

Sector concentration risk in SME credit portfolios :

A multifactor approach.

Michel DIETSCH^{1,2}
Joël PETEY²
LARGE, Université de Strasbourg

December 2009

Abstract: In large portfolios of small and medium-sized businesses (SME), which are highly granular, concentration risk arises from correlated defaults among groups of borrowers. Consequently, measurement of concentration risk needs to take into account borrowers' heterogeneity. One way to proceed is to extend the standard asymptotic single factor framework by introducing additional factors of systematic risk varying between groups of borrowers. Using a generalized linear mixed model, the paper extends the standard one factor credit risk model to the multi-factor framework taking into account industry effects. The paper uses a large database containing ratings history of more than 600.000 French SME over the 1999-2008 period. Results show that the standard one factor model and the IRB regulatory formula largely fail capturing potential risk concentration. Moreover, loans to the real estate industry may be a first order determinant of concentration risk, even when considering loans to small and very small businesses.

¹ Corresponding author : michel.dietsch@unistra.fr

² Université de Strasbourg, Institut d'Etudes Politiques, 47 avenue de la Forêt-Noire 67072 Strasbourg Cedex France

1. Introduction

As emphasized recently by the Basel Committee (BCBS, 2008), a main challenge of portfolio credit risk modelling is the modelling of dependency across obligors. Especially, internal measurement of portfolio credit risk must take into account concentration risk. In retail banking, where most portfolios are granular, concentration refers to a potential situation of correlated defaults. Concentration risk in credit portfolios could come from name concentration (the incomplete diversification of idiosyncratic borrower risk) and sector concentration (the existence of multiple systematic risk factors, generally related to industry effects, but also to other sources such as location or explicit macroeconomic factors). BCBS (2006) underlines that “concentration of credit risk in asset portfolios has been one of the major causes of bank distress”. The calibration of the IRB formulas was allegedly chosen to match the economic risk in a credit portfolio that should be very-well diversified across industries. However, departures from the underlying assumptions of the ASRF model, i.e. perfect granularity and a single source of systematic risk, could result in substantial deviations of economic capital from AIRB regulatory capital charges. This paper investigates the level and the sources of such deviations considering a large database of SME rating histories.

Correlated defaults can be attributed to the dependency of credit exposures to common factors that are specific to some segments of the portfolio or to particular banks’ clienteles. Indeed, SME may not be homogeneous firms. If firm heterogeneity is defined as heterogeneity of risk factor loadings across firms, it characterizes SME loan portfolios due to differences in size, sector or localization. Recent research shows that concentration effects exist in SME’ loans portfolios and that credit risk varies in portfolios according to their industry and size composition (Carling, Ronnegard and Roszbach, 2004, Dietsch and Petey, 2004, Duellmann

and Scheule, 2003, Heitfield, Burton and Chomsisengphet, 2006, Duellmann and Masschelein, 2006).

One recent strand of research expands the asymptotic single risk factor (ASFR) model to consider additional risk factors (Lucas et al., 2001, Tasche, 2006). However, it might be difficult to identify the risk factors leading to SME defaults. Two strategies could be followed in searching for explicit risk factors. First, micro data on businesses might be incorporated in a structural credit risk model, by introducing accounting or qualitative information or a credit score. Second, one could try to make explicit the latent factor in the Merton-Vasicek framework by introducing a set of macroeconomic variables. However, SME credit markets are local by nature and data availability concerns may limit the implementation of such an approach in retail banking. More specifically, lack of time series data on potential risk factors at the local level (for example, regional economic activity or regional prices data), which might be the relevant data for retail clienteles, impedes using the same methodologies which are commonly used to compute dependency structure parameters in corporate portfolios. Moreover, the two types of data may be combined (as for instance in Carling et al., 2007), which raises the question of the dependence between the two data sources. At this time, there is no real consensus either on methodologies to compute risk parameters in retail loans portfolios, or on the set of relevant risk factors that impact obligors in such portfolios.

This paper proposes an alternative approach that consists in expanding the one-factor model by adding new latent factors that are linked to observable characteristics of the firm, such as size or industry. Our observations show in fact that the portfolio credit risk of French SME is significantly linked to such factors. These factors are expected to reinforce or attenuate the effect of general economic conditions on portfolio losses. Thus, the extension of the single-

factor model to a multi-factor model could improve substantially the computation of the dependency structure across exposures in a typical retail banking portfolio.

In this paper, we use credit ratings history of a large database of French SME and apply the methodology of generalized linear mixed models (GLMM) to produce estimates of portfolio's credit risk parameters in a multi-factor context. This model implements in a coherent way the latent factor default model. It produces estimates of default thresholds considered as fixed effects and covariance matrixes of a set of latent random effects corresponding to the set of systematic factors. The estimation of such parameters allows computing economic capital as buffer of losses in portfolios exposed to different systematic risk factors.

Measuring concentration risk calls for allocating economic capital between segments of borrowers, i.e. the computation of marginal contributions of different segments to total loss. A segmentation of the portfolio can be built by identifying groups of borrowers with the same observable characteristics which expose them to the same risk factors. In a multi-factor context, capital allocation can be implemented at the segment level such that it is possible to investigate the heterogeneity in capital allocation induced by the various risk factors. Thus, a single factor homogeneous framework could induce a misrepresentation of the concentration risk even in large portfolios of retail exposures. This could also lead to biased risk adjusted performance measures that are used in the capital allocation process to business lines or decentralized decisions units. Moreover, while they are calibrated using a single factor framework, Basel 2 IRB regulatory formulas of capital requirements could be of limited interest in allocating capital. The computation of the portfolio's value-at-risk (VaR) and marginal risk contributions are made by using a methodology proposed by Tasche (2008), which grounds on an importance sampling based simulation of expected conditional losses.

This methodology has the advantage to take into account the impact of borrowers' heterogeneity on economic capital charges and capital allocation.

Section 2 presents the asymptotic multifactor credit risk model, its economic specification as a general linear mixed model (GLMM) and the chosen dependence structure of risk factors. Section 3 presents portfolio credit risk parameters results as well as the estimated economic capital levels. Section 4 presents the methodology we use for the computation of marginal contributions and analyses the concentration risk through a characterization of the heterogeneity in capital allocations. Section 5 presents a simulation that evaluates the impact of risk concentration on capital levels. Section 6 concludes.

2. Credit risk model specification

2.1 The asymptotic multi-factor credit risk model

Measures of portfolio credit risk require estimates of dependency across assets. The credit risk literature has emphasized the importance of obligors' exposure to common factors in determining the shape of the distribution of losses. Losses at the portfolio level can be defined as the sum of individual losses on defaulting expositions, given the individual severity of losses, i.e. loss given default rates. Hence, defining u_i as obligor's i loss given default (LGD) and $\mathbf{1}_{D_i}$ the default indicator variable of obligor i :

$$L = \sum_{i=1}^n u_i \mathbf{1}_{D_i}$$

In structural credit risk models, default occurs when the value of one obligor's assets become smaller than the value of debt at maturity date, what can be modelled as crossing a default threshold. Moreover, the financial health (or final asset value) of obligor i is represented by a

latent (unobservable) variable U_i . The level of U_i is determined by the realizations of a set of risk factors such that:

$$U_i = \mathbf{w}_i' s + \sqrt{1 - \mathbf{w}_i' R \mathbf{w}_i} \varepsilon_i \quad (1)$$

where S is a vector of “systematic” risk factors with realization s , \mathbf{w}_i is the vector of sensitivities (or factor loadings) of the i -th borrower to the set of factors and ε_i is a specific risk factor for borrower i . R is the correlation matrix of the risk factors. Assuming risk factors are multivariate Gaussian, the sensitivity to specific risk in equation (1) ensures that U_i is also standard normal. Specific risk factors are assumed to be uncorrelated among obligors and also independent from the systematic factors. In this framework, default occurs when the latent variable U_i falls below a default threshold calibrated according to the stationary (long term) default probability \bar{p}_i of obligor i . Noting Φ the standard normal cdf, default occurs when:

$$U_i = \mathbf{w}_i' s + \sqrt{1 - \mathbf{w}_i' R \mathbf{w}_i} \varepsilon_i < \Phi^{-1}(\bar{p}_i) \Leftrightarrow 1_{D_i} = 1$$

Moreover, assuming specific risk can be entirely diversified away, the default indicator variable of a given obligor can be approximated by its expected value, that is by its default probability conditional to the realization of systematic risk factors. Portfolio losses are then defined as:

$$L \approx \sum_{i=1}^n u_i \Phi \left[\frac{\Phi^{-1}(\bar{p}_i) - \mathbf{w}_i' s}{\sqrt{1 - \mathbf{w}_i' R \mathbf{w}_i}} \right] \quad (2)$$

This framework is called the asymptotic multi-factor framework of credit risk (e.g. Lucas et al., 2001). Equation (2) assumes that each obligor can be characterised by an individual

default threshold and factor sensitivities. However, in retail loans portfolios, default rates are computed by rating classes and sensitivities to risk factors cannot be computed on an individual basis. Thus, assumptions are required in order to reduce the number of parameters of the loss variable. Following previous literature on large portfolios of loans, we assume that obligors belonging to the same rating notch r will share the same default threshold.³ We further assume that the vector of risk factor sensitivities is the same for obligors belonging to the same segment of a portfolio. Hence, assuming a portfolio composed of K segments, losses can be rewritten as:

$$L \approx \sum_{i=1}^n u_i \Phi \left[\frac{\Phi^{-1}(\bar{p}_r) - \mathbf{w}_{ki}' s}{\sqrt{1 - \mathbf{w}_{ki}' R \mathbf{w}_{ki}}} \right]$$

Now, considering a multi-factor framework calls for the specification of the dependence structure of risk factors as well as for the estimation of the default thresholds and factor sensitivities.

2.2 Estimation of credit risk parameters using Generalized Linear Mixed Models.

The single factor model is expanded by adding new latent factors that can be linked to easily observable characteristics of businesses. These are sector and geographic location. In addition, to highlight the role of size, we run separate sets of estimation in two sub-populations of very small, on the one side, and small and medium-sized businesses, on the other side. In order to estimate the impact of additional risk factors on portfolio credit risk, we use an econometric

³ Recent results have highlighted the importance of firm level heterogeneity in the measurement of credit risk. Pesaran, Schuermann, Treutler (2005) illustrate the importance of the dispersion of default probabilities within ratings on the level of credit risk. Duellmann, Scheicher and Schmieder (2007) show the prominence of name concentration over sector concentration. These results are obtained considering larger (mostly) quoted companies for which equity price changes can be related to explicit risk factors. Nevertheless, it could remain in our portfolio a share of heterogeneity in risk attributable to the firm level which is not accounted for.

model belonging to the class of generalized linear mixed models. GLMM models mix fixed and random effects for observable and unobservable factors. Detailed presentations of GLMM models in credit risk modelling can be found in Frey and McNeil (2003) and Wendin and McNeil (2007).

Suppose Y represents the $(N \times 1)$ vector of observed default data and γ is a $(K \times 1)$ vector of random effects. In this framework the expected default probability of obligor i is given by:

$$E[Y_i = 1|\gamma] = G^{-1}(X\beta + Z\gamma)$$

where $G(\cdot)$ is a differentiable monotonic link function, $G^{-1}(\cdot)$ is its inverse and Y_i the default indicator variable for obligor i . In the following applications, we will focus on the probit response because the normal distribution is also the underlying response function assumed by the Basel 2 framework of credit risk, thus $G(x) = \Phi^{-1}(x)$. X is a $(N \times P)$ matrix containing the (observed) fixed effects, and Z is a $(N \times K)$ design matrix for the random effects. The random effects are assumed to be normally distributed with mean 0 and variance matrix G . β is the vector of parameters associated to the fixed effects.

Considering a portfolio of N obligors dispatched in $r = 1, \dots, R$ rating classes and given a (latent) vector γ_t of random effects, we estimate different GLMM models which share the same structure in order to measure the effects of additional risk factors in the borrowers population. The default probability of borrower i at time t is given by:

$$P(Y_{it} = 1|\gamma_t) = \Phi(\mu_0 + x'_{it} \mu_r + z'_{it} \gamma_t)$$

Where:

- μ_r denotes the vector of parameters of the fixed effect of rating. Hence, the different models try to capture the heterogeneity in default rates not explained by the affectation of each obligor to a rating class.
- $x'_{it} = [0, \dots, 1, \dots, 0]$ is a $(1 \times R)$ vector defining the rating of borrower i in time period t . As we assume exchangeability of borrowers within segments, the estimation is not done on individual borrowers, but on annual default rates within segments. This leads assuming homogeneity with respect to credit rating and any segmentation variables.
- z_{it} is the design matrix of random effects. Its form depends on the assumed dependence structure of random effects, as developed in the following section.

Here, we consider one fixed effect: the borrower's credit quality – measured by its rating - defining default thresholds - and two different random effects: the firm's industry, in addition to a general latent risk factor. Random effects are introduced through a segmentation of rating histories according to each factor.

2.3 Dependence structure of random effects.

Extending the one factor model calls also for a specification of the risk factors' dependence structure. By assuming that the general risk factor (the risk factor of the one factor model) represents the impact on default rates of variations in general economic conditions, it seems straightforward to consider that additional risk factors can reinforce or weaken the sensitivity of a given subset of firms in the portfolio to general economic conditions. This corresponds to the idea that a given sector or region can be either procyclical, cycle neutral or countercyclical. In order to capture these effects, we estimate the correlation between the general risk factor and a set of additional factors associated to a given segmentation of the

portfolio. In order to keep the model tractable, we further assume that the additional factors, i.e. shocks that affect subgroups of the portfolio, are independent. This specification implies in particular that the inter-segment correlation is not directly attributable to the segments' risk factors but rather to the dependence between these latter factors and the general economic factor. The covariance structure we will focus on is of the form:

$$G = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \sigma_{1,q+1} \\ 0 & \ddots & 0 & \sigma_{2,q+1} \\ 0 & 0 & \sigma_q^2 & \vdots \\ \sigma_{1,q+1} & \sigma_{2,q+1} & \cdots & \sigma_{q+1}^2 \end{bmatrix}$$

considering q latent segment factors and one systematic factor (denoted $q+1$, thus $q+1=K$). The last column of G defines a random effect common to all obligors and reflects the heterogeneity in default rates related to time. This random effect corresponds to the heterogeneity in default rates attributable to time-heterogeneity, which is assumed to be related to general economic conditions. In this specification, the linear predictor in the logistic regression contains an intercept term that randomly varies at the year level, the highest level in the modelling, where all other effects are nested in. In other words, a random intercept is drawn separately and independently for each year. This structure implies that a given obligor is affected by two factors: the factor representative of general economic conditions and its industry risk factor.

3. Estimation results and computation of economic capital

3.1 Data and choice of additional risk factors

We use a large database of historical default data provided by COFACE Services, the largest French financial information provider. This database contains around 600.000 very small (with turnover less than 750 k€) and small and medium-sized (with turnover between 750 k€ and 50 millions euros). The database covers a ten years' period (end of 1999 to end of 2008) and provides information about firm's size, industry and credit quality measured by a rating. We grouped COFACE ratings in a simplified seven grades system, in which firms are ranked from rating 1 (high risk) to rating 7 (low risk) and 8 (failure). The last rating grade corresponds to legal bankruptcy⁴.

However, the introduction of a large number of, possibly correlated, random effects makes trickier the estimation of the model's parameters. Methods based on the approximation of the objective function (like Laplace or Gaussian quadrature) can be implemented as long as the number of random effects remains low. Consequently, the models are estimated by restricted pseudo-likelihood, which is based on the linearization of the likelihood function (Wolfinger and O'Connell, 1993) where the optimization is based on some linear expansion of the objective function. The drawback of this approach is the absence of a true objective function for the overall optimization process, which can lead to biased estimates of covariance parameters.⁵

Several studies have underlined the differences in credit risk parameters when considering borrowers of different size. Indeed, larger firms have lower default probabilities and are often

⁴ This is clearly a limitation of the data, as it leads to an underestimation of the frequency of default events.

⁵ We estimate the models using the SAS 9.1 Glimmix procedure.

assumed to be more sensitive to overall economic conditions, i.e. to have higher asset correlations (Lopez, 2004, Dietsch and Petey, 2004). This observation underlies the inverse relationship between asset correlation and default probability in the IRB regulatory formula. In order to account for such possible relationship, the industry multifactor credit risk model is estimated on two subpopulations: firms with an annual turnover lower than 0.75 millions €, which we term as very small firms and firms with a turnover between 0.75 and 50 millions euros which will be denoted as SME. Firms with a turnover over 50 millions euros are excluded. Table 1 shows the distribution of the firms across industries for the two size classes using the number of firms and the industry shares in outstanding debt.

Table 1 Industry repartitions of firms and credit exposures

Industry	Number of firms		Total banking debt	
	SME		SME	Very small firms
Manufacturing	10,5%	10,5%	13,5%	6,4%
Construction	13,2%	14,4%	8,2%	8,1%
Retailing	36,1%	26,1%	21,7%	14,7%
Transport & logistics	5,1%	3,4%	4,1%	1,8%
Hotels & restaurants	3,7%	10,6%	4,7%	13,8%
Information & communication	3,9%	4,3%	2,6%	1,9%
Real estate	3,2%	6,3%	33,1%	36,5%
Scient. & tech. services to businesses	7,5%	11,1%	6,6%	10,0%
Administrative services	4,2%	4,6%	4,7%	3,3%
Education	0,7%	3,1%	0,2%	0,5%
Arts, culture, entertainment	0,8%	1,7%	0,5%	0,9%
Other services	0,7%	3,7%	0,2%	2,2%

Table 1 shows that there is a considerable potential for credit risk concentration when considering the importance of the real estate industry in terms of credit exposures. This suggests that the evolution of potential credit losses in that industry might be of utmost importance when considering concentration risk. Indeed, credit exposures are concentrated in the real estate sector, with a share of more than 30% of banking debt while they represent only about 3% of SME and 6% of very small firms. The second most important industry in terms of

debt exposure (and the first sector when considering the number of firms) is retailing. As considering two size classes and excluding large firms helps to a certain extent to control for size effects, table 1 thus suggests that there are considerable discrepancies in leverage across industries. These differences could lead to concentration of losses in highly leveraged industries if industry-specific shocks explain an important share of credit risk and/or the industry asset correlations widely differ across industries.

3.2 Estimation results and the computation of economic capital

In addition to the industry multifactor model, we estimate a single factor homogeneous model. This specification assumes that all obligors share the same sensitivity (asset correlation) to a unique ‘general’ risk factor. In this specification, obligors only differ according to their default thresholds (estimation details are shown in Appendix A.1). Goodness-of-fit statistics of the GLMM models (in particular the generalized chi-square to degrees of freedom ratio) indicate that the single factor model has a poor fit as the residual heterogeneity remains high. Adding industry random effects lowers considerably residual heterogeneity. Table 2 shows the covariance matrices of the industry model for the two size classes. First, table 2 shows that in some industries asset correlations are primarily determined by the industry risk factor, as the variance of the corresponding random effect is higher than the variance of the general risk factor, i.e. the variance of the random intercept of the GLMM model. The industry with the largest random effect is real estate both for SME and very small firms. This strengthens the potential impact on credit risk of the observed concentration of credit exposures in that industry (see table 1). Moreover, the magnitude of the general factor is comparable between the two size classes whereas the magnitude of industry random effects seems to be higher for SMEs than for very small businesses. As the estimated stationary default probabilities are comparable (see table C in appendix A.2 for details) between the two size classes and

notwithstanding the differences in the industry composition of the two sub-populations, the estimation results also suggest that credit risk may be higher for SME than for very small firms. Finally, the covariances of the random effects allow indentifying industries that are positively, negatively or independent from general economic conditions. For the very small firms, the corresponding correlations range from -0.59 (retailing) to 0.34 (transports and logistics) . For SME, correlations are bounded by -0.43 (construction) and 0.50 (arts, culture, entertainment). Thus, there are considerable disparities across industries in terms of correlations to general economic conditions. The association of an important industry risk factor and a high correlation to overall economic conditions could further be a source of concentration risk.

Table 2: Covariance matrix of random effects

Panel A: Very small firms													
	Sect1	Sect2	Sect3	Sect4	Sect5	Sect6	Sect7	Sect8	Sect9	Sect10	Sect11	Sect12	Syst
Sect1	0.006258												-0.00093
Sect2		0.009980											-0.00563
Sect3			0.01414										-0.00832
Sect4				0.02018									0.005808
Sect5					0.01401								0.004705
Sect6						0.006526							-0.00229
Sect7							0.09277						-0.00505
Sect8								0					0
Sect9									0.002610				-0.00017
Sect10										0.01155			-0.00050
Sect11											0.008196		0.003027
Sect12												0.003200	0.001089
Syst	-0.00093	-0.00563	-0.00832	0.005808	0.004705	-0.00229	-0.00505	0	-0.00017	-0.00050	0.003027	0.001089	0.01407

Panel B, SME													
	Sect1	Sect2	Sect3	Sect4	Sect5	Sect6	Sect7	Sect8	Sect9	Sect10	Sect11	Sect12	Syst
Sect1	0.03102												0.01006
Sect2		0.009317											-0.00491
Sect3			0.02814										-0.00505
Sect4				0.004891									-0.00092
Sect5					0.1287								0.01601
Sect6						0.01754							-0.00425
Sect7							0.2372						0.007441
Sect8								0					0
Sect9									0.001050				-0.00011
Sect10										0.06322			0.01477
Sect11											0.008169		-0.00065
Sect12												0.04488	-0.00024
Syst	0.01006	-0.00491	-0.00505	-0.00092	0.01601	-0.00425	0.007441	0	-0.00011	0.01477	-0.00065	-0.00024	0.01370

Total portfolio's economic capital is computed by simulation of the risk factors given default thresholds and risk factor sensitivities, which are the outputs of the GLMM models⁶. Given the credit risk parameters and a set of simulated risk factors, defaults in each sub-portfolio defined by crossing the 12 industries with the 7 rating grades are produced by drawing from a binomial probability with the number of firms in each sub-portfolio and the conditional default probability defined within equation (2) as parameters. Exposures are then defined as the average bank debt within classes crossing rating and industries, i.e. we assume at this stage homogeneity in exposures within sub-portfolios. Finally, losses are aggregated over the complete portfolio assuming a 45% LGD. Table 3 shows economic capital (% of total exposure) for various quantiles of the distribution of losses and for the two sub-samples of very small firms and small and medium-sized firms. It shows clearly the existence of considerable divergences between models when considering the level of economic capital. Indeed, the introduction of industry factors leads to a considerable increase in the level of economic capital as the corresponding random effects increase the heterogeneity in annual default rates captured by the credit risk model. As suggested by the results, the credit risk of SMEs turns out to be higher than the credit risk of very small businesses.

⁶ At this stage, a technical difficulty arises as estimated covariance matrices are rarely definite positive. Yet, definite positivity being a necessary condition for correctly specified correlation matrices, we correct the estimated matrices by applying the correction suggested by Higham (2002).

Table 3 Economic capital (% of total exposure)

Quantile	90%	95%	99%	99.5%	99.9%	99.97%
Panel A Very small firms						
One factor	0,85%	0,89%	0,98%	1,01%	1,08%	1,13%
Industry	0,96%	1,06%	1,28%	1,38%	1,61%	1,82%
Panel B SME						
One factor	0,70%	0,73%	0,82%	0,85%	0,91%	0,96%
Industry	0,73%	0,88%	1,33%	1,56%	2,14%	2,67%

4. Capital allocation and credit risk heterogeneity.

The objective of this section is to assess the ability of the credit risk models to detect potential concentration. In fact, if obligors were homogenous in terms of credit risk, capital ratios should not differ across obligors and/or portfolios' segments. On the contrary, if obligors are heterogeneous, a higher capital ratio in a given segment would indicate potential risk concentration. A straightforward source of heterogeneity is credit rating which is accounted for by the estimation of default thresholds. Moreover, industry could be an additional source of credit risk heterogeneity in SME portfolios. If there is concentration risk, capital ratios should vary along this source of risk. In fact, it is likely that the portfolio under consideration is highly granular due to its size. Thus, the heterogeneity in capital charges will mainly come from the risk factors affecting the different portfolios' segments.

4.1 Computing marginal contributions by Monte Carlo simulation

The detection of potential concentration (given the structure of the portfolio) requires allocating portfolio-wide economic capital to sub-portfolios or individual assets. It is the (relative) importance of economic capital "consumed" by given sub-portfolios that highlights

the potential concentration of losses on these sub-portfolios and could lead to investment and restructuration decisions. Grounding on the works of Tasche (1999) and Gouriéroux and al. (2000), marginal contributions to portfolio VaR at the α level can be expressed as the expected loss on a given exposure conditional on losses reaching VaR:

$$RCVAR_i = E[L_i | L = VaR_\alpha(L)] = \frac{E[L_i \mathbf{1}_{VaR_\alpha(L)=L}]}{P[L = VaR_\alpha(L)]} \quad (3)$$

This last definition shows that, given a positive probability for losses reaching VaR, the computation of marginal contributions relies heavily on the ability to estimate individual losses when aggregate losses approach VaR. Thus, when considering Monte Carlo simulation, the conditional mean might be based on a limited number of simulations, leading to unreliable estimates. In order to improve estimation, some authors (Tasche, 2008, Glasserman and Li, 2005, Egloff et al., 2005) propose to use importance sampling, which consists in shifting the distribution's parameters in order to increase the likelihood of observing desired realizations of the risk factors. However, the difficulty in importance sampling lies in the choice of the alternative distribution. We follow Tasche (2008) in shifting only the risk factors (S) means such as:

$$S_i^* = S_i - E_F[S_i] + \mu_i, \mu_i = E[S_i | L = VaR_\alpha(L)]$$

The next step is the computation of conditional expectation (3). As the computation of VaR is done by Monte Carlo simulation, one has as well the realizations of each risk factor as the resulting credit losses. This allows using the non parametric Naradaya-Watson estimator for conditional expectations. Using the standard normal density as kernel and denoting h the bandwidth of the kernel, the estimator of the conditional expectation for risk factor k is defined as:

$$\hat{E}[S_k | L = VaR_\alpha(L)] = \frac{\sum_{t=1}^T S_k \phi\left(\frac{VaR - L_t}{h}\right)}{\sum_{t=1}^T \phi\left(\frac{VaR - L_t}{h}\right)} \text{ where } h = 1.06\sigma_L T^{-1/5}$$

Tasche (2008, proposition 4.2) using Klebaner (2007) establishes that the expected losses conditional to VaR under the natural distribution can be defined as, where F^* denotes the shifted distribution of risk factors:

$$E_F[L_i | L = VaR_\alpha(L)] = \frac{E_{F^*}[L_i R | L = VaR_\alpha(L)]}{E_{F^*}[R | L = VaR_\alpha(L)]}$$

As previously noted, these conditional expectations can be computed using the Naradaya-Watson estimator and the simulations of risk factors and losses under the shifted distribution.

4.2 Heterogeneity in capital allocations

The output of the capital allocation procedure is the representative capital charge for a given obligor of a given rating and exposed to an additional industry risk factor. These capital charges are merged with the entire portfolio what allows aggregating economic capital across sub-portfolios according to any segmentation variable, here industry. We do not consider here capital allocation at the asset level (as would be the case for loan pricing or RAROC computation purposes). We focus on an intermediate allocation level in order to characterize the risk's structure within the portfolio. Table 4 shows the capital ratios by industry for the two size classes considered here. It appears that there is a considerable heterogeneity across industries in terms of credit risk measured by the economic capital at the 99.97% confidence level. While ranging from a low 0.87% to a high 3.16% for very small firms, the ratio goes from 0.59% to 6.74% for SMEs. Moreover, the riskiest industry, consistent with observations made in the previous sections, is real estate. Nevertheless, excluding real estate, credit risk in SMEs is finally lower than for very small firms as capital ratios are lower for latter category

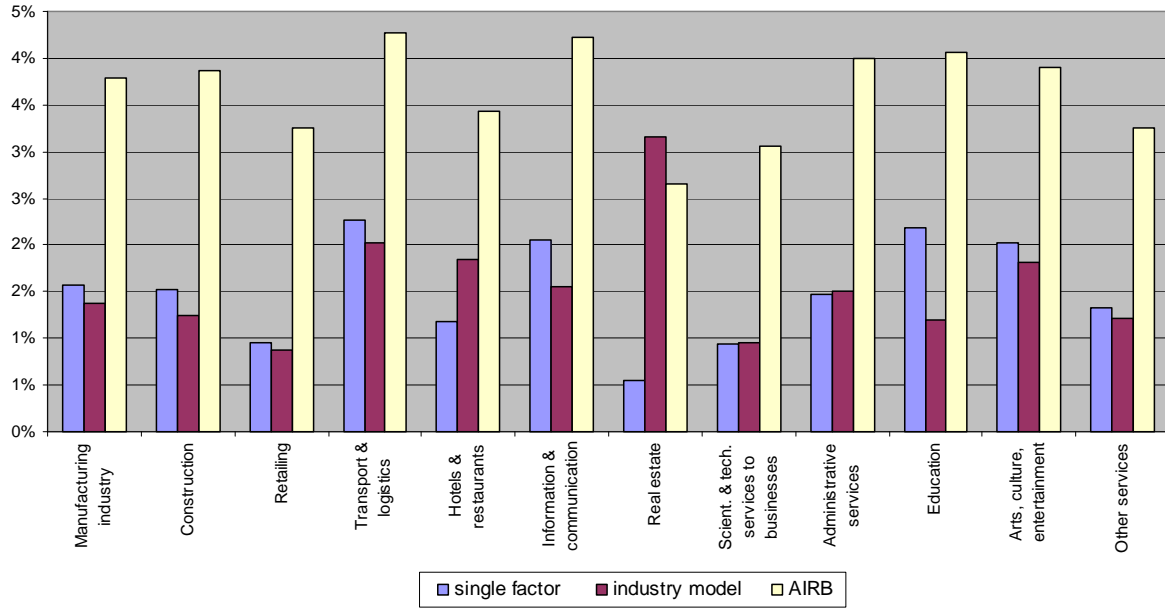
for all remaining industries. Thus, disaggregating the portfolio economic capital at the industry level helps unveiling its heterogeneous repartition of credit risk. Moreover, it highlights the potential for credit risk concentration, here most likely through the potential accumulation of losses on real estate exposures, even if these exposures are on small businesses.

Table 4: Capital ratios by industry (99.97% confidence level).

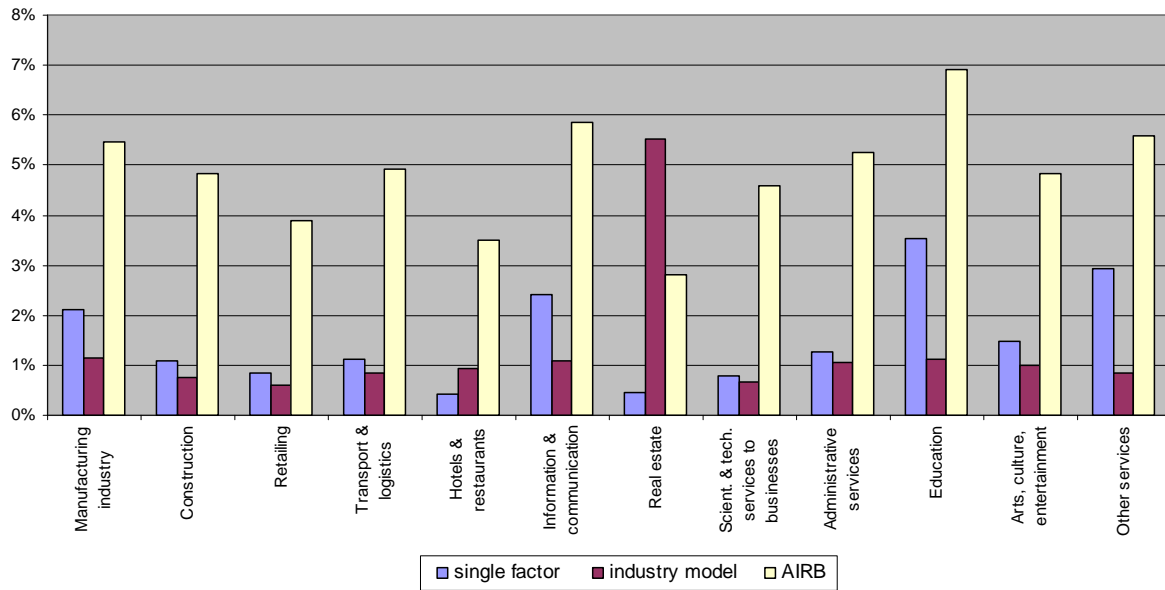
Industry	Very small firms	SME
Manufacturing	1,37%	1,14%
Construction	1,24%	0,75%
Retailing	0,87%	0,59%
Transport & logistics	2,03%	0,84%
Hotels & restaurants	1,84%	0,94%
Information & communication	1,56%	1,10%
Real estate	3,16%	5,53%
Scient. & tech. services to businesses	0,96%	0,67%
Administrative services	1,51%	1,06%
Education	1,20%	1,13%
Arts, culture, entertainment	1,81%	0,99%
Other services	1,21%	0,84%

In order to further highlight this large variation of capital ratios across industries, graph 1A and 1B show the capital ratios shown in table 4 in comparison with the capital ratios derived on the one hand from the homogeneous one factor model and on the other hand from the IRB regulatory formula (in both cases always considering à 45% LGD). Moreover, table 5 provides statistics on these capital ratios. To compute the regulatory capital ratios, we do not consider estimated stationary default probabilities, but instead the empirical average default rates, and we take into account the size adjustment for businesses with turnover between 5 and 50 millions euro (BCBS, 2006, §273).

Graph 1A: Very small firms - Comparison of capital ratios single factor, multi factor and IRBA



Graph 1B: Small and medium-sized firms - Comparison of capital ratios single factor, multi factor and IRBA



In the following, we will focus on the 99.97% quantile of economic capital. Indeed, this confidence level tends to become the financial industry standard when considering internal credit risk models, reflecting AA target rating. Despite the difference in confidence level between the two measures of capital, we will compare the resulting capital ratios from the

multifactor internal model and the regulatory capital requirement. The objective of this comparison is not to determine differences in the level of asset correlations between the two models, which would require considering identical confidence levels. The comparison rather tries to show the potential divergence due to concentration risk between two measures used in the financial industry. If a multifactor internal model calibrated at the industry standard level could lead to higher capital levels than the regulatory formula, there could appear an incentive to use the regulatory formula as a capital target. In other words, there could be an incentive for some regulatory arbitrage. Moreover, this incentive could be related to the heterogeneous repartition of risk, i.e. concentration risk within the credit portfolio.

Table 5 Capital ratio distributions (%).

Panel A Very small businesses (turnover < 750.000 euros)					
Model	Mean	Std	Cv (%)	Min	Max
One factor	0,99%	0,18%	18,2	0,63%	1,34%
Industry	1,56%	0,61%	39,1	0,87%	3,16%
IRB	3,39%	0,31%	9,14	2,72%	3,84%
Panel B SME (750.000 < turnover < 50 millions euros)					
Model	Mean	Std	Cv (%)	Min	Max
One factor	0,86%	0,25%	29,8	0,47%	1,30%
Industry	1,40%	1,69%	120,7	0,59%	6,74%
IRB	4,68%	0,70%	14,9	3,54%	5,69%

Note: Cv is the percent coefficient of variation. Mean is the unweighted mean of capital ratios according to a given credit risk model and a given aggregation variable. All computations are based on a fixed *LGD* of 45%.

Table 5 confirms the observation of graphs 1.A and 1.B. The industry model leads to much more volatile industry capital ratios than the two other models. The model that shows the lowest variability in industry capital ratios is the regulatory model. The lower variability in comparison with the homogeneous one factor model can be attributed to the inverse relationship between the default probability and the model-based asset correlation. Thus, the regulatory model structurally dampens the variability in capital requirements as any increase of default probabilities across industries (in this case attributable in different repartitions of

internal ratings across industries) is at least partially compensated by a lower computed asset correlation. This suggests that the regulatory formula may not be able to capture the impact of industry shocks on capital allocation. In other words, the IRB scheme leads to a homogeneous capital allocation. This feature, intentional in the conception of the ratio in order to reduce the incentive of banks to discriminate across borrowers in access to and pricing of credit, limits the usefulness of the regulatory model in the context of Pillar 2 of Basle 2. These results moreover illustrate that the introduction of additional risk factors has not only an impact on the level of economic capital, but also an important effect on the representation of the risk's distribution within the portfolio.

5. Evaluating the impact of concentration risk: a simulation.

In order to evaluate the ability of the multifactor model and the regulatory formulas to capture concentration risk, this section performs a simulation exercise based on the results of the preceding sections. For each size class, we construct random portfolios of, respectively, 10.000 and 20.000 exposures of SMEs and very small businesses. For each simulated portfolio, we compute the corresponding regulatory and economic capital ratios at the portfolio and the industry levels. Moreover, the concentration of each portfolio in terms of credit exposures and capital are characterized by computing Herfindahl-Hirschman indexes using alternatively the industry shares in exposures and capital (economic and regulatory) This procedure is repeated 10 000 times.

Concentration risk corresponds to the potential concentration of losses in some sub-portfolios due to industry specific shocks and/or different sensitivities with respect to aggregate (economy-wide) shocks. Thus, concentration risk is not solely linked to (relatively) large

industry credit exposures, but also to the industry credit risk. A portfolio's credit risk, measured by the level of its capital requirement could stem from industries representing relative small shares of total credit exposures. Consequently, the industry concentration of a portfolio might by itself not represent its credit concentration risk. In addition to the capital ratios, we also compute the industry capital concentration. One would expect the overall capital level to be related to the portfolio's concentration risk. In other words, we would expect the capital HH index to be positively to the capital level whereas the link between the exposure HH index and the capital level could not ex ante be specified. In fact, portfolios could be concentrated on low risk industries as far these industries are the most important in the economy leading to a possible negative relationship between capital and the industry concentration. On the contrary, if a high risk industry, i.e. with higher marginal risk contributions represents a large share of capital, we would expect a positive relationship between the capital ratio and the HH index of capital. Table 6 shows the correlations across the different measures of capital and concentration.

Table 6 Simulation results

Panel A SME (750.000 < turnover < 50 millions euros)					
	IRB	EC	HHI IRB	HHI EC	HHI
IRB	1				
EC	0.46	1			
HHI IRB	0.26	0.44	1		
HHI EC	-0.21	0.66	0.45	1	
HHI	-0.20	0.28	0.74	0.50	1
Panel B Very small businesses (turnover < 750.000 euros)					
	IRB	EC	HHI IRB	HHI EC	HHI
IRB	1				
EC	0.83	1			
HHI IRB	0.05	0.30	1		
HHI EC	0.36	0.78	0.55	1	
HHI	-0.43	-0.22	0.81	0.17	1

Note: Table shows the correlation matrices between regulatory capital requirements (IRB), the 99,97% economic capital level (EC), the corresponding industry concentration of capital (HHI_IRB and HHI_EC) as well as the

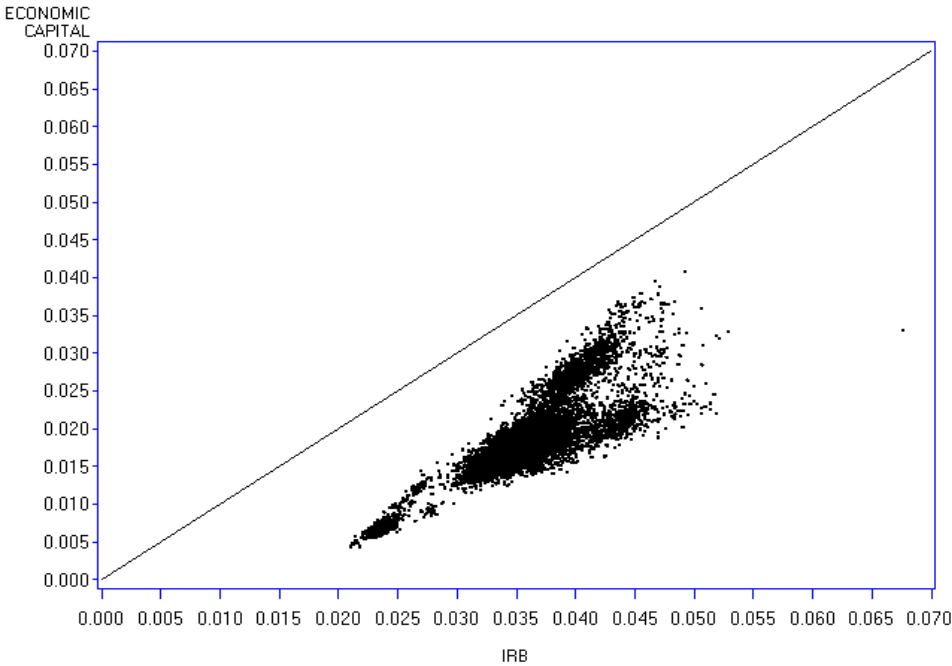
industry concentration based on exposures (HHI). All correlations are statistically significant at the 1% level and are computed using 10000 simulated portfolios.

First, the simulation shows that there is a strong link between the concentration of exposures and the concentration of regulatory capital requirements (respectively 0.74 for small and 0.81 for very small businesses). These correlations are lower when considering the internal model based capital (respectively 0.50 and 0.17). Thus, the regulatory formula reflects to a larger extent the industry composition of the portfolio. This might be the result of lower potential dispersion of asset correlations in the regulatory formula than in an internal model, as differences in asset correlations follow differences in default probabilities. Indeed, the higher the correlation between the concentration of exposures and the concentration of capital, the higher will be the homogeneity of the credit risk model underlying the capital ratio. However, assuming that name concentration is not a concern when considering large retail portfolios, concentration risk does not pertain solely to the concentration of exposures, but should be evaluated based on the repartition of credit risk within the portfolio. Thus, this result also suggests that the level of industry heterogeneity could be higher for very small businesses as the link between the industry composition of the simulated portfolios and the distribution of capital across industries is smaller (0.17 for very small firms against 0.5 for SME). This observation is coherent with the observation of section 4 that SME may be more homogeneous in terms of credit risk when excluding the real estate industry.

Moreover, the link between the concentration of capital and the level of capital appears to be stronger for the internal multifactor model than for the regulatory formula. Thus, the explicit introduction of industry heterogeneity in the modelling of credit risk leads to capital requirements that allows for a stronger recognition of the potential concentration of losses when exposures on riskier sub-portfolios (industries) gain more importance. Indeed, the expansion on more risky sub-portfolios should be related to an increase in the overall level of

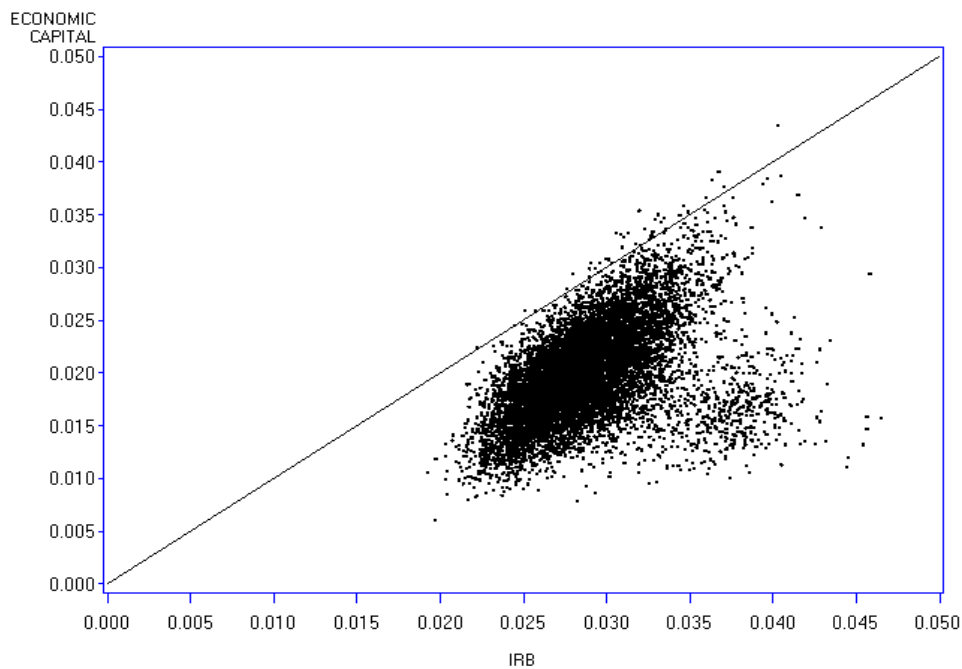
capital in order to reflect concentration risk. Results also underline the central role of the real estate industry on the credit risk of business retail portfolios. This sector distinguishes itself sharply by a far larger leverage level than other industries, even when considering only small and medium firms, all firms with turnover larger than 50 millions € having been removed from the database. Although the average default frequency remains low⁷, real estate exhibits the most important asset correlation, i.e. the highest sensitivity to the business cycle. Despite its specification of asset correlation based on an inverse relationship to default probabilities, the regulatory formula of capital requirements is only imperfectly able to capture the specificity of the real estate industry in terms of asset correlation. This highlights the importance to reinforce complementary models and approaches in addition to regulatory formulas in order for supervisory processes (especially within the Pillar 2 of Basel 2) to be meaningful.

Figure 2.A Economic capital against IRB capital for very small firms



⁷ For very small firms, the average default rate is about 0.7% against 1.8% for all industries and for SME, figures are respectively 0.4% and 1.2%

Figure 2.B Economic capital against IRB capital for SME



Figures 2.A and 2.B show the simulation results by plotting the economic capital against the minimum regulatory capital requirements for respectively very small firms and SME. Figure 2.A shows the conservative level for very small firms, as there subsists a considerable gap between the IRB requirements and the computed economic capital even when considering the higher 99.97% quantile. On the contrary, when considering SME (with turnover between 0.75 and 50 millions euros), figure 2.B shows that in some cases (about 0.3% of simulated portfolios), economic capital appears to be higher than regulatory requirements and that former is near latter in many cases. This suggests that the regulatory model may, for some concentrated portfolios, understate the portfolio's credit risk using an industry standard for the confidence level of economic capital. Again, this result can be partly attributed to the importance of credit risk in real estate. This further underlines that specializations on credit markets may lead to highly heterogeneous credit risk levels across banks, challenging the ability of the regulatory framework to insure financial stability. Finally, the differences in the

sensitivity of capital with respect to capital concentration may further suggest the possibility of capital arbitrage if banks exploit the risk differences across industries not reflected in the regulatory framework. Simulations nevertheless show that in most cases concentration may not threaten the stability of banks by inducing a considerable gap between economic capital and regulatory capital levels (these results are in line with e.g. Duellmann and Masschelein, 2006). However, the ignored heterogeneity in industry credit risk parameters may lead to an increase in credit risk by concentrating exposures on risky segments without inducing a corresponding increase in capital.

5. Conclusion

In the context of Pillar 2 of Basel II, the detection and measurement of credit risk concentration in large portfolios needs the extension of the regulatory framework in order to introduce additional sources of systematic risk in the modelling of credit risk. This article gives some qualitative insights on the concentration potential within large portfolios of small and medium businesses by extending the standard one factor credit default model to a multi-factor framework. Results show that the one factor model largely fails in capturing potential risk concentration as characterised by capital allocation. This shows that the heterogeneity captured by credit ratings, the only source of heterogeneity in the asymptotic one factor framework, fails to describe the effective heterogeneity in default rates within large portfolios and that other factors might be at play. Moreover, industry effects appear to have considerable effects on the heterogeneity of credit risk. More specifically, concentration or specialization on certain industries may even challenge regulatory minimum capital requirements. Finally, capital allocation based on the regulatory framework has only a limited ability to reflect concentration risk.

References

- BCBS, 2006, Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version, Basel Committee on Banking Supervision.
- BCBS, 2006, Studies on credit risk concentration, Bank for International Settlements, Basel Committee on Banking Supervision, WP No. 15.
- BCBS, 2008, Range of practices and issues in economic capital modelling, Consultative document, Basel Committee on Banking Supervision.
- Carling K., T. Jacobson, J. Linde, K. Roszbach, 2007, Corporate credit risk modelling and the macroeconomy, *Journal of Banking and Finance* 31, 845-868.
- Carling K., L. Ronnegard, K. Roszbach, 2004, Is firm interdependence within industries important for portfolio credit risk?, *Sveriges Riksbank working paper series*, n° 168.
- Dietsch M., J. Petey, 2004, Should SME exposures be treated as retail or corporate exposures? A comparative analysis of default probabilities and asset correlations in French and German SMEs, *Journal of Banking and Finance* 28, 773-788.
- Duellmann K., H. Scheule, 2003, Asset correlation of German corporate obligators: Its estimations, its drivers and implications for regulatory capital. Paper presented at Banking and Financial Stability: A Workshop on Applied Banking Research, Banca d'Italia, Rome, March.
- Duellmann, K., N. Masschelein, 2006, The Impact of Sector Concentration in Loan Portfolios on Economic Capital, *Financial Stability Review*, ed. National Bank of Belgium, 175-187.
- Duellmann K., M. Scheicher, C. Schmieder, 2007, Asset correlations and credit portfolio risk: An empirical analysis, Working paper, Deutsche Bundesbank.
- Egloff D., M. Leippold, S. Jöhri, C. Dalbert, 2005, Optimal importance sampling for credit portfolios with stochastic approximation, working paper.
- Frey, R., A. McNeil, 2003, Dependent defaults in models of portfolio credit risk. *Journal of Risk* 6 (1), 59-92.
- Glasserman P., J. Li, 2005, Importance sampling for portfolio credit risk. *Management Science*, 51(11), 1643-1656.
- Gouriéroux C., J.P. Laurent, O. Scaillet, 2000, Sensitivity analysis of Values at Risk, *Journal of Empirical Finance* 7, 225-245.
- Heitfield N., S. Burton, S. Chomsisengphet, 2006, Systematic and idiosyncratic risk in syndicated loan portfolios, *Journal of Credit Risk* 2 (3), 3-31.
- Higham N., 2002, Computing the nearest correlation matrix: a problem from finance, *IMA Journal of Numerical Analysis* 22, 329-343.

Klebaner F. C., *Introduction to Stochastic Calculus with Applications*. Imperial College Press, second edition, 2005.

Lopez J.A., 2004, The empirical relationship between average asset correlation, firm probability of default and asset size, *Journal of Financial Intermediation* 13, 265-283.

Lopez J.A., M.R. Saldenber, 2000, Evaluating credit risk models, *Journal of Banking and Finance* 24, 151-165.

Lucas A., P. Klaassen, P. Spreij, S. Straetmans, 2001, An analytic approach to credit risk of large corporate bond and loan portfolios, *Journal of Banking and Finance* 25, 1635-1664.

McNeil A., J. Wendin J, 2007, Bayesian inference for generalized linear mixed models of portfolio credit risk, *Journal of Empirical Finance* 14, 131-149.

Pesaran M.H., T. Schuermann, B.J. Treutler, 2005, The role of industry, geography and firm heterogeneity in credit risk diversification, IEPR working paper 05.25.

Tasche, D., 1999, Risk contributions and performance measurement, Working paper, Technische Universität München.

Tasche D., 2006, Measuring sectoral diversification in an asymptotic multi-factor framework, *Journal of Credit Risk* 2 (3), 33-55.

Tasche D., 2008, Capital allocation for credit portfolios with kernel estimators, working paper.

Wolfinger R., M. O'Connell, 1993, Generalized linear mixed models: A pseudo-likelihood approach, *Journal of Statistical Computation and Simulation* 4, 233-243.

Appendix A.1 Estimation results: Homogeneous one factor model

A. Goodness-of-fit measures

	Very small firms	SME
-2 Res Log Pseudo-Likelihood	184.02	12.95
Pseudo AIC	186.02	14.95
Pseudo BIC	186.21	15.15
Generalized Chi-Square	468.32	261.52
Gener. Chi-Square / DF	8.36	4.67

B. Covariance parameters

	Very small firms		SME	
Parameter	Mean	Std error	Mean	Std error
Intercept	0.003691	0.001869	0.005669	0.002888

C. Default thresholds and probabilities

	Very small firms			SME		
Rating	Estimate	Std error	Default probability	Estimate	Std error	Default probability
1	-0.8078	0.02263	0.20960	-0.7183	0.02819	0.23627
2	-1.4100	0.02114	0.07927	-1.2872	0.02659	0.09902
3	-1.7908	0.02082	0.03666	-1.7312	0.02632	0.04170
4	-1.9942	0.02081	0.02306	-1.9877	0.02607	0.02342
5	-2.2423	0.02111	0.01247	-2.3197	0.02643	0.01018
6	-2.4666	0.02164	0.00682	-2.5738	0.02728	0.00503
7	-2.8220	0.02203	0.00239	-3.0283	0.02743	0.00123

Appendix A.2 Estimation results: Industry risk factors

Industries are 1: Manufacturing industry; 2: Construction ; 3: Retailing ; 4 : Transport et logistics; 5: Hotels and restaurants; 6: Information and communication; 7: Real estate; Scientific and technical services to businesses; 8: Administrative services; 9: Education ; 10: Arts, culture, entertainment; 12: Other services. Other industries were excluded due to the absence of defaults over the period and the very low number of small and medium businesses in these sectors.

A. Goodness-of-fit measures

	Very small firms	SME
-2 Res Log Pseudo-Likelihood	-355.44	283.43
Pseudo AIC	-309.44	329.43
Pseudo BIC	-304.90	333.97
Generalized Chi-Square	1677.99	1512.18
Gener. Chi-Square / DF	2.24	2.02

B. Covariance parameters

Parameter	Very small firms		SME	
	Mean	Std error	Mean	Std error
CHOL(1,1)	0.07910	0.02146	0.1761	0.04334
CHOL(2,2)	0.09990	0.02562	0.09652	0.02720
CHOL(3,3)	0.1189	0.02934	0.1678	0.04133
CHOL(4,4)	0.1420	0.03613	0.06993	0.02724
CHOL(5,5)	0.1183	0.03012	0.3587	0.08907
CHOL(6,6)	0.08079	0.02340	0.1325	0.03788
CHOL(7,7)	0.3046	0.07353	0.4870	0.1203
CHOL(9,9)	0.05109	0.01833	0.03241	0.03275
CHOL(10,10)	0.1075	0.03066	0.2514	0.06666
CHOL(11,11)	0.09053	0.02971	0.09038	0.05310
CHOL(12,12)	0.05657	0.02274	0.2118	0.06580
CHOL(13,1)	-0.01173	0.03463	0.05714	0.04894
CHOL(13,2)	-0.05639	0.04358	-0.05082	0.02932
CHOL(13,3)	-0.06998	0.03852	-0.03013	0.05791
CHOL(13,4)	0.04089	0.04738	-0.01321	0.02198
CHOL(13,5)	0.03976	0.04820	0.04463	0.05967
CHOL(13,6)	-0.02834	0.02820	-0.03211	0.03520
CHOL(13,7)	-0.01658	0.07820	0.01528	0.06471
CHOL(13,8)	1.37E-23	0.01394	1.03E-23	0.01616
CHOL(13,9)	-0.00332	0.01967	-0.00354	0.01764
CHOL(13,10)	-0.00466	0.02278	0.05874	0.02780
CHOL(13,11)	0.03343	0.01812	-0.00719	0.01914
CHOL(13,12)	0.01926	0.02054	-0.00115	0.02623

Note: CHOL(,) denotes the position (row, column) of the estimated parameter in the Cholesky root of the covariance matrix of random effects. Non-shown elements of the matrix are zero. For all models, the covariance matrix of random effect is parameterized through its Cholesky root A . This parameterization ensures that the resulting variance-covariance matrix is at least positive semi-definite. If all diagonal values are nonzero, it is positive definite. The variance-covariance matrix of random effects C is then simply defined as $C = A'A$.

C. Default thresholds and probabilities

Rating	Very small firms				SME			
	Estimate	Std error	P-value	Default probability	Estimate	Std error	P-value	Default probability
1	-0.7856	0.01376	<.0001	0.21604	-0.7478	0.01711	<.0001	0.22729
2	-1.3903	0.01123	<.0001	0.08222	-1.3165	0.01422	<.0001	0.09401
3	-1.7571	0.01082	<.0001	0.03945	-1.7493	0.01390	<.0001	0.04012
4	-1.9660	0.01064	<.0001	0.02465	-2.0254	0.01346	<.0001	0.02141
5	-2.2200	0.01110	<.0001	0.01321	-2.3661	0.01412	<.0001	0.00899
6	-2.4442	0.01210	<.0001	0.00726	-2.6220	0.01563	<.0001	0.00437
7	-2.7928	0.01286	<.0001	0.00261	-3.0690	0.01594	<.0001	0.00107