MAGYAR NEMZETI BANK

MNB WORKING PAPERS

2006/11

ZOLTÁN VARSÁNYI

Pillar I treatment of concentrations in the banking book – a multifactor approach

Pillar I treatment of concentrations in the banking book – a multifactor approach

August 2006

MNB Working Papers 2006/11

Pillar I treatment of concentrations in the banking book – a multifactor approach

(Banki könyvi koncentrációk kezelése az első pillérben – egy többfaktoros megközelítés)

Written by: Zoltán Varsányi*

Magyar Nemzeti Bank

Szabadság tér 8–9, H–1850 Budapest

http://www.mnb.hu

ISSN 1585 5600 (online)

* Economist, Magyar Nemzeti Bank (the central bank of Hungary); varsanyiz@mnb.hu. The author wishes to thank Marianna Valentinyiné Endrész, Edit Horváth, Balázs Janecskó, Márton Radnai and Anikó Szombati for helpful discussions. All the remaining errors are those of the author.

Contents

Abstract

The present regulation of concentration risk does not take into consideration recent, sophisticated methods in credit risk quantification; the new Basle Capital Accord has left the regulatory treatment unchanged. Recently, substantial work has begun within the EU on this issue with the formation of the Working Group on Large Exposures within CEBS. The present paper is concerned with the models available under Basle 2 for credit risk quantification: it is searching for tools that can be applied in a new regime in general and that are capable of replicating the riskiness of credit portfolios with risk concentrations – an area that the original Basel model does not cover. The main idea of the paper is to disassemble non-granular portfolios into homogenous parts whose loss can then be directly simulated – taking into consideration the correlation between the parts – without the need to simulate single exposures. This makes the calculation of portfolio-wide loss very fast.

JEL Classification: C15, G28
Keywords: Basel regulation, multifactor model, NORTA, numerical integration.

Összefoglalás

A koncentrációs kockázat jelenlegi szabályozása nem veszi figyelembe a hitelkockázat-mérés új, fejlett módszereit, s ezek az új bázeli rezsimben sem jelennek meg. A változtatási igény és a lehetséges irányok feltérképezésére az utóbbi időben az Európai Unióban (például a CEBS-en belül működő Nagykockázati Munkacsoportban – Working Group on Large Exposures, WGLE) jelentős munka folyik. E tanulmány a hitelkockázat-mérésre alkalmazható modellek egy fajtáján keresztül azt próbálja meg számszerűsíteni, hogy mekkora veszteséget okozhatnak a kockázati koncentrációk a hitelportfólióban. Az elemzés során a homogén portfólióelemeket részportfóliókba soroljuk, majd ezekből – a közöttük lévő korrelációk figyelembevételével – közvetlenül szimulálunk veszteségeloszlást, így nem kell minden egyes kitettséget külön szimulálni. Ez az eljárás a portfóliószintű veszteség számítását rendkívül gyorssá teszi.

1. Introduction

Concentration risk, in short, refers to the risk of losses as a consequence of insufficient diversification. Although the new European Capital Requirements Directive (CRD, the European implementation of the Basel 2 regulation) devotes only a short paragraph to this issue explicitly and in general, for some time it has been an important aspect of banking regulation.[1] On the one hand, one specific type of concentrations, namely single name risk, is explicitly the subject even of the present regulation (paragraphs 106-118 of the CRD): the exposure towards single entities or groups is bounded by limits expressed as percentages of own funds. On the other hand, other types of concentrations have to be '…addressed and controlled by means of written policies and procedures' – these fall within national authorities' competence.

It is a widespread view that the present regulation is unsatisfactory: the evolution towards a more risk-sensitive regulatory regime has left it unchanged for a couple of years. Now the EU has allocated resources to this issue – for example, the Working Group on Large Exposures within the Committee of European Banking Supervisors has been formed. The changes are expected to be concerned with two aspects of the present regulation. First, in the model underlying the advanced (IRB) method of the CRD (and the Basel 2 regulation) there is one factor, the so called systemic factor, that connects asset values of obligors, and the rest of the obligors' asset value is explained by idiosyncratic shocks. Correlations between asset values are not taken into consideration beyond the level caused by the systemic component. Second, the regulation now only considers in more depth one particular part of concentration risk, client concentrations, or as it is more commonly referred to, large exposures. As far as regulation is concerned it is reasonable both to take correlated asset values into consideration and to deal with other types of concentrations.

The literature related to our topic is vast. First, it is worth reviewing the most common credit portfolio models shortly. Crouhy et al. [2000] introduces and compares three approaches. The first, which is followed by Creditmetrics and CreditVar I, and the second – that behind KMV's model – is based on Merton's model (Merton [1974]) where the asset value of firms follows a diffusion process. Originally, Merton examined the term structure of interest rates (and the pricing of risky bonds) but his model naturally extends to the analysis of credit risk in general. If the firm's asset value changes, the value of its (external) liabilities also changes reflecting the change in the (perceived) ability of the firm to fully service its debt obligations. In its simplest form the firm has two status in the model – default or non-default; the firm defaults when its asset value crosses a trigger downward. In more sophisticated forms – such as the above two or KMV's model – the non-default state is further divided into 'sub-states' enabling the modelling of downgrade risk: each sub-state is defined by an upper and a lower trigger and if the asset value is between these triggers the firm is assigned to that sub-state. For example, if the firms's asset value is described by a standard normal random variable over a one-year horizon and the actual asset value is between -2.04 and -1.23 at the end of the year a rating of 'B' is ordered to the firm; if the asset value is between -2.3 and -2.04 the rating is 'CCC'; and if it is below -2.3 the firm is regarded as in default (Crouhy et al. [2000], Fig. 8.). Correlations between exposures enter the model through common factors in the determination of asset values (see also Gordy [1998]). The major difference between KMV and the other two models seems to be that these latter two use simulation to arrive at the desired percentile of the portfolio loss distribution while KMV derives it analytically. The third approach is an 'actuarial' one and is followed by CreditRisk+. However, this approach is not really relevant for us because it uses a rather different framework: it models (at least in its original form) the loss distribution directly (rather than through asset value process) and there are only two states (default and non-default). In fact, the multifactor model I use in Section 3 (and later) is also based on the Merton-concept and is similar to those appearing in the dependency modelling in Creditmetrics or KMV. Further analysis of such models and their relation to the IRB approach in Basel 2 can be found in, for example, Hamerle et al. [2003].

Another aspect the importance of which is emphasised in this paper is, given a model for (individual and portfolio) credit losses, how we generate loss distributions from it. I argue that despite recent advances in analytic approaches (see below in this paragraph) finding an efficient method for simulation is desirable. Cao and Morokoff [2004] claim that a simulation model '…can capture the default correlation and credit migration in a practical manner, but requires a step-wise simulation

[1] According to Annex V, paragraph 7: 'The concentration risk arising from exposures to counterparties, …, counterparties in the same economic sector, geographic region…shall be addressed and controlled….'.

that may often be time-consuming'. Instead they describe the 'default time approach' in which the portfolio loss distribution can be expressed without simulation – still, later in the paper they use the asset-value approach with simulations. A more recent paper (Huang et al. [2006]) uses saddlepoint approximation of the loss percentile for the type of models that the IRB approach in Basel 2 is based on. Although they claim that '…the saddlepoint approximation technique can readily be applied to more general Bernoulli mixture models *(possibly multifactor)'* [2], later they acknowledge that in a multifactor setup Monte Carlo simulation or other complementary method is needed (page 13). An important reference for us is Pykhtin [2004]. Here, the author extends the granularity adjustment (see, for example, Martin and Wilde [2002]) approach to include more factors. His method is very flexible (allowing an arbitrary number of factors, sub-portfolios – buckets – and exposures in the buckets), fast (since the calculations result in closed-form solutions). At the same time, for relatively low levels (0.1-0.2) of asset-value correlations and a small number of independent factors the results seem to be somewhat inaccurate (cf. Table 2 and Table C in Pykhtin [2004]). Thus, in light of the fact that one way or another simulation might be needed to arrive at given portfolio loss percentiles (and still more for the whole distribution) I intended to find a method that makes portfolio loss simulation as fast as possible.[3] I arrived at the NORTA technique (see Cario and Nelson [1997]) which uses correlated normal variables to arrive at correlated random variables from arbitrary marginal distributions – I have not find any references to this method in the context of portfolio credit risk.

Finally, the simplest form of factor models has to be highlighted. It and its explanation appear in, for example, Gordy [2002] and Bank and Lawrenz [2003] and is an important one since the new Basel 2 regulation largely builds on it. Gordy [2002] examines portfolios of n exposures with the assumption that the asset value of obligors behind these exposures evolves according to a one-factor model. The term 'granularity' is used in essence to describe the number of (conditionally independent) exposures: the higher n is, the more granular the portfolio is, so the stronger the diversification effect on the risk of the portfolio is. It has to be noted that this term does not refer to what extent exposures are correlated, so the Gordy model, in its original form, is not capable of handling correlated exposures.[4] It also has to be noted that my approach to granularity is different from Gordy's (and most of the other papers that address this question): I do not adjust the level of granularity through the number of exposures but rather through the parameters that effect correlations across exposures and buckets. This approach seems to be more natural in some cases and works with the NORTA technique very well (it allows fast simulations). According to Martin and Wilde [2002] saddlepoint methods can also handle such situation: in their article the parameter *'u'* can take up every effect that drives a given percentile of the loss distribution in the presence of any 'additional' risks away from the percentile of the perfectly granular distribution. This additional risk can be, for example, risk concentration – however, in this case the interpretation of the model may become difficult, as well as the mapping of the true effect of concentration to the model parameters.

In what follows I first describe the model underlying the CRD (I will refer to it as the 'Basel model'). Then, in Section 3 I present a multifactor model – a natural extension to the Basel one-factor model. Subsequently I analyse the effects of a particular type of risk concentrations, single name risks, in Section 4. Here, I don't apply simulation but a different method – numerical integration – which might seem to cause a break in the logic of the build-up of the paper. However, I found it easier to introduce the problem of single name risks this way and, in very simple cases, numerical integration is an alternative to simulation and, moreover, can serve as benchmark for the results of later simulations. I found no previous references to the method I apply here. In Section 5 I bring together the concepts of multiple factors and risk concentrations using both methods – numerical integration and simulation – emphasising that even in less complicated cases only the latter one is applicable. Finally, I discuss some relevant issues and draw conclusions. Although I base the analysis on the model behind the advanced (IRB) method of the regulation the results are intended to have general validity, covering the Standardised Method, as well. Indeed, I regard the analysis here as exploring the 'true' risks of concentrations that have to be addressed both within the IRB and the Standardised approaches, in the latter one possibly in a simplified manner (at present there is no distinction in the large exposure regulation between IRB and Standardised).

---

[2] Page 2, highlighted by me.

[3] It is not clear to what extent the simulation technique in this paper performes better then the granularity adjustment method in Pykhtin [2004]. Asset correlations according to the model behind Basel 2 are at the lower end of the spectrum of the values Pykhtin [2004] examines (with a maximum of 0.24 and for worse quality assets it's even smaller) – in this range the analytic approximation is not very good for a model with 11 factors. And while the approximation worsens when there are fewer factors the more accurate simulation becomes since fewer sources of risk have to be simulated. However, in this paper the emphasis is on the presentation of the ideas and not on their comparison with other existing techniques.

[4] Although decreasing n and increasing the exposure size might be regarded as increasing the portfolio's concentration on n factors. However it seems to be useful to separate the problem of correlation from the question of the number of elements in the portfolio.

2. The Basel model and the RW function

In the Basel model it is assumed that the lending institution has a sufficiently fine-grained credit portfolio (i.e. one, in which no individual obligations 'dominate'). Moreover, it is assumed that the asset-value of the obligors is determined by a systemic risk factor and an idiosyncratic shock and the obligor defaults if the decline in its asset value crosses a trigger. Formally, the asset value change process of obligor *i* can be written as:

$$R_i = wX + \varphi\varepsilon_i ,$$

where X represents the systemic factor, ε the idiosyncratic shock – both are i.i.d. standard normal – and w and φ are parameters. It is further assumed that the obligor defaults if the fall in its asset-value is above (in absolute terms) a trigger level, say, γ. Thus, the probability of default can be expressed as:

$$P\left(R_i < \gamma\right) = P\left(\varepsilon_i < \frac{\gamma - wX}{\varphi}\right) = N\left(\frac{\gamma - wX}{\varphi}\right),$$

or, expressed conditional on X:

$$P\left(R_i < \gamma \mid X = x\right) = P\left(\varepsilon_i < \frac{\gamma - wx}{\varphi}\right) = N\left(\frac{\gamma - wx}{\varphi}\right). \qquad (2.1)$$

The formula in (2.1) forms the basis of risk weights for exposures to a given obligor in the Basel 2 regulation.[5] Actually, φ is set at $\sqrt{1 - w^2}$ so that the unconditional variance of the asset value equals 1.

The underlying idea behind moving from the individual to the portfolio level is that exposures are conditionally independent so conditional probabilities can be summed (more about this issue can be found in Bank and Lawrenz [2003]).

2.1 UNDERSTANDING THE BASEL MODEL

As was described earlier the central component of the risk weight an exposure receives is the conditional probability (given the realised value of the systemic factor, X). This is illustrated in Figure 1.

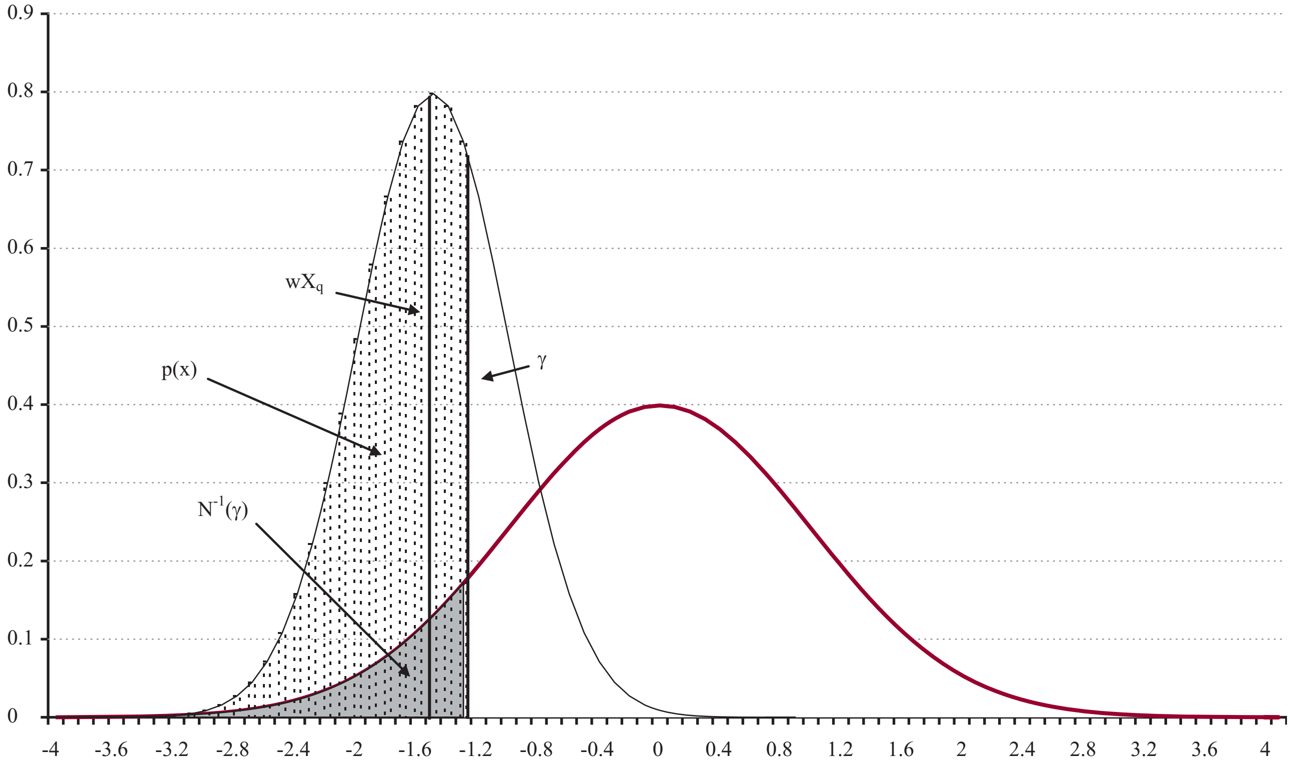
The bell-shaped curve with white fill is the standard normal density function. This represents both the distribution of the systemic factor and the unconditional distribution of the asset value changes of an obligor. $N^{-1}(\gamma)$ is the area in grey: this is the unconditional probability of default for the obligor. The corresponding trigger level (the change in the asset value below which the obligor defaults) is γ. On the figure γ equals 1.3 approximately, and so the unconditional PD is around 10%.

The realised value of X is shown by X_q and wX_q is its effect on the asset value. In the Basel model X_q is fixed at -3.09 (the inverse normal value of 0.1% or, put it in a different way, the value above which X takes on values with probability 99.9%). In this example w is around 0.5, so wX_q is around 1.5.

[5] Probably the largest difference between this formula and the actual Basel risk weight formula is that the latter contains a correction for the expected loss: banks are supposed to apply value corrections to loss levels up to the expected loss.

Figure 1

The Basel model represented visually



Now, if we have the realised value of *X*, the only uncertainty as to the change in the asset value stems from idiosyncratic, firm specific shocks. From the model it follows that these shocks will be normally distributed with expected value 0 and standard deviation less than 1.[6] As a consequence, the conditional (on *X*) distribution of the asset value is distributed normally, with expected value wX_q and standard deviation less than 1. This distribution is also shown in the Figure. The probability – given X_q – that the asset value will be less than γ is represented by the area of this conditional distribution to the left of γ. This probability is denoted by *p(x)* and in the Basel model it forms the basis of the risk weight laid on an exposure.

It can be shown that if there are enough number of exposures in the portfolio (none of which dominating the portfolio) the portfolio loss rate will be *p(x)* in each period, in the absence of correlation between the idiosyncratic shocks. This is true for any values of *X*; moreover, the relationship between *X* and *p(x)* is strictly monotone (a lower *X* implies a higher *p(x)*). As a consequence, a given quantile of the distribution of *X* corresponds to the same quantile of the *unconditional* distribution of *p(x)*. In short, *the regulation is looking for the q^{th} quantile of the unconditional loss distribution, which it finds through the conditional distribution belonging to the q^{th} quantile of the distribution of the systemic factor.*

2.1.1 Correlation with the systemic factor

In the one-factor model the conditional default probability depends on two variables: the unconditional default probability and the correlation with the systemic factor. This can be seen from equation (2.1) which I repeat here for convenience:

$$P\left(R_i < \gamma \mid X = x\right) = N\left(\frac{\gamma - w_i x}{\varphi_i}\right).$$

[6] In fact, the standard deviation equals $\sqrt{1-w^2}$.

The unconditional PD (PD_u) enters the picture through γ, since $\gamma = N^{-1}(PD_u)$. Moreover, since the volatility of the change in the asset value equals 1 in the model, φ equals $\sqrt{1-w^2}$.

A special feature of the Basel model is that γ and w are connected one-to-one through a function, commonly referred to as the correlation function:

$$w = \sqrt{\left(0.12(1-e^{-50PD})/(1-e^{-50}) + 0.24(1-(1-e^{-50PD}))/(1-e^{-50})\right)}\,.$$

As a consequence of this function lower unconditional PD results in a lower conditional default probability: the effect of increasing w is more than offset by the decrease of γ. So, while in the previous subsection an increasing w led to a higher conditional PD, in the Basel model – as the consequence of the correlation function – the opposite is true: higher w goes with lower conditional PD.

3. Correlation across the idiosyncratic shocks

When conditional distributions (corresponding to idiosyncratic shocks) are independent it is true that – as shown above – the Basel risk weights will reflect the risks appropriately. When the idiosyncratic shocks are correlated, the q[th] quantile of the unconditional loss distribution doesn't correspond any more to the same quantile of the systemic factor, so risk weights based on the latter will underestimate the true risk at the q[th] quantile. To introduce correlation between the idiosyncratic shocks I use a multifactor model that – beyond the systemic factor – contains an arbitrary number of 'secondary' factors. I present the model in the next subsection.

3.1 A MULTIFACTOR SOLUTION

The model is a straightforward extension of the basic one-factor model to more factors where the extra factors can represent regional, sectoral or other types of concentrations. I will refer to these factors as secondary factors.

Recall that the one-factor model has the form:

$$R_i = wX + \varphi\varepsilon_i \,,$$

where *X* is the systemic factor and ε is the idiosyncratic shock. Moreover, *j* is defined to equal $\sqrt{1-w^2}$ so the variance of *R* equals 1 (while the expected value is 0).

The extension of the model with secondary factors yields:

$$R_i = wX + b_1 F_1 + b_2 F_2 + \ldots + b_k F_k + d\eta \,, \qquad (2.2)$$

where Fis are correlated factors (e.g. regions, sectors), $b_i s$ are sensitivities and *d* is such that the variance of *R* equals 1.

The F factors are allowed to be correlated, which makes the model very difficult to apply for risk weight calculation. For this reason I rewrite the model with such a structure that only uncorrelated factors appear in the equation of *R*. This is done, first, by assuming that the F factors can be written as:

$$F_1 = \alpha_1 X + \beta_1 v_1 \,,$$

$$F_2 = \alpha_2 X + \beta_2 v_1 + \gamma_2 v_2 \,,$$

$$\ldots$$

$$F_k = \alpha_k X + \beta_k v_1 + \ldots + \delta_k v_k \,,$$

where the $v_i s$ are independent factors that have standard normal distribution and $\beta_i s$ are sensitivities to these independent factors. The coefficients are constructed in such a way that all $F_i s'$ variance equals 1 (for example, $\beta_1 = \sqrt{1-\alpha_1^2}$, $\gamma_2 = \sqrt{1-\alpha_2^2-\beta_2^2}$).

It is important to introduce an additional independent v_i factor for each F_i, otherwise (if each v_i were allowed in the equation in each F_i) the model would be impossible to identify. This way identification is possible if we know the correlations of the F_i factors with each other and the systemic factor.

Then it is possible to rewrite *R* in terms of the independent factors (by substituting the expressions for the $F_i s$ into (2.2)):

$$R_i = wX + \sum_{j=1}^{k} \delta_j v_j + \varphi_i \varepsilon_i \,, \tag{2.3}$$

where v_j *(=1,…,k)* denotes the j[th] secondary factor and δ_{ji} is its coefficient. The δ_{ji} coefficients are simple linear combinations of the bis and the coefficients of the independent factors in the equations of the F_i factors (for example, δ_1 equals $\sum_{i=1}^{k} b_i \beta_i$). None of the factors are correlated either with each other or with the systemic or idiosyncratic component.

The proposed solution for risk weight calculation, *within a bucket,* is as follows. Since the non-idiosyncratic variables are independent standard normal variables their weighted sum is also normal with expected value zero and variance $w^2 + \sum \delta_{ji}^2$. Thus, (2.3) can be compared to the following one-factor model:

$$R_i = \sqrt{w^2 + \sum_{j=1}^{k} \delta_{ji}^2}\, Y + \varphi_i \varepsilon_i = w_i^* Y + \varphi_i \varepsilon_i, \tag{2.4}$$

where *Y* is a standard normal random variable. As is obvious, the properties of the asset value are the same under the two models, (2.3) and (2.4): the setup of the multifactor model can be translated into a one-factor model.

For two secondary factors the table below shows the conditional probabilities:

Table 1

Conditional probabilities corresponding to the 0.1% percentile in a three factor model

		β_2/β_1	0	0.1	0.2	0.3	0.4	0.5	0.6
PD (%)	1	0	14.03	14.76	17.01	21.00	27.09	35.88	48.28
w	0.44	0.1	14.76	15.50	17.79	21.83	28.01	36.93	49.51
pd_cond_x	14.03	0.2	17.01	17.79	20.18	24.40	30.85	40.16	53.30
		0.3	21.00	21.83	24.40	28.94	35.88	45.88	59.99
		0.4	27.09	28.01	30.85	35.88	43.54	54.60	70.10
		0.5	35.88	36.93	40.16	45.88	54.60	67.13	84.05
		0.6	48.28	49.51	53.30	59.99	70.10	84.05	98.30

As can be seen from Table 1 increasing deltas – while holding *w* fixed – increases the default probability thereby implying higher risk weights. The reason is that the coefficient of the 'superfactor' *Y*, *w** in (7), increases. In Appendix 3 I demonstrate through a simulation how – as an example – a three-factor model and the corresponding multifactor model give the same results.

3.1.1 The order of the factors

It is an interesting question whether the results of the model are sensitive to the order of the factors in it. For example, does it make any difference if we assume that F_1 is the regional factor and F_2 is a sectoral factor instead of assuming that F_1 is the sectoral and F_2 is the regional factor?

To demonstrate the point I use a three-factor setup. *X* denotes – as usual – the systemic factor, and the two secondary factors are related to *X* and each other as follows:

$$F_1 = \alpha_1 X + \sqrt{1-\alpha_1^2}\,\varepsilon_1$$

$$F_2 = \alpha_2 X + \beta_2 \varepsilon_1 + \sqrt{1-\alpha_2^2-\beta_2^2}\,\varepsilon_2$$

In this model correlations between the factors are as follows:

$$\rho_{X,F_1} = \alpha_1$$
$$\rho_{X,F_2} = \alpha_2$$
$$\rho_{F_1,F_2} = \alpha_1\alpha_2 + \beta_2\sqrt{1-\alpha_1^2}$$

Given the above correlation structure the F_i factors are interchangeable in a sense: we simply change the α parameters and adjust the other parameters so that the correlations with the *systemic factor* and between the $F_i s$ remain unchanged.

At the same time, changing the order of the factors, it is not possible to ensure that their correlation with the independent ε_i factors remain unchanged. However, I don't think it is a problem because we care about the correlation of the *F* factors with each other and with the systemic factor (these remain unchanged with the changing of the order of the *F* factors) and not with the independent factor. Moreover, the interpretation of these latter ones is not straightforward and involves some flexibility when changing the order of the *F* factors.

3.2 THE MULTIFACTOR MODEL ACROSS BUCKETS

Unfortunately, the 'superfactor' approach, as presented above, only works when there is one bucket. The reason is that when there are more buckets, each bucket has its own superfactor and the conditional PDs cannot be summed to get the portfolio conditional PD: in the Basel model the systemic factor is common across the exposures, so the qth quantile is calculated with respect to the same variable, whereas in the superfactor model the quantiles are calculated with respect to different variables (so their correlation structure should be taken into account).

Before any formal analysis I consider what can be expected when there are different relationships between two or more buckets. Let's take two buckets, each depending on the systemic factor and some other factors.[7] The bigger the sum of squares of the coefficients is, the bigger is the probability of default *within the buckets.* At the same time, the more factors are common across the buckets and the higher the related coefficients are, the bigger the probability of default *in the portfolio of the two buckets.* This is because a stronger relation between asset-values across buckets leads to increased correlation of defaults between the buckets, so losses tend to occur simultaneously. For demonstration, I simulated 3 buckets, each containing 1000 exposures of equal size and repeated 20000 times. I used the following model:

$$b_1 : R_1 = 0.36X + 0.2v_1 + 0.4v_2 + 0.82\varepsilon_1$$
$$b_2 : R_2 = 0.36X + 0.2v_1 + 0.4v_2 + 0.82\varepsilon_2$$
$$b_3 : R_3 = 0.36X + 0.45v_1 + 0.82\varepsilon_3$$

The unconditional PD in each bucket equals 5% and the individual 99.9th percentile loss equals 56% (i.e. the coefficient of the 'superfactor' is the same for the buckets). However, if we create a portfolio from b_1 and b_2 the portfolio 99.9th percentile of the loss rate equals 54% (the difference compared to the theoretical value must be due to limited sample size), whereas the portfolio of b1 and b3 has around 48% as the corresponding percentile (this, again, is subject to some uncertainty due to the limited sample size). This is because the correlation between b_1 and b_3 is smaller then between b_1 and b_2.

3.2.1 A framework for fast simulation of portfolio returns with arbitrary (number of) buckets

In order to obtain the portfolio loss distribution (and, at the end, its 99.9th percentile) we need to know, at least, the (marginal) distribution of bucket losses and their dependency structure. This latter can be, for example, the correlation of losses across buckets. In what follows I first derive the distribution of (individual) bucket returns, then I show how to generate random variables from the distribution, finally I derive the correlation coefficient between two buckets.

[7] Actually, the asset values depend directly on the factors, but since buckets are assumed to be homogenous in terms of dependencies on factors, we may say that the buckets depend on the factors and may think of asset value equations as describing buckets, in a sense.

The probability that an exposure in bucket *i* defaults is (in terms of eq. (2.4)):

$$p = N\left(\frac{\gamma_i - w_i^* y}{\varphi_i}\right),$$

This is the single variable that is of interest for us in the Basel model, since the 99.9th percentile of its unconditional distribution is the basis of the risk weight. The probability distribution of *p* can be derived as follows:

$$\begin{aligned} P(p \le c) &= P\left(N\left(\frac{\gamma_i - w_i^* y}{\varphi_i}\right) \le c\right) \\ &= P\left(\frac{\gamma_i - w_i^* y}{\varphi_i} \le N^{-1}(c)\right) \\ &= P\left(-y \le \frac{N^{-1}(c)\varphi_i - \gamma_i}{w_i^*}\right), \\ &= 1 - P\left(y < \frac{\gamma_i - N^{-1}(c)\varphi_i}{w_i^*}\right) \\ &= 1 - N\left(\frac{\gamma_i - N^{-1}(c)\varphi_i}{w_i^*}\right) \end{aligned} \tag{3.1}$$

where *N(x)* denotes the standard normal distribution function. It can be shown that the density has the form:

$$\begin{aligned} P'(p) &= N\left(\frac{\gamma_i - \varphi_i N^{-1}(c)}{w_i^*}\right)\frac{\varphi_i}{w_i^*}\frac{1}{n\left(N^{-1}(c)\right)} \\ &= n\left(\frac{\sqrt{\gamma_i^2 - w_i^{*2}}\,N^{-1}(c) - \frac{\gamma_i}{\sqrt{\gamma_i^2 - w_i^{*2}}}\gamma_i}{w_i^*}\right)\sqrt{2\pi}\, e^{\frac{\gamma_i^2}{\gamma_i^2 - w_i^{*2}}}, \end{aligned} \tag{3.2}$$

where *n(x)* denotes the standard normal density.

(3.1) provides for a straightforward way to generate random variables from the bucket loss distribution: we generate random variables from the uniform distribution and substitute them into the inverse of the distribution function in the place of p_c:

$$\begin{aligned} p_c &= P(p \le c) = 1 - N\left(\frac{\gamma_i - N^{-1}(c)\varphi_i}{w_i^*}\right) \Rightarrow \\ c &= N\left(\frac{\gamma_i - N^{-1}(1 - p_c)w_i^*}{\varphi_i}\right). \end{aligned}$$

Finally, I show how to derive the correlation of bucket losses. Formally, the correlation can be expressed as:

$$\rho_{ij} = \frac{E\left(N\left(\frac{\gamma_i - w_i^* y_i}{\varphi_i}\right)N\left(\frac{\gamma_j - w_j^* y_j}{\varphi_j}\right)\right) - E\left(N\left(\frac{\gamma_i - w_i^* y_i}{\varphi_i}\right)\right)E\left(N\left(\frac{\gamma_j - w_j^* y_j}{\varphi_j}\right)\right)}{\sqrt{\operatorname{var}\left(N\left(\frac{\gamma_i - w_i^* y_i}{\varphi_i}\right)\right)\operatorname{var}\left(N\left(\frac{\gamma_j - w_j^* y_j}{\varphi_j}\right)\right)}} \tag{3.3}$$

The first term in the numerator, the expected value of the product of loss rates, can be calculated as follows:

$$\begin{aligned} E_{ij} &= E\left(N\left(\frac{\gamma_i - w_i^* y_i}{\varphi_i}\right)N\left(\frac{\gamma_j - w_j^* y_j}{\varphi_j}\right)\right) \\ &= E\big(E(D_i \mid y)E(D_i \mid y)\big) \\ &= \int\!\!\int D_1(y,\varepsilon) f_\varepsilon \, d\varepsilon \int D_2(y,\nu) f_\nu \, d\nu \, f_y \, dy, \\ &= \int\!\!\int\!\!\int D_1(y,\varepsilon) D_2(y,\nu) f_\nu f_\varepsilon \, d\nu \, d\varepsilon \, f_y \, dy \\ &= E\left(D_i D_j\right) \end{aligned}$$

where *D* denotes the individual exposures' default indicator (*D* equals 1 if the obligor defaults and 0 otherwise), ε is the idiosyncratic shock corresponding to exposure *i*, ν is the same for exposure *j* and *f* is the corresponding variable's probability density function. The probability that two exposures from different buckets simultaneously default is the same as the probability that the asset value behind the exposures reaches the default trigger, γ. Since asset values are jointly normally distributed, with correlation implied by the parameters in (3.1), this probability can be expressed by a two-variable normal distribution:

$$E\left(D_i D_j\right) = P\left(D_i = 1, D_j = 1\right) = N_2\left(\gamma_i, \gamma_j, \rho_{R_i,R_j}\right).$$

The second term in (3.3) is simply the product of the unconditional default probabilities of the buckets, while the denominator can be derived by the same logic as above (and also appeared, for example, in Bank and Lawrenz [2003]):

$$\operatorname{var}\left(N\left(\frac{\gamma_i - w_i^* y_i}{\varphi_i}\right)\right) = N_2\left(\gamma_i, \gamma_i, \rho_{R_i,R_i}\right) - PD_i^2,$$

In this expression the correlation is that of the exposures *within* the i^{th} bucket. Putting the parts together we obtain the inter-bucket correlation as:

$$\rho_{i,j} = \frac{N_2\left(\gamma_i, \gamma_j, \rho_{R_i,R_j}\right) - PD_i PD_j}{\sqrt{\left(N_2\left(\gamma_i, \gamma_i, \rho_{R_i,R_i}\right) - PD_i^2\right)\left(N_2\left(\gamma_j, \gamma_j, \rho_{R_j,R_j}\right) - PD_j^2\right)}} \tag{3.4}$$

3.2.2 Simulating portfolio losses with given bucket loss marginals and correlations[8]

The method I use is referred to as NORTA (normal-to-anything) and is described in, for example, Cario and Nelson [1997]. The method consists of generating multivariate standard normal variables with the desired correlation matrix, turning these variables into probabilities with the cumulative distribution function and turning the resulting probabilities into variables using the desired inverse distribution function. More formally, denoting the bucket loss distribution of bucket *i* by B_i and by *n* the number of buckets, the steps are:

[8] The Matlab code can be found in Appendix 2.

1. Generate $\{X_1,...,X_n\} \sim N_n(0,1,\Sigma)$,

2. Create $\{U_1,...,U_n\} = N_n(\{X_i\}_{i=1,...n})$,

3. Create the variables with marginal $B_i : \{b_i\} = \{B_i^{-1}(U_i)\}$.

In fact, as Cario and Nelson [1997] emphasizes, it cannot be ensured that the variables will have exactly the desired correlation structure. However, in most cases a very close approximation is achieved and, moreover, it is possible to adjust the correlation matrix of the original standard normal variables to improve the approximation.

Having simulated the correlated bucket returns it is easy to find the desired percentile of the distribution. In the rest of this subsection I show some examples.[9] The examples will differ in the number of buckets, the number of (systemic) factors and in the parameters. Denoting the ith example by Ei the following table summarizes the specifics of each example:

	Number of buckets	Number of factors in the buckets	Unconditional PD in the bucket (%)	Factor coefficients	Conditional PD (%)
E_1	1	1	10	w=0.348	p=41.24
E_2	3	1	{5, 10, 15}	w={0.36, 0.348, 0.347}	p={28.5, 41.24, 51.5}
E_3	3	2	{5, 10, 15}	w={0.36, 0.348, 0.347} δ_1={**0.6**, 0, 0}	p={**76.6**, 41.24, 51.5}
E_4	3	3	{5, 10, 15}	w={0.36, **0**, 0.347} δ_1={0.6, 0, 0} δ_2={0, **0.348**, 0}	p={76.6, 41.24, 51.5}
E_5	3	3	{5, 10, 15}	w={0.36, 0, **0.283**} δ_1={0.6, 0, 0} δ_2={0, 0.348, **0.3**}	p={76.6, 41.24, 51.5}
E_6	3	3	{5, **5**, 15}	w={0.36, **0.36**, 0.283} δ_1={0.6, 0, 0} δ_2={0, **0**, 0.3}	p={76.6, 28.5, 51.5}

In each example the weights of the buckets are the same. In the first example we simply expect to obtain the Basle conditional PD in the last column as a result of the simulation; in the second the simulation should result in the average of the conditional PDs. In E_3 the losses in the first bucket should increase and this should lead to an increased portfolio loss, as well. In E_4 the portfolio loss (the 99.9th percentile) should be smaller than in E_3 because the introduction of the second secondary factor decreases the correlation between the second bucket and the others. In E_5 the portfolio loss should be bigger than in E_4 but smaller than in E_3 because although the correlation between the second and the third buckets increases, correlation between the third and the others decreases. Finally, in E_6 the decrease in the unconditional PD of the second bucket points toward a decrease in the portfolio loss, but, at the same time, the increase in the correlation with the first bucket increases the loss.

The next table shows the 99.9th percentile of portfolio losses in each example:

	E_1	E_2	E_3	E_4	E_5	E_6
Empirical	41.35	40.56	48	41	43	44
Theoretical	41.24	40.38				

It can be seen from the table that relatively small values of coefficients don't change the results significantly, while the coefficient equaling 0.6 of the first secondary factor in bucket 1 in examples E_3-E_5 is high enough to increase the loss substantially. It is interesting, that in E_4 and E_5 the loss falls back to around 42%: this is due to the decreased correlation

[9] The simulations were carried out in Matlab. With my own code, I generated 1 000 000 portfolios for each example; the running time was around 6 seconds per example.

with the first bucket of the other buckets – if we set the coefficient of the systemic factor in the first bucket to zero and assign its previous value to the third factor we, again, increase the correlation of the first bucket with the others thereby increasing the loss. Decreasing the unconditional PD of the second bucket is not enough to compensate for the effect of the increased correlation with the loss of the first bucket.

***

It is worth summing up the simulation procedure and highlighting its key features that make it very fast. As a first step I create homogenous sub-portfolios (these can be buckets or smaller units within buckets). These consist of exposures that have the same asset-value equation. Next, I assume that the sub-portfolios are perfectly fine-grained, i.e. the loss formula for a single exposure conditional on the systemic and secondary factors applies to the portfolio.[10] This enables the use of equation (2.4), the 'superfactor' and its coefficient, by which I turn the multifactor problem to a one-factor problem for a single bucket. In the next step I calculate inter-bucket correlations using the coefficients in (2.3), cf. (3.4) – so in this case I can't use the 'superfactor' – and using the NORTA method I simulate directly from the bucket distributions. Through this procedure I avoid the need to simulate individual exposures (which is rather time-consuming – instead, I only have to simulate as many random variables as many sub-portfolios there are) and, still, can simulate correlated sub-portfolios. Simulating 1000000 (a million) outcomes for 6 buckets with 6 factors (systemic and secondary) took 15 seconds.

3.2.3 Granularity revisited

As already mentioned in the Introduction, granularity in this paper is not adjusted through the number of exposures in a given bucket or in the portfolio. Rather, I divide buckets into sub-portfolios of similar properties (parameter values) and for the simulation I always assume these sub-portfolios are perfectly fine-grained so that I can use the theoretical value of the distribution of losses for that given sub-portfolio (and do not need simulation). I can also calculate the correlation between sub-portfolios analytically (this is an important input to the NORTA simulation). At the end I only have to simulate a lot of outcomes and avoid the simulation of individual defaults in the sub-portfolios.

For example, considering one bucket with one concentration I consider the concentrated exposures as a separate sub-portfolio, calculate the conditional default probability analytically, simulate the factors (the systemic and the other(s) that cause concentration) and put the two sub-portfolios together to arrive at the distribution of bucket loss. The same procedure applies when there are more buckets even if concentrations *across which* are present.

3.2.4 The application of the model in practice

An important question is how such a multifactor model could work in practice. It is essential that the concept of sensitivities, factors and buckets in the multifactor model be clear so that the application of the model becomes straightforward.

I think that the best way is to regard different – though correlated – risk sources as separate factors. This means that region A and region B are separate factors; similarly, sector 1 and sector 2 are separate factors. On the other hand, there should be as many buckets as many combinations of factor-sensitivities are possible. Of course, it is impossible to handle factor-sensitivities on a continuous scale, so these could be assigned to the closest one-tenth between 0 and 1 (for example scale for the sensitivity to region A could be [0 0.1 0.3 0.5]).

With the above structure, the problem with the model is that the number of factors and buckets can grow very high. If we have a systemic factor with 7 possible sensitivities, three regional and three sectoral factors with two possible sensitivities for each then we have $7*2^6=448$ possible buckets. It can be argued though that in practice most of these buckets would not be used.

[10] Assuming that the sub-portfolios are perfectly granular, in my view, is reasonable since if there are really dominant exposures these can be taken out to separate sub-portfolios with 'superfactor' coefficient set at maximum (or, alternatively, one can regard the idiosyncratic shock as a secondary factor in this case).

4. Single name exposures

In the Introduction I suggested to separate the problem of granularity and idiosyncratic correlations. In fact, they are not very distinct concepts and if we seek a general model to handle different types of concentrations, we should clear their relationship. Granularity means (irrespective of how many factors we have) that all the exposures are sufficiently small so that their idiosyncratic part does not affect materially the properties of the whole portfolio. Consequently, the lack of granularity means that there is at least one 'too large' exposure in the portfolio. These exposures can be viewed as a group of smaller exposures which have very strong correlations with each other (because of regional or sectoral factors, for example); alternatively, they can be viewed as single name risks.

The problem of granularity is not related to the number of factors or the correlation between exposures in general (single name exposures and exposures to other factors) but to the exposure size behind these. One might say that the portfolio is not granular enough if there is a big exposure towards a single obligor; and also when the total exposure to obligors in a given region or sector is very big. On the other hand, even if there is correlation between some exposures in the portfolio it can be regarded as granular if these exposures, together, make up only a small portion of the total portfolio.

In the Basel regulation the smallest extent of an exposure towards a client or group of connected clients which, by definition, constitutes a 'large exposure' is 10% of the bank's own funds.[11] Moreover, no single large exposures can exceed 25% of own funds and the sum of the size of all the large exposures must be below 800% of own funds. For other types of concentrations the present regulation does not give such exact, quantitative rules.

As already discussed in Section 2.1, large exposures result in the portfolio failing to comply with the granularity condition. As discussed in, for example, Gordy [2002] and Bank and Lawrenz [2003] this results in the increased unconditional volatility of portfolio losses and in the fact that the q^{th} percentile of the distribution of the systemic factor won't correspond any more to the q^{th} quantile of the unconditional loss distribution. That is, by calculating the conditional default probability that belongs to the q^{th} quantile of the distribution of the systemic factor (in essence, the risk weight) we underestimate the true loss at the q^{th} quantile.

In the papers referred to above the emphasis is on the (unconditional) volatility of the loss distribution. This is useful, especially when the aim is to calculate granularity adjustment (as in Gordy [2002]) which requires this volatility. However, it doesn't tell much about the percentiles of the loss distribution – whereas this is what we are interested in. In what follows I show how I calculated such percentiles: first, assuming no secondary factors (the Basel one-factor model) and second, in Section 5, within the multifactor framework. The method relies on numerical approximations, but it works very fast and flexibly.

4.1 SINGLE NAME EXPOSURES IN THE ONE-FACTOR MODEL

Let's assume that there are no large exposures in the portfolio, so δ – representing the proportion of granular exposures – equals 1. In this case the Basel conditional default probabilities are correct and given by (2.1). Now let's assume that there are only one exposure in the portfolio, i.e. δ=0. This case can be referred to as a 'perfect' concentration. In this case, if the default probability is larger than the percentile at which we want to evaluate the conditional PD, the q^{th} percentile of the unconditional loss distribution will be one.[12] Next, let's assume that there are some large exposures (δ is neither 1 nor 0) *within a given bucket.* The question is, how the 'no-large-exposures' conditional default probability changes as we decrease δ. When we have one large exposure that has a proportion of (1-δ) percent then the loss on the portfolio will be as follows:

---

[11] In this section I use 'large exposures', 'single name exposures' and 'single name risks' interchangeably, although 'large exposures' generally (and in other parts of this study) include any other type of concentrations.

[12] If the unconditional default probability is 0.1 then the portfolio loss will be 100% in 10% of the cases and 0% in 90% of the cases, so the 0.1th percentile will be 100%.

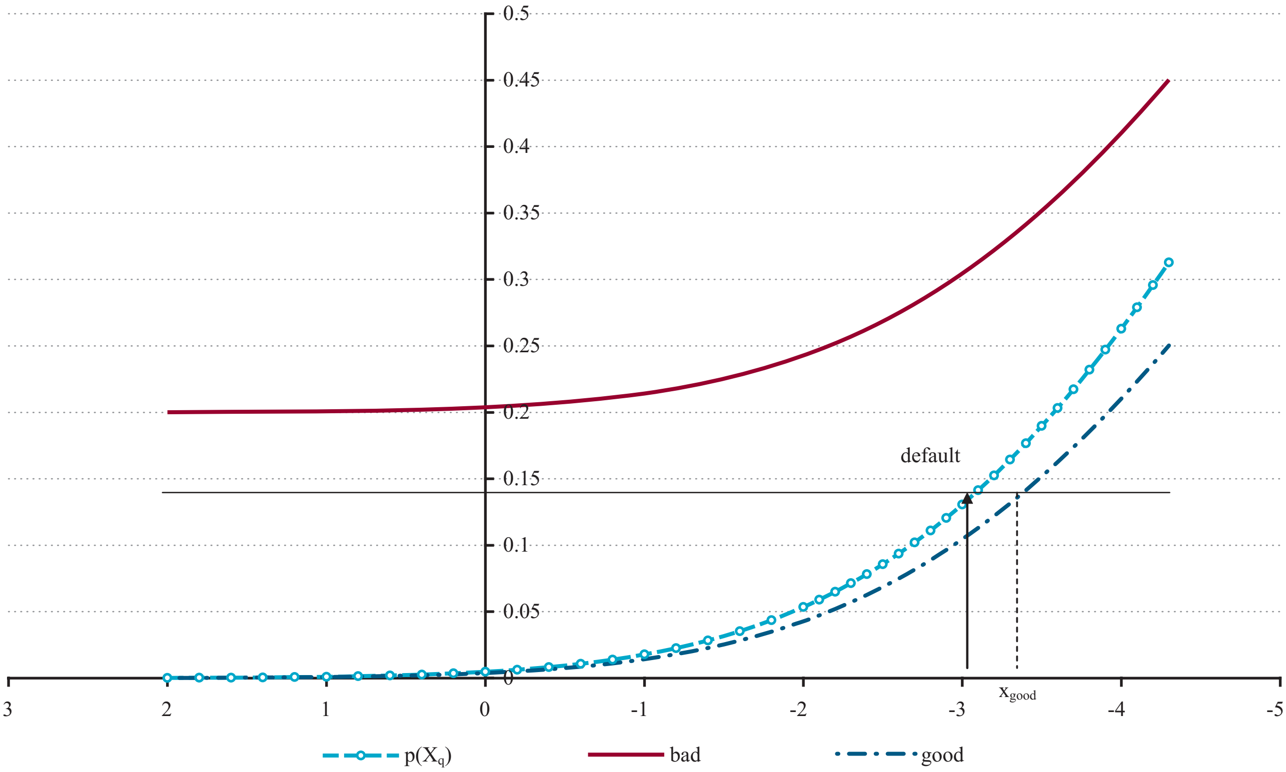
$L(x)=\delta p(x)+(1-\delta)$, with probability $p(x)$ – referred to as the 'bad' outcome;

$L(x)=\delta p(x)$, with probability $1-p(x)$ – referred to as the 'good' outcome, where $p(x)$ denotes the conditional default probability at x.

The portfolio loss has two components: one comes from the granular part (its proportion is δ) and it always equals to the conditional default rate ($p(x)$ – hence the product on the right hand side); the other comes from the large exposure part (with a proportion of $1-\delta$) and it equals to either 1 (with probability equal to the conditional default probability; the 'bad' outcome) or 0 (with one minus the conditional default probability; the 'good' outcome). It can be seen from the equations that for $x=x_q$ the loss of the portfolio with $1-\delta$ proportion of large exposure will be bigger than $p(x_q)$ with probability $p(x_q)$ and will be smaller with probability $1-p(x_q)$. In order to find out which quantile of the unconditional loss distribution of a portfolio with $1-\delta$ proportion of large exposures does $p(x_q)$ represent we have to examine outcomes of the above loss distribution for *all* values of x and not only for $x=x_q$. The next figure shows more clearly the task:

Figure 2

The possible outcomes of portfolio losses with and without the presence of large exposure, as a function of the value of the systemic factor



Practically, we must 'count' all bad and good cases above the default line. In the example above the bad outcomes are above the default line for all x; the good outcomes are above the line for $x>x_{good}$.

Counting, in this context, means integrating, in fact, where for all values of x for which the bad and/or the good outcome is above $p(x_q)$ we have to calculate the probability of that situation occurring. In the above example, when we have one large exposure, this can be written formally as follows:

$$p_\delta(x_k) = \int_A f(x)p(x)dx + \int_B f(x)(1-p(x))dx$$
$$A: \{x \mid \delta p(x) + 1 - \delta \geq p(x_k)\}$$
$$B: \{x \mid \delta p(x) \geq p(x_k)\}$$

where x_k is the kth quantile of x.

Given the above formula we can find the qth quantile of the unconditional loss distribution with a large exposure by adjusting x_k until $pd(x_k)$ equals q.

Next, I derive the formula for calculating the qth quantile of the unconditional loss distribution when there are an arbitrary number large of exposures. I will apply one restriction: the proportion is the same for all large exposures, i.e. equals $(1–\delta)/n$, where n is the number of large exposures.

With similar logic as in the one large exposure case we can write the loss of the portfolio as follows:

$$L(x) = \delta p(x) + n\frac{(1-\delta)}{n},$$ with probability $p(x)^n$, (4.1a)

$$L(x) = \delta p(x) + (n-1)\frac{(1-\delta)}{n},$$ with probability $p(x)^{n-1}(1–p(x))$, (4.1b)

$$L(x) = \delta p(x) + (n-2)\frac{(1-\delta)}{n},$$ with probability $p(x)^{n-2}(1–p(x))^2$, (4.1c)

…

$$L(x) = \delta p(x),$$ with probability $(1–p(x))^n$. (4.1d)

One additional point has to be considered: m out of n exposures can default in $\binom{n}{m}$ different combinations, so we have to multiply the probabilities with the number of these combinations (this procedure only applies when the proportion of different large exposures equal!). The appropriate formula for the calculation of the quantiles of the unconditional loss distribution, when the proportion of the large exposures together equals $(1–\delta)$, the share of one large exposure is $(1–\delta)/n$ and there are n large exposures, is thus:

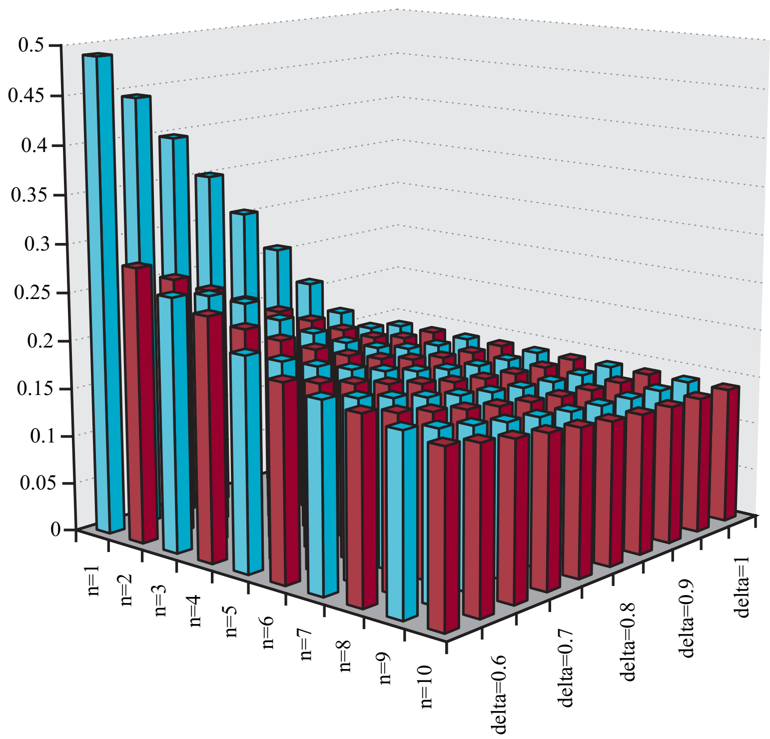
$$p_{\delta,n}(x_k) = \sum_{i=0}^{n}\binom{n}{i}\int_{A_i} f(x)p(x)^i(1-p(x))^{n-i}dx$$
$$A_i: \left\{x \mid \delta p(x) + i\frac{1-\delta}{n} \geq p(x_k)\right\}$$ (4.2)

The above formula can be calculated by numerical integration. As an example, I examine how the 99.9th percentile of the unconditional loss distribution changes when I change δ between 1 and 0.55 and n between 1 and 10. For the calculations I wrote a Matlab code which can be found in Appendix 3. The next figure shows unconditional PDs at the 99.9th percentile for the different values of δ and n:[13]

[13] In the examples in this section the unconditional PD equals 1%. This implies that w=0.439, γ=-2.326, $p_0(x)$=14%.

Figure 3

99.9th percentile portfolio losses with different number (n) and proportion (1-delta) of large exposures, PD=1%

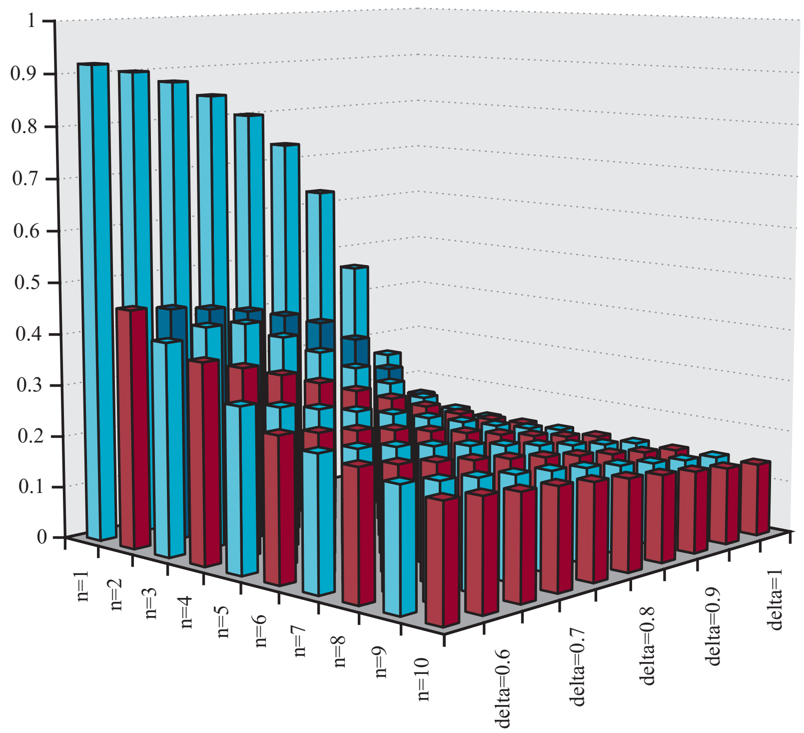


The results of the calculation are intuitive: as we increase the number of large exposures (while holding their overall proportion fixed) the probability of default decreases – this reflects the effect of diversification; and if we increase the proportion of large exposures the probability of default increases. It is very interesting that even two large exposures (with identical shares in the portfolio) can decrease the risk of the portfolio substantially as compared to only one large exposure (here the two exposures are assumed to be uncorrelated beyond the effect of the systemic factor).

The next step, after assessing the effect of large exposures on the q^{th} quantile of the unconditional loss distribution, is to create a rule according to which the portfolio loss should be apportioned between the exposures in the portfolio. I now propose a rule where the idea is to assign the increase in portfolio losses exclusively to the source of this increase – the large exposures – and leave the risk weight of the rest of the exposures unchanged. The next figure shows the risk weights that should be assigned to large exposures in such a regime:

Figure 4

Proposed risk weights for large exposures as a function of their number (n) and proportion (1-delta), PD=1%



The figure shows that, for example, when there is one large exposure having a proportion of 10 percent its risk weight should be about 46%!

A further important point to consider (and a further way to extend the model) is that large exposures don't fit into a 'bucketing framework': they should be examined from a portfolio-wide perspective. The reason for it is that just as the granular part of a bucket damps the effect of large exposures, other buckets' granular part damps, as well. So far in the paper, when I referred to portfolios I meant buckets – now we move to the aggregate portfolio level including all the buckets.

The above method can handle more buckets very easily (still remaining within the one-factor framework): when calculating the probabilities with which the different losses can occur *(L(x)* in (4.1(a)-(d)) we simply have to use the unconditional default probability that belongs to the single name exposure.

That is, instead of $p(x) = N^{-1}\left(\frac{\gamma - wx}{\sqrt{1-w^2}}\right)$ we use $p_{SN}(x) = N^{-1}\left(\frac{\gamma_{SN} - w_{SN}x}{\sqrt{1-w_{SN}^2}}\right)$, where the SN subscript is to

emphasize that the parameters (the unconditional default probability) of the single name exposure are used. As confirmed by calculations, when PD_{SN} (the unconditional default probability of the large exposure) is bigger than the PD (or some kind of average PDs) belonging to the granular part of the portfolio the risk weights increase further. Furthermore, although this complicates the matter further, the case when there are more large exposures situated in different buckets can be handled similarly, within the framework.

5. Large exposures in the multifactor framework

In the previous section I examined the case where there is a granular portfolio part and a large (single name) exposure both depending only on the systemic factor. Now, I substitute the latter one with a group of exposures depending on more (sectoral and/or regional) factors.

In Section 3 I analysed the multifactor model with simulations; in Section 4 I analysed single name exposures with (numerical) integration. When we mix the two features in a portfolio it turns out that (apart from an exceptional case) in the analysis we have to use simulation again. The reason is that integration requires that one uses the same density functions in (4.2), i.e. the same factors should be underlying all the buckets with identical sensitivities across buckets – if this condition fails to hold we can't 'run through' the values of a factor or 'superfactor' in (4.2).

The need to use simulation techniques carries over all of the difficulties that were faced in Section 3 – most notably, the intractability of the analysis when there are a lot of buckets. It is not possible to consider the 'granular' part of the portfolio separately and calculate the loss on the concentrated part independently because the portfolio loss depends on the correlation of the different parts. In the Basel model the loss of the buckets are perfectly correlated (all are determined by the systemic factor exclusively); in the single name exposure case in Section 4 the analysis was made tractable by the fact that both the granular part and the single name exposure part depended on the same (single) risk factor. In the multifactor (and 'multi-bucket') case neither loss correlation is perfect across buckets nor is there only one factor being common across buckets.

5.1 ONE CONCENTRATED GROUP OF EXPOSURES WITH ONE SECONDARY FACTOR

Let's consider first the case where there is a granular part depending only on the systemic factor and another granular part that depends on a secondary factor, as well. Conditional on *x*, the default loss can be written on the two parts, indexed by 'sys' and 'sec' respectively, as:

$$L_{sys} = N\left(\frac{\gamma - wx}{\sqrt{1-w^2}}\right)$$
$$L_{sec} = N\left(\frac{\gamma - wx - \beta v}{\sqrt{1-w^2-\beta^2}}\right) \qquad (5.1)$$

From L_{sec} it can be seen that the multifactor case is 'somewhere between' the one factor granular case and the single name exposure case: if $\beta=0$ then we have the Basel (granular) model; if β is such that the denominator approaches zero then, in effect, we have the single exposure case (the exposures in the bucket become perfectly correlated). For this second case the same treatment applies as presented in the previous section.

When β is between zero and its maximum (which is determined by the requirement that the denominator exists) we are not in such a simple situation as in the previous section that we either have 100% loss or 0% loss depending on the conditional probabilities: now we have different default rate for each value of v. As a consequence, while in the single name risk case we had a 'bad' and a 'good' case (depending on the value of the idiosyncratic shock of the single name exposure) here we have a continuum of cases. This, in turn, induces some complication to the extent that we do not simply have to add two (or in more complicated cases more) loss-outcomes (as in (4.1a-d)), but we have to integrate the losses according to v, for each value of *x*. This can be easily done with only a few factors but otherwise we have difficulties – for the same reason, in Section 3 I used simulation instead of integration.

For the model in (5.1) the portfolio loss is easily found to be:

$$L_{portf} \begin{aligned} &= \int \delta L_{sys} + (1-\delta) \int L_{sec} f(v)\,dv\,f(x)\,dx \\ &= \int \delta N\left(\frac{\gamma - wx}{\sqrt{1-w^2}}\right) + (1-\delta) \int N\left(\frac{\gamma - wx - \beta v}{\sqrt{1-w^2-\beta^2}}\right) f(v)\,dv f(x) dx\,. \end{aligned}$$

It follows that the problem is to find x_k, such that:

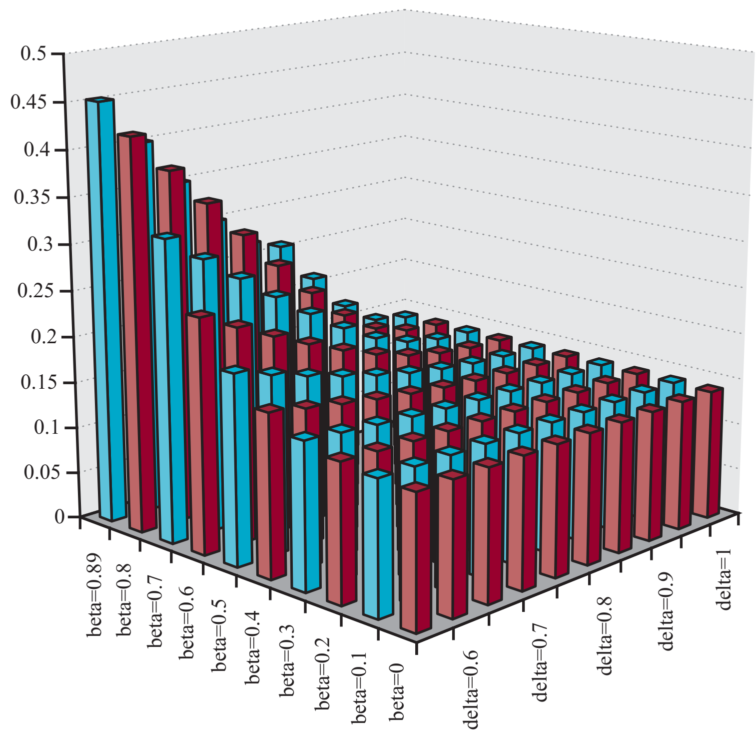
$$p_\delta(x_k) = \int_A f(x) f(v)\,dxdv = 0.001$$

$$A : \left\{ x, v \mid \delta N\left(\frac{\gamma - wx}{\sqrt{1-w^2}}\right) + (1-\delta) N\left(\frac{\gamma - wx - \beta v}{\sqrt{1-w^2-\beta^2}}\right) \geq N\left(\frac{\gamma - wx_k}{\sqrt{1-w^2}}\right) \right\}. \qquad (5.2)$$

I carried out the analysis in Matlab, using the 'dblquad' function (the code can be found in Appendix 4). The following figure shows how the 99.9th percentile of the portfolio loss distribution changes if we change the weight of the secondary factor in the bucket it affects:[14]

Figure 5

99.9th percentile portfolio loss with different weight *(1–δ)* and level *(β)* of secondary factor concentration (calculated using numerical integration)



[14] The first value of beta is 0.89, for which the reason is that for higher values the term under the square-root in the denominator in (5.2) would become negative. It is important to note that the calculations are not totally accurate: the loss levels belonging to beta=0.89 should be the same as the levels belonging to n=1 in Figure 3, however at delta=0.55 there is a 5 percentage point difference between the two. The error is purely due to technical reasons.

The results are intuitive: as we sensitivity to the second factor decreases the effect of concentration decreases and the portfolio loss percentile also decreases. At $\beta=0$ it is correctly shown that the loss percentile is independent of δ.

The analysis is similar in the case where there are more buckets depending on the same single secondary factor (with different sensitivities): one simply augments definition of *A* in (5.2) with the additional buckets.

5.2 MORE THAN ONE CONCENTRATIONS IN THE PORTFOLIO

However, computations become intractable with more factors (and buckets) – given such situations one has to use simulation techniques. The one presented in Section 3 consists of the following steps:

1. We create sub-portfolios (defined by exposures with the same model parameters so a bucket can contain more sub-portfolios if there are concentrations); asset value equations for the sub-portfolios can include several factors.

2. Next, we draw correlated random variables from the sub-portfolios using the NORTA technique.

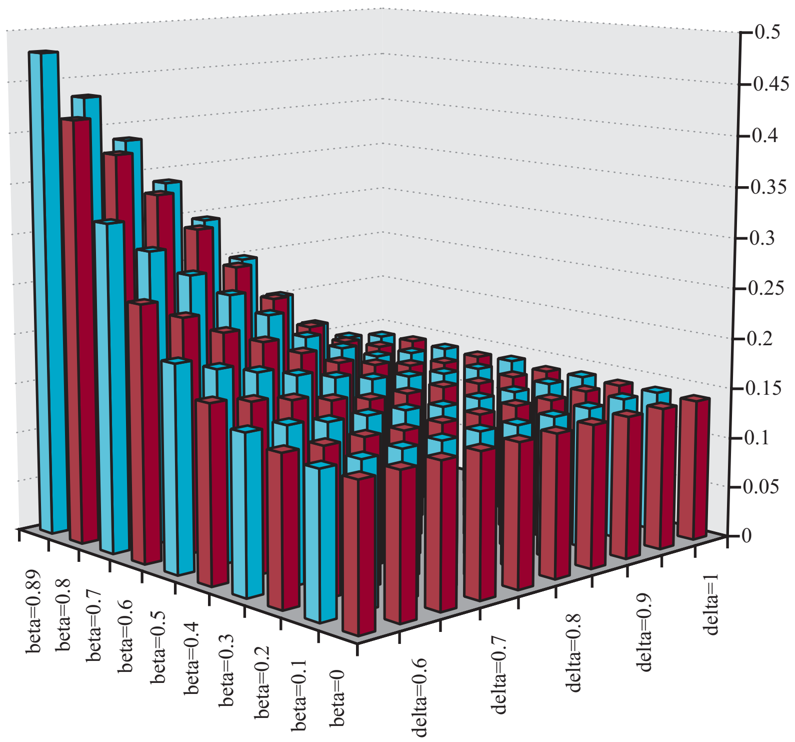
3. Finally, we add up (weight) losses to arrive at the bucket or portfolio loss.

Throughout the calculation we do not need to simulate the factors – these only enter the model through their coefficient and, consequently, the correlation of sub-portfolios. Along with the assumption that the loss distribution of all sub-portfolios can be perfectly characterised by their conditional loss distribution (i.e. we can use 'superfactors' as in (2.4)) this is the other feature of the framework used here that makes calculation rather fast.

As a demonstration of the simulation technique and its equivalence to the numerical integration in the above subsection I simulated the same model as above: there is a granular bucket in the portfolio depending only on the systemic factor and another granular part depending on a secondary factor, as well. I carried out the simulation for different sensitivities to the secondary factor and different weights of the second bucket in the portfolio. The results are shown inFigure 6:

Figure 6

99.9th percentile portfolio loss with different weight *(1–δ)* and level *(β)* of secondary factor concentration (calculated through simulation)



As can be seen, the results are identical to those obtained in the previous subsection. The only difficulty with the simulation process is that when generating correlated bucket losses starting from correlated normal variables it is not ensured that the resulting correlation of the bucket losses will equal to that of the normal variables. In the example above, without any correction to the 'input correlation' of the normal variables, there were as high as more than three-fold differences for certain parameter combinations. Unless one has an analytic solution to the problem these differences can only be handled by an iterative procedure (it took me three iterations).[15]

[15] The simulations revealed that the level of the correction needed is linked to the sensitivity to the secondary factor and is not affected by the weight of the buckets in the portfolio.

6. Discussion

In this section I highlight the most important issues that follow from the above analysis. First, I asses the present regulation in light of the above results and then I question an important property of the one-factor model (that is lost in the multifactor model).

6.1 CURRENT PRACTICE IN LE REGULATION IN THE LIGHT OF SECTION 3

The analysis above raises some interesting points in relation to the current LE regime. Before the discussion of these let's recall the main elements of the present regulation: an exposure to a single client or group of connected clients constitutes a large exposure as soon as its size reaches 10% of the institution's own fund; the maximum size of a single large exposure is allowed at 25% of the own funds; the total size of large exposures is maximised at 800% of the own fund; finally, these rules apply only to single name exposures and not to other types of concentrations.

6.1.1 Large exposures and credit quality

Now, let's examine first the definition of large exposures. This has not been the subject of this paper directly – I took the definition as given – but it has important implication for the risks of portfolios. If there are two institutions with the same balance sheet total (as regards credit granting activity) but with assets of different average level of riskiness then the minimum required level of own funds will differ: the level of own funds at the institution holding the riskier assets will be higher (assuming that both operates at the minimum required level). Consequently, this latter institution will be allowed to take on bigger large exposures and, since in general its assets are riskier, there is a chance that its large exposures will be riskier than those of the other institution. We can draw the conclusion that *it might be desirable to make LE prescriptions sensitive to the riskiness* (i.e. credit rating) *of the large exposures* (and, perhaps, also of the granular part of the portfolio).

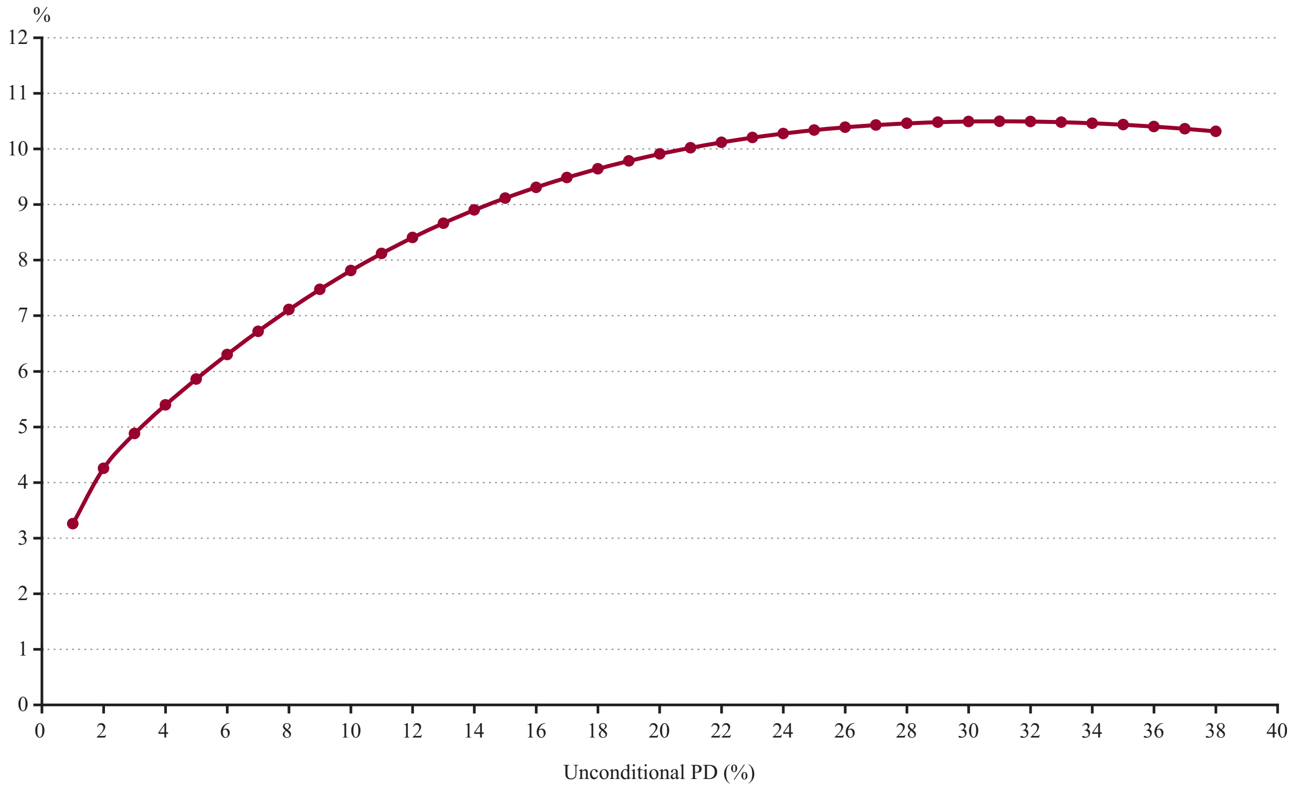
6.1.2 Limits on large exposures

The second point to explore is whether the 25% limit on a single large exposure can be regarded as appropriate. First of all, recall that the percentage limit in the present regulation applies in terms of the own funds of an institution. The level of own funds has to be at least 8% of the risk weighted assets (RWA) which, in turn, is determined in such a way that, at the end, the own funds cover the unexpected losses of the institution. This means that own funds (in the IRB approach) should equal to at least $p_0(x_q)–PD$ (where PD is the unconditional default probability, the expected value of the conditional default probabilities), the 25% of which is the upper limit on one single large exposure. For example, if PD=14.2%, the $p_0(x_{0.001})$=50% and the minimum level of own funds (for credit risk) is 35.8%. As a consequence, the upper limit on a large exposure is 8.95% in terms of the (nominal) value of the exposures! Figure 7 shows the maximum size of a large exposure – in terms of the (nominal) value of all the exposures and given that the institution holds the minimum amount of own funds – as a function of the unconditional PD.

The 25% limit can not be exceeded in most countries and where it can (in Hungary, for example) the part above the limit must be fully covered by own funds. This rule and, in general, the 25% limit are not shown by my calculations to be well-founded (at least, irrespective of the riskiness of the portfolio, cf. subsection 5.1.1). For an unconditional PD of 1% the maximum size of one large exposure is about 3.3% of the total nominal portfolio size (see Figure 7). Having one such large exposure in the portfolio increases the risk of the portfolio marginally (at the same time it is true that this exposure should receive a significantly increased risk weight if it were to be treated within the first Pillar, as described in Section 4.1, see Figure 4: the risk weight would be around 22% instead of 14%).

Figure 7

Maximum size of a large exposure in terms of total nominal assets, as a function of unconditional PD (approximate results)



6.1.3 Maximum LE risk in the present regime

As described in the previous subsection, at present, institutions are not allowed to have any exposures the size of which exceeds 25% of their own funds. In other words, since the minimum level of own funds corresponds to the (estimated) unexpected loss of the portfolio, assuming that institutions hold the minimum level means that they are not allowed to have any exposures in excess of 25% of the unexpected loss. Moreover, such exposures are not allowed to exceed 800% of own funds. This implies that institutions can have 32 large exposures of size 25% of own funds. However, for riskier portfolios the unexpected loss is higher and, consequently, the 25% size is also higher in absolute terms and also relative to nominal assets. Since the total nominal value of large exposures can't exceed the total nominal value of the portfolio the riskier a portfolio is the smaller the possible number of large exposures is (and it can be below 32).

In this subsection I present results demonstrating what maximum levels of risk (as expressed by the 99.9th percentile of the loss distribution) is implied by the present regime. To carry out the simulations I first set the unconditional PD of each bucket (now a bucket is equivalent to a LE); here I assign the same value to each exposure. Five values are considered: 0.1%, 0.5%, 1%, 5% and 10%. Based on these PDs I set the number of large exposures by first dividing the unexpected loss with 4 (so that I consider LEs of maximum allowed size) and then applying the rule that the total nominal size of LEs can't exceed the total nominal size of the portfolio. For example, for PD=10% the unexpected loss is 31.24% leading to a maximum LE size of 7.81% (in terms of total assets). This means that 12 LEs should be considered here (since 12x7.81<100 but 13x7.81>100) and the rest of the portfolio is assumed to be granular and dependant only on the systemic factor.

A very important question, after setting the number of large exposures, is how many secondary factors we have and how the LEs depend on these. The less secondary factors we have, the higher the correlations between bucket losses are and, thus, the higher the 99.9th percentile loss can be expected to be. In the present simulations I applied 9 secondary factors in addition to the systemic factor. Another question is which factors affect which exposure: the more common factors LEs have, the higher their correlation is and the larger unexpected loss can be expected. In this analysis I assumed that secondary factors affect exposures evenly, that is, all coefficients are such that they are equal and the 10 factors (including the systemic one) together fully explain the variance of the loss on the exposures (idiosyncratic shock is virtually zero). This set-up results in perfect correlation across bucket- (LE-) losses.

The following table shows the 99.9^{th} loss percentiles for the 5 different unconditional PDs:

Table 2

Portfolio 99.9^{th} loss percentile as a function of unconditional PD of the buckets

Unconditional PD	10%	5%	1%	0.5%	0.1%
99.9^{th} loss percentile	96.23%	95.45%	95.1%	76%	28.1%

The results show that, in general, the risk of large losses decreases as the unconditional PD decreases. However, in the region where the 800% limit is not binding (above a certain level of unconditional PD) the decrease in the 99.9^{th} percentile is very slow. Where the limit is binding, a decrease in the unconditional PD leads 1. to a more granular portfolio (since the 25% limit leads to smaller LEs) and 2. to LEs with smaller risk – these two effects contribute to the sharp decrease in the 99.9^{th} loss percentile.

6.1.4 Large exposure and client diversification

In the simulation procedure used in Section 3.2.1 I assigned the large exposure to a separate sub-portfolio whose correlation to other sub-portfolios is determined by the asset-value coefficients. Since higher correlation leads to higher variance and, thus, to higher (potential) portfolio loss it can be argued that the regulation of large exposures could depend on the features (sectoral and/or geographical specifics) of the client. For example, if a given geographical region is overrepresented in the portfolio a large exposure from another geographical region might even decrease the riskiness. Of course, in such a case the properties of the current portfolio also matter and this leads us to the question whether we need (and should restrict ourselves to) portfolio invariant rules.

6.1.5 Common denominator: potential loss

The present system applying limits is not really risk sensitive. It is true that a higher level of concentration is riskier, so in this sense limits are sensible; however, the system is quite crude in that a large exposure of 25% is allowed, but 25.5% is not allowed at all. Moreover, concentrations other than single name risks are not addressed quantitatively: while one obligor or group of connected obligors is bound by the limits, sectoral concentration is not limited. These anomalies could be solved if the regulation were concerned with the true risks concentrations cause, i.e. potential losses to the portfolio. This would make even a simple limit system more consistent. For example, limits could depend on the number of large exposures: if there is one large exposure the limit could be higher than with two large exposures (it is easy to calculate at which level does a given percentile of the loss of the portfolio with one large exposure equals to that of a portfolio with, say, two large exposures).

6.1.6 Open issues: Standardized method

From a practical point of view, large exposure regime for exposures under the Standardised method can be expected to be at least as an important issue as for exposures under the IRB method.[16] The reason is that 1, a large proportion of institutions may stick to the Standardised approach and 2, even IRB-bank may use such approach in the framework of partial use and roll-out. This has several implications. First, at least some flexibility could be given to the regulation by, for example, making it sensitive to the number and quality of (single name) large exposures. Even in this case other types of concentrations would not be considered which should also be addressed. A second problem might be the contra-selective nature of the present regulation, i.e. the fact that unrated exposures receive a fixed and a relatively not too high

[16] This argument as well as 6.1.5 was pointed out by Márton Radnai.

risk weight (e.g. credits to unrated central governments receive 100%). Since if a bank introduces the IRB approach it has to apply it to almost all of its exposures this might lead to a situation where 'Standardised' banks will collect bad quality but unrated exposures (exposures are not rated so that in the Standardised method the relatively advantageous regulatory risk weights apply) while 'IRB' banks will collect better quality exposures resulting in relatively high and uncovered risks in Standardised banks' portfolios. Although it is a more general issue, the large exposure regime should reflect this situation.

6.1.7 Open issues: Size and capital

Another issue that can be examined on a 'potential loss basis' and that is also relevant when we consider the treatment of large exposures in the Standardised approach is the relatively high level of capital of smaller banks as a result of the need to cover large exposures. If a smaller bank wants to extend a given (nominal) amount of credit then the amount of own funds that it should hold is a higher percentage of its risk weighted assets compared to a larger bank – according to anecdotal information this is an important reason why smaller banks, on average, operate at a higher solvency ratio (e.g. around 20% or higher). Without knowing the true risk of large exposures (in terms of losses) it is difficult to interpret the capital level of banks in terms of solvency and to calculate and compare the highest percentile of the loss distribution that is still covered.

6.2 DO WE NEED PORTFOLIO-INVARIANT RULES?

It is often emphasised in connection with the Basel model that it leads to 'portfolio-invariant' capital rules. Two points are worth discussing here. First, it is true that the rules are portfolio invariant – but in exchange for the generality of the model we have to make serious restrictions. Two of these restrictions are the central issue of this study: the granularity of portfolios and the one-factor nature of the model's world. So, when applying rules based on such a portfolio-invariant model one must not forget about the restrictions (conditions) behind the model.

Second, it may be questioned whether we need portfolio-invariant rules. To my understanding, this term refers to the property of the rules that if we add an exposure to an existing portfolio the risk weight to this exposure will not depend on the composition of the portfolio. Thus, in the lack of portfolio-invariance a new exposure in a portfolio may change the portfolio risk disproportionately to its Basel risk weight. This is undesirable, because at every significant change in a bank's portfolio the capital should be re-allocated between the exposures in the whole portfolio. However, this would really be a problem if institutions' portfolio would often change significantly – if it is not very typical then it might be possible to assign the increase in the portfolio risk to the new (significant) exposure(s) and calculate risk weights accordingly (as was done in Section 4.1 for a large exposure, see Figure 4). Moreover, as the major concern of regulators, the overall capital requirement might be found through daily evaluations even in cases where the portfolio composition changes rapidly (without the *allocation* of capital to sub-portfolios).

7. Conclusion

In this paper I presented a multifactor approach to calculate the effect of concentrations on portfolio losses and to assign risk weights to exposures in the portfolio. The method is a multifactor extension of the one-factor model currently underlying the Basel regulation.

In the absence of closed form solutions there are basically to ways to calculate portfolio loss percentiles with the model: integration (by the uncorrelated factors) and simulation. The major problem with integration is that it becomes intractable even for a relatively small number of factors; simulation can handle a higher level of 'dimensionality'.

In the study I carried out analysis of both types. In the simulation part I modelled portfolio losses directly and not through the individual exposures. I used bucket-loss distributions to calculate portfolio level losses with the NORTA simulation technique. Results proved to be accurate and intuitive: the key determinants of portfolio losses are the factor-sensitivities of exposures, the correlation matrix of bucket losses (which can be derived from sensitivities), the unconditional default probabilities in the buckets and the number of independent factors underlying the portfolio. Increasing a factor-sensitivity increases the loss of that bucket and the portfolio and – because of the potential increase in the portfolio variance through correlations – it may increase the portfolio loss further. The methods (models and simulations) I used here are flexible and easy to understand. Through their use it is possible not only to give solutions to certain problems (e.g. calculation of risk weights) but to raise important issues (e.g. the inclusion of credit quality in the large exposure regime).

In my opinion there are two problems with the simulation approach. First, it can't handle a large number of buckets. This would require decreasing the 'dimensionality' of the problem by compressing the information on the portfolio distribution. I couldn't find any solutions to it. The second, less important, problem relates to the NORTA technique I used: the correlation of the starting normal variables might change so much in the transformation to the bucket loss variables that results become highly distorted. This problem can be solved by iterating the starting correlation matrix so that we get the desired result.

As for the analysis through integration, I showed a method for fast calculation of the loss percentile when there is one single name exposure in the portfolio. The method then was extended to an arbitrary number of large exposures with the restriction that they are independent, dependant on the same single systemic factor and have the same weight. I also showed how to carry out the calculation for one large exposure when there is an additional factor.

The major drawback of the integration approach is that it can handle only very 'small scale' problems in terms of the number of factors.

Using the techniques above I examined a lot of different scenarios. An important result of this analysis is that even two independent large exposures can decrease the risk compared to only one large exposure (with the size of the two together). Furthermore, calculations suggest that risks posed by concentrations other than single name exposures can be constrained either by limiting the weight of the concentration in the portfolio or by limiting the allowed sensitivity of exposures to the factors underlying such concentrations.

Though this aspect is more important from banks' rather than regulators' perspective, it is worth noting that having calculated the portfolio loss percentile it may be necessary to allocate the losses between sub-portfolios (buckets) in order to calculate how much capital should be allocated to the exposures. I proposed, in the simplest case where there is only one large exposure in the portfolio and a granular part, to assign the increase in the portfolio loss percentile to the large exposure and showed that it leads to a relatively high increase in the capital requirement of the large exposure. Appropriate rules should be found in more complicated cases.

References

BANK, M. AND J. LAWRENZ (2003): "Why simple when it can be difficult?", *Kredit und Kapital,* 36., p. 534-556.

BCBS (2005): "An Explanatory Note on the Basel II IRB Risk Weight Functions", BIS.

CAO, M. AND W. J. MOROKOFF (2004): "Calibrating Credit Portfolio Loss Distributions", Proceedings of the 2004 Winter Simulation Conference, R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.).

CARIO, M. C. AND B. L. NELSON (1997): "Modeling and Generating Random Vectors with Arbitrary Marginal Distributions and Correlation Matrix", *Technical Report,* Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois.

CROUHY, M., D. GALAI, AND R. MARK (2000): "A comparative analysis of current credit risk models", *Journal of Banking and Finance* 24 (2000) p. 59-117.

GORDY, M. (1998): "A Comparative Anatomy of Credit Risk Models", Board of Governors of the Federal Reserve System.

GORDY, M. (2002): "A Risk Factor Model Foundation for Ratings-Based Bank Capital Rules", Board of Governors of the Federal Reserve System.

HAMERLE, A., T. LIEBIG, AND D. RÖSCH (2003): "Credit Risk Factor Modelling and the Basel II IRB Approach", *Deutsche Bundesbank Discussion Paper* No. 02/2003.

HUANG, X., C. W. OOSTERLEE, AND J. A. M. VAN DER WEIDE (2006): "Higher order saddlepoint approximations in the Vasicek Portfolio Credit Loss Model", Delft University of Technology, Report 06-08.

LINDSKOG, F. (2000): "Modeling Dependence with Copulas and Applications to Risk Management", Master Thesis, ETH.

MARTIN, R., T. WILDE (2002): "Unsystematic credit risk", *Risk,* Vol. 15. No. 11. p. 123-128.

MERTON, R. C. (1974): "On the Pricing of Corporate Debt: the Risk Structure of Interest Rates", *Journal of Finance* 29 (2) p. 449-470.

PYKHTIN, M. (2004): "Multi-factor adjustment", *Risk,* March 2004, p. 85-90.

Appendix 1: The multifactor model and the equivalent "superfactor model"

Let's assume there are three factors denoted by X, F_1 and F_2. The factors are correlated in the following way. X is a standard normal random variable; F_1 has the form:

$$F_1 = \alpha_1 X + \beta_1 \varepsilon_1 ,$$

where $\beta_1 = \sqrt{1-\alpha_1^2}$, i.e. F_1 is again standard normal; finally, F_2 has the form:

$$F_2 = \alpha_2 X + \beta_2 \varepsilon_1 + \gamma_2 \varepsilon_2 ,$$

$\gamma_2 = \sqrt{1-\alpha_2^2-\beta_2^2}$, and F_2 is also standard normal. ε_1 and ε_2 are standard normal variables and are uncorrelated.

Using the three factors the asset value change of an obligor can be written as:

$$R_i = aX + bF_1 + cF_2 + d\eta ,$$

where η is the idiosyncratic shock and d is its standard deviation.

It can be easily shown that R_i can be expressed with the uncorrelated factors, X, ε_1 and ε_2 as:

$$R_i = \left(a + b\alpha_1 + c\alpha_2\right)X + \left(b\beta_1 + c\beta_2\right)\varepsilon_1 + \left(b\gamma_1 + c\gamma_2\right)\varepsilon_2 + d\eta . \qquad \text{(A1.1)}$$

To assure that R_i is standard normal, d should equal

$$\sqrt{1-\left(a + b\alpha_1 + c\alpha_2\right)^2 - \left(b\sqrt{1-\alpha_1^2} + c\beta_2\right)^2 - \left(c^2(1-\alpha_2^2-\beta_2^2)\right)} .$$

The last equation is equivalent (in its distributional properties) to:

$$R_i = w^* Y + d\eta , \qquad \text{(A1.2)}$$

where

$$w^* = \sqrt{\left(a + b\alpha_1 + c\alpha_2\right)^2 + \left(b\sqrt{1-\alpha_1^2} + c\beta_2\right)^2 + \left(c^2(1-\alpha_2^2-\beta_2^2)\right)} .$$

Extending the model with more factors is straightforward. Factor i has the form:

$$F_i = \alpha_i X + \beta_i \varepsilon_1 + \ldots + \delta_i \varepsilon_i + \sqrt{1-\alpha_i^2-\beta_i^2-\ldots\delta_i^2}\chi ,$$

where ε_i is the ith independent factor, δ_i is F_i's sensitivity to this factor, and χ is F_i's idiosyncratic component.[17]

The identification of the model can be done using correlations and asset-value sensitivities to the F factors (not the independent factors): for example to obtain the coefficients in (A1.1) we need the pair-wise correlations of X, F_1 and F_2, and R_i's sensitivity to these factors (a, b and c, respectively).

[17] The model could also be formed in such a way that each independent factor influences each F factor. However, in this case the model would be impossible to identify since there would be more unknown parameters than correlations (which can be either measured or assumed to be at a given level).

Appendix 2: The Matlab code for the calculation of quantile of the loss distribution using simulation of portfolio loss for two buckets and one factor, the systemic factor

```
clear all;

estim=0;    %1 if we want to estimate model parameters from a sample

n_bct=2;            %number of buckets
pd(1,1)=0.01;       %unconditional PD for exposures in the buckets
pd(2,1)=0.01;

gamma=norminv(pd); %default trigger for exposures in the buckets
w=(0.12*(1-exp(-50*pd))/(1-exp(-50))+0.24*(1-(1-exp(-50*pd)))/(1-exp(-50))).^0.5;     %sesitivity to the systemic factor
fi=(1-sumw).^0.5;  %coefficient of the idisyncratic shock
p=normcdf((gamma-sumw.^0.5*norminv(0.001,0,1))./fi,0,1); %conditional PD

sumw=sum(w.^2,2);         %square of the coefficient of the 'superfactor' if there are more factors

b_korrekt(1,1)=2.6;     %correlation correction for NORTA simulation
b_korrekt(2,1)=2;       %..
b_korrekt(3,1)=1.7;     %since the number of buckets is 2 the rest is not needed in this case
b_korrekt(4,1)=1.5;
b_korrekt(5,1)=1.35;
b_korrekt(6,1)=1.2;
b_korrekt(7,1)=1.1;
b_korrekt(8,1)=1.05;
b_korrekt(9,1)=1;

for k=1:n_bct     %—- calculating the bucket covariance matrix ——-
   korr_k=[1 sum(w(k,:).^2);sum(w(k,:).^2) 1] ;

for j=1:n_bct
   korr_kj=[1 sum(w(j,:).*w(k,:));sum(w(j,:).*w(k,:)) 1];
   korr_j=[1 sum(w(j,:).^2);sum(w(j,:).^2) 1];
   bucket_corr(k,j)=(mvncdf([gamma(j,1) gamma(k,1)],0,korr_kj)-pd(j,1)*pd(k,1));
   bucket_corr(k,j)=bucket_corr(k,j)/(mvncdf([gamma(j,1) gamma(j,1)],0,korr_j)-pd(j,1)*pd(j,1))^0.5;
   bucket_corr(k,j)=bucket_corr(k,j)/(mvncdf([gamma(k,1) gamma(k,1)],0,korr_k)-pd(k,1)*pd(k,1))^0.5;
   bucket_covar(k,j)=(mvncdf([gamma(j,1) gamma(k,1)],0,korr_kj)-pd(j,1)*pd(k,1));
end
end           %————————————————————————————

b_corr=bucket_corr;

b_corr(1,2)=b_corr(1,2)*b_korrekt(2,1);    %applying correlation correction for NORTA
b_corr(2,1)=b_corr(1,2);
```

```
smpl_size=1000000;  %——- simulating correlated bucket losses ———

   x_sim=mvnrnd(zeros([n_bct 1]),b_corr,smpl_size);
   u=normcdf(x_sim,0,1);
for i=1:n_bct
   bucket(:,i)=normcdf((gamma(i,1)-norminv(1-u(:,i),0,1)*sum(w(i,:).^2)^0.5)/fi(i,1),0,1);
end     %————————————————————————————————

b_corr_emp=corrcoef(bucket);
bucket_p=mean(bucket,2);    %calculating portfolio loss (here assuming equal, 50% weights)
p_999=prctile(bucket_p,99.9);   %taking out the 99.9th percentile

if estim>0 %———————— ML estimation of parameters ——————
pca_g=@(y) norminv(y,0,1);

pca_fi=@(y) (1-y.^2).^0.5;
maxli=@(x,y)                                                  log(normpdf((pca_g(y(1))-
pca_fi(y(2))*norminv(x,0,1))./y(2),0,1).*pca_fi(y(2))./(y(2)*normpdf(norminv(x,0,1),0,1)));
MXL=@(z) -sum(maxli(bucket_p,z));
pd_estim =fminsearch(MXL,[0.1 0.1]);

end %————————————————————————————————————
```

Appendix 3: The Matlab code for the calculation of quantile of the loss distribution in the presence of large exposures

```
%This function have to be made equal to the desired probability level (here
%0.1%) as a function of 'xq'

function y = px_le_f(xq,pd,delta,n)

gamma=norminv(pd,0,1); %———- setting the model parameters ———
w=(0.12*(1-exp(-50*pd))/(1-exp(-50))+0.24*(1-(1-exp(-50*pd)))/(1-exp(-50)))^0.5;
pxq=normcdf((gamma-w*xq)/(1-w^2)^0.5);

pd_le=0.01;
gamma_le=norminv(pd_le,0,1);
w_le=(0.12*(1-exp(-50*pd_le))/(1-exp(-50))+0.24*(1-(1-exp(-50*pd_le)))/(1-exp(-50)))^0.5;
%————————————————————————————————

up=10;
down=-30;

for k=1:n
delt(k,1)=(1-delta)/n;
end
cs=cumsum(delt);

for k=1:n %————- determining the domain of integration ————
   if (pxq-cs(n+1-k))/delta<1  %if this condition fails, 'BAD' outcomes are never below the loss of the granular case
     x_(k,1)=(gamma_le-(1-w_le^2)^0.5*norminv((pxq-cs(n+1-k))/delta))/w_le;   %finding 'x' where the 'BAD' outcome crosses the loss of the grnular case
   if isnan(x_(k,1))   %if no crossing, the the 'BAD' outcome is always above the loss of the granular case
   x_(k,1)=up;
   end
   end
end
   if pxq/delta<1  %the same as above, only for the 'GOOD' outcome
   x_(n+1,1)=(gamma_le-(1-w_le^2)^0.5*norminv(pxq/delta))/w_le;
   if isnan(x_(n+1,1))
   x_(n+1,1)=up;
   end
   end     %————————————————————————————

for k=1:n   %integration over the domain determined above for the 'BAD' outcomes
   if (pxq-cs(n+1-k))/delta>1
   p2x_(k,1)=0;
   else
  p2x_(k,1)=quadl(@(x)normpdf(x,0,1).*normcdf((gamma_le-w_le*x)/(1-w_le^2)^0.5).^(n+1-k).*(1-normcdf((gamma_le-w_le*x)/(1-w_le^2)^0.5)).^(k-1),down,x_(k,1),0.000000001);
   end
```

```
end
    if pxq/delta>1   %integration over the domain determined above for the 'GOOD' outcomes
    p2x_(n+1,1)=0;
    else
        p2x_(n+1,1)=quadl(@(x)normpdf(x,0,1).*(1-normcdf((gamma_le-w_le*x)/(1-
w_le^2)^0.5)).^n,down,x_(n+1,1),0.000000001);
    end

for k=1:n+1
    poli(1,k)=factorial(n)/(factorial(k-1)*factorial(n+1-k));
end

y=poli(1,1:n+1)*p2x_(1:n+1,1); %summing the integrals to get the probability at Xq.
```

Appendix 4: The Matlab code for the calculation of quantile of the loss distribution in the presence of a secondary factor

```
function y = px_le_f(xq,pd,delta,d)

gamma=norminv(pd,0,1);
w=(0.12*(1-exp(-50*pd))/(1-exp(-50))+0.24*(1-(1-exp(-50*pd)))/(1-exp(-50)))^0.5;
pxq=normcdf((gamma-w*xq)/(1-w^2)^0.5);
fi=(1-w^2)^0.5;

pd_le=0.01;
gamma_le=norminv(pd_le,0,1);
w_le=(0.12*(1-exp(-50*pd_le))/(1-exp(-50))+0.24*(1-(1-exp(-50*pd_le)))/(1-exp(-50)))^0.5;
delta_le=d;
fi_le=(1-w^2-delta_le^2)^0.5;

meet=@(x,v)  delta*normcdf((gamma-w*x)/fi,0,1)+(1-delta)*normcdf((gamma_le-w_le*x-delta_le*v)/fi_le,0,1);  %This function is used to check whether the for given ‘x’ and ‘v’ the portfolio
%loss is above ‘pxq’, the latter being a function of ‘xq’

y=dblquad(@(x,v)  normpdf(x,0,1).*normpdf(v,0,1).*(meet(x,v)>pxq),-20,20,-20,20,0.000000001);  %carrying out the integration to find the probability that losses for given ‘x’ and ‘v’ are above pxq
```

MNB Working Papers 2006/11
Pillar I treatment of concentrations in the banking book – a multifactor approach

Print: D-Plus
H–1037 Budapest, Csillaghegyi út 19–21.