—classical and Bayesian approaches. The classical technique is covered in Section 2.2, which has two subsections: one for reduced form models and the other for structural models. The Bayesian technique is covered in Section 2.3, which is broken into two subsections: prior distributions and posterior distributions, the latter of which encompasses both reduced form models and structural models. In Section 2.4, a reasonable criterion is used to connect the classical and Bayesian approaches, allowing one to distinguish between imprudence, conservatism, and exaggeration in terms of default probability.

The outputs of such models and approaches are provided in Chapter 3. In the last section of this chapter, the main conclusions drawn from the comparison of those models and approaches are presented.

Finally, some closing remarks are made. They address subjects like uncertainty, mix of prior functions, and other open issues that need to be investigated further.

---

[1]Prudential requirements for credit institutions and investment firms—Regulation 575/2013 of the European Parliament and the Council of 26 June 2013 and its several amending.

2. Formulae and Other Technical Issues

2.1. General Considerations

Reduced form models are based on the default independence hypothesis, which indicates that asset correlation, "ρ", is assumed to be zero. The use of these models to assess banking regulatory capital for credit risk should not be acceptable because the reality follows a completely different pattern, defined by interdependence among default occurrences.

Because loss defaults are not independent, reduced form models must be replaced with structural models. In these last models, the default events are assumed to be positively connected. The stronger the link to the systematic risk, the larger "ρ". When $\rho > 0$ the default probabilities are computed using an integrating algorithm, which requires a stochastic treatment or a simulation procedure for the classical (or frequentist) approach. The algorithm takes into account all possible values "y" of a standard and normally distributed random variable "Y" that represents the systematic risk's realization range. Under the Bayesian (or subjective) approach, the stochastic treatment or the simulation procedure must be doubled: one for "y" and the other for "λ", the default probability.

The frequentist approach and the subjective approach are used to discuss both reduced form models and structural models. The trapezoidal rule for numerical integration approximation is used to obtain outputs when an analytical solution is not attainable.

For each risk class, the binomial and Poisson distributions are used to simulate the default probability random variable. Nonetheless, unlike the binomial distribution, using the Poisson distribution to represent the default probability is not entirely accurate because the size of the risk class, "n", is not fixed in this type of distribution.

The posterior distribution, according to Bayesian inference, corresponds to the conditional distribution of the default probability random variable, "Λ", given a set number of borrowers, "n", and a fixed number of defaults, "k", as well as the previous distribution of the default probability. The posterior density of the default probability is obtained by matching the likelihood function and the prior function.

The Bayesian approach provides another conceptual distinction. In the classical approach, there are frequentist confidence intervals, while in the Bayesian approach, there are posterior credible intervals. Despite the fact that the latter are commonly conceived of as a Bayesian variant of confidence intervals used in classical probability, they have different meanings. The highest density interval existing a unimodal posterior density[2] is related to the shortest possible interval, determined by numerical calculation. Because "n", "k", and "ρ" are regarded as constants, there is a $(100 - \delta)\%$ probability that the true value of the unidi-

---

[2]If the posterior density is a multimodal distribution, "the highest density region" should be used, instead of "the highest density interval".

mensional parameter, "λ", falls inside the credible interval, being "δ" the risk level (equivalent to the significance level of classical probabilities)[3].

2.2. Classical Approach

2.2.1. Classical Approach with Reduced Form Models

Being "K" the random variable that represents the number of defaults, "k" the number of defaults, "n" the number of obligors, "λ" the default probability and "ω" the confidence level; and assuming that "K" follows a standard binomial distribution, the probability of having no more than "k" defaults inside a risk class with "n" obligors is given by:

$$\mathrm{P}\left(K \le k\right) = \sum_{i=0}^{k} \binom{n}{i} \lambda^{i} \left(1-\lambda\right)^{n-i} \ge 1-\omega \tag{1}$$

If a standard Poisson distribution is used to represent "K", the probability is:

$$\mathrm{P}\left(K \le k\right) = \sum_{i=0}^{k} \mathrm{e}^{-n\lambda} \left(n\lambda\right)^{i} \big/ i! \ge 1-\omega \tag{2}$$

When default events are completely independent (ρ = 0%), "upper confidence bounds"[4] for both binomial and Poisson distributions can be computed analytically through beta and gamma approximations or numerically. They provide the same outputs because the binominal distribution is proportional to the beta distribution and the Poisson distribution is proportional to the gamma distribution:

$$\mathrm{Binomial}\left(n,k\right) \propto \mathrm{Beta}\left(\alpha,\beta\right), \ \alpha = k+1 \text{ and } \beta = n-k$$

$$\binom{n}{k} \lambda^{k} \left(1-\lambda\right)^{n-k} \propto \Gamma\left(\alpha+\beta\right) \big/ \left[\Gamma\left(\alpha\right)\Gamma\left(\beta\right)\right] \lambda^{\alpha-1} \left(1-\lambda\right)^{\beta-1} \tag{3}$$

and

$$\mathrm{Poisson}\left(n\lambda\right) \propto \mathrm{Gamma}\left(\alpha,\beta\right), \ \alpha = k+1 \text{ and } \beta = 1/n$$

$$\mathrm{e}^{-n\lambda} \left(n\lambda\right)^{k} \big/ k! \propto \beta^{\alpha} \big/ \Gamma\left(\alpha\right) \mathrm{e}^{-\beta\lambda} \lambda^{\alpha-1} \tag{4}$$

2.2.2. Classical Approach with Structural Models

The theoretical environment must be drastically changed by positive asset correlation values, "ρ". Instead of basic and unrealistic reduced form models that rely on default independence, complex and adequate structural models should be used.

$J\left(\lambda, y, \rho\right)$ represents the probability function of the sample data resulting from the binomial function of "λ" or the Poisson function of "λ", $\phi\left(y\right)$ represents the standard normal probability density function of "Y", $\Phi\left(.\right)$ represents the standard normal cumulative distribution function, and $\Phi^{-1}\left(\lambda\right)$ represents the inversed standard normal cumulative distribution function for "λ". The mean-

---

[3]The true value of "λ" has a $\left(1-\delta\right)\%$ probability of belonging to the shortest possible interval. If $\left[\lambda_{\mathrm{lower}}, \lambda_{\mathrm{upper}}\right]$ is the 90% credible interval of "λ" then $\mathrm{P}\left(\lambda_{\mathrm{lower}} < \lambda < \lambda_{\mathrm{upper}}\right) = \left(100-10\right)\% = 90\%$.

[4]This term is tied to Katja Pluto and Dirk Tasche's "the most prudent estimation" concept that they established in 2005 and applied to the classical default probability. Each risk class contains not only "n" and "k" from that particular class, but also "n" and "k" from other classes with lower rating grades.

ings of "K", "k", "n", "λ", and "ω" are the same above mentioned.

Therefore, the probability of having no more than "k" defaults inside a risk class with "n" obligors is provided by:

$$\mathrm{P}\left(K \le k\right) = \int_{-\infty}^{+\infty}\left[\sum_{i=0}^{k}J\left(\lambda, y, \rho\right)_{i}\right]\phi\left(y\right)\mathrm{d}y \ge 1 - \omega, \tag{5}$$

with

$$J\left(\lambda, y, \rho\right)_{i} = \binom{n}{i}\left[G\left(\lambda, y, \rho\right)\right]^{i}\left[1 - G\left(\lambda, y, \rho\right)\right]^{n-i} \tag{6}$$

or

$$J\left(\lambda, y, \rho\right)_{i} = \mathrm{e}^{-nG\left(\lambda, y, \rho\right)}\left[nG\left(\lambda, y, \rho\right)\right]^{i}\Big/i!, \tag{7}$$

if the probability function of the sample data follows a binomial distribution or a Poisson distribution, respectively, and

$$G\left(\lambda, y, \rho\right) = \Phi\left\{\left[\Phi^{-1}\left(\lambda\right) - y\sqrt{\rho}\right]\Big/\sqrt{1 - \rho}\right\} \tag{8}$$

The function $G\left(\lambda, y, \rho\right)$ distinguishes the formulae of reduced form models from the structural models. A brief explanation of that crucial function and its origin can be found in Appendix A.

2.3. Bayesian Approach

2.3.1. Prior Distributions

The default probability of the posterior distribution is derived from the likelihood and prior densities matching, as previously indicated. Multiple prior functions, ranging from less prudent to over conservative, are identified in this subsection. They reflect various elements or beliefs about the effective default probability.

When there is no understanding of the behavior of posterior default probability, it is common to use a non-informative prior. The most non-informative prior is a flat prior, concretely a uniform distribution between 0 and 1. However, prior functions are useful in most circumstances since they represent the default risk profile, in which case distributions need to be parametrized.

Because of its versatility in expressing the uncertainty of the default probability, the beta distribution is a widely used parametrized prior distribution. Let "Λ" be the random variable of the default probability "λ". Assuming "Λ" follows a beta distribution, $\Lambda \sim \mathrm{Beta}\left(\alpha, \beta\right)$, with $\alpha, \beta > 0$ and $\lambda \in \left]0, 1\right[$, it is simple to adapt this distribution to subjective information about the mean (or average) and variance of default probability using the hyperparameters "α" and "β". The mean and variance of a beta distribution are respectively:

$$\mathrm{E}\left(\Lambda\right) = \alpha\Big/\left(\alpha + \beta\right) \tag{11}$$

$$\mathrm{V}\left(\Lambda\right) = \alpha\beta\Big/\left[\left(\alpha + \beta\right)^{2}\left(\alpha + \beta + 1\right)\right] \tag{12}$$

When no (objective or subjective) information about the default probability is available, a beta distribution with $\alpha = \beta = 1$ can be taken because it represents a

uniform distribution between 0 and 1, making it a non-informative prior.

The set of prior functions, $f(\lambda)$, addressed in this document is listed below.

1) Uniform distribution

$$f(\lambda) = 1/u, \quad \lambda \in \,]0, u], \tag{13}$$

being "u" the maximum limit of "λ". Four possibilities of "u" are tested: 1, 0.25, 0.1, and 0.01—the same values evaluated by Dirk Tasche in his work[5]. These values are also utilized in two other types of priors: linear growth and linear decrease, as shown below. A tighter representation should be used because when $u = 1$ the condition $\lambda < u$ must be met, not $\lambda \leq u$. This note is also applicable to prior functions with linear growth and linear decrease.

2) Linear growth

$$f(\lambda) = \lambda, \quad \lambda \in \,]0, u] \tag{14}$$

3) Linear decrease

$$f(\lambda) = 1 - \lambda/u, \quad \lambda \in \,]0, u] \tag{15}$$

4) Conservative[6]

$$f(\lambda) = 1/(1 - \lambda), \quad \lambda \in \,]0, 1[\tag{16}$$

5) Immoderate

$$f(\lambda) = 1/\lambda, \quad \lambda \in \,]0, 1[\tag{17}$$

6) Expert judgement

6.1) Base scenario (linear growth and linear decrease)

$$f(\lambda) = \begin{cases} 2(\lambda - m)/\left[(M - m)(Mo - m)\right] & \lambda \in \,]m, Mo] \\ 2(\lambda - M)/\left[(M - m)(Mo - M)\right] & \lambda \in \,]Mo, M], \\ 0 & \text{others } \lambda \end{cases} \tag{18}$$

where "m" stands for the minimum of "λ", "Mo" stands for the mode of "λ", and "M" stands for the maximum of "λ", as assumed by the expert who defines the prior distribution. The values for "m", "Mo", and "M" were initially set to be conservative: $m = 0.01$, $Mo = 0.025$, and $M = 0.075$[7].

6.2) Beta distribution as a proxy

For the beta distribution, $\alpha = 6.67$ and $\beta = 175.31$ were used. These two parameters were set to ensure that the beta distribution's mean and variance matched those of the base scenario, as decided by expert opinion—prior 6.1—, resulting in $\mu = 0.03667$ and $\sigma^2 = 0.00019$, respectively.

6.3) Normal distribution as a proxy

To check that the normal distribution's mean is equal to the mean of the prior 6.1 and the variance is equal to 0.03667/1.645, one used $\mu = 0.03667$ and $\sigma^2 = $

---

[5]Tasche (2012).

[6]Term used by Tasche (2012).

[7]In the second stage, another hypothetical scenario with $m = 0.003$, $Mo = 0.01$, and $M = 0.02$ was found considerably more suitable to the real-world issue of low default portfolios.

0.02229 for normal distribution[8].

7) Beta distribution based on empirical default rate

The prior's mean is assumed to be the observed default rate for each combination of "n" and "k", and the prior's variance is expected to be equal to the ratio between that rate (and thus the mean) and the number 1.645. These mean and variance, on the one hand, and Equation (11) and Equation (12), on the other hand, are used to compute the beta distribution's parameters[9]. This prior will only be used in Section 3.4 to compare the classical and Bayesian approaches.

2.3.2. Posterior Distributions

1) Bayesian Approach with Reduced Form Models

Let's use "K" to represent the random variable of the default number once more. For any potential default probability, "λ", each value "y" of a standard and normally distributed random variable, "Y", and a specified asset correlation, "ρ", the posterior probability computation based on the Bayes' theorem[10] of the occurrence exactly "k" and no more than "k" as for Equation (1) and Equation (2) defaults is as follows:

$$
\begin{aligned}
P(K=k) &= H(\lambda, y) \\
&= f(\lambda) \int_{-\infty}^{+\infty} \binom{n}{k} \lambda^k (1-\lambda)^{n-k} \phi(y) \mathrm{d}y \\
&\quad \cdot 1 \bigg/ \left[\int_0^u f(\lambda) \int_{-\infty}^{+\infty} \binom{n}{k} \lambda^k (1-\lambda)^{n-k} \phi(y) \mathrm{d}y \mathrm{d}\lambda \right],
\end{aligned}
\tag{19}
$$

if the likelihood function follows a binomial distribution, or

$$
\begin{aligned}
P(K=k) &= H(\lambda, y) \\
&= f(\lambda) \int_{-\infty}^{+\infty} \mathrm{e}^{-n\lambda} (n\lambda)^k \big/ k! \, \phi(y) \mathrm{d}y \\
&\quad \cdot 1 \bigg/ \left[\int_0^u f(\lambda) \int_{-\infty}^{+\infty} \mathrm{e}^{-n\lambda} (n\lambda)^k \big/ k! \, \phi(y) \mathrm{d}y \mathrm{d}\lambda \right],
\end{aligned}
\tag{20}
$$

if the likelihood function follows a Poisson distribution.

It's worth noting that "u" is the upper limit of "λ"[11], $f(\lambda)$ is the prior probability density function of "λ" as described in 2.3.1, and $\phi(y)$ is the standard normal probability density function of "Y".

The $H(\lambda, y)$ denominator, also known as prior predictive distribution, nor-

---

[8]The 1.645 denominator corresponds to a 90% confidence level.

[9]When there are no defaults, it is used at $k = 0.000000001$ instead of $k = 0$ to ensure that the analytical solution of the beta distribution is valid.

[10]For events "X" and "Z", the conditional probability of "X" given the occurrence of "Z", representing "s" the number of disjoint events, is computed as follows:

$$P(X \mid Z) = P(X \cap Z) / P(Z) = P(Z \mid X) P(X) / P(Z) = P(Z \mid X) P(X) \bigg/ \sum_{j=1}^{s} P(Z \mid X_j) P(X_j)$$

If the probability $P(Z \mid X)$ is derived from a statistical model $L(z \mid x)$ that describes the likelihood function and the probability $P(X)$ is derived from a prior function $\phi(x)$, the posterior density function is obtained by:

$$L(x \mid z) = L(z \mid x) \phi(x) \bigg/ \int L(z \mid x) \phi(x) \mathrm{d}x$$

[11]It should be remembered that four different values of "u" were tested: 1, 0.25, 0.1, and 0.01.

malizes the posterior distribution function.

Analytical outputs can be generated if no correlation of defaults among borrowers is assumed—reduced form models—, as mentioned in Subsection 2.2.1. Only a few circumstances in Bayesian estimation have an explicit analytical solution: when there is also independence among borrowers, and simultaneously when joining the prior function with the likelihood function yields a standard distribution.

Concretely, analytical forms occur in the following special cases: when the likelihood is a binomial distribution and the prior is a beta distribution, on the one hand, and when the likelihood is a Poisson distribution and the prior is a gamma distribution, on the other hand. The beta-binomial distribution and the gamma-Poisson distribution derive from those joining.

2) Bayesian Approach with Structural Models

It should be noted that the existence of correlation requires numerical outputs from stochastic treatments or simulation procedures. Being $\rho > 0$, the formulae stated in the preceding subsection must be changed:

$$
\begin{aligned}
P(K = k) &= H(\lambda, y, \rho) \\
&= f(\lambda) \int_{-\infty}^{+\infty} L(\lambda, y, \rho) \phi(y) \, dy \\
&\quad \cdot 1 \Big/ \left[\int_0^u f(\lambda) \int_{-\infty}^{+\infty} L(\lambda, y, \rho) \phi(y) \, dy \, d\lambda \right],
\end{aligned}
\tag{21}
$$

with

$$
L(\lambda, y, \rho) = \binom{n}{k} \left[G(\lambda, y, \rho) \right]^k \left[1 - G(\lambda, y, \rho) \right]^{n-k}
\tag{22}
$$

or

$$
L(\lambda, y, \rho) = e^{-nG(\lambda, y, \rho)} \left[nG(\lambda, y, \rho) \right]^k \Big/ k!,
\tag{23}
$$

if the likelihood function is represented respectively by a binomial or a Poisson distribution, and $G(\lambda, y, \rho)$ has the meaning expressed in Equation (8).

$L(\lambda, y, \rho)$ denotes the probability of the sample data generated from binomial or Poisson likelihood functions of "λ", for a risk class with "n" counterparties and "K" defaults. $\Phi(.)$ stands for the standard normal cumulative distribution function, and $\Phi^{-1}(\lambda)$ is the inversed standard normal cumulative distribution function for "λ" (see 2.2.2).

The $H(\lambda, y, \rho)$ denominator ensures that the posterior distribution function is normalized once again.

2.4. Conservative Zone

The posterior default probabilities differ significantly depending on the priors used. Some of those probabilities are unwise, while others are overblown. The key goal is to establish limits that will allow for the identification of a conservative zone.

Drift and volatility should be included in every estimate of a random variable

to explain the uncertainty. The historical default data for the drift and the standard deviation of the likelihood distribution for the volatility are adopted to ensure an acceptable margin of conservatism, avoiding unwanted levels of imprudence and exaggeration. The coefficient of skewness is also included in the rule because default distributions are heavily skewed toward positive asset correlations.

On the one hand, if at least one default occurrence is recorded, the following rule is applied to create imprudent, conservative, and exaggerated zones:

$\lambda < \lambda_{97.5\%}$ Imprudent zone,

$\lambda \in \left[\lambda_{97.5\%}, \lambda_{99.9\%}\right]$ Conservative zone,

$\lambda > \lambda_{99.9\%}$ Exaggerated zone.

The default probability "λ_c" is determined as follows for a given confidence level "c":

$$\lambda_c = \left[k/n + \sigma_L \big/ \sqrt{n}\, \Phi^{-1}(c) \right] \gamma_L \qquad (24)$$

The ratio k/n represents the empirical default experience, σ_L represents the standard deviation of the correlated likelihood function, γ_L represents the coefficient of skewness of this function, and $\Phi^{-1}(c)$ represents the inversed standard normal cumulative distribution function for "c". The likelihood function is used to indicate the level of conservatism since that function depicts the group's intrinsic risk and so eliminates the need for any prior risk data. Furthermore, to improve the required conservatism, volatility is computed using the standard deviation of the likelihood function at 97.5% and 99.9% confidence levels.

The expression inside square brackets in Equation (24) is the result of a binomial test of significance adaptation[12]. It should be noted that the conventional binomial test assumes mutual independence of events, which is an incorrect assumption in the default probability models. Aside from the standard deviation of the correlated binomial distribution—i.e., the correlated likelihood function—, the conservative margin should also account for the asymmetry of the same distribution.

On the other hand, when no default event is identified, $\lambda = \mu_L$ is assumed, reflecting μ_L the mean of the correlated likelihood distribution, instead of the conservative zone $\lambda \in \left[\lambda_{97.5\%}, \lambda_{99.9\%}\right]$. One keeps in mind that when $\rho > 0$ the posterior distribution's mean is much greater than when default events independence is assumed. As a result, when no past defaults have been recorded, the mean will be an overly cautious estimator of default probability.

It is important to note that the suggested practical rule should not be used to calculate default probabilities. Its advantage is that it provides for a more impartial comparison between an expected default probability and a conservative threshold based on the likelihood distribution.

---

[12]It's worth noting that the rule is an approximation, thus it does not exactly match the conventional binomial test. The mean for drift and the standard deviation for volatility from the same distribution, the ratio of the standard deviation to the number of observations, a two tailed method, and the lack of skewness are all used in that conventional test.

3. The Application of the Methodology

3.1. Risk Classes and Asset Correlation

The figures in the tables of this paper were made for three different risk categories described in Table 1[13].

The results are provided from two perspectives: an individual approach by credit risk class, in which each class is seen as a separate group; and an integrated approach combining two or more classes[14], in which the upper confidence bound concept[15] is used assuming that the classes combined have the same rating category. When a unique rating grade is assigned to homogenous classes, it is presumed that default risks for all counterparties within the integrated group are exposed to the same default probability (and the same asset correlation).

As a result of regulatory regimes, banks are frequently subjected to a constant "ρ". The Basel Committee on Banking Supervision recognizes that major corporations are more dependent on systematic risk than small firms and retail counterparties because they are more exposed to overall economic conditions. Small firms and retail counterparties are less affected by economic cycles, therefore their defaults are more idiosyncratic rather than systematic.

It will always be set to ρ = 12%, as this is one of the standards available to banking regulators.

Four risk criteria are used to calculate the classical confidence intervals and the Bayesian credible intervals: 50%, 25%, 10%, and 5%.

3.2. Classical Approach

The binomial and Poisson distributions describe the counting of defaults. Table 2 and Table 3 show that those distributions produce effectively identical outcomes. In fact, it is expected that the means of both distributions are identical, $n\lambda = k$, and the variances are likewise quite similar[16]. In the tables, the terms "basic binomial" and "basic Poisson" are used to describe the corresponding distributions where default occurrences are assumed to be completely independent. It is also worth noting that the default probability increases as the confidence level rises. The observed default rates are presented in the tables so that the results from those distributions can be quickly compared to these rates.

Assuming ρ = 0%, upper confidence bound computations can be done analytically (through beta and gamma approximations) or numerically, as described in Subsection 2.2.1. For the five scenarios assumed, the largest discrepancy between the numerical simulation method and the analytical alternative approximation is 0.0003%.

---

[13]Similar to Pluto and Tasche (2005).

[14] $A + B + C = (150,0) + (500,1) + (350,2) = (1000,3)$; and $B + C = (500,1) + (350,2) = (850,3)$.

[15]See also footnote 4.

[16]Because the default probability is believed to be low, the binomial variance, $n\lambda(1-\lambda)$, is marginally lower than the Poisson variance, $n\lambda$.

Table 1. Number of obligors (n) and defaults (k) for each risk class.

Risk class	n	k
A	150	0
B	500	1
C	350	2

Table 2. Classical default probability for basic binomial (asset correlation = 0%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
Confidence level	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
50%	0.46%	0.34%	0.76%	0.37%	0.43%
75%	0.92%	0.54%	1.12%	0.51%	0.60%
90%	1.52%	0.78%	1.51%	0.67%	0.78%
95%	1.98%	0.95%	1.79%	0.77%	0.91%
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%

Table 3. Classical default probability for basic Poisson (asset correlation = 0%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
Confidence level	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
50%	0.46%	0.34%	0.76%	0.37%	0.43%
75%	0.92%	0.54%	1.12%	0.51%	0.60%
90%	1.53%	0.78%	1.52%	0.67%	0.79%
95%	2.00%	0.95%	1.80%	0.78%	0.91%
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%

When the independence assumption is replaced with the correlation assumption and thus standard distributions are turned into correlated distributions, the greater the confidence level, the greater the difference between reduced form and structural models' results. Table 4 and Table 5 shows that at a 50% confidence level, from ρ = 0% to ρ = 12%, default probabilities for binomial distribution range from 0.34% - 0.76% to 0.53% - 1.11% (depending on "n" and "k"). At a 95% confidence level, one rises from 0.77% - 1.98% to 2.89% - 5.16%. Therefore, at a 50% confidence level, default probabilities increase by 46% - 59%[17], and at a 95% confidence level, they increase by 138% - 274%[18]. At 75% and 90% confidence levels, the growth ranges by 81% - 130%[19], and 115% - 214%[20], respectively.

Table 4. Classical default probability for correlated binomial (asset correlation = 12%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
Confidence level	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
50%	0.69%	0.53%	1.11%	0.58%	0.67%
75%	1.66%	1.14%	2.20%	1.18%	1.35%
90%	3.28%	2.13%	3.81%	2.10%	2.38%
95%	4.71%	3.01%	5.16%	2.89%	3.26%
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%

Table 5. Classical default probability for correlated Poisson (asset correlation = 12%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
Confidence level	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
50%	0.69%	0.53%	1.12%	0.58%	0.67%
75%	1.67%	1.14%	2.20%	1.18%	1.35%
90%	3.29%	2.13%	3.81%	2.10%	2.38%
95%	4.73%	3.01%	5.16%	2.89%	3.26%
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%

There are significant capital savings when an integrated approach is used rather than an individual approach[21]. The greater the confidence level, the greater the spread of savings: from 27%[22] at a 50% confidence level to 45% at a 95% confidence level when three risk classes are aggregated as one homogeneous group, and from 16% at a 50% confidence level to 30% at a 95% confidence level when risk classes B and C are aggregated as one homogeneous group. These savings were obtained with ρ = 0%.

When reduced form models are replaced with structural models and asset cor-

---

[17]46% = 1.11% (Table 4)/0.76% (Table 2) – 1, for (350, 2), and 59% = 0.58% (Table 4)/0.37% (Table 2) – 1, for (1000, 3).

[18]138% = 4.71% (Table 4)/1.98% (Table 2) – 1, for (150, 0), and 274% = 2.89% (Table 4)/0.77% (Table 2) – 1, for (1000, 3).

[19]81% = 1.66% (Table 4)/0.92% (Table 2) – 1, for (150, 0), and 130% = 1.18% (Table 4)/0.51% (Table 2) – 1, for (1000, 3).

[20]115% = 3.28% (Table 4)/1.52% (Table 2) – 1, for (150, 0), and 214% = 2.10% (Table 4)/0.67% (Table 2) – 1, for (1000, 3).

[21]For the sake of simplicity, one assumes that the exposure amounts for all loan contracts are always the same, regardless of the rating grade associated with each risk class.

[22] $27\% = 1 - \left[0.37\% \cdot (150 + 500 + 350) \right] / \left[0.46\% \cdot 150 + 0.34\% \cdot 500 + 0.76\% \cdot 350 \right].$

relation is assumed to be uniform, with ρ = 12%, the savings are lower: from 23% at a 50% confidence level to 28% at a 95% confidence level, and from 13% at a 50% confidence level to 16% at a 95% confidence level, respectively for three and two aggregated classes.

3.3. Bayesian Approach

3.3.1. Likelihood Functions

Table 6 and Table 7, on the one hand, and Table 8 and Table 9, on the other hand, show that there are no significant differences between the binomial distribution and the Poisson distribution used as the likelihood function of the Bayesian default probability, with the outcome being roughly the same for both the basic and correlated techniques (similar to the classical approach). Nonetheless, the Poisson distribution's means are marginally higher than the binomial distribution's because the Poisson distribution is a little skewer than the binomial distribution[23]. Hence the matching percentiles associated with the mean in the binomial distribution are immaterially higher than the equivalent percentiles in the Poisson distribution.

Table 6. Bayesian default probability—Likelihood: basic binomial (asset correlation = 0%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%
Mean	0.67%	0.40%	0.85%	0.40%	0.47%
(matching percentile)	63.59	61.70	57.41	53.46	58.90
Standard deviation	0.65%	0.28%	0.49%	0.20%	0.23%
Coefficient of skewness	1.96	1.41	1.14	0.99	0.99

Table 7. Bayesian default probability—Likelihood: basic Poisson (asset correlation = 0%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%
Mean	0.68%	0.40%	0.86%	0.40%	0.47%
(matching percentile)	63.21	61.53	57.12	53.35	58.76
Standard deviation	0.67%	0.28%	0.49%	0.20%	0.24%
Coefficient of skewness	2.00	1.42	1.15	1.00	1.00

---

[23]The largest deviation in the mean of the default probability (0.014%) occurs with $n = 150$, $k = 0$ and ρ = 12%.

Table 8. Bayesian default probability—Likelihood: correlated binomial (asset correlation = 12%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%
Mean	1.87%	1.51%	2.64%	1.62%	1.80%
(matching percentile)	66.73	65.16	63.50	64.04	64.14
Standard deviation	2.21%	1.55%	2.34%	1.49%	1.64%
Coefficient of skewness	2.66	2.42	2.03	2.16	2.12

Table 9. Bayesian default probability—Likelihood: correlated Poisson (asset correlation = 12%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
Empirical default rate (k/n)	0.00%	0.20%	0.57%	0.30%	0.35%
Mean	1.88%	1.52%	2.64%	1.62%	1.80%
(matching percentile)	66.51	65.09	63.37	64.00	64.09
Standard deviation	2.23%	1.56%	2.35%	1.49%	1.64%
Coefficient of skewness	2.67	2.42	2.03	2.16	2.12

When comparing Table 6 and Table 8 (or Table 7 and Table 9), it is clear that (regardless of "n" and "k") the larger "ρ", the greater the likelihood function's mean and standard deviation. When the asset correlation is introduced, the likelihood function becomes considerably skewer than when no correlation is employed. Furthermore, one verifies the rule that the larger the risk group, the higher the coefficient of skewness and the smaller the mean and the standard deviation[24].

There are capital savings with the integration of risk categories, as seen in the classical approach. Savings in the Bayesian context, assuming ρ = 12%, equate to 9% or 18% whether two risk classes (B + C) or three risk classes (A + B + C) are aggregated. For ρ = 0%, corresponding savings increase to 20% or 33%—34% (rather than 33%) regarding that the likelihood function follows a Poisson distribution—if two or three risk classes are aggregated[25].

[24]This rule may also be shown comparing the values of the posterior probabilities in the last two columns of Table 10 and Table 11 which will be presented later. They are both tied to the same number of defaults, k = 3.

The lowering effect on the standard deviation when "n" grows is explained by the fact that the percentage increase in "n" is smaller than the modulus of the percentage decrease in "λ".

3.3.2. Posterior Distributions

A binomial distribution is suitably better than a Poisson distribution for depicting the number of defaults within a risk class containing a fixed number of obligors, as stated in Section 2.1. As a result, for the sake of simplicity, only outcomes based on the binomial distribution are now reported.

The prior functions discussed in 2.3.1 are used to find a group of statistical values for the posterior distributions: mean, median, mode, standard deviation and coefficient of skewness as well as four highest density intervals. Table 10 displays the outcomes of a 12% asset correlation, while Table 11 in Appendix B shows the results of no correlation[26]. These tables indicate that different prior functions and asset correlations have a big impact on posterior probabilities[27].

Table 10. Default probability distribution—Posterior using the correlated binomial (asset correlation = 12%) as likelihood, for several priors.

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
Types of prior function	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
1) Uniform distribution					
1.1) $u = 1$					
Mean	1.87%	1.51%	2.64%	1.62%	1.80%
Median	1.07%	0.98%	1.87%	1.10%	1.24%
Mode	0.03%	0.30%	0.79%	0.45%	0.51%
Highest density interval					
50%	0.03 - 1.06	0.07 - 0.99	0.24 - 2.02	0.15 - 1.18	0.17 - 1.36
75%	0.02 - 2.43	0.03 - 1.97	0.12 - 3.51	0.09 - 2.13	0.10 - 2.34
90%	0.01 - 4.51	0.02 - 3.35	0.06 - 5.57	0.03 - 3.47	0.07 - 3.86
95%	0.00 - 6.17	0.01 - 4.51	0.03 - 7.25	0.01 - 4.56	0.03 - 5.02
Standard deviation	2.21%	1.55%	2.34%	1.49%	1.64%
Coefficient of skewness	2.66	2.42	2.03	2.16	2.12
1.2) $u = 0.25$					
Mean	1.87%	1.51%	2.64%	1.62%	1.80%
Median	1.07%	0.98%	1.87%	1.10%	1.24%
Mode	0.03%	0.30%	0.79%	0.45%	0.51%

[25]When capital savings derived from the classical approach—see the last two paragraphs of 3.2—are compared to capital savings derived from the Bayesian approach, the former are higher (than the latter) at a 50% confidence level when ρ = 12%—13% > 9% and 23% > 18%—, and lower (at the same confidence level) when ρ = 0%—16% < 20% and 27% < 33%.

[26]Although both priors 6.2 and 7 are connected to beta distribution, only the first one is included in Table 10 and Table 11.

[27]With ρ = 12%, the means for likelihood distribution—or posterior with uniform distribution and u = 1 as a prior—are three or four times bigger than those with ρ = 0%, depending on the pair "n" and "k" considered.

Continued

Highest density interval					
50%	0.03 - 1.06	0.07 - 0.99	0.24 - 2.02	0.15 - 1.18	0.17 - 1.36
75%	0.02 - 2.43	0.03 - 1.97	0.12 - 3.51	0.09 - 2.13	0.10 - 2.34
90%	0.01 - 4.51	0.02 - 3.35	0.06 - 5.57	0.03 - 3.47	0.07 - 3.86
95%	0.00 - 6.17	0.01 - 4.51	0.03 - 7.25	0.01 - 4.56	0.03 - 5.02
Standard deviation	2.20%	1.55%	2.34%	1.49%	1.64%
Coefficient of skewness	2.61	2.42	2.02	2.16	2.12
1.3) $u = 0.1$					
Mean	1.74%	1.49%	2.47%	1.60%	1.77%
Median	1.04%	0.97%	1.85%	1.08%	1.22%
Mode	0.03%	0.30%	0.79%	0.44%	0.50%
Highest density interval					
50%	0.03 - 1.05	0.07 - 0.97	0.29 - 1.99	0.16 - 1.17	0.18 - 1.34
75%	0.02 - 2.35	0.03 - 1.93	0.14 - 3.37	0.09 - 2.10	0.10 - 2.29
90%	0.01 - 4.22	0.02 - 3.31	0.07 - 5.28	0.03 - 3.42	0.08 - 3.78
95%	0.00 - 5.67	0.01 - 4.47	0.05 - 6.57	0.01 - 4.48	0.04 - 4.90
Standard deviation	1.83%	1.46%	1.97%	1.44%	1.55%
Coefficient of skewness	1.96	2.00	1.31	1.90	1.78
1.4) $u = 0.01$					
Mean	0.44%	0.50%	0.59%	0.55%	0.57%
Median	0.37%	0.42%	0.61%	0.49%	0.53%
Mode	0.03%	0.30%	0.77%	0.46%	0.49%
Highest density interval					
50%	0.03 - 0.35	0.17 - 0.51	0.63 - 1.00	0.29 - 0.64	0.37 - 0.72
75%	0.02 - 0.61	0.12 - 0.73	0.42 - 1.00	0.23 - 0.83	0.28 - 0.90
90%	0.01 - 0.82	0.09 - 0.89	0.26 - 1.00	0.20 - 0.95	0.21 - 1.00
95%	0.00 - 0.90	0.04 - 0.93	0.19 - 1.00	0.15 - 1.00	0.17 - 1.00
Standard deviation	0.28%	0.27%	0.25%	0.26%	0.26%
Coefficient of skewness	0.32	0.14	-0.21	0.01	0.07
2) Linear growth					
2.1) $u = 1$					
Mean	4.47%	3.10%	4.72%	2.99%	3.29%
Median	3.43%	2.42%	3.84%	2.40%	2.61%
Mode	1.52%	1.20%	2.23%	1.33%	1.49%

Continued

Highest density interval					
50%	0.52 - 3.71	0.48 - 2.65	1.06 - 4.33	0.57 - 2.67	0.68 - 2.94
75%	0.23 - 6.10	0.25 - 4.23	0.63 - 6.52	0.36 - 4.11	0.40 - 4.50
90%	0.09 - 9.31	0.13 - 6.32	0.34 - 9.31	0.19 - 5.98	0.23 - 6.55
95%	0.07 - 11.76	0.07 - 7.89	0.23 - 11.36	0.11 - 7.34	0.17 - 8.06
Standard deviation	3.66%	2.42%	3.33%	2.18%	2.37%
Coefficient of skewness	1.66	1.61	1.42	1.45	1.44
2.2) $u = 0.25$					
Mean	4.46%	3.10%	4.72%	2.99%	3.29%
Median	3.43%	2.42%	3.84%	2.40%	2.61%
Mode	1.52%	1.20%	2.23%	1.33%	1.51%
Highest density interval					
50%	0.52 - 3.71	0.48 - 2.65	1.06 - 4.33	0.57 - 2.67	0.68 - 2.94
75%	0.23 - 6.10	0.25 - 4.23	0.63 - 6.52	0.36 - 4.11	0.40 - 4.50
90%	0.09 - 9.29	0.13 - 6.32	0.34 - 9.31	0.19 - 5.98	0.23 - 6.55
95%	0.07 - 11.71	0.07 - 7.89	0.23 - 11.30	0.11 - 7.34	0.17 - 8.06
Standard deviation	3.61%	2.42%	3.32%	2.18%	2.37%
Coefficient of skewness	1.56	1.61	1.40	1.45	1.44
2.3) $u = 0.1$					
Mean	3.67%	2.93%	4.04%	2.89%	3.13%
Median	3.10%	2.34%	3.59%	2.36%	2.58%
Mode	1.51%	1.17%	2.22%	1.33%	1.50%
Highest density interval					
50%	0.58 - 3.45	0.53 - 2.63	1.17 - 4.13	0.62 - 2.64	0.71 - 2.87
75%	0.29 - 5.34	0.27 - 4.07	0.71 - 5.85	0.39 - 4.03	0.43 - 4.40
90%	0.16 - 7.39	0.17 - 5.98	0.44 - 7.66	0.22 - 5.77	0.25 - 6.21
95%	0.09 - 8.42	0.10 - 7.23	0.38 - 8.63	0.15 - 6.97	0.20 - 7.38
Standard deviation	2.42%	2.07%	2.34%	1.98%	2.07%
Coefficient of skewness	0.68	1.06	0.59	1.10	0.99
2.4) $u = 0.01$					
Mean	0.63%	0.65%	0.71%	0.67%	0.68%
Median	0.67%	0.68%	0.77%	0.70%	0.72%
Mode	1.00%	1.00%	1.00%	1.00%	1.00%
Highest density interval					
50%	0.69 - 1.00	0.16 - 0.51	0.75 - 1.00	0.72 - 1.00	0.70 - 1.00
75%	0.44 - 1.00	0.10 - 0.73	0.60 - 1.00	0.51 - 1.00	0.57 - 1.00

Continued

90%	0.30 - 1.00	0.07 - 0.90	0.42 - 1.00	0.36 - 1.00	0.38 - 1.00
95%	0.17 - 1.00	0.05 - 0.93	0.31 - 1.00	0.29 - 1.00	0.30 - 1.00
Standard deviation	0.26%	0.24%	0.22%	0.23%	0.22%
Coefficient of skewness	-0.34	-0.37	-0.59	-0.44	-0.48
3) Linear decrease					
3.1) $u = 1$					
Mean	1.82%	1.49%	2.58%	1.59%	1.77%
Median	1.05%	0.94%	1.87%	1.10%	1.26%
Mode	0.06%	0.29%	0.80%	0.45%	0.53%
Highest density interval					
50%	0.00 - 1.04	0.07 - 0.98	0.25 - 1.97	0.16 - 1.20	0.19 - 1.34
75%	0.00 - 2.37	0.02 - 1.91	0.12 - 3.43	0.10 - 2.09	0.09 - 2.33
90%	0.00 - 4.36	0.02 - 3.32	0.07 - 5.47	0.03 - 3.38	0.06 - 3.80
95%	0.00 - 6.00	0.01 - 4.46	0.02 - 7.07	0.01 - 4.45	0.03 - 4.91
Standard deviation	2.14%	1.52%	2.28%	1.47%	1.61%
Coefficient of skewness	2.65	2.41	2.03	2.16	2.12
3.2) $u = 0.25$					
Mean	1.66%	1.41%	2.39%	1.52%	1.69%
Median	0.96%	0.93%	1.74%	1.05%	1.20%
Mode	0.04%	0.27%	0.76%	0.43%	0.49%
Highest density interval					
50%	0.00 - 0.95	0.06 - 0.96	0.24 - 1.90	0.17 - 1.11	0.20 - 1.30
75%	0.00 - 2.14	0.03 - 1.80	0.12 - 3.00	0.10 - 1.95	0.10 - 2.25
90%	0.00 - 3.97	0.02 - 3.11	0.07 - 5.02	0.03 - 3.20	0.08 - 3.60
95%	0.00 - 5.40	0.01 - 4.20	0.03 - 6.45	0.01 - 4.23	0.05 - 4.61
Standard deviation	1.89%	1.42%	2.06%	1.38%	1.51%
Coefficient of skewness	2.47	2.35	1.94	2.13	2.08
3.3) $u = 0.1$					
Mean	1.33%	1.24%	1.96%	1.35%	1.48%
Median	0.83%	0.80%	1.51%	0.97%	1.09%
Mode	0.03%	0.25%	0.72%	0.41%	0.46%
Highest density interval					
50%	0.00 - 0.80	0.07 - 0.84	0.28 - 1.66	0.15 - 1.07	0.19 - 1.20

Continued

75%	0.00 - 1.77	0.03 - 1.61	0.12 - 2.70	0.10 - 1.79	0.10 - 1.97
90%	0.00 - 3.11	0.02 - 2.68	0.05 - 4.02	0.07 - 2.83	0.06 - 3.10
95%	0.00 - 4.17	0.01 - 3.57	0.03 - 5.01	0.05 - 3.67	0.03 - 3.92
Standard deviation	1.37%	1.16%	1.51%	1.15%	1.23%
Coefficient of skewness	1.93	1.93	1.38	1.81	1.71
3.4) $u = 0.01$					
Mean	0.30%	0.35%	0.47%	0.40%	0.41%
Median	0.21%	0.27%	0.42%	0.32%	0.36%
Mode	0.01%	0.14%	0.33%	0.25%	0.27%
Highest density interval					
50%	0.00 - 0.20	0.10 - 0.31	0.22 - 0.49	0.16 - 0.40	0.18 - 0.47
75%	0.00 - 0.39	0.07 - 0.48	0.15 - 0.62	0.11 - 0.54	0.13 - 0.60
90%	0.00 - 0.59	0.03 - 0.63	0.10 - 0.71	0.07 - 0.68	0.10 - 0.72
95%	0.00 - 0.70	0.02 - 0.70	0.07 - 0.79	0.04 - 0.75	0.06 - 0.78
Standard deviation	0.22%	0.22%	0.22%	0.21%	0.21%
Coefficient of skewness	0.81	0.58	0.25	0.43	0.38
4) Conservative					
Mean	1.92%	1.54%	2.69%	1.64%	1.83%
Median	1.10%	0.97%	1.94%	1.12%	1.30%
Mode	0.05%	0.30%	0.78%	0.45%	0.51%
Highest density interval					
50%	0.03 - 1.09	0.07 - 1.03	0.26 - 2.04	0.15 - 1.20	0.20 - 1.40
75%	0.02 - 2.47	0.03 - 1.98	0.11 - 3.60	0.10 - 2.17	0.10 - 2.42
90%	0.01 - 4.61	0.02 - 3.43	0.06 - 5.74	0.03 - 3.48	0.07 - 3.93
95%	0.00 - 6.40	0.01 - 4.62	0.03 - 7.45	0.01 - 4.62	0.03 - 5.11
Standard deviation	2.28%	1.58%	2.41%	1.52%	1.67%
Coefficient of skewness	2.67	2.42	2.03	2.16	2.12
5) Immoderate					
Mean	0.38%	0.49%	1.09%	0.66%	0.75%
Median	0.10%	0.19%	0.62%	0.35%	0.42%
Mode	0.02%	0.02%	0.13%	0.10%	0.11%
Highest density interval					
50%	0.03 - 0.08	0.03 - 0.19	0.03 - 0.60	0.03 - 0.37	0.03 - 0.41

Continued

75%	0.02 - 0.32	0.02 - 0.54	0.02 - 1.36	0.02 - 0.80	0.02 - 0.90
90%	0.01 - 0.90	0.01 - 1.17	0.01 - 2.53	0.01 - 1.47	0.01 - 1.70
95%	0.00 - 1.53	0.00 - 1.72	0.00 - 3.51	0.00 - 2.13	0.00 - 2.38
Standard deviation	0.75%	0.71%	1.30%	0.79%	0.89%
Coefficient of skewness	5.84	4.07	3.09	3.31	3.21
6) Expert judgement					
6.1) Base scenario					
Mean	2.84%	2.66%	2.94%	2.64%	2.70%
Median	2.57%	2.40%	2.64%	2.39%	2.46%
Mode	2.50%	2.33%	2.50%	2.31%	2.44%
Highest density interval					
50%	1.67 - 2.87	1.60 - 2.72	1.76 - 3.03	1.62 - 2.73	1.67 - 2.78
75%	1.33 - 3.54	1.30 - 2.99	1.41 - 3.73	1.27 - 3.25	1.30 - 3.36
90%	1.17 - 4.48	1.11 - 4.10	1.20 - 4.58	1.12 - 4.07	1.13 - 4.14
95%	1.10 - 5.07	1.07 - 4.69	1.10 - 5.21	1.07 - 4.62	1.09 - 4.80
Standard deviation	1.13%	1.03%	1.14%	1.02%	1.04%
Coefficient of skewness	0.98	1.09	0.92	1.10	1.06
6.2) Beta distribution as a proxy					
Mean	2.86%	2.67%	2.97%	2.66%	2.72%
Median	2.66%	2.47%	2.79%	2.48%	2.53%
Mode	2.40%	2.22%	2.51%	2.21%	2.27%
Highest density interval					
50%	1.80 - 3.16	1.64 - 2.90	1.88 - 3.25	1.66 - 2.92	1.71 - 2.98
75%	1.40 - 3.80	1.31 - 3.57	1.53 - 3.92	1.30 - 3.53	1.34 - 3.57
90%	1.10 - 4.55	1.02 - 4.26	1.20 - 4.67	1.01 - 4.19	1.04 - 4.30
95%	0.90 - 5.05	0.87 - 4.76	1.00 - 5.14	0.86 - 4.70	0.90 - 4.79
Standard deviation	1.13%	1.06%	1.12%	1.04%	1.06%
Coefficient of skewness	0.79	0.81	0.77	0.80	0.80
6.3) Normal distribution as a proxy					
Mean	2.10%	1.86%	2.61%	1.95%	2.09%
Median	1.73%	1.48%	2.32%	1.61%	1.76%
Mode	0.04%	0.50%	1.67%	0.74%	0.97%

Continued

Highest density interval					
50%	0.03 - 1.70	0.16 - 1.59	0.83 - 2.77	0.28 - 1.77	0.40 - 1.99
75%	0.02 - 3.02	0.07 - 2.61	0.44 - 3.76	0.16 - 2.75	0.19 - 2.97
90%	0.01 - 4.33	0.03 - 3.80	0.21 - 4.83	0.10 - 3.84	0.10 - 4.07
95%	0.00 - 5.11	0.01 - 4.55	0.13 - 5.52	0.07 - 4.56	0.07 - 4.78
Standard deviation	1.60%	1.40%	1.56%	1.36%	1.40%
Coefficient of skewness	0.94	1.10	0.78	1.07	0.99

Highest density intervals expressed in percentage.

It's worth noting that, according to the concept of expected value, the mean of the posterior probability of "λ", μ_λ, is computed by:

$$\mu_\lambda = \int_0^u \lambda f(\lambda) \int_{-\infty}^{+\infty} L(\lambda, y, \rho) \phi(y) \mathrm{d}y \mathrm{d}\lambda$$
$$\cdot 1 \Big/ \int_0^u f(\lambda) \int_{-\infty}^{+\infty} L(\lambda, y, \rho) \phi(y) \mathrm{d}y \mathrm{d}\lambda, \tag{25}$$

having $f(\lambda)$, $L(\lambda, y, \rho)$, and $\phi(y)$ the same meaning as before.

The effect described in the last phrase of the penultimate paragraph of 3.3.1 about the growth of "n" is validated for all priors used: the larger the risk group, the smaller the mean and the standard deviation, and the greater the coefficient of skewness.

When the prior function is a uniform distribution spanning from 0 to 1 (i.e., the same figures as the default probability assumes), the values for posterior probabilities are obviously the same as those provided by the likelihood function. This can be seen by comparing Table 8 and Table 10 likewise Table 6 and Table 11. By definition, default probabilities are tiny; empirical rates for the five scenarios range from 0% to 0.57%. As a result, a uniform distribution with values between 0 and 0.25 is expected to yield identical results (about the likelihood function and the posterior with uniform distribution as a prior). Even if the top limit of the uniform distribution is set to 0.1, there are no discernible differences.

Table 10 shows that the posterior with an immoderate prior—the polar opposite of the conservative prior—produces the lowest default probability, with the mean of 4.3 to 11.9 times lower than the posterior resulting from linear growth as a prior (being "u" constrained between 0 and 1) and between 2.4 and 5 times lower than the posterior resulting from the uniform distribution as a prior (being "u" between 0 and 1 too). Therefore, the most cautious or conservative prior is the linear growth function—the theoretical polar opposite of the linear decrease prior—, not the conservative prior itself.

The uniform distribution, linear decrease (both with "u" between 0 and 1), and conservative priors generate comparable posterior default probabilities for all the pairs of "n" and "k" studied. Those three sorts of functions will appear to be the most appropriate and beneficial priors, in contrast to immoderate and linear growth priors. Comments on expert knowledge priors—base scenario as well as

beta and normal distributions as proxie—will be addressed later.

One returns to capital savings through risk group integration. Using the linear growth function with "u" between 0 and 1 as a prior and $\rho = 12\%$, savings are substantially equivalent to savings at a 50% confidence level under the classical approach: 13% for $n = 850$ and 23% for $n = 1000$. Furthermore, a 95% confidence level comparison is required when $\rho = 0\%$. This is because, at a 95% confidence level, savings obtained using the linear growth function as a prior are similar to those obtained using the classical approach: 28% for $n = 850$ and 44% for $n = 1000$, which are close to the 30% and 45% mentioned in 3.2, respectively.

Because the uniform distribution is a non-informative prior, posterior distribution outputs match to likelihood function outputs as aforementioned; corresponding capital savings have already been presented (in the last paragraph of 3.3.1). Being the means of posterior probabilities identical for uniform distribution, linear decrease, and conservative priors (ranging from 0 to 1), the savings are also equivalent. Immoderate prior yields the smallest savings.

Table 10 and Table 11 show a simple rule: the stronger the asset correlation, the wider the tail of the right side of the distribution (or the higher the coefficient of skewness)[28]. However, this rule is not established in two instances: when $u = 0.01$ (and in some cases, when $u = 0.1$ for linear growth prior), and when expert information is the base of the prior function. The first exception is self-evident, as "λ" has a smaller upper limit—0.01 (or 0.1 is some cases)—than the broad range of values that default probabilities can tolerate (based on the binomial distribution).

The second exemption is explained by the nature of the prior subjected to expert judgment. The previous point concerning the 0.01 threshold also applies to priors based on that judgment. These judgments typically specify a much narrower range of values than the likelihood function's available values[29].

It is important to remember that priors that provide information have a stronger influence on the posterior than priors that do not. Furthermore, as previously stated, a prior based on expert opinion is excessively rigid because the posterior probability density is constrained by the range of values for expert knowledge as a prior.

Table 12 compares the cumulative densities of prior and posterior distributions linked to expert information thresholds to the densities of different kinds of priors and their corresponding posterior distributions. The mean is utilized as a reference point for prior and posterior distributions. In the case of employing the normal distribution as a proxy for the base scenario of expert knowledge, $\mu = 0.03912$ rather than 0.03667 because only positive values less than 1 were used[30].

---

[28]The second paragraph of 3.3.1 came to a similar conclusion.

[29]This is clear in the base scenario since the full mass of probability density is fixed between 0.01 and 0.075, the minimum and maximum values of the expert prior function. However, it can be seen in any other priors that are used as a proxy for the base scenario.

[30]One notes that unlike the beta distribution, which has an incognita that ranges from 0 to 1—the same range as the default probability—, the normal distribution's domain is between $-\infty$ and $+\infty$, whereas the prior function's domain may only allow values between 0 and 1.

Table 12. Prior and posterior distributions—Densities by intervals of default probabilities (λ), for 350 obligors, 2 defaults, and a 12% asset correlation.

	]0%; 1%[	[1%; 7.5%]	]7.5%; 100%[	Total
Binomial likelihood				
Cumulative density	24.1%	71.3%	4.6%	100.0%
Mean	0.14%	2.04%	0.45%	2.64%
Uniform distribution, λ ∈]0,1[				
Posterior				
Cumulative density	24.1%	71.3%	4.6%	100.0%
Mean	0.14%	2.04%	0.45%	2.64%
Prior				
Cumulative density	1.0%	6.5%	92.5%	100.0%
Mean	0.0%	0.3%	49.7%	50.0%
Linear growth, λ ∈]0,1[				
Posterior				
Cumulative density	5.3%	77.4%	17.2%	100.0%
Mean	0.04%	2.87%	1.81%	4.72%
Prior				
Cumulative density	0.0%	0.6%	99.4%	100.0%
Mean	0.0%	0.0%	66.6%	66.7%
Linear decrease, λ ∈]0,1[				
Posterior				
Cumulative density	24.6%	71.2%	4.2%	100.0%
Mean	0.14%	2.02%	0.42%	2.58%
Prior				
Cumulative density	1.9%	12.5%	85.6%	100.0%
Mean	0.0%	0.5%	32.8%	33.3%
Conservative				
Posterior				
Cumulative density	23.6%	71.5%	4.9%	100.0%
Mean	0.14%	2.06%	0.49%	2.69%
Prior				
Cumulative density	0.1%	0.8%	99.1%	100.0%
Mean	0.0%	0.0%	88.3%	88.4%
Immoderate				
Posterior				
Cumulative density	63.7%	35.8%	0.5%	100.0%
Mean	0.26%	0.78%	0.05%	1.09%

Continued

Prior				
Cumulative density	46.2%	23.7%	30.1%	100.0%
Mean	0.1%	0.8%	10.8%	11.6%
Expert judgement				
Base scenario (#)				
Posterior				
Cumulative density	0.0%	100.0%	0.0%	100.0%
Mean	0.00%	2.94%	0.00%	2.94%
Prior				
Cumulative density	0.0%	100.0%	0.0%	100.0%
Mean	0.0%	3.7%	0.0%	3.7%
Beta distribution as a proxy				
Posterior				
Cumulative density	0.9%	98.9%	0.1%	100.0%
Mean	0.01%	2.95%	0.01%	2.97%
Prior				
Cumulative density	0.4%	98.5%	1.1%	100.0%
Mean	0.0%	3.6%	0.1%	3.7%
Normal distribution as a proxy				
Posterior				
Cumulative density	14.6%	84.9%	0.5%	100.0%
Mean	0.09%	2.48%	0.04%	2.61%
Prior				
Cumulative density	6.7%	88.9%	4.4%	100.0%
Mean	0.0%	3.5%	0.4%	3.9%

(#) minimum of $\lambda = 0.01$, mode of $\lambda = 0.025$, and maximum of $\lambda = 0.075$.

Figure 1 graphically depicts the pattern of eleven types of default rates as a function of the number of obligors, when $k = 2$ and $\rho = 12\%$: posterior default probabilities for nine prior functions (UD, LG, LD, C, I, EJ-BS, EJ-BD, EJ-ND, and EJ-ABS), probabilities for the likelihood of the default distribution (BL), and historical default occurrences (EDR). The most important topics of that figure are described below.

When expert information is used as a prior, the general rule that the larger the risk group, the less relevant the prior distribution in defining the posterior distribution, and the stronger the convergence of the posterior distribution to the likelihood function is not obvious. That rule and, in the case of expert judgment as a prior, the exception as well as the trend associated with an alternative base scenario for expert judgment (EJ-ABS) are depicted in the figure.
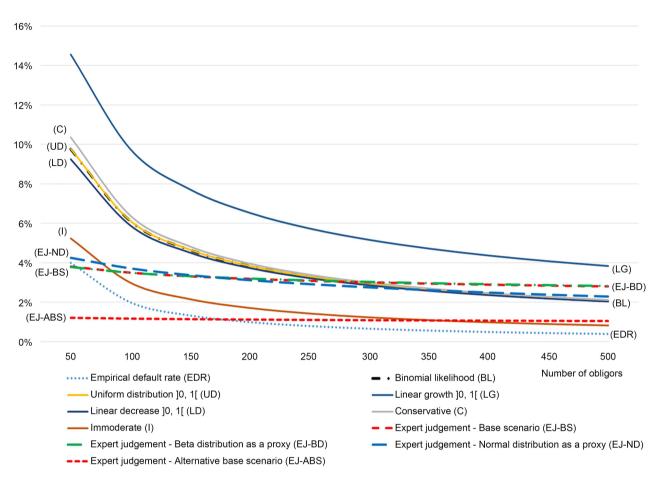
Figure 1. Posterior default probability for different prior distributions (number of defaults = 2; asset corelation = 12%).

In other words, when the size of the risk class, "n", grows, the ratio between posterior probability with prior based on expert opinion and likelihood probability climbs significantly. The likelihood probabilities decrease greatly when "n" rises, which is not the case with the posterior based on expert information as a prior. By contrast, when posterior probability is calculated from any other prior, the ratio remains roughly constant because both posterior and likelihood probabilities are sensitive to "n" in almost equal proportions. Similarly, the ratio of empirical default rate to likelihood probability is quite stable across "n".

When "k" and "ρ" are fixed, the figure shows that posterior default probabilities tend to empirical rates as the number of obligors grows, as expected by the law of large numbers. There is no such coherence when posterior probabilities are generated from any prior that relies on expert judgment.

From $n = 50$ to $n = 500$, with $k = 2$ and $\rho = 12\%$, the empirical default rate falls by 360 basis points (bp), from 4% to 0.4%. The likelihood decreases by 770 bp, from 9.77% to 2.07%, corresponding to same drop as the posterior with uniform distribution between 0 and 1 as a prior. For posteriors with linear growth, conservative, linear decrease, and immoderate functions, the respective declines are 1,072, 825, 721, and 441 bp. For posteriors with hypothetical scenarios related to 1%/2.5%/7.5% (EJ-BS) and 0.3%/1%/2% (EJ-ABS)—for the minimum/mode/

maximum values of the prior default probability—, the decreases are only 101 and 16 bp, respectively.

Finally, Figure 1 shows that the concavity verified in general for posterior functions of default probabilities does not exist when expert prior knowledge is provided, regardless of the values taken by the base scenario as a prior. To demonstrate the (in)flexibility of expert distributions, one notes that coefficients of variation—the mean and variance of these coefficients are measured via the 10 means (one for each n-value[31])—in two aforesaid expert scenarios are vastly smaller: 0.1 or 0.04 regarding respectively EJ-BS or EJ-ABS. The coefficients of variation for posteriors with other priors are substantially greater, ranging from 0.48 (linear growth) to 0.7 (immoderate).

The prior density of the default probability for EJ-BS has a mean of 3.67% and a standard deviation of 1.39%, whereas EJ-ABS as a prior has a mean of 1.1% and a standard deviation of 0.35%. Although these two scenarios are quite different, they both display the same inflexibility: large risk class sizes are insufficient to ensure a substantial convergence of the posterior distribution to the likelihood function, and thus the prior is stronger than the likelihood. In light of the foregoing concerns, selecting a prior based on expert information should be approached with caution.

The 0.01 and 0.1 thresholds are retaken. The two formers of the four thresholds of "u" used for uniform, linear growth, and linear decrease priors—1, 0.25, 0.1, and 0.01—will be sufficient because they reflect the default probability profile. The others, particularly the latter (0.01), have some difficulties due to differences in probability profiles. The following is a quick rundown of these difficulties.

There are a number of inconsistencies when the upper threshold assumed for "λ" is believed to be low. As previously mentioned (after the analysis of Table 10 and Table 11), the stronger the asset correlation, the higher the coefficient of skewness. However, this rule does not work with $u = 0.01$ and (for linear growth function as a prior) with $u = 0.1$.

The default distribution is right-skewed, regardless of the asset correlation percentages that are connected. Nevertheless, the coefficient of skewness might be negative when $u = 0.01$ (either for $\rho = 0\%$ or $\rho = 12\%$).

The higher the asset correlation, the broader the range of the highest density interval; this is, as correlation increases, the distance between the upper and lower bounds of the shortest credible interval growths. At times, such a rule is not verified when $u = 0.01$. Furthermore, in some circumstances with $u = 0.01$, that range is wider with $\rho = 0\%$ than with $\rho = 12\%$.

3.4. Comparison of the Results

Table 13 compares the classical and Bayesian approaches. For each value of the mean linked to the posterior default probability of Table 10 and Figure 1, the corresponding matching percentile was found using the classical binomial ap-

---

[31]Multiples of 50, from $n = 50$ to $n = 500$.

proach, assuming a correlation of 12%. That table also provides the matching percentile associated with the beta prior for the empirical default rate as well as the matching percentile related to this empirical rate. It also provides minimum and maximum default probabilities, which are derived using the practical rule described in 2.4.

Table 13. Mean of the Bayesian default distributions and corresponding matching percentiles based on the classical binomial approach (asset correlation = 12%).

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(n, k)		
Types of prior function	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
1) Uniform distribution					
1.1) u = 1					
Mean	1.87%	1.51%	2.64%	1.62%	1.80%
Matching percentile	77.91	82.62	80.71	84.17	83.62
1.2) u = 0.25					
Mean	1.87%	1.51%	2.64%	1.62%	1.80%
Matching percentile	77.89	82.62	80.70	84.17	83.62
1.3) u = 0.1					
Mean	1.74%	1.49%	2.47%	1.60%	1.77%
Matching percentile	76.00	82.19	78.76	83.91	83.19
1.4) u = 0.01					
Mean	0.44%	0.50%	0.59%	0.55%	0.57%
Matching percentile	38.05	48.06	28.29	47.85	43.61
2) Linear growth					
2.1) u = 1					
Mean	4.47%	3.10%	4.72%	2.99%	3.29%
Matching percentile	94.38	95.34	93.76	95.41	95.11
2.2) u = 0.25					
Mean	4.46%	3.10%	4.72%	2.99%	3.29%
Matching percentile	94.33	95.34	93.75	95.41	95.11
2.3) u = 0.1					
Mean	3.67%	2.93%	4.04%	2.89%	3.13%
Matching percentile	91.72	94.68	91.12	95.01	94.48
2.4) u = 0.01					
Mean	0.63%	0.65%	0.71%	0.67%	0.68%
Matching percentile	47.19	56.67	33.88	55.27	50.52
3) Linear decrease					

Continued

3.1) $u = 1$					
Mean	1.82%	1.49%	2.58%	1.59%	1.77%
Matching percentile	77.22	82.21	80.06	83.81	83.21
3.2) $u = 0.25$					
Mean	1.66%	1.41%	2.39%	1.52%	1.69%
Matching percentile	74.81	80.83	77.70	82.59	81.81
3.3) $u = 0.1$					
Mean	1.33%	1.24%	1.96%	1.35%	1.48%
Matching percentile	68.75	77.24	71.00	79.31	78.01
3.4) $u = 0.01$					
Mean	0.30%	0.35%	0.47%	0.40%	0.41%
Matching percentile	28.91	37.15	21.48	36.75	32.93
4) Conservative					
Mean	1.92%	1.54%	2.69%	1.64%	1.83%
Matching percentile	78.61	83.03	81.36	84.54	84.02
5) Immoderate					
Mean	0.38%	0.49%	1.09%	0.66%	0.75%
Matching percentile	34.17	47.49	49.17	54.52	53.96
6) Expert judgement					
6.1) Base scenario (a)					
Mean	2.84%	2.66%	2.94%	2.64%	2.70%
Matching percentile	87.35	93.45	83.80	93.82	92.30
6.2) Beta distribution as a proxy					
Mean	2.86%	2.67%	2.97%	2.66%	2.72%
Matching percentile	87.45	93.51	84.05	93.91	92.43
6.3) Normal distribution as a proxy					
Mean	2.10%	1.86%	2.61%	1.95%	2.09%
Matching percentile	80.76	87.38	80.41	88.52	87.33
6.4) Alternative base scenario (b)					
Mean	1.01%	1.01%	1.08%	1.02%	1.04%
Matching percentile	60.96	71.12	48.77	70.38	66.02
7) Beta distribution based on EDR					
Mean	0.38%	0.50%	1.11%	0.67%	0.76%
Matching percentile	34.54	47.88	49.87	54.95	54.43
EDR	0.00%	0.20%	0.57%	0.30%	0.35%

Continued

Matching percentile	0.05	22.51	27.39	27.92	28.18
Practical rule					
$k = 0$					
Mean	1.87%	-	-	-	-
Matching percentile	77.91	-	-	-	-
$k = 1, 2, 3$					
Minimum bound (97.5%)	-	0.81%	1.66%	0.85%	0.98%
Matching percentile	-	64.31	65.03	63.83	64.03
Maximum bound (99.9%)	-	1.00%	1.95%	0.96%	1.12%
Matching percentile	-	71.02	70.83	68.35	68.65

(a) minimum of $\lambda = 0.01$, mode of $\lambda = 0.025$, and maximum of $\lambda = 0.075$; (b) minimum of $\lambda = 0.003$, mode of $\lambda = 0.01$, and maximum of $\lambda = 0.02$; EDR - Empirical default rate.

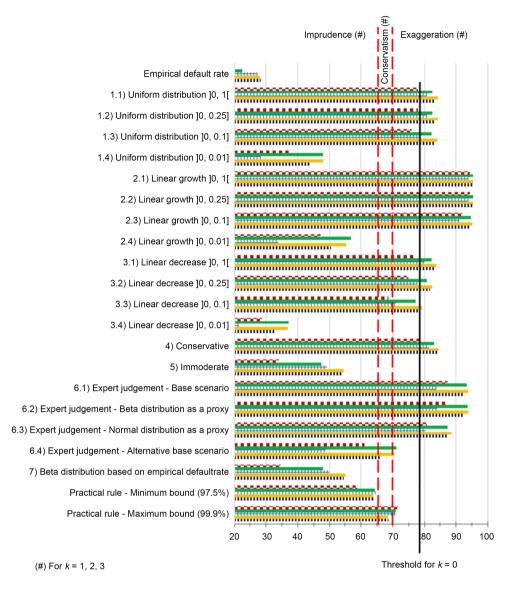Because the data in Table 13 shows a wide range of percentiles, it is necessary to distinguish between imprudence, conservatism, and exaggeration. Applying the practical rule aforementioned to the figures in the table, one may deduce that the 65th percentile can be used to separate imprudence from conservatism, and the 70th percentile can be used to separate conservatism from exaggeration. Other caps may be considered, in part because evaluating probabilities in low default portfolios requires a great deal of personal judgment. Given this, a 70th minimum percentile would be dangerous for some people, whereas a 75th maximum percentile would not be excessive. Over and above percentiles, the most significant consideration is the need to objectively discern different levels of safety or prudence, which demand the use of a practical rule like the one stated above.

Figure 2 shows the percentiles for all posterior distributions established in this paper (also assuming 12% for asset correlation).

When there are one or more defaults, the 65th and 70th percentiles—dashed vertical lines—correspond to the lower and upper limits of conservatism. When using the practical rule as a reference to choose the model to estimate the default probability, one notes that only the prior related to the alternative base scenario of the expert judgement may be deemed acceptable to establish a prudent approach estimating. In that scenario, it is reliable combining empirical data (integrated into the likelihood function) with expert knowledge (integrated into the prior function). The other priors do not deliver such reliable outputs because they produce either imprudent or exaggerated default probabilities.

When there is no default event (therefore, only for $n = 150$), the reference corresponds to the 77.91th percentile—solid vertical line. The priors 1.1 and 1.2 (uniform distribution), 3.1 (linear decrease), and 4 (conservative) are suitable since they correspond to the intended conservatism[32].

---

[32]The probability indicated by the prior 6.3 might also be adequate. It is, nevertheless, connected to an exaggerated base scenario of expert knowledge (i.e., the prior 6.1).

Figure 2. Percentiles of the classical binomial approach corresponding to mean of several Bayesian default distributions (asset correlation = 12%).

Because default distributions are heavily skewed to the right, using the mean to distinguish between conservatism and exaggeration in defaulted portfolios looks to be unnecessarily cautious. On the other hand, if there are no defaults in the portfolios, using the mean of the default distributions is a good strategy. When n = 150 and k = 0, the mean of default probability is 1.87%, which corresponds to the 77.91th percentile in a classical approach (see the prior 1.1 in **Table 13**)[33]. The corresponding upper limits of the highest density interval for 75%, 90%, and 95% are 2.43%, 4.51%, and 6.17%, respectively (as shown in **Table 10**), which are clearly unrealistic default probabilities for a low default portfolio.

---

[33]If there is no default occurrence, one expects that the prudence level given by the percentile is higher than the one when there are historical defaults.

Table 14 compares the effective number of observed defaults in each risk class to the projected number of defaults using the practical rule. Even with 65th and 70th percentiles, which would appear to be insufficiently cautious at first glance, default estimates are significantly higher than empirical data, a condition that any risk management tool for capital buffer measurement must take into account.

Table 14. Effective and estimated number of defaults.

	(n, k)				
	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
k effective	0	1	2	3	3
k estimated					
Minimum bound		4.1	5.8	8.5	8.3
	2.8				
Maximum bound		5.0	6.8	9.6	9.5

The disparities between effective defaults and estimated defaults for risk groups having at least one default (corresponding to the 63.83th - 71.02th percentiles) are larger than for risk group with no default (corresponding to the 77.91th percentile). Adaptations to the standard binomial test result in more prudent results, ensuring that the prudence margin demanded by internal models is sufficient to absorb losses in low default portfolios.

4. Final Thoughts

Both classical and Bayesian techniques have a degree of uncertainty in default probabilities, as seen in the figures. The estimation of default probabilities for low default portfolios is a wide topic with many questions. Each question has a number of possible answers. This diffusion process of answers is mainly intensified in the Bayesian approach because there are a very broad array of possible prior functions, ranging from the non-informative flat prior distribution (where there is no information about the parameter representing the default probability) to the most informative priors based solely on expert opinion (where it will have to be put in place all data about that parameter, both quantitative and qualitative elements).

It is easier to apply a non-informative prior if there is no idea about the default probability. Nonetheless, it is beneficial to make efforts to get information about that probability. Although the prior function is just one of many assumptions in the entire complex model, it is desirable if it reflects knowledge or feelings about the default probability. In low default portfolios with a lack of loss observations, additional information and expertise become particularly crucial.

The more informative the prior function, the worse the convergence of the posterior distribution to the likelihood function. This resistance holds true for both the base scenarios (regardless of how conservative they are) and the theoretical distributions generated by these scenarios, particularly the beta distribution.

The figures also show that the prior distribution can be chosen with a lot of freedom. Thus, it is possible to use a combination of priors, i.e., an average prior rather than simply one as a way of attempting to reflect the uncertainty degree of the default probability. However, differentiations between imprudence, conservatism, and exaggeration must first be made.

Subjectivity is present in both "the Bayesian (or subjective) approach" and "the classical (or frequentist) approach"—terminology stated in the second paragraph of 2.1. Indeed, the classical approach requires the subjective selection of confidence levels (as well as upper confidence bounds), whereas the Bayesian approach requires the selection of prior functions. A realistic rule was proposed to deal with such a wide variety of arbitrary choices. The rule will have the benefit of validating default financial models in general and Bayesian options in particular. As a result, unduly optimistic or overly pessimistic default probabilities are eliminated, allowing banks' pricing to reflect adequate and consistent levels of provisioning and economic capital.

Practical experts, theoretical academics and wary regulators cannot agree on methods and strategies for estimating default probabilities in portfolios with scarce historical data. Some topics were addressed in this paper; others, such as the model calibrations, and the use of a multi-period method by estimating models, are difficult open issues that require further exploration.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Pluto, K. and Tasche, D. (May 2005). *Estimating Probabilities of Default for Low Default Portfolios*. The Basel II Risk Parameters 2006—Estimation, Validation and Stress Testing (2006). Springer.

Tasche, D. (November 2012). Bayesian Estimation of Probabilities of Default for Low Default Portfolios. *Journal of Risk Management in Financial Institutions, 6,* 302-326. https://doi.org/10.2139/ssrn.2048818

Vasicek, O. A. (December 2002). The Distribution of Loan Portfolio Value. *Risk, 15,* No. 12.

Appendix A. A Brief Summary of Vasicek's Model (See 2.2.2)

The Basel Committee on Banking Supervision's structural models for credit risk are based on an approximation of Vasicek's equations, which use one-factor Gaussian copula to incorporate asset correlation between each pair of counterparties. Merton pioneered that type of models in 1974, when he developed the default theory using the Black-Scholes option pricing and the associated stochastic process.

The value of an asset changes over time and follows a dynamic evolution marked by drift and volatility. The systematic and idiosyncratic features incorporated into Vasicek's (2002) portfolio value model in 2002 correspond to the Merton model's concepts of drift and volatility.

A corporation "j" breaks its payment obligation if the probability of a random variable "X_j", which reflects its asset value, falls below a specific default threshold "m_j", according to this model. In addition, under the same model, over a specified length of time, "t"—usually a one year period of time—, it comes:

$$X_{jt} = S_t\sqrt{\rho} + Z_{jt}\sqrt{1-\rho}, \tag{9}$$

where "S" and "Z_j" are the systematic and idiosyncratic components of the asset value, respectively, and "ρ" is the constant asset correlation between two separate corporations.

Despite the fact that each corporation "j" has its own individual and idiosyncratic risk (represented by $Z_{jt}\sqrt{1-\rho}$), corporations are all exposed to a common and systematic risk, namely the overall status of the economy in which they operate (represented by $S_t\sqrt{\rho}$)[34].

"S" and "Z_j" are standardized normal random variables that are equicorrelated. Being these variables mutually independent, "X" is also a standardized normal random variable.

As $P(X < m) = \Phi(m) = p$, it equals:

$$
\begin{aligned}
P\left[X < \Phi^{-1}(p)\right] &= P\left[s\sqrt{\rho} + z_j\sqrt{1-\rho} < \Phi^{-1}(p)\right] \\
&= P\left\{z_j < \left[\Phi^{-1}(p) - s\sqrt{\rho}\right]\big/\sqrt{1-\rho}\right\} \\
&= \Phi\left\{\left[\Phi^{-1}(p) - s\sqrt{\rho}\right]\big/\sqrt{1-\rho}\right\} \\
&= G(p, s, \rho),
\end{aligned}
\tag{10}
$$

the main theoretical axis in structural default models.

---

[34]For instance, an economic index could be a random variable "S" that reflects the portfolio's exposure to a common factor.

Appendix B. Posterior Default Probabilities Assuming Independence of Events (See 3.3.2)

Table 11. Default probability distribution—Posterior using the basic binomial (asset correlation = 0%) as likelihood, for several priors.

	Individual approach (by risk class)			Integrated approach	
	A	B	C	A + B + C	B + C
			(*n*, *k*)		
Types of prior function	(150, 0)	(500, 1)	(350, 2)	(1000, 3)	(850, 3)
1) Uniform distribution					
1.1) *u* = 1					
Mean	0.67%	0.40%	0.85%	0.40%	0.47%
Median	0.43%	0.28%	0.70%	0.33%	0.39%
Mode	0.02%	0.17%	0.57%	0.31%	0.34%
Highest density interval					
50%	0.03 - 0.43	0.12 - 0.33	0.35 - 0.89	0.18 - 0.37	0.26 - 0.50
75%	0.02 - 0.89	0.07 - 0.53	0.23 - 1.14	0.15 - 0.53	0.18 - 0.63
90%	0.01 - 1.47	0.03 - 0.73	0.13 - 1.51	0.13 - 0.66	0.12 - 0.79
95%	0.00 - 1.91	0.02 - 0.88	0.07 - 1.77	0.09 - 0.74	0.10 - 0.88
Standard deviation	0.65%	0.28%	0.49%	0.20%	0.23%
Coefficient of skewness	1.96	1.41	1.14	0.99	0.99
1.2) *u* = 0.25					
Mean	0.67%	0.40%	0.85%	0.40%	0.47%
Median	0.43%	0.28%	0.70%	0.33%	0.39%
Mode	0.02%	0.17%	0.57%	0.31%	0.34%
Highest density interval					
50%	0.03 - 0.43	0.12 - 0.33	0.35 - 0.89	0.18 - 0.37	0.26 - 0.50
75%	0.02 - 0.89	0.07 - 0.53	0.23 - 1.14	0.15 - 0.53	0.18 - 0.63
90%	0.01 - 1.47	0.03 - 0.73	0.13 - 1.51	0.13 - 0.66	0.12 - 0.79
95%	0.00 - 1.91	0.02 - 0.88	0.07 - 1.77	0.09 - 0.74	0.10 - 0.88
Standard deviation	0.65%	0.28%	0.49%	0.20%	0.23%
Coefficient of skewness	1.96	1.41	1.14	0.99	0.99
1.3) *u* = 0.1					
Mean	0.67%	0.40%	0.85%	0.40%	0.47%

Continued

Median	0.43%	0.28%	0.70%	0.33%	0.39%
Mode	0.02%	0.17%	0.57%	0.31%	0.34%
Highest density interval					
50%	0.03 - 0.43	0.12 - 0.33	0.35 - 0.89	0.18 - 0.37	0.26 - 0.50
75%	0.02 - 0.89	0.07 - 0.53	0.23 - 1.14	0.15 - 0.53	0.18 - 0.63
90%	0.01 - 1.47	0.03 - 0.73	0.13 - 1.51	0.13 - 0.66	0.12 - 0.79
95%	0.00 - 1.91	0.02 - 0.88	0.07 - 1.77	0.09 - 0.74	0.10 - 0.88
Standard deviation	0.67%	0.28%	0.49%	0.20%	0.23%
Coefficient of skewness	1.75	1.41	1.14	0.99	0.99
1.4) $u = 0.01$					
Mean	0.40%	0.37%	0.57%	0.39%	0.45%
Median	0.27%	0.28%	0.60%	0.33%	0.40%
Mode	0.01%	0.20%	0.63%	0.30%	0.35%
Highest density interval					
50%	0.03 - 0.27	0.10 - 0.32	0.41 - 0.76	0.20 - 0.37	0.27 - 0.48
75%	0.02 - 0.54	0.05 - 0.49	0.36 - 0.90	0.17 - 0.53	0.18 - 0.59
90%	0.01 - 0.79	0.03 - 0.64	0.30 - 1.00	0.13 - 0.63	0.13 - 0.73
95%	0.00 - 0.86	0.02 - 0.77	0.21 - 1.00	0.10 - 0.70	0.11 - 0.81
Standard deviation	0.27%	0.23%	0.24%	0.19%	0.20%
Coefficient of skewness	0.52	0.72	−0.10	0.70	0.50
2) Linear growth					
2.1) $u = 1$					
Mean	1.31%	0.60%	1.13%	0.50%	0.59%
Median	1.07%	0.51%	0.97%	0.43%	0.49%
Mode	0.64%	0.38%	0.85%	0.40%	0.46%
Highest density interval					
50%	0.29 - 1.23	0.27 - 0.58	0.54 - 1.17	0.30 - 0.53	0.35 - 0.63
75%	0.16 - 1.79	0.17 - 0.78	0.37 - 1.53	0.23 - 0.64	0.25 - 0.76
90%	0.07 - 2.53	0.11 - 1.03	0.29 - 1.93	0.16 - 0.78	0.21 - 0.95
95%	0.03 - 3.03	0.06 - 1.23	0.23 - 2.21	0.13 - 0.92	0.15 - 1.07
Standard deviation	0.92%	0.34%	0.56%	0.22%	0.26%
Coefficient of skewness	1.37	1.14	0.98	0.89	0.88

Continued

2.2) $u = 0.25$					
Mean	1.31%	0.60%	1.13%	0.50%	0.59%
Median	1.07%	0.51%	0.97%	0.43%	0.49%
Mode	0.64%	0.38%	0.85%	0.40%	0.46%
Highest density interval					
50%	0.29 - 1.23	0.27 - 0.58	0.54 - 1.17	0.30 - 0.53	0.35 - 0.63
75%	0.16 - 1.79	0.17 - 0.78	0.37 - 1.53	0.23 - 0.64	0.25 - 0.76
90%	0.07 - 2.53	0.11 - 1.03	0.29 - 1.93	0.16 - 0.78	0.21 - 0.95
95%	0.03 - 3.03	0.06 - 1.23	0.23 - 2.21	0.13 - 0.92	0.15 - 1.07
Standard deviation	0.92%	0.34%	0.56%	0.22%	0.26%
Coefficient of skewness	1.37	1.14	0.98	0.89	0.88
2.3) $u = 0.1$					
Mean	1.31%	0.60%	1.13%	0.50%	0.59%
Median	1.07%	0.51%	0.97%	0.43%	0.49%
Mode	0.64%	0.38%	0.85%	0.40%	0.46%
Highest density interval					
50%	0.29 - 1.23	0.27 - 0.58	0.54 - 1.17	0.30 - 0.53	0.35 - 0.63
75%	0.16 - 1.79	0.17 - 0.78	0.37 - 1.53	0.23 - 0.64	0.25 - 0.76
90%	0.07 - 2.53	0.11 - 1.03	0.29 - 1.93	0.16 - 0.78	0.21 - 0.95
95%	0.03 - 3.03	0.06 - 1.23	0.23 - 2.21	0.13 - 0.92	0.15 - 1.07
Standard deviation	0.92%	0.34%	0.56%	0.22%	0.26%
Coefficient of skewness	1.37	1.14	0.98	0.89	0.88
2.4) $u = 0.01$					
Mean	0.58%	0.51%	0.69%	0.48%	0.54%
Median	0.56%	0.47%	0.70%	0.43%	0.52%
Mode	0.64%	0.39%	0.85%	0.38%	0.46%
Highest density interval					
50%	0.49 - 0.87	0.24 - 0.55	0.73 - 1.00	0.31 - 0.53	0.36 - 0.59
75%	0.38 - 1.00	0.19 - 0.73	0.57 - 1.00	0.23 - 0.63	0.28 - 0.73
90%	0.27 - 1.00	0.17 - 0.87	0.43 - 1.00	0.17 - 0.77	0.23 - 0.87
95%	0.17 - 1.00	0.13 - 0.93	0.33 - 1.00	0.17 - 0.87	0.20 - 0.93
Standard deviation	0.25%	0.23%	0.21%	0.20%	0.21%

Continued

Coefficient of skewness	−0.17	0.23	−0.41	0.44	0.22
3) Linear decrease					
3.1) $u = 1$					
Mean	0.67%	0.40%	0.85%	0.40%	0.47%
Median	0.43%	0.30%	0.70%	0.33%	0.39%
Mode	0.02%	0.18%	0.57%	0.27%	0.37%
Highest density interval					
50%	0.00 - 0.43	0.11 - 0.33	0.33 - 0.84	0.22 - 0.39	0.25 - 0.48
75%	0.00 - 0.89	0.06 - 0.53	0.23 - 1.17	0.16 - 0.53	0.21 - 0.63
90%	0.00 - 1.48	0.03 - 0.73	0.13 - 1.47	0.13 - 0.65	0.13 - 0.80
95%	0.00 - 1.93	0.01 - 0.89	0.08 - 1.78	0.11 - 0.76	0.09 - 0.92
Standard deviation	0.65%	0.28%	0.49%	0.20%	0.23%
Coefficient of skewness	1.96	1.41	1.14	0.99	0.99
3.2) $u = 0.25$					
Mean	0.66%	0.40%	0.84%	0.40%	0.47%
Median	0.43%	0.29%	0.70%	0.32%	0.39%
Mode	0.02%	0.18%	0.56%	0.27%	0.35%
Highest density interval					
50%	0.00 - 0.41	0.11 - 0.33	0.34 - 0.83	0.22 - 0.39	0.25 - 0.48
75%	0.00 - 0.84	0.06 - 0.53	0.25 - 1.15	0.16 - 0.53	0.21 - 0.63
90%	0.00 - 1.39	0.03 - 0.73	0.14 - 1.42	0.13 - 0.65	0.13 - 0.80
95%	0.00 - 1.81	0.01 - 0.89	0.09 - 1.68	0.11 - 0.76	0.09 - 0.92
Standard deviation	0.64%	0.28%	0.48%	0.20%	0.23%
Coefficient of skewness	1.96	1.41	1.13	0.99	0.99
3.3) $u = 0.1$					
Mean	0.63%	0.39%	0.83%	0.40%	0.46%
Median	0.39%	0.28%	0.70%	0.32%	0.39%
Mode	0.01%	0.18%	0.54%	0.27%	0.31%
Highest density interval					
50%	0.00 - 0.40	0.11 - 0.33	0.34 - 0.82	0.22 - 0.39	0.23 - 0.46
75%	0.00 - 0.83	0.06 - 0.53	0.23 - 1.11	0.16 - 0.53	0.18 - 0.60
90%	0.00 - 1.38	0.03 - 0.73	0.12 - 1.36	0.13 - 0.65	0.11 - 0.77

Continued

95%	0.00 - 1.79	0.01 - 0.89	0.08 - 1.61	0.11 - 0.76	0.08 - 0.89
Standard deviation	0.61%	0.27%	0.47%	0.20%	0.23%
Coefficient of skewness	1.85	1.40	1.13	0.99	0.99
3.4) $u = 0.01$					
Mean	0.27%	0.29%	0.45%	0.34%	0.37%
Median	0.19%	0.23%	0.40%	0.29%	0.33%
Mode	0.03%	0.17%	0.38%	0.27%	0.30%
Highest density interval					
50%	0.00 - 0.18	0.10 - 0.27	0.28 - 0.53	0.17 - 0.37	0.23 - 0.40
75%	0.00 - 0.37	0.07 - 0.39	0.18 - 0.63	0.13 - 0.43	0.17 - 0.47
90%	0.00 - 0.53	0.03 - 0.50	0.13 - 0.73	0.10 - 0.53	0.13 - 0.63
95%	0.00 - 0.67	0.01 - 0.58	0.10 - 0.77	0.09 - 0.63	0.08 - 0.67
Standard deviation	0.20%	0.18%	0.20%	0.16%	0.17%
Coefficient of skewness	0.93	0.86	0.27	0.69	0.57
4) Conservative					
Mean	0.68%	0.40%	0.85%	0.40%	0.47%
Median	0.43%	0.31%	0.73%	0.33%	0.42%
Mode	0.02%	0.20%	0.54%	0.27%	0.35%
Highest density interval					
50%	0.03 - 0.43	0.09 - 0.33	0.36 - 0.90	0.20 - 0.39	0.27 - 0.48
75%	0.02 - 0.87	0.07 - 0.53	0.23 - 1.16	0.15 - 0.53	0.20 - 0.63
90%	0.01 - 1.49	0.03 - 0.73	0.13 - 1.48	0.13 - 0.64	0.13 - 0.77
95%	0.00 - 1.97	0.02 - 0.90	0.08 - 1.74	0.09 - 0.75	0.10 - 0.89
Standard deviation	0.66%	0.28%	0.49%	0.20%	0.23%
Coefficient of skewness	1.96	1.41	1.14	0.99	0.99
5) Immoderate					
Mean	0.21%	0.22%	0.57%	0.30%	0.35%
Median	0.07%	0.11%	0.43%	0.22%	0.24%
Mode	0.03%	0.03%	0.29%	0.20%	0.22%
Highest density interval					
50%	0.03 - 0.05	0.03 - 0.12	0.11 - 0.48	0.12 - 0.26	0.17 - 0.34
75%	0.02 - 0.17	0.02 - 0.25	0.07 - 0.76	0.10 - 0.38	0.10 - 0.46

Continued

90%	0.01 - 0.47	0.01 - 0.42	0.02 - 1.04	0.06 - 0.51	0.07 - 0.60
95%	0.00 - 0.75	0.00 - 0.54	0.01 - 1.28	0.02 - 0.59	0.05 - 0.73
Standard deviation	0.31%	0.20%	0.40%	0.17%	0.20%
Coefficient of skewness	3.74	2.00	1.40	1.15	1.15
6) Expert judgement					
6.1) Base scenario					
Mean	2.08%	1.45%	1.76%	1.26%	1.31%
Median	1.90%	1.32%	1.61%	1.19%	1.23%
Mode	1.67%	1.20%	1.44%	1.11%	1.15%
Highest density interval					
50%	1.36 - 2.12	1.13 - 1.37	1.20 - 1.71	1.10 - 1.22	1.10 - 1.23
75%	1.20 - 2.48	1.07 - 1.54	1.13 - 2.02	1.05 - 1.31	1.06 - 1.40
90%	1.10 - 2.97	1.03 - 1.82	1.07 - 2.36	1.03 - 1.47	1.03 - 1.54
95%	1.05 - 3.33	1.01 - 1.98	1.01 - 2.59	1.02 - 1.56	1.01 - 1.67
Standard deviation	0.68%	0.30%	0.46%	0.17%	0.21%
Coefficient of skewness	1.12	1.25	0.98	1.31	1.27
6.2) Beta distribution as a proxy					
Mean	2.01%	1.13%	1.63%	0.82%	0.94%
Median	1.84%	1.03%	1.53%	0.75%	0.86%
Mode	1.70%	0.97%	1.42%	0.71%	0.83%
Highest density interval					
50%	1.29 - 2.20	0.77 - 1.21	1.12 - 1.80	0.58 - 0.87	0.69 - 1.01
75%	1.03 - 2.66	0.60 - 1.45	0.91 - 2.08	0.50 - 1.02	0.54 - 1.20
90%	0.80 - 3.17	0.49 - 1.73	0.77 - 2.47	0.41 - 1.18	0.46 - 1.34
95%	0.65 - 3.47	0.42 - 1.88	0.67 - 2.69	0.36 - 1.27	0.39 - 1.50
Standard deviation	0.77%	0.40%	0.55%	0.26%	0.30%
Coefficient of skewness	0.75	0.71	0.66	0.63	0.63
6.3) Normal distribution as a proxy					
Mean	0.91%	0.45%	0.98%	0.43%	0.51%
Median	0.67%	0.30%	0.84%	0.36%	0.41%
Mode	0.02%	0.22%	0.70%	0.33%	0.35%

Continued

Highest density interval					
50%	0.03 - 0.67	0.12 - 0.43	0.41 - 1.03	0.23 - 0.42	0.26 - 0.51
75%	0.02 - 1.24	0.06 - 0.59	0.30 - 1.35	0.16 - 0.54	0.18 - 0.67
90%	0.01 - 1.97	0.03 - 0.82	0.17 - 1.70	0.11 - 0.73	0.12 - 0.83
95%	0.00 - 2.46	0.01 - 1.03	0.12 - 2.01	0.10 - 0.80	0.07 - 0.95
Standard deviation	0.80%	0.31%	0.54%	0.21%	0.25%
Coefficient of skewness	1.47	1.34	1.01	0.98	0.97

Highest density intervals expressed in percentage.