

# New Approaches to Banking Analysis in Austria

Evelyn Hayden  
Jürgen Bauer<sup>1</sup>

This paper describes three new off-site monitoring tools recently developed by the Oesterreichische Nationalbank (OeNB) and the Austrian Financial Market Authority (FMA) with scientific support from the University of Vienna.<sup>2</sup> As logit models currently represent the state of the art in credit score modeling both in the academic literature and in practice, in a first step a logit model was estimated, for which an AUROC<sup>3</sup> of more than 80% was achieved. Next, the logit model was complemented by a Cox model to learn more about the time structure of default probabilities. Finally, a structural model was developed with the aim of showing clear causal connections between a bank's risks and default probabilities. Hence, a system of value-at-risk (VaR) models was constructed for the main risk factors faced by banks, i.e. credit risk, market risk and operational risk, which was brought in relation to banks' potential to cover losses.

## Introduction

The OeNB and the FMA attach great importance to the development and use of powerful off-site analysis tools. Therefore, the two institutions recently began to cooperate in developing a new system of off-site monitoring tools (to be implemented in addition to the systems currently in use).<sup>4</sup> The new tools can be divided broadly into “statistical” and “structural” models. In this paper, the term “statistical model” denotes systems which exclusively use econometric methods to find powerful predictors of bank distress, while “structural” approaches do not only investigate variables highly correlated with default to identify troubled banks but aim to explain banks' risk via economic models, thus offering clear causal connections.

In the category of statistical models, the project team has developed (and still works on improving) logit and Cox models, while on the structural side a system of value-at-risk (VaR) models has been constructed for the main risk factors faced by

banks (credit risk, market risk and operational risk). An overview of all these models is given below, with a special focus on the innovative aspects of the various approaches.

## Statistical Models – The Logit Model

The project team chose a logit model as the principal statistical off-site analysis tool because the results of this type of model can be directly interpreted as default probabilities. Besides, logit models currently represent the state of the art in credit score modeling both in the academic literature and in practice; logit models can easily test whether the empirical dependence between the potential input variables and the default risk is economically reasonable.<sup>5</sup>

The major challenge in developing this type of model was to identify the “correct” definition of default. In Austria, about 1,100 banks existed in the past ten years, and a wide range of quarterly information<sup>6</sup> on these banks has been available in most cases

<sup>1</sup> Evelyn Hayden, OeNB; Jürgen Bauer, FMA (Austrian Financial Market Authority).

<sup>2</sup> Evgenia Glogova, Markus Hameter, Andreas Höger, Johannes Turner, OeNB; Wolfgang Errath, Stephan Unterberger, FMA; Engelbert Dockner, Michael Halling, Alfred Lehar, Josef Zechner, University of Vienna.

<sup>3</sup> The AUROC measures the predictive power of a rating model. A model that discriminates perfectly between defaulters and non-defaulters would receive an AUROC of 100%. For details see Sobehart and Keenan (2001) or Engelmann et al. (2003).

<sup>4</sup> See Turner (2000).

<sup>5</sup> See e.g. Hayden (2003).

<sup>6</sup> The available information includes typical information like balance sheet data, detailed information about large exposures (i.e. exposures of more than EUR 350,000) and various regional macroeconomic indices.

since December 1995. However, since then there have hardly been any cases of actual bank default in Austria, at least far too few to base any statistical model on the relating observations. What is more, all of the few actual defaults were traceable to events that were probably not reflected in the available data before the event of default. Therefore, the project team did not develop a model for true bank defaults, but defined the default event as a situation where a bank was facing such serious trouble that it seemed unlikely to have been able to cope without any kind of intervention (usually in the form of mergers with, or allowances from, affiliated banks). Besides, given this kind of default criterion, the project team found it was unrealistic to declare a bank to be entering into the state of default at the time of intervention, but assumed that the bank must have been in difficulties for at least two quarters before an intervention occurred. Similarly, it seemed probable that troubled banks needed at least two quarters to recover completely after an intervention took place, implying that these banks should have been marked as defaulters for five consecutive quarters. Bearing this in mind, the project team was able to construct a data set of about 33,000 quarterly observations with 750 problematic bank events covering a time period of more than seven years.

In the view of the project team, the number of observable defaults was now large enough to split the available data into an estimation and

a validation sample. To guarantee that the structure of the Austrian banking system – i.e. a few large and many small banks and a concentration in certain banking sectors – was reflected accordingly in both data sets, the 33,000 quarterly observations were split into seven sector groups. Within each such group, large and small banks were separated. Thus, a total of 14 subsamples were generated. In a next step, two thirds of the observations marked as defaulting respectively non-defaulting from each of the 14 subgroups was randomly drawn for the estimation sample, while the remaining observations from all groups formed the validation sample.

Using this estimation sample, 280 candidate model input ratios were constructed. These 280 ratios can be classified according to the 11 risk categories displayed in the table below. After eliminating outliers, testing for the linearity assumption implicit in the logit model<sup>7</sup> and checking whether the univariate relationships between the candidate input ratios and the default event were economically plausible,<sup>8</sup> all variables were tested for their univariate power to identify troubled banks one year before the default event took place. Only those ratios that had an Accuracy Ratio<sup>9</sup> of more than 5% were considered for further analysis. As these included still more than 200 variables, it was not feasible to test all possible model specifications. Hence, another procedure had to be found for selecting the final logit model.

<sup>7</sup> Some input variables had to be linearized using the Hodrick-Prescott filter, see Hodrick and Prescott (1997).

<sup>8</sup> The exact procedures are analogous to those described in Hayden (2003).

<sup>9</sup> The Accuracy Ratio is another measure for the predictive power of rating models; see e.g. Keenan and Sobehart (1999). As illustrated in Engelman et al. (2003), the Accuracy Ratio and the AUROC measure exactly the same information.

One possibility to proceed would have been to determine the most powerful univariate ratio of each of the 11 risk categories and to combine them to form a multivariate model for further analysis. However, when looking at the correlation between the variables of one group, the project team found that – for most categories – not all variables were highly correlated, but that there existed correlation subgroups. This implied that if

only the candidate input ratio with the highest Accuracy Ratio (or the largest area under the ROC<sup>10</sup> curve – AUROC) from each risk category had been included in the model-building process, there would probably have been the risk of ignoring important variables. Instead, the best variable from each correlation subgroup was selected, reducing the list of candidate input ratios to 83.

Table 1

	Number of ratios		
	Original	After univariate and correlation analysis	Final
Bank characteristics	38	7	1
Credit risk	52	15	3
Credit risk based on large exposures	21	5	1
Capital structure	22	7	2
Profitability	41	19	4
Market risk	12	3	–
Liquidity risk	15	5	–
Operational risk	11	1	–
Reputation risk	6	2	–
Management quality	13	5	–
Macroeconomic factors	49	14	1
Total	280	83	12

Then, after further reducing the list of ratios to 56 by eliminating variables that were highly correlated between different risk categories, backward and forward selection methods<sup>11</sup> could be applied to check whether all remaining input ratios were statistically significant or whether the logit model could be reduced to a lower number of input variables. Indeed, the final model only consists of 12 input ratios. Their distribution among the risk categories is displayed in the table above.

The fit of the final model as well as its predictive power with a view to

new data were tested in several ways. To check the fit of the model, the typical statistical tests for logit models like deviance, leverage or the Hosmer-Lemeshow goodness-of-fit tests were applied. The popular concept of the Accuracy Ratio respectively the AUROC mainly served to assess the model's power; however, the project team also adopted the latest procedures to calculate confidence intervals for the above measures and implemented rigorous statistical tests to ensure the superiority of the final logit model compared to other rating methodologies.<sup>12</sup> Besides, the project

<sup>10</sup> The Receiver Operating Characteristic (ROC) curve is a plot of the fraction of defaulters predicted correctly versus the fraction of non-defaulters incorrectly specified as defaulters for all possible cut-off values of the tested model. See Sobehart and Keenan (2001) or Engelmman et al. (2003) for details.

<sup>11</sup> The significance levels were set at 10%.

<sup>12</sup> See Engelmman et al. (2003).

team did not only rely on the results of the validation sample, but also randomly drew further test samples from the total data pool and evaluated the performance of the logit model for these samples, too. The result was that

both the fit and the power of the model were satisfactory and very stable over various data samples. By way of example, the table below shows the AUROC for the estimation and the original validation sample.

Table 2

**The Power of the Logit Model Measured by AUROC**

	AUROC in %	$\sigma_{\text{AUROC}}$	95% confidence interval
In sample	82.87	0.0129	[0.8034, 0.8539]
Out of sample	80.63	0.0210	[0.7651, 0.8475]

Finally, the project team calibrated the estimated model probabilities to obtain “severe problem” and “true default” probabilities. Moreover, these values were mapped into a rating scheme to reduce the variability in the time series of the results for individual banks.

**Statistical Models –  
The Cox Model**

The OeNB and the FMA decided that – using the same data base and the same candidate input variables as described in the previous section – a Cox Proportional Hazard Rate Model should be developed in addition to the logit model. Originally, the reasons for this decision were, on the one hand, the desire to learn more about the time structure of default or problem probabilities (i.e. the survival function of the average “defaulting” bank) and on the other hand, the idea to use the output of the Cox model as a robustness check for the results of the logit model. Later on, however, the project team found an innovative solution of building the Cox model in such a way that it truly complements the logit model.

Whenever Cox Proportional Hazard Rate models are applied to predict bank failures<sup>13</sup> in the academic literature<sup>14</sup>, they are usually set up in such a way that the observation periods for all banks start at the same point in time. In the case of the Austrian data set, where information about banks is available since December 1995, this would imply that the observation periods for all Austrian banks should begin with exactly this date. In this set-up, the Cox model would – based on the available input data – try to separate banks that experience problems at an early stage from those that face difficulties at a later stage (or not at all). Thus, the model would indicate whether banks are “at risk.”

An alternative procedure could be to define a certain cutoff rate for the output of the logit model and to classify Austrian banks as probable “defaulters” and “non-defaulters” (or problem banks and non-problem banks) accordingly.<sup>15</sup> Now only the banks identified as being at risk by the above procedure should enter into the Cox model, where the observation period starts at the point in time when the bank hits the defined cutoff rate for the first time. As the logit

<sup>13</sup> The same is true for the prediction of nonfinancial business defaults.

<sup>14</sup> See e.g. Henebry (1996).

<sup>15</sup> Here, the cutoff rate should be set to a level where (almost) all troubled banks are correctly classified.

model misclassifies some banks (as a statistical model will do per definition), some non-defaulting banks would also enter into the Cox model. This means that in this set-up, those parameters would enter into the Cox model that best predict whether banks that are at risk really default later on. Therefore, as all banks identified as being at risk by the logit model are then reclassified by the Cox model, the combined output of both models probably has a higher Accuracy Ratio<sup>16</sup> than the logit model by itself.

To develop the two types of Cox models, the project team applied procedures similar to those described for the logit model. While the classical version of the Cox Proportional Hazard Rate model is ready for implementation, the development process for the advanced Cox model type has not been completed yet.

### Structural Models

In addition to the statistical models described above, the OeNB and the FMA decided to develop a structural model that should show clear causal connections between a bank's risks and default probabilities. Hence, a system of value-at-risk (VaR) models was constructed for the main risk factors faced by banks, i.e. credit risk, market risk and operational risk, which was brought in relation to banks' potential to cover losses. The individual model components are summarized below.

### Credit VaR

When the project team began to develop a credit value-at-risk model, it examined the usefulness of the three most popular credit risk portfolio models – KMV, CreditMetrics and CreditRisk+<sup>17</sup>. Finally, the team decided in favor of the latter, basing its decision primarily on input data restrictions, as the market input data required for the KMV model are only available for the largest two Austrian banks and the rating data for all individual loans – a crucial input for CreditMetrics – have not yet been made available to the Austrian regulators in the desired quality. However, the OeNB has collected data on large exposures for a number of years, and this information was used as the major input for the following approach.

The project team decided to implement a CreditRisk+ model, in which the available information about the distribution of banks' exposures across various industries was to be utilized. Besides, while the project team felt that under the assumption of fixed default frequencies per industry, the CreditRisk+ approach was too unrealistic, implementing a CreditRisk+ model based on many industries with stochastic default rates seemed, by contrast, too cumbersome for a first version. Therefore, the following type of mixed procedure was adopted.

All large exposures were allocated to 11 broadly defined industries in which the respective borrowers were operating. Besides, as historical default data were available for these industries, the project team was able to calculate individual empirical default frequen-

<sup>16</sup> The output of the Cox Proportional Hazard Rate model comprises relative hazard rates for the observed banks. Just like the default probabilities of the logit model, they can be used to rank banks according to their perceived riskiness; thus, they may serve as the basis for calculating the Accuracy Ratio.

<sup>17</sup> See Crouhy et al. (2000).

cies and standard deviations for all 11 industries. The loss given default, however, had to be set to a fixed percentage for all branches. Next, for each bank all large exposures were allocated to different loss-given-default buckets. By interpreting the historical default rates per industry as expected future default frequencies and by taking the industry composition of the exposures per bucket into consideration, the project team was able to derive the expected number of defaults (and standard deviations) per bucket. Given that since 2003, rating information has been reported for each loan within the monthly statement of large exposures, the project team decided to use this rating information to adapt the probabilities resulting from the information gathered on the individual industries. The adjusted figures were summed up over all buckets, and thus the expected number of defaults of one “meta”- industry was identified. This procedure produced all the information needed to allow for analyzing a CreditRisk+ model with one stochastic process.

However, as small banks usually grant only few large-scale loans, the above procedure was improved, taking small-scale loans into account by way of approximating the total exposure of small-scale loans from the balance sheet data reported in the monthly return. The resulting approximated total volume of small-scale loans per bank was then allocated to the lowest bucket per bank. Finally, under the assumption that all small-scale loans were of similar magnitude, the number of small-scale loans per bank could be approximated by dividing the volume of small-scale loans by the size of the respective bucket. The rest of the approach remained unchanged and was executed as described above.

### Market VaR

The Austrian market value-at-risk model focuses on interest rate position risk, equity position risk and foreign exchange risk. It was implemented as a standard delta-normal approach based on daily variance-covariance matrices for the risk factors. The major challenge concerning this model was collecting the necessary input data for test calculations, as the information banks currently are required to report (especially concerning the equity position risk of large trading books) was not sufficiently detailed.

### Operational VaR

Although the Austrian banks have already begun to collect the data on operational losses necessary to quantify this risk properly, these data have not been available to the regulators yet. However, as the project team agrees with international studies which claim that operational risk is an important risk factor calculating that banks hold up to 30% of their economic capital to cover operational risk, the following work-around based on the Basel II basic indicator approach was developed to include at least a crude approximation of this risk factor in the first version of the structural model.

If one assumes that the frequency of operational loss events is geometrically distributed and if one approximates the loss given event via an exponential distribution, then the total losses attributable to operational risk are also exponentially distributed and can hence be described completely via the identification of only one parameter. As a consequence, the operational VaR can be calculated for any confidence level once this parameter is known. This calculation is based

on the fact that advanced measurement approaches according to Basel II require a 99.9% confidence level in order to calculate the minimum capital requirement and on the assumption that also the basic indicator approach, which is easy to implement, has been calibrated to that confidence level.

### **Aggregation of VaRs**

Once the individual VaRs are calculated, they have to be aggregated to derive one total VaR for each bank. This probably represented the biggest challenge for the project team.

In a first step, the individual VaRs had to be adjusted to represent risk measures for equal time periods, as the credit VaR and the operational VaR were derived for a yearly time horizon, while the market VaR was derived for a daily horizon. As rating agencies usually quote yearly default probabilities and also Basel II favors this time horizon, the project team decided to adjust the market VaR accordingly. To do so, the daily market VaR was scaled up by the square root of 250. The project team felt that this was the best and most consistent procedure, although it realized that the chosen approach probably overestimated market risk, given that banks can easily restructure their trading portfolio in a much shorter time period.

Concerning the actual aggregation of the individual VaR components, the project team mainly evaluated two approaches – aggregation via using a variance-covariance matrix, and application of copulas. However, both methods did not seem to be convincing. On the one hand, the use of a variance-covariance matrix is only theoretically sound if the risk factors are normally distributed, which appears to be ques-

tionable particularly for credit risk and operational risk; moreover, it seems unclear how the covariances can be estimated, especially when taking into account that the composition of market portfolios can be very volatile. On the other hand, the application of copulas is rather cumbersome and it remains questionable whether this level of precision is necessary for the aggregation, given the approximations needed to calculate the individual VaRs. Given these considerations and the view that in case of doubt, an overestimation of banks' default probabilities was preferable to an underestimation, the project team decided in favor of a "conservative" approach, where the overall VaR was defined as the simple sum of the individual VaRs.

### **Banks' Capacity to Cover Losses**

The last step in the structural model is to relate the total bank VaR to the bank's capacity to cover losses. Given the total VaR distribution, one can identify the significance level for which the bank's covering funds are exactly equal to its value at risk. The bank's default probability is then just one minus this figure.

The project team has already performed some test calculations for the structural model for a number of selected banks including large banks which are highly relevant for the entire banking sector as well as smaller network banks. In all cases, the results are of plausible magnitude and hence support the chosen model specifications.

All in all, the project team is convinced that, although the structural model is currently based on a set of simplifying assumptions, the foundations have been laid for a comprehen-

sive model that is able to explain and predict the risks banks face via clear causal connections. The modular structure of the approach favors further improvements of this model as specific components can be updated whenever new data or insights are available without the need to adjust the system as a whole.

### Conclusion and Outlook

The OeNB and the FMA made great efforts to develop a set of modern, powerful off-site analysis tools. Al-

though the predictive power of these new models is already very satisfactory at the current stage, further work will be carried out to improve the results and to keep the statistical tools in line with the latest state of the art. Further details regarding the discussed models will be published in autumn 2004 and the newly revised Austrian Off-Site Analysis System (which will incorporate the models discussed above) will be presented in spring 2005.

### References

- Crouhy, M., D. Galai and R. Mark. 2000.** A Comparative Analysis of Current Credit Risk Models. In: *Journal of Banking and Finance* 24. 59–117.
- Engelmann, B., E. Hayden and D. Tasche. 2003.** Testing Rating Accuracy. In: *Risk* 16. 82–86.
- Hayden, E. 2003.** Are Credit Scoring Models Sensitive to Different Default Definitions? Evidence from the Austrian Market. SSRN Working Paper.
- Henebry, K. 1996.** Do Cash Flow Variables Improve the Predictive Accuracy of a Cox Proportional Hazards Model for Bank Failure? In: *The Quarterly Review of Economics and Finance* 36. 395–409.
- Hodrick, R. and C. Prescott. 1997.** Post-War U.S. Business Cycles: An Empirical Investigation. In: *Journal of Money, Credit and Banking* 29. 1–16.
- Keenan, S. and J. Sobehart. 1999.** Performance Measures for Credit Risk Models. Moody's Risk Management Services.
- Sobehart, J. and S. Keenan. 2001.** Measuring Default Accurately. In: *Credit Risk Special Report*, Risk 14. March. 31–33.
- Turner, J. 2000.** Das österreichische Bankenanalysesystem. In: *Berichte und Studien* 1. OeNB, 90–100.