Kredit und Kapital, 40. Jahrgang, Heft 4, Seiten 527–551
Abhandlungen

Calibration of Internal Rating Systems: The Case of Dependent Default Events

By André Güttler, Oestrich-Winkel,
and Helge G. Liedtke, Frankfurt/M.*

I. Introduction

Banks' internal rating systems have gained considerable importance in recent years. This is due to regulatory pressure imposed by the new Basel II framework, and to economic reasons such as the imperative of managing a credit portfolio according to the principles of economic capital or risk adequate pricing. Given the increased significance of internal rating systems, banks and regulatory authorities are becoming more and more interested in assessing their quality. In other words, banks must frequently review their rating systems, a process which is referred to as "validation". According to *Deutsche Bundesbank* (2003), the quantitative validation of rating systems can be separated into an assessment of two of their attributes: their discriminatory power, which denotes their ability to discriminate ex ante between defaulting and non-defaulting debtors; and the accuracy of their calibration, which is high if the estimated probabilities of default (PD) deviate only slightly from the observed default rates.[1] The maximization of the discriminatory power is guaranteed by the bank's own economic incentives, since otherwise risk inadequate pricing occurs. Incorrect calibration on the other hand, which means assigning too low PDs, would lead to lower regulatory equity requirements.[2] Therefore, banking regulatory authorities concentrate on calibration. In testing the quality of calibration, the default correlation plays a decisive role. In this paper we present four approaches to testing the

* We thank Michael Gordy and one anonymous referee for their helpful comments. All errors and opinions expressed in this paper are of course our own.

1 As a third criterion, the discriminatory power should be stable over time.

2 Of course, assigning too high probabilities of default would lead to higher regulatory equity requirements. We regard this as the less realistic and impacting case.

quality of calibration of internal rating systems with (positive) default correlations. We extend the existing literature since this paper is the first to compare these approaches.

We find that multi-factor models generate more precise results through lower upper bound default rates and narrower confidence intervals. For confidence levels of 95 %, the approximation approaches overestimate the upper bound default rates. For asset correlation of less than 0.5 %, the granularity adjustment approach does not deliver reasonable results. For low numbers of debtors in a given rating class (or credit portfolio), the approximation approaches sharply overestimate the upper bound default rates. Using empirical inter-factor correlations we find that confidence intervals of two-factor models are much tighter compared with the one-factor model.

The study is organized as follows. Section II. provides a brief review of the literature. The following section presents four different approaches in the case of dependent default events. First, a one-factor simulation approach for default probabilities is demonstrated. Then, two approximation approaches to determining confidence intervals analytically are described: the granularity adjustment approach and the moment matching approach. Fourth, a multi-factor model for calculating confidence intervals is shown. Section IV. presents a comparative analysis of the four methods. Section V. provides a test for a two-factor model with heterogeneous default correlations. The last section summarizes.

II. Literature Review

Common factor models used in practice are CreditMetrics™ (*Gupton, Finger* and *Bhatia* (1997)), CreditRisk+™ (*CSFB* (1997)), PortfolioManager™ (*Crosbie* and *Bohn* (2003), and *McQuown* (1993)), CreditPortfolioView™ (*Wilson* (1998)), and the model used for the calculation of the minimum capital requirements according to Basel II (*BCBS* (2004)).[3] The first credit risk models were introduced by *Merton* (1974) and *Black* and *Scholes* (1973). Among others, *Black* and *Cox* (1976), *Geske* (1977), as well as *Longstaff* and *Schwartz* (1995) advance the basic asset value model that assumes a default event to occur if the value of an obligor's assets falls below the value of its liabilities. *Vasicek* (1997) introduces a one-factor model based on the previous research, which incited many

[3] Among others, *Frey/McNeil* (2001) and *Crouhy/Galei/Mark* (2000) review and analyze credit risk modeling approaches and methodologies.

authors to extend the model structure. *Tasche* (2003) recommends a traffic lights approach incorporating extensions of Vasicek's one-factor model. To test the quality of calibration given a certain correlation of defaults, he calculates confidence intervals for the number of defaulting firms using two approximation approaches, since no closed solution is available. He compares the results to upper bound default rates calculated with a binomial test (assuming independent default events). *Blochwitz, Wehn,* and *Hohl* (2005) further extend Tasche's approach by incorporating correlation over time and correlation between several rating grades into the model. Further studies focusing on the approximation approaches are *Gordy* (2003), *Martin* and *Wilde* (2002), *Gouriéroux, Laurent,* and *Scaillet* (2000), and *Rau-Bredow* (2002).

A different line of the literature focuses on the importance of the incorporation of macroeconomic factors on PD estimation. *Helwege* and *Kleiman* (1996) and *Alessandrini* (1999) show that default rates depend on the phase of the business cycle, i. e. defaults are more likely in economic downturns than in economic booms. *Nickell, Perraudin* and *Varotto* (2000) present a probit model for the estimation of rating transition probabilities considering macroeconomic factors such as the industry, the business cycle, and the country of establishment. *Hamerle, Liebig* and *Scheule* (2004) derive factor models for the PD estimation which include positive default correlations. As a result of empirical analyses of more than 50,000 German firms they find that the incorporation of macroeconomic factors improves the forecasts of default probabilities. They also show that default rates can be forecasted by including those factors. A factor model presented in this context allows the forecasting of PDs for individual debtors by considering their dependency structures. *Huschens* and *Stahl* (2005) propose a test framework for general factor models (although they only present Vasicek's one-factor model). They indicate that the assumption of independent default events, e. g. zero correlation, as well as assuming too high asset correlations of over 20 %, yields wrong PD forecasts. Finally, *Schönbucher* (2000) presents conditionally independent models, ranging from the simple case of a homogeneous portfolio to the complex structures of a multi-factor model.

III. Calibration Approaches

The assignment of default probabilities to a rating model's output is referred to as calibration (*OeNB/FMA* (2004)). The rating model's output may be a grade or other score value. The probability of default p of a portfolio of debtors can also be denoted as a vector of probabilities of defaults considering multiple rating classes, mainly in order to facilitate reporting. The internal rating system may consist of R rating grades. The vector $p = (p_1, \ldots, p_R)$ denotes the corresponding PDs. In the following we focus on either one rating class or the whole portfolio at once, denoting the PD as p.

The quality of calibration depends on the degree to which the PDs forecasted by the rating model match the default rates actually realized.[4] The basic data used for calibration are:

- The PD forecasts over a rating class and the credit portfolio for a specific forecasting period.
- The number of obligors assigned to the respective rating class by the model.
- The default status of the debtors at the end of the forecasting period.

In practice, realized default rates are subject to huge fluctuations. Thus, it is necessary to develop indicators to show how well a rating model estimates the PDs, i.e. to check the significance of deviations in the default rate. Therefore, we calculate confidence intervals at two confidence levels: 95 % and 99.9 %. These levels correspond to a traffic lights approach for practice in Germany for the purpose of interpreting confidence levels proposed by *Tasche* (2003). We calculate confidence intervals (upper bounds and lower bounds) for the two confidence levels such that the probability that the true number of defaults does not exceed the confidence intervals' upper bounds will equal 95 % (low) and 99.9 % (high) respectively.

Tasche (2003) recommends using traffic lights as indicators of whether deviations of realized and forecasted default rates should be regarded as significant or not as follows:

- Green traffic light: The true default rate is equal to or lower than the upper bound default rate at a low confidence level. The PD forecast seems to be appropriate as there is no significant deviation of the forecasted default rate from the true default rate.
- Yellow traffic light: The true default rate is higher than the upper bound default rate at a low confidence level and equal to or lower

[4] The calibration of a rating model is often referred to as back testing.

than the upper bound default rate at a high confidence level. This case needs to be subjected to further analyses.

- Red traffic light: The true default rate is higher than the upper bound default rate at a high confidence level. This case is regarded as being significant, and has to be emended as the likelihood of a wrong PD forecast is too high.

In section IV., we use two different approximation approaches and two simulation models to determine confidence intervals depending on the level of confidence. Testing the different methods we determine confidence intervals for the number of defaults and the default rates that are still acceptable in terms of the calibration of the rating model.

The simplest way to compute confidence intervals can be used if one assumes uncorrelated default events. Under this assumption, confidence intervals based on the standard normal distribution are employed for the purpose of comparison (*OeNB/FMA* (2004)). As the banking industry experiences dependent default events in loan portfolios, methods that assume independence are not appropriate for validating probabilities of default. The dependence between defaults by different obligors may have multiple different causes. Correlations among debtors exist directly and indirectly via economic factors. Direct relations exist already if one obligor is a debtor of the other, or one is the other's customer. Even if direct dependency structures are not obvious, indirect dependency can be caused by the influence of macroeconomic or industry-specific factors that both debtors have in common.

Düllmann and *Scheule* (2003) find asset correlations ranging between 0.9 % and 9.4 % for German companies.[5] Default correlations that are not equal to zero increase fluctuations in PDs (*OeNB/FMA* (2004)). The assumption of uncorrelated defaults generally yields an overestimate of the significance of deviations of the true default rate from the forecast rate. If the true default rate is higher than the forecast rate the true risk will be underestimated. From a conservative risk assessment standpoint, overestimating credit risk significance is not critical in the case of risk underestimates. This implies that it might be possible to operate under the assumption of uncorrelated defaults. But continuous overestimates of significance will lead to more frequent recalibration of the rating model,

[5] Please note that in the one-factor model that has been used to derive Basel II's IRB risk-weight functions exists a one-to-one mapping between default correlation and asset correlation for a given probability of default.

which can also have negative effects on the model's stability over time. It is therefore necessary to determine as precisely as possible the extent to which default correlations influence PD estimates.

1. *One-factor Model*

Factor models are used to provide correlated defaults by using Monte Carlo simulations.[6] Factor models serve to calculate confidence intervals depending on the probability of error α in a simulation process determining the number of defaults that can still be tolerated in terms of the calibration of the model.

Vasicek (1997) introduced a widely used one-factor model. In this one-factor model all probabilities of default depend on a single random variable X that models correlation or systematic risk. The PD is decomposed into a monotonic function of the one factor X and a residual ξ. The factor X models the systematic risk whereas ξ represents the idiosyncratic risk of every individual debtor. These variables are assumed to be standard normally distributed. Thus, every single debtor i will be assigned an individual idiosyncratic risk $\xi_1, \ldots, \xi_n$. These obligor-specific risk variables ξ_i can be interpreted as the debtor's key financial figures such as the return on equity.

In order to calculate the confidence intervals, the one-factor model will serve as a stochastic model to determine the number of defaults D_n. The one-factor model can be written as:

$$D_n = \sum_{i=1}^{n} 1_{E_i} = \sum_{i=1}^{n} 1_{\{\sqrt{\rho}X+\sqrt{1-\rho}\xi_i \le t\}}. \tag{1}$$

The number of defaults in a certain period of time equals the sum of default events E that are expected among the n debtors. 1_E denotes a binary indicator function. If a specific debtor i is expected to default, the indicator function is one, and zero otherwise. Dependent default events are modeled by using a uniform asset correlation ρ that is addressed to the independent standard normal random variables X and ξ_i. The question whether a default occurs or not can be addressed by defining a threshold t in such way that a default occurs if $\sqrt{\rho}X + \sqrt{1-\rho}\xi_i$ falls short of the threshold.

6 For an overview of Monte Carlo simulations see *Jäckel* (2001).

As it should always be possible to calibrate the factor model to mirror specific PDs, the threshold t has to be chosen in such a way that the expectancy of the number of defaults of a portfolio equals the probability of default times the number of obligors in the respective portfolio. As an example among others, *Tasche* (2003) recommends setting t to $\Phi^{-1}(p)$, where Φ^{-1} denotes the inverse of the standard normal distribution function and p denotes the realized PD.

2. *Granularity Adjustment Approach*

The Basel Committee on Banking Supervision conceived the granularity adjustment approach as part of the Basel II proposals for reforming the calculation of regulatory capital for credit risk (*BCBS* (2001)). Originally, *Gordy* (2003) invented the approach using CreditRisk+$^{\text{TM}}$. *Wilde* (2001) enhanced the approach and derived it theoretically for any one-factor model. The granularity approach is a formula for risk effects in a portfolio of loans assessing the change of percentiles when enhancing the risk of the portfolio. This risk can be interpreted as all the concentration risk comprised in the portfolio (*Martin* and *Wilde* (2002)).

First, economic or systematic risk has to be assessed, which is defined as the risk attributed to the portfolio if all loans could be subdivided into infinitely many infinitesimal loans. In this hypothetical case the portfolio would be infinitely granular. Second, the granularity adjustment can be conducted as an adjustment for concentration risk. Infinitesimal loans are only hypothetical as is an infinitely granular portfolio. Thus, the actual portfolio typically comprises a higher level of risk than this hypothetical portfolio. The granularity adjustment proves to be appropriate though, as the marginal impact of additional risk on the percentiles of the distribution of default rates can analytically be evaluated quite exactly. To approximate the confidence intervals of the distribution of the default rate, i.e. the quantiles $q((1-\alpha), R_n)$, the granularity adjustment approach can be used as follows:

The substructure of the approximation is a second order Taylor expansion of $q((1-\alpha), R_n)$ that can be written as follows:

$$q((1-\alpha), R_n) = q((1-\alpha), R + h(R_n - R))|_{h=1} \tag{2}$$

A parameter h is introduced for the second order Taylor expansion to be multiplied with the difference $(R_n - R)$. Thus, h can be interpreted as a measure of the difference of R_n to R. At first, h is set to be one, addressing R_n directly. R_n is not yet known, thus R_n will be approximated by $R + h(R_n - R)$.

For $h = 1$, (2) is an equation based on the fixed default rate R. (2) can then be approximated as

$$q\big((1-\alpha), R\big) + \frac{\partial}{\partial h} q\big((1-\alpha), R + h(R_n - R)\big)\Big|_{h=0} + \frac{1}{2}\frac{\partial^2}{\partial h^2} q\big((1-\alpha), R + h(R_n - R)\big)\big|_{h=0}\,. \tag{3}$$

(3) is a Taylor series in h, which is set to be zero now. Remaining on R the term serves to determine the quantile $q((1-\alpha), R_n)$.

For plausibility reasons, R in (2) and (3) can be determined as follows:

$$R = \lim_{n\to\infty} R_n = \Phi\left(\frac{t - \sqrt{\rho}X}{\sqrt{1-\rho}}\right). \tag{4}$$

Because R is known, the second order Taylor expansion can be based on R. Based on the determined rate R in equation (4), the quantile $q((1-\alpha), R_n)$ can be written as

$$q\big((1-\alpha), R_n\big) = \Phi\left(\frac{\sqrt{\rho}\Phi^{-1}(1-\alpha) + t}{\sqrt{1-\rho}}\right). \tag{5}$$

As a consequence, the quantile $q((1-\alpha), R_n)$ can be calculated (*Tasche* (2003)). For computational reasons one can also write (5) as

$$q\big((1-\alpha), R_n\big) = \Phi\left(\frac{t - \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right). \tag{6}$$

The unknown quantile (2) can now be calculated using the approximation (6). Now, the derivatives can be computed. As (4) defines R, one minor weakness of the granularity adjustment may occur, due to the fact that the partial derivatives in (3) may not be calculated because the distribution of D_n is entirely discrete but was derived for smooth distributions (*Tasche* (2003)). Based on the analysis of *Martin* and *Wilde* (2002) this weakness might be negligible as they found the approximation to generate significant results although the distributions considered have not been perfectly smooth.

Now, the known quantile $q((1-\alpha), R_n)$ of equation (6) can be applied to equation (2) to calculate equation (7). *Tasche* (2003) uses the formulas for the derivatives determined by *Martin* and *Wilde* (2002) to create the following granularity adjustment formula:

$$q\big((1-\alpha),D_n\big)\approx nq\big((1-\alpha),R\big)+\frac{1}{2}$$

(7)

$$\left(2q\big((1-\alpha),R\big)-1+\frac{q\big((1-\alpha),R\big)\big(1-q\big((1-\alpha),R\big)\big)}{\phi\left(\frac{\sqrt{\rho}q(\alpha,X)-t)}{\sqrt{1-\rho}}\right)}\left(\frac{\sqrt{\rho}q(\alpha,X)-t}{\sqrt{1-\rho}}-\sqrt{\frac{1-\rho}{\rho}}q(\alpha,X)\right)\right),$$

where $\phi(x)=(2\pi)^{-\frac{1}{2}}e^{\frac{-x^2}{2}}$ denotes the standard normal density.

3. *Moment Matching Approach*

Moment matching is based on the moment-generating function and has no direct resemblance to the granularity adjustment approach. It approximates the distribution of the default rate R_n with a Beta-distribution (*Tasche* (2003)). The one-factor model is the foundation of the moment matching approach as well. The parameters of the Beta-distribution are determined by matching the expectation and the variance of R_n (*Overbeck* and *Wagner* (2000)). From this procedure, the approximation deduces its name “moment matching”, because the first and the second central moments are being matched. The density of a $B(a,b)$-distributed random variable Z is defined by

$$\beta(a,b,x)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1},0<x<1, \tag{8}$$

The Gamma-function Γ expands the factorial function to the positive reals. The first central moment, the expectation, of a random variable Z is given by

$$E[Z]=\frac{a}{a+b}, \tag{9}$$

(*Abramowitz* and *Stegun* (1972)). Accordingly, the second central moment, the variance of Z, is defined as

$$\mathrm{var}[Z]=\frac{ab}{(a+b)^2(a+b+1)}. \tag{10}$$

Now, let the random variable Z be replaced by the default rate R_n. The moments can now be matched by equating the right-hand sides of (9) and (10) with $E[R_n]$ and $var[R_n]$ respectively to yield the following relations for the parameters a and b of the Beta-distribution:

$$a = \frac{E[R_n]}{\mathrm{var}[R_n]}(E[R_n](1 - E[R_n]) - \mathrm{var}[R_n]) \tag{11}$$

and

$$b = \frac{1 - E[R_n]}{\mathrm{var}[R_n]}(E[R_n](1 - E[R_n]) - \mathrm{var}[R_n]). \tag{12}$$

Based on the assumptions made for the one-factor model, the expectancy of the default rate R_n equals the probability of default (PD):

$$E[R_n] = p \tag{13}$$

The following equation represents the variance of the default rate R_n accordingly

$$\mathrm{var}[R_n] = \frac{n-1}{n}\Phi_2(t, t, \rho) + \frac{p}{n} - p^2, \tag{14}$$

with $\Phi_2(t, t, \rho)$ denoting the distribution function of the bivariate standard normal distribution with asset correlation ρ. Unfortunately common tools like MS Excel™ do not provide algorithms for the calculation of the bivariate standard normal distribution function Φ_2. Thus, we use

$$\Phi_2(t, t, \rho) \approx \Phi(t)^2 + \frac{e^{-t^2}}{2\pi}(\rho + \frac{1}{2}\rho^2 t^2) \tag{15}$$

as an approximation for Φ_2 in (14) following *Tasche* (2003).

Thus, the moment matching approach serves to determine confidence intervals and upper bound default rates analytically at good accuracy, too. To deliver appropriate results, the approximation requires the same set of input parameters as does the granularity adjustment approach: an asset correlation ρ, the realized PD p, the number of debtors n considered, and a requested confidence level. Based on the above equations, the confidence interval (quantile) $q((1-\alpha), D_n)$ can be determined via the approximation

$$q((1-\alpha), D_n) \approx nq((1-\alpha), Z). \tag{16}$$

with Z as a $B(a, b)$-distributed random variable.

4. Multi-factor Model

The general framework of the factor models can be adopted for a multi-factor model, too. A multi-factor model can be defined as a latent variable model. A default occurs if a latent variable falls below a defined threshold t. Accordingly, a term of variables can be interpreted as the value of the obligor's assets, and the threshold t as the value of the obligor's liabilities. The dependence between the latent variables causes dependency of defaults. Driven by these specific variables, the model may incorporate macroeconomic factors like the business cycle and industry-specific factors.[7]

Whereas a uniform asset correlation for all debtors in a portfolio is assumed in the one-factor model, this assumption will be given up and extended to a matrix of individual debtor-specific asset correlations. For every debtor considered in the portfolio, individual factor weights and idiosyncratic attributes can be considered, yielding an asset correlation matrix $\boldsymbol{P}$ reflecting individual dependency structures.

To ensure the comparability of the simulation results of the multi-factor model with the simulation results of the one-factor model the asset correlation matrix of a multi-factor model should also allow the calculation of a uniform asset correlation under specific conditions. The asset correlation matrix $\boldsymbol{P}$ can be calculated using a correlation matrix of estimated correlations among all macroeconomic or industry-specific factors in addition to a matrix of the estimated influences of every factor on every single debtor in the selected credit portfolio.

We consider a two-factor model as an example of a multi-factor model. We assume dependency structures between the two factors and the influences of the two factors on every debtor. For each debtor we assign a weight to every factor in addition to an idiosyncratic weight. If this set of correlation structures yields a uniform asset correlation ρ comparable to the asset correlations assumed in the one-factor model, the simulation results can be compared. In general, the asset correlation matrix $\boldsymbol{P}$ of the multi-factor model is an output of the dependency structures of the factors and the obligors considered and thus can vary infinitely. Only assigning for all debtors the same weights to the respective factors and the same idiosyncratic weight yields a uniform asset correlation ρ compar-

[7] Our multi-factor model is based on the CreditMetrics™ framework (*Gupton/Finger/Bhatia* (1997)) and the results of *Tasche* (2003), *Hamerle/Liebig/Scheule* (2004), *Lucas* (1995), *Huschens/Stahl* (2004) and *Schönbucher* (2000).

able to the asset correlation ρ assumed in the one-factor model. In this case the asset correlation matrix $\boldsymbol{P}$ consists of equal correlations ρ (except for the main diagonal that assumes the value 1).

A multi-factor model can be expressed as a model which includes idiosyncratic obligor-specific and statistical risk drivers and in our case two systematic risk drivers. A debtor i defaults in the observed period of time if:

$$\beta_{1i}X_1 + \beta_{2i}X_2 + \beta_{3i}\xi_{\mathrm{i}} \leq t \tag{17}$$

To determine the number of defaults for the calculation of confidence intervals and the PD forecast we annotate the two-factor model as follows:

$$D_n = \sum_{i=1}^{n} 1_{E_i} = \sum_{i=1}^{n} 1_{\{\beta_{1i}X_1 + \beta_{2i}X_2 + \beta_{3i}\xi_{\mathrm{i}} \leq t\}}. \tag{18}$$

As asset correlation models the co-movement of two obligors' (i and $j \neq i$) asset values V_i, ρ can be defined as:

$$\rho_{ij}(V_i, V_j) = \frac{Cov(V_i, V_j)}{\sqrt{Var(V_i)}\sqrt{Var(V_j)}}, i \neq j \tag{19}$$

Accordingly, default correlation can be derived from the asset correlation:

$$Corr(D_i, D_j) = \frac{\Phi_2\big(t, t; \rho_{ij}(V_i, V_j)\big) - p^2}{p(1-p)}, \tag{20}$$

with D_i and D_j denoting the number of defaults of two correlated obligors i and j respectively, and $\Phi_2(t, t, \rho_{ij})$ denoting the distribution function of the bivariate standard normal distribution.

To determine the weight vectors β_{fi} of the multi-factor model and to extend the assumption of a uniform asset correlation ρ made in the one-factor model for all debtors of a portfolio into a matrix $\boldsymbol{P}$ of individual debtor-specific asset correlations in a multi-factor model, standard weights have to be assigned to every debtor, concerning its dependency on the macroeconomic factors and the idiosyncratic risk.[8]

[8] A further model type is feasible, but not shown here. A single-factor model with heterogeneous factor loadings would be an "intermediate case" between the one-factor and the multi-factor model.

To calculate obligor-specific weight vectors, consider an obligor i with dependencies on 2 macroeconomic factors X_1 and X_2 and with a weight on idiosyncratic risk of η_i. The factors account for $(1-\eta_i)$ of the movements of a debtor's equity or asset values. The volatility of the weighted index for an obligor i considering weights w_{1i} and w_{2i} on the two systematic factors respectively can be written as:

$$\overset{\wedge}{\sigma}_i = \sqrt{w_{1i}^2\sigma_1^2 + w_{2i}^2\sigma_2^2 + 2w_{1i}w_{2i}\rho_{12}\sigma_1\sigma_2}, \tag{21}$$

with volatilities σ_1 and σ_2 for factors X_1 and X_2 respectively, and inter-factor correlation ρ_{12} between the two factors X_1 and X_2. The three vectors $\beta_{fi}(f=1,2,3)$ alter for every debtor $i(i=1,\dots,n)$ accordingly. We annotate the two macroeconomic weight vectors as follows:

$$\beta_{1i} = (1-\eta_i)\frac{w_{1i}\sigma_1}{\overset{\wedge}{\sigma}_i}, \beta_{2i} = (1-\eta_i)\frac{w_{2i}\sigma_2}{\overset{\wedge}{\sigma}_i}, i=1,\dots,n. \tag{22}$$

The idiosyncratic weight vector β_{3i} for the debtor-specific idiosyncratic risk factor ξ_i can be written as follows:

$$\beta_{3i} = \sqrt{1-(1-\eta_i)^2}, i=1,\dots,n \tag{23}$$

Considering various macroeconomic or industry-specific factors we continue with the treatment of asset correlation in the multi-factor model. The asset correlation matrix $\boldsymbol{P}$ considering all debtors and all systematic and idiosyncratic factors will be the product of the matrices $\boldsymbol{W}$ and $\boldsymbol{C}$. $\boldsymbol{C}$ denotes the correlation matrix covering both the correlation matrix for the macroeconomic factors $\boldsymbol{F}$ and the idiosyncratic risk factors ξ_i of every debtor. $\boldsymbol{W}$ denotes the weights matrix for every debtor. The result is the asset correlation matrix consisting of the correlations between all debtors. In order to compare the simulation approaches of the one-factor model and the multi-factor model a uniform asset correlation is required. Thus, we choose for any debtor the same weights w_1, w_2, and η in the input matrix yielding a uniform asset correlation matrix with a uniform asset correlation ρ for every debtor.

IV. Comparison of the Different Approaches

The first analysis focuses on determining confidence intervals for the default rate R_n, dependent on different asset correlations. While the highest asset correlation occurring in the Basel II framework is 24 %, *Tasche* (2003) recommends an asset correlation of 5 % for Germany. To include all plausible values, we vary the asset correlation from 0 to 10 %. Besides different asset correlations, we analyze the quality of different methods and calculate confidence intervals and upper bound default rates in dependence on the different model parameters as follows:

– confidence level (95 % and 99.9 %),

– PD (1 % to 10 %),

– and the size of rating classes or the size of the portfolio (50 to 1,000 debtors).

Comparing the simulation results of the one-factor and the multi-factor model and the analytical results of the two approximation methods using the moment matching and the granularity adjustment approaches, we first create confidence intervals for a credit portfolio consisting of 1,000 debtors. All simulation results, i. e. for the one-factor model and the multi-factor model, are based on 5,000 scenarios using Halton random numbers.[9]

Comparing factor-model and approximation method results for a (uniform) PD of 2 % at confidence levels of 95 % and 99.9 % we analyze the dependence of the confidence intervals on asset correlation. From a conservative point of view and considering the threat when underestimating the true PD significantly, we consider narrower confidence intervals and lower upper bound default rates to indicate a higher level of backtesting precision. The consecutive figures present the results graphically and depict either the width of the confidence interval or the upper bound of the confidence interval on default rates in dependence on different model parameters.

Figure 1 shows that with increasing asset correlation, the upper bound default rates rise. This tendency is also documented by *Tasche* (2003) and by *Schönbucher* (2000) who observe that asset correlation significantly affects the magnitude of the PD. Second, with increasing asset correlation the differences between the one-factor model and the multi-factor results increase. For 1 % asset correlation the upper bound default rates are equal for the approximation approaches. For less than 1 % asset correlation, the approximation results of moment matching generate in-

[9] Halton numbers are quasi-random numbers which avoid any asymptotic periodicity; see *Jäckel* (2001). Sensitivity analyses show that even with 1,000 scenarios we could produce robust results.

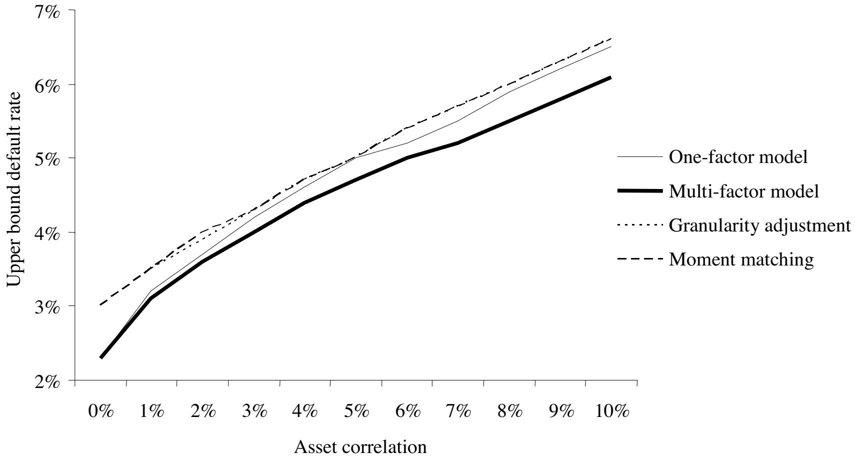

Figure 1: Upper Bound Default Rates in Dependence on Different Asset Correlations (95% Confidence Level)

creasingly higher upper bound default rates. At asset correlations close to 0%, the granularity adjustment does not calculate reasonable results as it is not defined for an asset correlation of zero. Thus, in the case of asset correlations less than 0.5%, we recommend that only the moment matching method be used when approximating upper bound default rates analytically.[10] In general, the approximation approaches overestimate the upper bound default rates compared to the simulation results. The multi-factor model results are more precise considering the multiple factors, invariably generating lower upper bound default rates than the one-factor model does.

For regulatory purposes, only the upper bound default rates are important. But for banks' internal use, the lower bound default rates are also important; e.g. for pricing purposes the PDs should not be too conservative. Hence, looking at the width of the confidence intervals determined by the factor models (difference between upper bound and lower bound default rates), the results of both simulation methods show increasing intervals with rising asset correlations (see Figure 2). The multi-factor

10 However, asset correlations of less than 0.5% are quite unrealistic, although close to this level *Düllmann/Scheule* (2003) find asset correlations of about 0.9%.

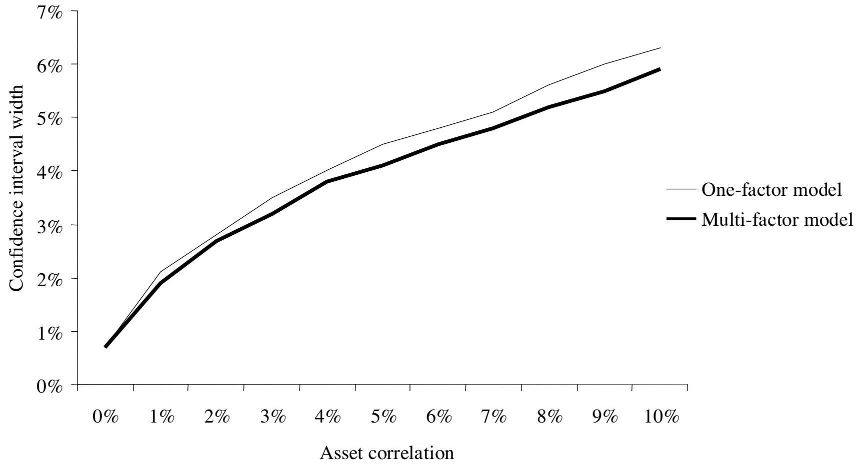

Figure 2: Width of Confidence Intervals in Dependence on Different Asset Correlations (95% Confidence Interval)

model intervals are tighter, underlining the higher precision and accuracy of the multi-factor model due to its advantages in assessing a larger set of variables considering multiple factors. Compared to the structures of the 95% confidence intervals, the 99.9% confidence intervals are wider.[11] The tendency of the one-factor model to slightly overestimate the multi-factor model confidence intervals observed at a 95% confidence level still occurs, but at this very high confidence level, we find only a slight difference. Thus, both factor-models generate adequate forecasts.

We now analyze the upper bound default rate in dependence on different asset correlations for the 99.9% confidence level (see Figure 3). The differences in the determined bounds between simulation and approximation results become more pronounced. Now, the moment matching approach underestimates the true upper bounds of the confidence intervals as this method faces more and more restrictions for extreme sets of parameters. However, in line with Tasche's observations, we find that the granularity adjustment excels the moment matching results at high confidence levels with an increasing asset correlation as the upper bound default rates determined by moment matching deviate increasingly from

11 Results are not shown here but are available upon request.

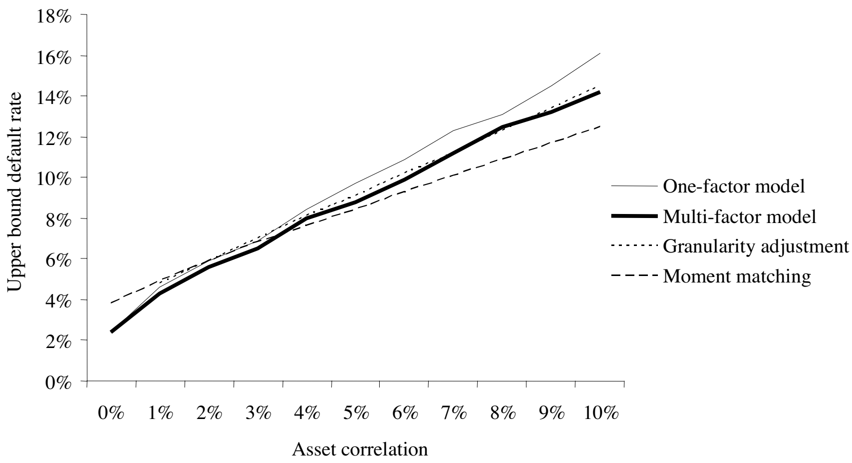

Figure 3: Upper Bound Default Rates in Dependence on Different Asset Correlations (99.9% Confidence Level)

the granularity adjustment bounds. These results confirm Tasche's assumption that the granularity adjustment is the more reliable of the two approximation approaches. A weakness of the granularity adjustment may be the fact that calculated percentiles may not be the sums of their Taylor series (*Martin* and *Wilde* (2002)). This disadvantage can be debilitated otherwise, considering that using an approximation is not supposed to yield a perfect copy of simulation results, but rather an appropriate and fast approximation to percentiles, which can always be achieved using the granularity adjustment.

Next, we perform an analysis of the model validity in dependence on the true PD. We choose a low uniform asset correlation level of 2% as the effect of the asset correlation on the results is very minor at this level, because simulation results and approximation values are approximately equal, ceteris paribus. The model results are very similar with only marginal differences (see Figure 4). In dependence on the true PD, the factor models as well as the approximation methods deliver stable and precise results. Independent of the model, the upper bound default rates increase as the probability of default increases. For every percent increase in PD, the upper bound default rate increases by approximately 1.3% to 1.5%. The relationship is quite linear in all four models. Thus, we conclude that in dependence on the true PD, the factor-models as

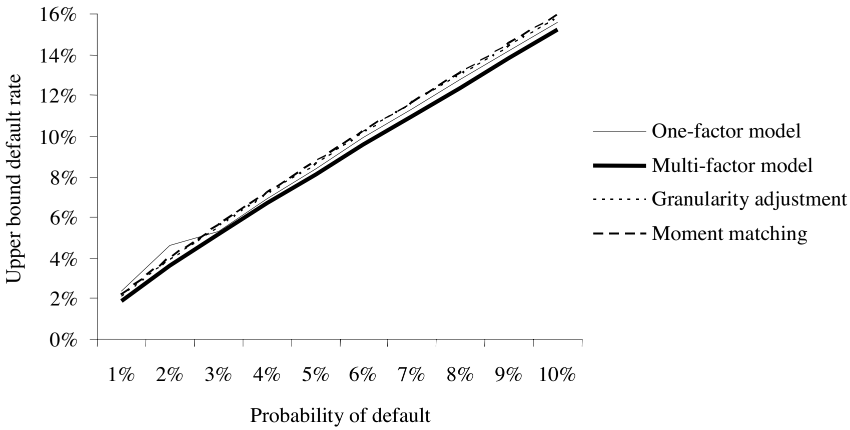

Figure 4: Upper Bound Default Rates in Dependence on True PD (95% Confidence Level)

well as the approximation methods deliver comparable results. Independent of the model, the upper bound default rates increase in a linear relationship as the probability of default increases.

Next, we subdivide the portfolio and build smaller portfolios to analyze the effect of the number of debtors on the model results. Exemplarily, we focus on portfolios consisting of 50 to 950 debtors in steps of 50 debtors {50, 100, 150, …, 950}. Finally the whole portfolio of 1,000 debtors will be added. To separate the effect of the number of debtors from the other parameters' influences we perform all test sets assuming an asset correlation of 2% and a PD of 2% at a 95% confidence level. For any set of parameters, every model determines at high levels of precision confidence intervals that narrow (marginally) as the number of debtors considered per portfolio increases (see Figure 5). The asset correlation, the confidence level, and the true PD mainly affect the default rate level of the curve. For the set of parameters chosen the slope is approximately a 0.07% decrease in the upper bound default rate for any additional debtor considered. The smaller the number of debtors per portfolio is, the more the approximation methods overestimate the real upper bound default rates. In general, for a decreasing number of debtors fewer than 150 to 250 (independent of the other model parameters), the simulation and approximation upper bound default rates increase at higher rates due to the small number of debtors available to determine precise de-

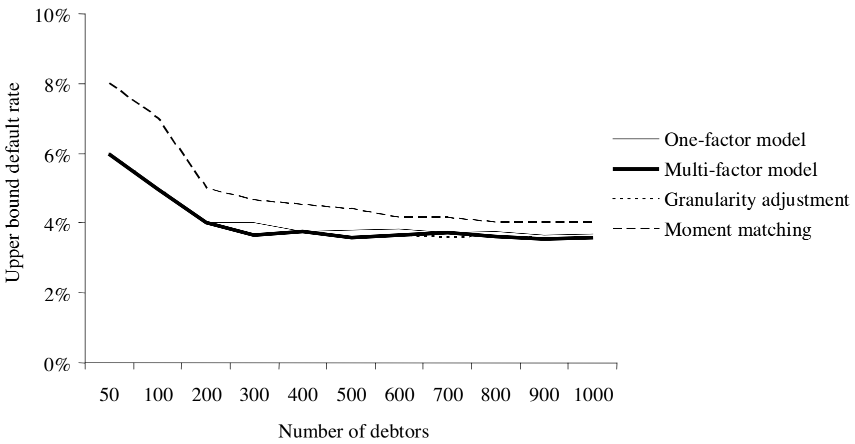

Figure 5: Upper Bound Default Rates in Dependence on the Number of Debtors (95% Confidence Level)

fault rates. Thus, independent of the method used and also independent of the set of the other model parameters, the upper bound default rates decrease as the number of debtors increases, as a higher number of debtors enhances the ability to calculate more precise default rates.

At this point we would like to give further insights into the way in which the results differ between the one-factor and the multi-factor models. One main distinction between the multi-factor model and the one-factor model is the treatment of asset correlation. In the one-factor model a uniform asset correlation is assumed for the whole portfolio, whereas the multi-factor model presented allows the calculation of obligor-specific asset correlations depicted in an asset correlation matrix. The multi-factor model offers the user a higher level of precision that can be improved the more detailed the available historic and current market data are, reflecting complex interdependencies among the different model parameters. Even under the assumption of uniform asset correlation, the multi-factor model allows tighter and more precise confidence intervals to be calculated. Therefore, we will test the impact of considering individual weights for every single debtor and the use of complex dependency structures in the next section.

V. Multi-factor Model Test with Heterogeneous Asset Correlation

To compare the multi-factor with the one-factor model further, we first give up the assumption of a uniform asset correlation. For every debtor, individual weights on the macroeconomic and idiosyncratic risk factors have to be considered to generate heterogeneous asset correlations. Second, we use one specific two-factor model as the simplest form of a multi-factor model. In our case, the two factors are the insolvency rate and the growth in new orders. However, to simulate our models we need the inter-factor correlation and variances for the two factors. We employ the parameters of *Hamerle, Liebig* and *Scheule* (2004), which use a data-set with 195,476 observations and 1,391 defaults for the period 1991 to1999. They find that the correlation between the insolvency rate and the growth in new orders equals –0.7851, the variance of the insolvency rate equals 0.00000169, and the respective value for the growth in new orders equals 0.00042436.[12] Third, we need a representative set of companies for our simulations. We assume 2,715 European companies assigned to 7 rating classes ranging from AAA to C. The cumulative historic default rates over 7 years for companies domiciled in the European Union are assumed to represent the true PDs of the respective rating classes (*Standard & Poor's* 2005, see Table 1). We use the cumulative default rates over 7 years for two reasons:

1. Otherwise, i.e. by using one year default rates, the small number of defaulting debtors would not serve to generate robust results in a statistical sense.
2. Seven years is the maximum period for which S&P provides default data for European companies.

Even using cumulative default rates over 7 years, for AAA rated debtors this value equals zero for the given period. To employ our simulation approach, we set this value to 0.0001. We argue that this is appropriate, given the fact that Basel II requires a minimum PD of 0.003.

The next two figures show exemplarily the distribution of default rates for one rating class. The set of underlying parameters is as follows: the (7 year) PD equals 31.64 %, 940 individual debtors of rating class B with different idiosyncratic weights for each debtor, the insolvency rate and

[12] Since we do not focus on detecting variables for estimating default rates in this paper, and since *Hamerle/Liebig/Scheule* (2004) provide appropriate parameters, we do not conduct an own empirical assessment.

Table 1

Seven Year Cumulative Default Rates

Rating class	Cumulative default rate	Number of debtors
AAA	0.0001	39
AA	0.0071	84
A	0.0082	416
BBB	0.0139	572
BB	0.1040	552
B	0.3164	940
C	0.5294	112
Total		2,715

Source: Standard & Poor's (2005)

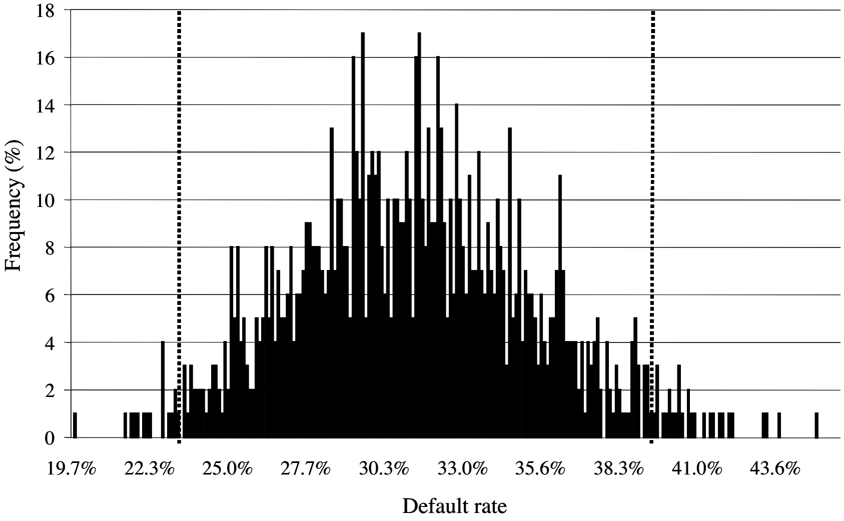

Figure 6: Frequency Distribution of Default Rates for a Two-factor Model

the growth in new orders are the macroeconomic factors for the two-factor model (Figure 6).

The vertical lines show the confidence interval bounds for the 95 % confidence level. Compared to the multi-factor model results in Figure 6,

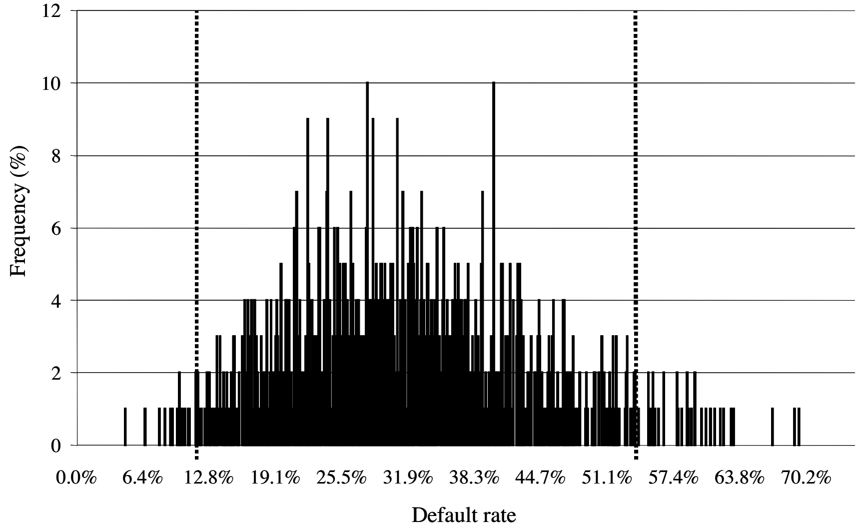

Figure 7: Frequency Distribution of Default Rates for a One-factor Model

the one-factor model results in Figure 7 are less precise because the one-factor model does not allow different macroeconomic factors to be considered, but only one systematic factor, and because it cannot consider individual debtor-specific weight vectors and asset correlations. Thus, the confidence intervals determined by the multi-factor model are much tighter and more precise and the determined frequency distribution is subject to less volatility and follows the assumed normal distribution more eloquently.[13] In conclusion, the multi-factor model test confirms the improved quality of the multi-factor model compared to the one-factor model and shows the dependency of the model results on the chosen macroeconomic factors and the sample of debtors with their respective characteristics. The results of this section show that the opportunity to reflect individual asset correlations among all debtors further

13 As the same set of data was used when parameterizing the models, the model characteristics can always be compared at the best, when applying identical sets of data to both models; the set chosen can strengthen the different model characteristics and outline the respective advantages of a model. In this case the more detailed set of data applied to both models allows to further assess the opportunities of the multi-factor model as it incorporates the different data characteristics in a more comprehensive way.

increases the precision and the explanatory power of the multi-factor model.

VI. Summary

This paper has compared four different approaches to testing the quality of calibration. We find that the higher the asset correlation, the higher the upper bound default rates are and the wider the confidence intervals become. This is in line with previous results (e.g. *Tasche* (2003)), which is interesting, since we are the first to analyze multi-factor models in addition. The multi-factor model generates more precise results given the lower upper bound default rates and narrower confidence intervals. For confidence levels of 95 %, the approximation approaches overestimate the upper bound default rates. On the other hand, for confidence levels of 99.9 %, the moment matching approach underestimates the upper bound default rates. For low asset correlation, especially for less than 0.5 %, the granularity adjustment approach does not deliver reasonable results. All four approaches give comparable results if we vary the true PD for a given asset correlation. For low numbers of debtors in a given rating class (or credit portfolio), the approximation approaches sharply overestimate the upper bound default rates. This result becomes even more pronounced for a higher confidence level of 99.9 %. Using empirical inter-factor correlations for an illustrative two-factor model we find that confidence intervals of this two-factor model (as they are in general for multi-factor models) are much tighter compared with the one-factor model.

References

Abramowitz, Milton and *Stegun,* Irene A.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 1972, New York. – *Alessandrini,* Fabio: Credit Risk, Interest Rate Risk, and the Business Cycle, in: Journal of Fixed Income, Vol. 9 (2), 1999, pp. 42–53. – *BCBS*: The Internal Ratings-Based Approach: Supporting Document to the New Basel Capital Accord, Technical Report, Bank for International Settlements, 2001, Basel. – *BCBS*: International Convergence of Capital Measurement and Capital Standards, A Revised Framework, 2004, Basel. – *Black,* Fischer and *Cox,* John C.: Valuing Corporate Securities: Some Effects of Bond Indenture Provisions, in: Journal of Finance, Vol. 31 (2), 1976, pp. 351–367. – *Black,* Fischer and *Scholes,* Myron: The Pricing of Options and Corporate Liabilities, in: Journal of Political Economy, Vol. 81 (3), 1973, pp. 637–654. – *Blochwitz,* Stefan; *Wehn,* Carsten S. and *Hohl,* Stefan: Reconsidering Ratings, Working Paper, Deutsche Bundesbank and Bank for International Settlements, 2005. – *CSFB*:

CreditRisk+^{TM} – Technical Document, A Credit Risk Management Framework, 1997, London. – *Crosbie,* Peter and *Bohn,* Jeff: Modeling Default Risk – Modeling Methodology, 2003, San Francisco. – *Crouhy,* Michael; *Galai,* Dan and *Mark,* Robert: A Comparative Analysis of Current Credit Risk Models, in: Journal of Banking and Finance, Vol. 24, 2000, pp. 59–117. – *Deutsche Bundesbank*: Approaches to the Validation of Internal Rating Systems, in: Deutsche Bundesbank Monthly Report, Vol. 55 (9), 2003, pp. 59–71. – *Düllmann,* Klaus and *Scheule,* Harald: Determinants of the Asset Correlations of German Corporations and Implications for Regulatory Capital, Working Paper, Deutsche Bundesbank, 2003. – *Frey,* Rüdiger and *McNeil,* Alexander J.: Modeling Dependent Defaults, Working Paper, University of Zürich and Swiss Federal Institute of Technology, 2001. – *Geske,* Robert: The Valuation of Corporate Liabilities as Compound Options, in: Journal of Financial and Quantitative Analysis, Vol. 12 (4), 1977, pp. 541–552. – *Gordy,* Michael B.: A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules, Journal of Financial Intermediation, Vol. 12 (3), 2003, pp. 199–232. – *Gouriéroux,* Christian; *Laurent,* Jean-P. and *Scaillet,* Oliver: Sensitivity Analysis of Values at Risk, in: Journal of Empirical Finance, Vol. 7 (3–4), 2000, pp. 225–245. – *Gupton,* Greg M.; *Finger,* Christopher C. and *Bhatia,* Mickey: Credit Metrics^{TM} – Technical Document, 1997, New York. – *Hamerle,* Alfred; *Liebig,* Thilo and *Scheule,* Harald: Forecasting Credit Portfolio Risk, Deutsche Bundesbank Discussion Paper, No. 01/2004. – *Helwege,* Jean and *Kleiman,* Paul: Understanding Aggregate Default Rates of High-Yield Bonds, Current Issues in Economics and Finance, Federal Reserve Bank of New York, Vol. 2 (6), 1996, pp. 1–6. – *Huschens,* Stefan and *Stahl,* Gerhard: A General Framework for IRBA Backtesting, 2005, Bankarchiv, Zeitschrift für das gesamte Bank- und Börsenwesen, Vol. 53, pp. 241–248. – *Jäckel,* Peter: Monte Carlo Methods in Finance, 2001, John Wiley & Sons. – *Longstaff,* Francis A. and *Schwartz,* Eduardo S.: A Simple Approach to Valuing Risky Fixed and Floating Rate Debt, in: Journal of Finance, Vol. 50 (3), 1995, pp. 789–819. – *Lucas,* Douglas J.: Default Correlation and Credit Analysis, in: The Journal of Fixed Income, March, 1995, pp. 76–87. – *Martin,* Richard and *Wilde,* Tom: Unsystematic Credit Risk, in: Risk, November, Vol. 15 (12), 2002, pp. 123–128. – *McQuown,* John A.: A Comment on Market vs. Accounting-Based Measures of Default Risk, 1993, San Francisco. – *Merton,* Robert C.: On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, in: Journal of Finance, Vol. 29 (2), 1974, pp. 449–470. – *Nickell,* Pamela; *Perraudin,* William and *Varotto,* Simone: Stability of Rating Transitions, in: Journal of Banking and Finance, Vol. 24 (1/2), 2000, pp. 203–227. – *OeNB/FMA*: Rating Models and Validation, 2004, Vienna. – *Overbeck,* Ludger and *Wagner,* Christoph: Term Structure of Loss Cascades in Portfolio. – Securitisation, 2000, Frankfurt. – *Rau-Bredow,* Hans: Credit Portfolio Modeling, Marginal Risk Contributions, and Granularity Adjustment, 2002, Würzburg. – *Schönbucher,* Philipp J.: Factor Models for Portfolio Credit Risk, 2000, Bonn. – *Standard & Poor's*: Annual Global Corporate Default Study, 2005, New York. – *Tasche,* D.: A Traffic Lights Approach to PD Validation, 2003, Frankfurt. – *Vasicek,* Oldrich A.: The Distribution of Loan Portfolio Value, Working Paper, Moody's KMV, 1997. – *Wilde,* Tom: Probing Granularity, Risk, 2001, Vol. 14 (8), pp. 103–106. – *Wilson,* Thomas C.: Portfolio Credit Risk, FRBNY Economic Policy Review, New York: Federal Reserve Board, 1998, pp. 71–82.

Summary

Calibration of Internal Rating Systems: The Case of Dependent Default Events

We compare four different test approaches for the calibration quality of internal rating systems in the case of dependent default events. Two of them are approximation approaches and two are simulation approaches of one- and multi-factor models. We find that multi-factor models generate more precise results through lower upper bound default rates and narrower confidence intervals. For confidence levels of 95 %, the approximation approaches overestimate the upper bound default rates. For low asset correlation, especially for less than 0.5 %, the granularity adjustment approach does not deliver reasonable results. For low numbers of debtors, the approximation approaches sharply overestimate the upper bound default rates. Using empirical inter-factor correlations we find that confidence intervals of two-factor models are much tighter compared with the one-factor model. (JEL C6, G21)

Zusammenfassung

Kalibrierung interner Ratingsysteme bei korrelierten Ausfallereignissen

In dieser Arbeit vergleichen wir vier verschiedene Testverfahren für die Qualität der Kalibrierung interner Ratingsysteme bei korrelierten Ausfallereignissen. Zwei der Ansätze sind approximativer Natur und die anderen zwei stellen Simulationsansätze auf Basis von Einfaktoren- bzw. Mehrfaktorenmodellen dar. Wir finden, dass die Mehrfaktorenmodelle präzisere Ergebnisse in Form niedrigerer, oberer Grenzen der Ausfallraten und engerer Konfidenzintervalle liefern. Für ein Konfidenzniveau von 95 % überschätzen die approximativen Ansätze die oberen Grenzen der Ausfallraten. Für niedrige Assetkorrelationen, vor allem für solche unter 0,5 %, liefert der Granularitätsansatz keine belastbaren Ergebnisse. Für kleine Portfoliogrößen überschätzen die approximativen Ansätze zudem die oberen Grenzen der Ausfallraten deutlich. Bei einer Anwendung empirischer Faktorkorrelationen finden wir außerdem, dass die Konfidenzintervalle eines Zweifaktorenmodells im Vergleich zum Einfaktorenmodell erkennbar enger sind.