Calibration of Rating Systems – A First Analysis

Luise Breinlinger,
Evgenia Glogova,
Andreas Höger[1])

1 Introduction

The need for adequate calibration is an issue that necessarily arises in the process of constructing a rating system as well as in its ongoing maintenance. Owing, among other reasons, to the implementation of the New Basel Capital Accord (Basel II), this issue will increasingly attract attention in the near future. The present study was based on credit data made available by the credit information bureau Creditreform consisting of some 10,000 data sets for each of the years 1996 through 2001. In this initial attempt to explore the issue of calibration, we restricted our research to static methods. This implies that the estimates of probabilities of default are based on one-year transitional rates, and the classification into rating classes relates to a single point in time in each of the years. The dynamics created by an intertemporal approach are, for the time being, taken into account only where our considerations refer to the conceptual framework. In general, this study focuses on methodological aspects. As to the results of the analyses, it is interesting to note that – given our static approach – the ceteris paribus increase in the number of rating classes generated using the calibration methods applied is paralleled by a decrease in the capital requirement. However, once an intertemporal approach is chosen, the demand for monotonicity in the structure of default probabilities imposes a natural limit on the maximum possible number of rating classes. Moreover, capital sensitivity to default rate changes further corroborates the need for intertemporal modeling (as well as sufficiently long data histories).

The study is composed as follows: Section 2 provides an outline of the database. Chapters 3.1 and 3.2 are dedicated to the empirical calibration analyses estimating default probabilities on the basis of the relative frequency of defaults (Chapter 3.1) on the one hand, and by means of logistic regression (Chapter 3.2) on the other. Chapter 3.3 focuses on integrative considerations. A summary and an outlook conclude the present paper.

2 Database

Creditreform delivered the data it made available to the Oesterreichische Nationalbank (OeNB) in two sets: one set containing data as of year-ends 2000 and 2001 and the other containing data as of year-ends 1996 to 2001. The two data sets were extracted from the database at different points in time (in July and in August 2002) and are not immediately comparable with regard to the period covered by both sets, December 31, 2000, to December 31, 2001. Sample checks showed that scores and other key data were changed between the two extractions from the database (sometimes retroactively), and data had been added or deleted, etc. Finally, the data obtained from Creditreform were complemented by data from the Major Loans Register of the OeNB.

The data sets used in this study are structured as follows:

Data set 1: 9,752 observations, with characteristics including com-

pany register number, postal code (region), sector code (as assigned by Creditreform), sales volume as of December 31, 2001, credit risk score assigned by Creditreform as of December 31, 2000 and December 31, 2001, Major Loans Register data (above all credit line utilization) as of December 31, 2001. Of the 9,752 companies reported solvent as of December 31, 2000, 196 defaulted in the course of 2001, which translates into an average probability of default of about 2% for the respective year.

Data set 2: 10,273 observations including the same characteristics as data set 1 for the period December 31, 1996, to December 31, 2001. This data set provides sales volume figures as of December 31, 2000, and December 31, 2001.

The Creditreform credit standing index is a key component of both data sets. This credit risk score reflects Creditreform's credit assessment of the individual companies. It may vary between 100 (highest creditworthiness) and 600 (lowest creditworthiness, default) and is based on 15 criteria. These criteria include, among others, payment status, credit decision, company development, order book and sector development. Accounting for about 50% of the weighting, payment status and credit decision are crucial to determining the credit risk score.

3   Analyses and Results

3.1 Formation of Classes Based on Relative Default Frequencies

Classification by relative frequency of default (hereinafter referred to as frequency analysis) denotes the construction of rating classes through counting

processes that are directly based on the credit standing index (score). For this purpose, the enterprises are sorted by score and then assigned to a specified number of classes (e.g. class 1: score 100 to 200 etc.). One possible classification approach is to keep the number of companies per class more or less constant (uniform distribution of companies), with operationalization, for example, being effected by applying the following standard: Given ten rating classes, 10% of the companies are assigned to each rating class, while in the case of five rating classes each class contains 20% of the companies, etc. This approach has been adopted in the present study. Another possible variant is the approach pursued in the study by Lawrenz and Schwaiger (2002), which the present study builds on. It requires ex ante definition of the share in total defaults of each rating category (predefined default profile). Both methods use the relative frequency of defaults in the respective class as the estimator for the probability of default (PD).

Our preparatory research comprised the following steps, with the computations based on the update on the New Basel Capital Accord of October 2002 (which incidentally applies with regard to all calculations made): IRB[1]) foundation approach with 45% LGD[2]) and an adjustment for small and medium-sized enterprises (SME) depending on sales volume figures, the latter having been filtered out from the database. Credit line utilization data taken from the Major Loans Register were used as proxy for actual drawings and the default probabilities, owing to the lack of an adequate data

1   *Internal ratings-based.*
2   *Loss given default.*

source and contrary to the requirements of Basel II (see chapter 3.3.), were calculated on the basis of one-year observation periods.

Calculation 1: The capital requirement and the PD structure (i.e. a rating system's structure of default probabilities) were calculated for different numbers of rating classes, leaving conditions otherwise unchanged if possible. The objective was to analyze the effects of a systematic variation in the number of rating classes on both capital requirement and PD structure. These calculations were based on data set 1.

Calculation 2: The capital requirement was calculated and the PD structure was examined over time, based on data set 2. Again, default probabilities were in each case computed on the basis of a one-year observation period.

The method of uniform distribution of companies was applied to ensure comparability of the results: this means that the number of rating classes was varied while an approximately equal number of companies was assigned to each of the classes. The approach based on a predefined default profile (instead of a uniform distribution of companies), by contrast, would in a first step require the definition of a method allowing for a consistent ceteris paribus variation in the number of rating classes.

As to calculation 1, we examined the capital requirement and the PD structure for 5, 7, 10, 12 and 15 classes. As these calculations show, the capital requirement declines steadily as the number of rating classes increases, namely from 6.18% (5 classes) to 5.88% (15 classes).

The PD structure, however, exhibits a monotonically rising pattern only with a system of five classes (with the probabilities of default increasing monotonically from the highest to the lowest class), whereas monotonicity is absent given a system of seven or more classes. This implies that, other things being equal, the number of rating classes cannot be increased arbitrarily without forfeiting certain desirable characteristics of the PD structure, such as monotonicity.

As to calculation 2, based on data set 2, we investigated the behavior of the capital requirement and the PD structure over time. From the total number of observations made we filtered out the individually relevant data sets for each of the years, eliminating, among other things, enterprises that had become insolvent by the beginning of the respective year. We thus obtained differing numbers of data sets for each of the years, with the number of observations increasing from 6,137 for 1996/97 to 9,419 for 2000/01. We set up seven classes, each comprising approximately the same number of companies. The average probability of default increases across the entire observation period from 1.16% in 1996/97 to 1.93% in 2000/01.[1])

As the sales figures required for calculating the capital requirement over time were not available from Creditreform for the years 1997, 1998 and 1999, the SME adjustment for all of these years was invariably computed on the basis of the 2001 sales volume. Drawings were again estimated on the basis of the Major Loans Register data (credit line utilization) for the individual years. The following picture emerges with regard to the capital requirement over time (see table 1):

---

1   As already mentioned, the differences between data sets 1 and 2 with regard to the year 2000/01 are attributable to the different dates at which the data were extracted from the Creditreform database.

Table 1

Capital Requirement Over Time					
	1996/97	1997/98	1998/99	1999/2000	2000/01
Number					
Companies	6,176	6,883	7,653	8,527	9,419
%					
Capital requirement	4.68	5.32	5.40	5.29	5.86
Default rate	1.16	1.23	1.33	1.37	1.93

Source: Creditreform, OeNB, own calculations.

It should be noted in this context that the development of the capital requirement reflects the marked rise in the average probabilities of default in the observation period: for the years 1997/98 to 1999/2000 default rates equaled between 1.2% and 1.4%, while the capital requirement varied between just below 5.3% and 5.4%. By comparison, the default rate for the year 2000/01 was substantially higher at 1.93%, resulting in a corresponding increase in the capital requirement to 5.86% against the previous years. The same applies vice versa for the year 1996/97 by comparison to the subsequent years or for a direct comparison of the years 1996/97 and 2000/01, subject to the reservation that the figures for 1996/97 have to be interpreted with a certain degree of caution owing to the smaller sample size.

The average capital requirement for the five years observed is 5.31%, given an average default rate of 1.41% for the period as a whole. Factoring out 1996/97 from the calculation, the average capital requirement is shown to be 5.47% at an average default rate of 1.47%.

The changes in the PD structure computed on the basis of the frequency analysis show that the form varies over time. The only uniform aspect is a monotonic increase in the probabilities of default across all years from class five downwards. The better classes fail to exhibit a constant structure over time. These observations lead to the conclu-sion that frequency analysis, if calculated on the basis of a one-year observation period, allows the characteristic of monotonicity to be generated in a PD structure just once, but that this feature will be lost over time, all other things remaining equal.

This applies both for the uniform distribution of companies used in calculation 1 and in calculation 2, and the approach based on a predefined default profile. With the latter, a consistent monotonicity over time will only be obtainable if the predefined class-specific default rates are modified for each of the years to reflect the changes in the database, but not if these rates are assumed to be constant. Such an approach would also require the continuous adjustment of the (score) thresholds for the individual rating classes, thus increasing the erratic migration of some companies across different rating classes quite independent of their actual economic situation.

3.2 Logistic Regression

The logistic regression model estimates the relationship between a linear combination of impact factors $\beta_0 + \beta_1 x_1 + ... + \beta_n x_n$ and a dependent variable which may assume only one of two values (in our case default / no default). Furthermore, it is important that the score values (the values of the index $\boldsymbol{\beta'x}$, which can serve as measure for the credit quality of a company) are mapped on the interval [0,1], in which way each score value

relates to a number from the interval [0,1] that can be interpreted as probability of default (n being the number of companies, $\mathbf{x}$ the vector of the independent variable, and $\boldsymbol{\beta}$ the vector of the parameters). The index $\boldsymbol{\beta}'\mathbf{x}$ may serve as a measure of the creditworthiness of a borrower.

For the nondirectly observable variable y^* we have:

$$y* = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n + \varepsilon$$

Let ε have a logistic distribution, with the median 0 and the variance 1 of the distribution being nonrestrictive assumptions. The distribution function of the logistic distribution reads:

$$\Lambda(\mathbf{x}, \boldsymbol{\beta}) =$$
$$= \frac{\exp(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n)}{1 + \exp(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n)}$$

We can directly observe:

$$y = 1 \text{ if } y^* > 0$$
$$y = 1 \text{ if } y^* <= 0$$

The probability that $y = 1$ is hence:

$$P(y = 1) = P(y^* > 0) =$$
$$= P(\beta_0 + \beta_1 x_1 + ... + \varepsilon > 0) =$$
$$= P(\varepsilon > -\beta_0 - \beta_1 x_1 - ... - \beta_n x_n) =$$
$$= P(\varepsilon < \beta_0 + \beta_1 x_1 + ... + \beta_n x_n) =$$
$$= \Lambda(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n)$$

as the logistic distribution is a symmetrical one.

The maximum likelihood optimization method is used to estimate the coefficient vector $\boldsymbol{\beta}$. This method maximizes the probability (L) that the estimated model will reproduce the values observed for y. The logistic function Λ indicates the probability of default, whereas $1-\Lambda$ indicates the probability of survival.

$$MaxL = P(Y_1 = y_1, ..., Y_n = y_n) =$$
$$= \prod_{y_1=0} [1 - \Lambda(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n)]$$
$$\prod_{y_1=1} \Lambda(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n)$$

The logistic model is used to assign a probability of default to each company, to set up classes and to calculate the capital requirement for Austrian SMEs. It is implemented in four steps:

3.2.1 Selection of the Explanatory Variables Based on Data Set 1 with a View to Maximizing the Model's Robustness and Power

Model specification (MS) 1: A constant and the Creditreform credit risk score were included as explanatory terms. In a next step we examined whether the inclusion of other available variables with a potential impact on the credit standing of a borrower had the effect of improving the quality of the model.

Model specification 2: Constant, Creditreform credit risk score and ln(sales 2001) as proxy for the company size.

Model specification 3: Constant, Creditreform credit risk score and a dummy variable indicating the Austrian province to which it pertains.

Model specification 4: Constant, Creditreform credit risk score, ln(sales 2001) and a dummy variable indicating the Austrian province to which it pertains.

Sector membership was also tested on the basis of data set 2 using dummy variables. Application of the same maximum likelihood optimization routine failed to produce a satisfactory solution. One reason for this is that the effect of sector membership is already accounted for by the industry risk variable included in the Creditreform scoring model, which actually obviates the explicit inclusion of industry dummies as it only results in a less satisfactory model specification.

Table 2

Estimates Obtained on the Basis of Data Set I: Coeffizients with the Appertaining Values for t- and p (in parenthesis)

	Constant	Score 2000	ln(sales 2001)	Vienna	Lower Austria	Tyrol	Upper Austria	Salzburg	Vorarlberg	Styria	Burgenland	Capital requirement %
MS 1	− 9.85 − 24.84 (0.0)	0.02 17.2 (0.0)	. .	. .	. .	. .	. .	. .	. .	. .	. .	6.04
MS 2	14.23 − 9.63 (0.0)	0.02 17.14 (0.0)	0.26 3.12 (0.002)	. .	. .	. .	. .	. .	. .	. .	. .	6.50
MS 3	10.38 − 20.43 (0.0)	0.02 16.83 (0.0)	0.73 2.34 (0.02)	0.60 1.78 (0.08)	0.68 2.00 (0.05)	0.66 2.00 (0.05)	0.25 0.71 (0.48)	− 0.41 − 1.09 (0.28)	1.00 2.35 (0.02)	0.13 0.39 (0.70)	. .	5.94
MS 4	14.27 − 9.37 (0.0)	0.02 16.7 (0.0)	0.23 2.75 (0.006)	0.67 2.13 (0.03)	0.58 1.70 (0.09)	0.69 2.02 (0.04)	0.64 1.95 (0.05)	0.25 0.70 (0.48)	− 0.39 − 1.03 (0.30)	1.00 2.34 (0.02)	0.11 0.33 (0.74)	6.27

Quelle: Creditreform, OeNB, own calculations.

It is evident from the t statistic, which shows whether a coefficient is significantly different from 0 ("discriminatory power"), and its p value (the probability that this t value will be observed) that not only the score, but also ln(sales 2001) has a high information content. The joint significance of the province dummies was tested by means of the Wald test (Greene, 1993): The χ^2 statistic yields 1.12, with the critical value of the χ^2 distribution with eight degrees of freedom at the 95% confidence level being 15.50; this implies that the hypothesis that the coefficients of all province dummies are 0 cannot be dismissed. Creditreform does not take into account the province effects in calculating the credit risk score (a certain degree of significance is therefore observed in the above estimates), but a province-specific model specification still requires some further considerations.

In order to ensure a meaningful application of the goodness-of-fit measures in the further process, it is necessary to verify the robustness of the estimation model in the first place. Most of the problems arising with regard to the robustness of a logit-type model are related to heteroskedasticity, as it results in inconsistencies in the estimated coefficients (implying that the precision of the parameter estimate decreases as the size of the sample increases). We applied the statistical test by Davidson and MacKinnon (1993) to test for the hypothesis H_0 of homoskedasticity. The results of this test show that H_0 cannot be dismissed for the model specifications 1 and 3: In the first case, the χ^2 statistic is 0.08, in the second, 14.15, with the critical value of the χ^2 distribution with nine degrees of freedom being 14.68 at the 10% confidence level. In the case of model specifications 2 and 4, heteroskedasticity can be dismissed only at the 10% confidence level.

Model specification 1 hence represents the best model in terms of robustness and significance of the explanatory variables.

The goodness-of-fit tests, by contrast, provide no information on the model specifications, but only with regard to the information contained in the explanatory variables. One of the goodness-of-fit measures implemented in the present study is that of McKelvey and Zavoina (1975):

$$R_{MZ}^2 = \sum_{i=1}^{n}(\hat{y}_i^* - \bar{\hat{y}}_i^*)^2 /$$

$$(\sum_{i=1}^{n}(\hat{y}_i^* - \bar{\hat{y}}_i^*)^2 - n)$$

with n representing the number of observations, $\hat{y}_i^*$ the estimated value of y_i^*, and $\bar{\hat{y}}_i^*$ the corresponding mean value. According to the test results, R_{MZ}^2 being 59%, model specification 4 shows the highest information power.

Another goodness-of-fit measure implemented is the Gini coefficient derived from Gini curves. The Gini curve is created by sorting the companies by their risk, i.e. by their credit standing index. The fraction of defaulted companies $y(x)$ is established for each fraction x of companies with the highest risk score. The Gini coefficient is defined as the area between the model's Gini curve and a random model Gini curve ($y(x) = x$ applying for the random model) divided by the area between the Gini curve of a perfect model and a random model Gini curve. The higher the value of the Gini coefficient, the higher the model's power of differentiating companies according to their creditworthiness. The Gini coefficients vary between 62.08% for model 1 and 66.56% for model 3.

A conspicuous feature is that the models containing the explanatory variable ln(sales 2001) are less well specified, but exhibit a higher information content. The estimation re-sults of model specification 1 (see table 3) also indicate that average sales per rating class generally decrease as the probability of default increases. These findings suggest either that the company size has an impact on the probability of default – but this is already accounted for in the Creditreform score – or that other size indicators such as the number of employees might actually prove to be a more suitable proxy for the company size instead of sales.

3.2.2 Calibration and Formation of Classes

Calibration was based on the assumption that our sample exhibits the same default level as the universe, i.e. that it is sufficiently large and representative.

The classification into classes using the logistic approach was also subject to the rule that approximately the same number of companies had to be assigned to each rating class. Our objective was to calculate the capital requirement in line with the Basel requirements in terms of monotonicity and an adequate number of companies per class, but always restricted to one year at a time. This means that in calibrating the model we renounced the objective of temporal stability (and the implied homogeneity of rating classes).

Table 3 provides a list of the estimated probabilities of default set against the observed relative fre-

Table 3

Results Based on Model Specification I and Data Set I

	Probability of default	Default rate	Average sales volume	Average drawings
	%		EUR	
Rating class 1	0.303	0.359	12,369,335	661,056
Rating class 2	0.454	0.574	10,978,920	447,701
Rating class 3	0.553	0.431	655,008	225,184
Rating class 4	0.737	0.790	582,990	285,106
Rating class 5	1.298	0.790	767,034	343,341
Rating class 6	2.323	2.728	527,660	198,157
Rating class 7	8.402	8.399	394,803	173,116

Source: Creditreform, OeNB, own calculations.

Table 4

Five-Year-Development – Estimates Based on Data Set 2					
	1996/97	1997/98	1998/99	1999/2000	2000/01
%					
Capital requirement	5.29	5.66	5.44	5.19	6.12

Source: Creditreform, OeNB, own calculations.

quency of default (default rates). An interesting observation is that, as seen in columns four and five, average sales volume figures and average drawings decline as creditworthiness decreases.

3.2.3 Variation in the Number of Rating Classes

Based on data set 1, we calculated the capital requirement for different numbers of classes – 5, 7, 10, 12 and 15 classes – for model specification 1 as defined in chapter 3.2.1. We formed classes by assigning an equal number of companies to each class. The results show that the capital requirement declines continuously from 6.08% in the case of five rating classes to 5.96% given 15 rating classes.

3.2.4 Five-Year Development

By analogy with calculation 2, we examined the behavior of the capital requirement and the PD structure over time.

The capital requirement varies between 5.29% in 1997 and 6.12% in 2001, with the average value for the five-year period equaling 5.54%.

3.3 Integrative Considerations

3.3.1 Basel II Minimum Requirements and Their Implications for the Present Study

The basic principles behind the IRB minimum requirements are that rating and risk assessment systems should ensure:
– a well-founded assessment of the debtor and transaction characteristics,
– a meaningful risk differentiation,

– adequately accurate and consistent quantitative risk estimates.

3.3.1.1 Rating Structure

With explicit reference to corporates, banks and sovereign exposures in the present context, a bank must, both in terms of its debtor rating and its facility rating, exhibit a meaningful distribution of exposures across the different grades (without excessive concentrations) – i.e. sufficient to enable adequate risk differentiation. To ensure compliance with this requirement, a bank must have a minimum of seven debtor grades for nondefaulted borrowers and one for defaulted borrowers. The supervisory authority may (in the case of banks with borrowers of heterogeneous debtor quality) demand a higher degree of differentiation.

Banks with credit portfolios concentrating on a specific market segment and a certain default risk range must have a sufficient number of grades within this range to prevent an excessive debtor concentration within any specific grade. Significant concentrations within individual grades must be accompanied by sound empirical evidence to the effect that the grade covers a tolerably narrow PD bandwidth.

Worth discussing in this context is the Basel requirement regarding the minimum number of seven rating grades for nondefaulted borrowers. We compared this minimum number requirement with the results of the present study by mapping rating

classes in conjunction with the monotonicity feature in the PD structure. We observed that the Creditreform data for mapping classes on the basis of the relative frequency of default, given an equal number of companies per class, only yield a monotonically increasing PD structure if the computations are made on the basis of five rating classes. The assumption is that the approach of modeling on a single reference point in time for each of the years would be more likely to yield a monotonically increasing PD structure than a dynamic approach. The observation derived from the Creditreform data is hence evidence of the fact that the number of possible and assignable rating classes is a function of the discriminatory power of the rating system and cannot be treated as a separate feature. The degree of subdivision into rating classes is limited by the discriminatory power of the rating system.

3.3.1.2 Evaluation Horizon and Stress Scenarios

The time horizon used for computing PD estimates was one year. Banks, however, must base their rating assignments on a longer time horizon. The PD estimates must represent a long-term average of realized one-year default rates for the borrowers in the respective grade. A bank may use a simple average of PD estimates for individual borrowers in a given debtor class. Irrespective of whether a bank uses external, internal or pooled data sources or a combination of the three, the series of observations for at least one source must cover a minimum period of five years. If the available observational data from one of the three sources cover a longer period of time and the data are relevant, this longer period must be used. In order to avoid unjustified optimism, the bank must increase its estimates by a conservative margin, the size of which depends on the probable range of the estimation errors.

Given the difficulties of forecasting future events and their effects on the financial situation of specific borrowers, the bank must adopt a conservative stance with regard to forecast information. Furthermore, in the event that only a restricted database is available, the bank must adopt a conservative bias in its analyses.

Bank-internal assessments of the performance of their internal rating systems must be based on a long data history and cover a range of basic economic conditions ideally one or several economic cycles. A debtor rating must reflect the bank's assessment with regard to the capability and readiness of the borrower to meet its obligations even in an adverse economic environment or despite the occurrence of unexpected events.

To comply with this requirement, the bank must base its rating assignments on specific adequate stress scenarios. Alternatively, the bank may meet this requirement without explicit specification of a stress scenario by taking into account the debtor characteristics that reflect the borrower's vulnerability to adverse economic conditions and unexpected events. The range of economic conditions considered in the assessment must make allowance for both the current conditions and the conditions likely to prevail across an economic cycle within a specific sector or geographical region.

The bank must have a regular process cycle of model validation in place, including monitoring of the model's performance and stability as well as testing of model outputs against out-

comes. Model validation must include out-of-time and out-of-sample tests. Moreover, it must state the circumstances under which the model fails to generate efficient results.

As the subject of the present study does not consist in dynamic (i.e. intertemporal) PD estimates, but rather in the analysis of specific situations with reference to a specific point in time, the rules prescribed by Basel II were not applicable in the present context. Still, it is worth mentioning that the long-term averages of one-year default rates required with regard to PD estimates are inconsistent with the approach of mapping classes in accordance with the principle of predefined class-specific default rates. The difference between the two approaches basically manifests itself in the fact that the Basel provisions, owing to the envisaged inclusion of the default rates across the economic cycle, offset the procyclical PD trend (and the implied cyclical fluctuation in the capital requirement). By contrast, if based on predefined class-specific default rates, a constant PD structure for the rating classes and the individual companies cannot be obtained over time for two reasons: on the one hand, the PDs of the individual rating classes vary depending on the number of borrowers assigned to a class and the realized one-year default rates attributed to them. Assignments – based on the ranking derived from the score – are made in accordance with predefined threshold values for the cumulative default rates. Given a classification based on cumulative default rates, on the other hand, companies may be seen to migrate between adjoining rating classes over time.

At this point, we would like to underline that the issue of model validation was not subject of the present study, nor could it have been owing to the insufficient data source. However, there is no doubt that the issue of validating rating systems will assume a crucial role within the context of further system developments and the implementation process.

In the same vein, an IRB bank must have sound stress testing processes in place for determining capital adequacy. Stress testing must involve identifying possible events or future changes in economic conditions that could have unfavorable effects on a bank's credit exposures, as well as assessment of the bank's ability to withstand such changes.

3.3.2 Embedding the Data in a Macroeconomic Framework

In chart 1 the default rates of the Creditreform data used are set against the capital requirement computed by applying the frequency analysis method and given uniform company distribution, and annual GDP growth. The calculations were based on the period from 1993 to 2001, allowing a representation of the data embedded in the cyclical development. What immediately catches the eye is the fact that the real-term decline in growth of about 3 percentage points in the period from 2000 to 2001 entailed an increase by ½ percentage point in the computed capital requirement.

However, as suggested above, the change in the capital requirement is not to be interpreted as an indicator of procyclical behavior, because, since the PD estimates are based on one-year default rates, they fail to meet the required smoothing across the entire economic cycle as required under the New Basel Capital Accord. Chart 1 nevertheless clearly demonstrates that in the absence of this smoothing function, the period 2000 to 2001 would exhibit

Chart 1

Default Rates over the Economic Cycle

Source: Creditreform, OeNB, own calculations.
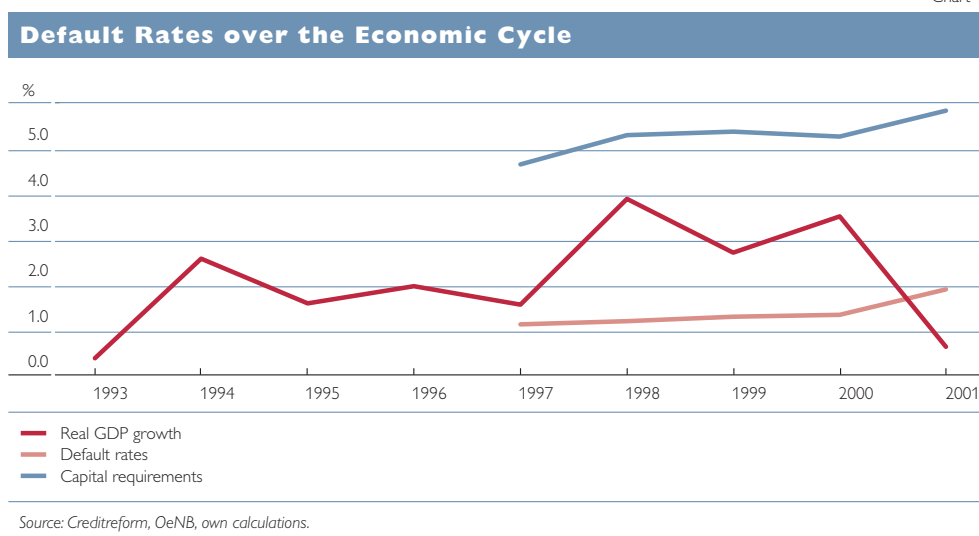
procyclical effects, since the decline in real GDP growth and increased default rates would result in a higher capital requirement.

4 Summary

The computations illustrate the sensitivity of the annual capital requirement to the respective average annual default rate. Any substantial increase in actual defaults, as observed in 2000/01 against the previous years, entails a marked increase in the capital requirement. In order to forestall undesirable procyclical effects, the New Basel Capital Accord prescribes that the default probabilities per rating class must be calculated on a data set comprising several years (a minimum of five years). The frequency analysis method is not a suitable instrument for generating stable and monotonic PD structures over time, hence other methods must be applied. A promising approach consists in the application of logistic regression.

The structure of a rating system's default probabilities (PD structure), generated on the one hand by varying the number of rating classes and, on the other hand, by testing the PD structure over time, yields the follow-

ing picture: The maximum number of rating classes depends on the structure of the underlying data and the goodness of fit (discriminatory power) of the rating. The greater the number of representative data that are available in adequate quality and the higher the discriminatory power of the rating, the larger the number of rating classes with a monotonically increasing PD structure that can be created.

Further steps suggesting themselves for the sequel to the research conducted within the context of the present study include:
1) calibration on the basis of a longer time series,
2) validation of rating systems and
3) dynamic (intertemporal) modeling of the PD structure.

Of these three items, the wide area of validation in particular should be seen as a rather complex task. Although the working group on validation appointed by the Basel Committee on Banking Supervision has been dealing with this issue, it is to be expected that there will be ample scope for discretionary action at the national level.

References

Davidson, R. and J. G. MacKinnon. 1993. Estimation and Inference in Econometrics. Oxford: Oxford University Press.

Greene, W. 1993. Econometric Analysis. New York: Prentice-Hall International.

Kaiser, U. and A. Szczesny. 2002. Logit- und Probit-Modelle für Kreditrisiken. In: Schroeder, M. (ed.). Finanzmarkt-Ökonometrie – Basistechniken, fortgeschrittene Verfahren, Prognosemodelle. Stuttgart: Schäffer-Poeschl.

Lawrenz J. and W. S. A. Schwaiger. 2002. Basel II: Quantitative Impact Study für Österreich (QIS2.5). In: BankArchiv – Zeitschrift für das gesamte Bank- und Börsenwesen 50. 77–89.

McKelvey R. and W. Zavoina. 1975. A Statistical Model for the Analysis of Ordinal Level Dependent Variables. In: Journal of Mathematical Sociology 4. 103–104.