Modelling Corporate Probability of Default – A Possible Supervisory Benchmark Model*

Márk Szenes – Zsófia Dabi

*In recent years, supervisory bodies around the world have lost some of their confidence in the estimations of credit risk parameters at banks applying the internal ratings-based methodology. Supervisory experience shows that differences in risk metrics and ultimately in regulatory capital requirement levels stem primarily from inconsistencies in the modelling techniques applied and the various methodological approaches, rather than from any actual differences between the inherent risks of bank portfolios. To avoid this unwanted effect, in its supervisory review of banks' internal capital adequacy assessment process, the Central Bank of Hungary (Magyar Nemzeti Bank, MNB) aims at specifying the necessary capital requirements by developing and applying harmonised benchmark models. This study shows how it is possible to estimate a probability of default (*PD*) for corporate portfolios, which is based on large banks' corporate default rate data series and available corporate financial data, uses a harmonised methodology that factors in differences between the credit quality ratings of various customers, and is suitable for the supervisor's calculation of the capital requirement for any given bank. Nonetheless, there may also be other factors in addition to individual financial data (*e.g. qualitative expert elements, sector information*) that may affect credit quality; identifying these may be one of the objectives of benchmark model development.*

Journal of Economic Literature (JEL) codes: C51, G21, G32

Key words: credit risk, probability of default, rating systems, supervisory benchmark model, PD

1. Internal models and supervision

In developing the Basel II framework, one of the primary goals pursued by the supervisory authorities was to strengthen the risk sensitivity of the regulations governing banks' capital requirement calculations (*BCBS 2006*); in addition to certain minimum requirements, the framework permitted institutions to develop their own

* The papers in this issue contain the views of the authors which are not necessarily the same as the official views of the Magyar Nemzeti Bank.

Márk Szenes is a Supervisory Advisor at the Magyar Nemzeti Bank. Email: szenesm@mnb.hu
Zsófia Borbála Dabi is a Financial Modeller at the Magyar Nemzeti Bank. Email: dabizs@mnb.hu

The Hungarian manuscript was received on 16 March 2020.

DOI: http://doi.org/10.33893/FER.19.3.5277

internal models to more precisely pinpoint their own risks, in the hope that this would improve the standards of risk management as well. Institutions which are permitted by their supervisory authorities to use the internal rating-based (IRB) method may calculate their credit risk capital requirement using their own rating systems.

The Basel II system was already criticised by many during its development, but the competent authorities voiced increasingly strong criticism as experience with the actual use of the framework accumulated. As Basel III laid down the liquidity requirements which had hitherto been absent from the regulation, tightened the requirements pertaining to institutions' own funds and introduced the macro-prudential and capital conservation buffers (*BCBS 2010*), the attention of supervisory authorities turned to internal models and specifically to uncertainties in credit risk parameter estimation.

While banks quite naturally called for the option of using as sophisticated methods as possible and freedom in the choice of methodology, supervisory authorities were making efforts to coordinate and strike a balance between the requirements of risk sensitivity, simplicity and comparability (*BCBS 2013*, *EBA 2013*). In recent years it has become clear that the existing regulatory framework allows banks too much leeway in their choice of methodology, as a result of which differences between risk levels based on internal models stem from the differences between the methodologies and approaches applied and the differences between time series available for different institutions, and not from the differences between the risk profiles of the institutions or their portfolios. It is no exaggeration to say that the competent authorities' confidence in the reliability of banks' internal models has been profoundly shaken.

Efforts to harmonise across methodologies and achieve comparability have been and are being made at all levels of supervision. With the finalisation of the Basel 3 package (*BCBS 2017*) at the international level, the Basel Committee no longer allows modelling of the loss given default (LGD) and the exposure at default (EAD) for the segments with the highest model risk (those with low observed default), such as large corporate and bank exposures; in the case of equity exposures, it has retained only the methodology based on simple weighting and has set a lower limit ("output floor") of 72.5 per cent of the standard methodology's capital requirement level for the capital requirement quantified with the IRB methodology.

Indeed, at the European level, the primary task of the European Banking Authority (EBA) is to achieve harmonisation of the prudential rules across Europe, such as the application of the Basel capital rules as well as banks' and competent authorities' practices. In a report published in 2013 (*EBA 2013*), the EBA identified differences between supervisory requirements pertaining to rating systems used by institutions applying the IRB methodology and formulated supervisory guidelines and technical standards for regulatory practices to coordinate such requirements [RTS on assessment methodology (*EBA 2016*), PD/LGD guidelines (*EBA 2017a*)]. Moreover, since 2015 the EBA has been collecting annual benchmark data from banking groups using the IRB approach (*EBA 2015*; *EBA 2017b*; *EBA 2017c*; *EBA 2019*; *EBA 2020*), and in its summary report prepared on the basis of the benchmark data it shows differences in IRB approach to capital requirement levels in a breakdown by portfolio segments. In the analysis, the EBA makes an attempt to pinpoint the possible causes of the differences, using a variety of techniques to differentiate the effects stemming from differences between portfolio compositions and risk profiles from effects that may result primarily from differences between the methodologies applied.

Within the framework of a comprehensive project started in 2016[1] (*ECB 2017*; *ECB 2019*), the European Central Bank, which is responsible for the supervision of the euro area's banking groups, prepares an assessment and a revision of the IRB models of the banking groups under its supervision, in order to minimise differences between modelling methodologies.

In the context of the internal capital adequacy assessment process (ICAAP), the amount of economically necessary capital is determined in Pillar 2 supplementing the minimum regulatory capital requirement of Pillar 1, in order to cover risks stemming from institutions' business activities by estimating possible future losses. Institutions quantify their Pillar 2 capital requirement in accordance with their own internal techniques by calculating the capital requirement for all relevant risk types (including those not handled under Pillar 1). Each year, in the context of the SREP,[2] the competent authority reviews the adequacy of the capital requirement levels calculated by the supervised institutions under the ICAAP. The primary aim of the review process is to scrutinise institutions' risk processes in full detail and to identify all material risk exposures, and thus determine the capital level ensuring solvent operation (*MNB 2020*).

The MNB develops and uses a variety of benchmark models to determine the Pillar 2 capital requirement of the banks present in the Hungarian market (*MNB 2020*), in order to make it possible to measure domestic banks' risks in a risk-sensitive way and by rendering them comparable with one another. The purpose of developing supervisory benchmarks is to enable the MNB to measure banks' inherent risks regardless of banks' definitions, modelling approaches and the data quality of the available historical time series, and thereby to adjust capital requirement levels in the Basel Pillar 2 wherever necessary. Finally, supervisory benchmarks provide the only possibility for determining the risk-sensitive capital requirement

[1] TRIM: Targeted Review of Internal Models
[2] Supervisory Review and Evaluation Process

for institutions operating without advanced and reliable internal models, most of which are spillover institutions.

Lessons drawn from losses resulting from retail loans failing in huge numbers in connection with economic downturns showed how the inhomogeneous modelling techniques had been causing unwanted differences between risk levels, but it was the same large number of observations that enabled the development of the MNB's retail PD and mortgage LGD benchmark models.

2. Review of the relevant literature

From the supervisory perspective, the assessment of corporate portfolio risks is just as important as that of retail portfolio risks. In this study, corporate portfolio risks are approached from the aspect of the probability of default. Hungarian and international literature both feature a wide variety of scientific articles and papers dealing with the modelling of corporate default. The methodology of bankruptcy prediction has developed considerably in recent decades; while earlier on, analysts used to apply various discriminant analysis models (initially univariate, later multivariate), as the years went by logit (logistic regression) and probit regression analyses gained popularity. Logit and probit models are also widely used in the development of rating systems meeting the requirements of the Basel "through-the-cycle" approach. Mention should also be made of the most recent methodologies used in bankruptcy prediction, such as decision trees, neural network, machine learning, artificial intelligence and hybrid models which combine the advantages of various models, thereby improving model performance (*Kristóf – Virág 2019*).

Our overview of the relevant literature focuses on the Hungarian models and studies which are most pertinent in relation to this paper, without aiming to present an exhaustive review. In Hungary, the first corporate model which used time series input variables and met the requirements of the through-the-cycle (TTC) approach was published by *Imre* (*2008*), who modelled the occurrence of 90-day-past-due defaults using actual Hungarian corporate data observed between 2002 and 2006, with the help of the decision tree, logistic regression and neural network methodologies. The logit model developed by Imre uses 11 variables including – similarly to the model discussed herein – indicators relating to capital structure, debt servicing, liquidity, profitability, as well as working capital and asset turnover.

Madar (*2014*) also applied logistic regression in developing his corporate rating model, which – in line with the Basel requirements – is also suitable for estimating the long-term probability of default (PD) and for capital requirement calculation. Data from domestic SMEs that prepared balance sheets during the period between 2007 and 2012 were used in the modelling process (not including false businesses formed out of necessity or other technical types of businesses). After a review of the strength of more than 40 financial indicators and ratios, the Weight of Evidence (WoE) transformed versions of 6 indicators came to be finally used as variables. Indicators describing capital structure were the strongest variables in the observed sample. Liquidity and profitability indicators also showed significant discriminatory power and were thus also incorporated in the model. In the study, the author describes how the rating system developed for the given population provides a stable PD value that is crisis resistant in terms of its discriminatory power and stable over the long-term and also presents proof of the fact that the more accurate discriminatory power a given rating system has, the more closely it will follow the varying default rate values, as a consequence of which the PD increases during a crisis and thus has a pro-cyclical, crisis-aggravating effect.

Banai et al. (*2016*) connected and used data from the Central Credit Information System (KHR) and businesses' financial reports for 2007–2014 to model the probability of default of micro, small and medium-sized enterprises. The banks' default is the model's dependent variable, i.e. in the analysis the authors examined 30-day-past-due items that had been so for at least 60 days (90+ days past due). In addition to company-specific variables and category variables, the model also includes macro-variables capturing unexplained heterogeneity over time, along with a trend for adjusting default events. In addition to modelling separated by size categories, the authors specified separate models for certain high priority sectors of the national economy with the aim of analysing relationships produced for functioning companies. Their results show that most variables behave in a similar way in these models as well, but the focuses are shifted by the effects of industry or sector characteristics. The estimated PDs show that agricultural companies have the lowest credit risks and construction has the highest credit risks among the sectors reviewed, in line with the results presented in *Section 6* of this study.

Bauer and Endrész (*2016*) estimate probability of bankruptcy for Hungarian companies using a probit model, combining micro and macro variables. Macro information needs to be integrated into the model, in order to capture the aggregated dynamics and the risk level. The estimate was based on the complete 1996–2012 time series of all domestic businesses applying double-entry bookkeeping (approx. 1.5 million observations). The model's target variable is bankruptcy from a legal aspect, identified on the basis of information available in the Opten database. Similarly to the studies discussed so far and the model presented in this article, the model of Bauer and Endrész also includes profitability, liquidity and debt servicing indicators and it takes account of heterogeneity in terms of company size. It is, however, different for instance on account of using foreign ownership and exporting activity as dummies, along with the inclusion of macro variables (GDP growth, credit growth) in the model. The need for the latter is explained by the authors' claim that macro variables can capture shock effects that are not reflected by company level variables, along with spillover effects.

A similar approach was used by *Inzelt et al.* (*2016*), who also estimated bankruptcy from a legal perspective, using the same set of data (Opten and NTCA databases). One major difference is, however, that while Bauer and Endrész aimed to develop a model with strong predictive power, suitable for forecasting future negative events, Inzelt et al. wished to present a simple, stable and easy-to-use corporate monitoring framework for comparing internal models used by credit institutions. This paper aims to improve on the model presented in the study published by Inzelt et al., in a way that it provides a reliable input PD parameter for determining the IRB-based capital requirement of corporate loan portfolios. To this end, in Section 3 we present a detailed discussion of the similarities and the differences between the two models.

3. The corporate PD model framework

Inzelt et al. (*2016*) presented a possible approach for the corporate portfolio, a kind of PD model for monitoring and measuring inherent corporate credit risks. To continue with its development, we changed their models in a number of aspects, as we wish to use our model to estimate banks' long-term probabilities of default regarding their corporate portfolio. For this very reason, we focused on modelling not the entire domestic non-financial corporate sector, but only corporate customers with bank loans/limits, excluding project financing companies and micro-enterprises in the retail segment.[3] In the case of projects, the project asset's[4] cash-flow generating capability and the sponsor's strength need to be explored, while in the case of financial enterprises, modelling would require the identification of the risks associated with the underlying portfolio in particular, but that is not possible from financial data.

One of the most important elements of the PD model is the definition of default. Since bankruptcy, liquidation, etc. procedures do not cover – in terms of timing or definition – the default definition of banks, and since banks also develop their capital models on the basis of the Basel default definition, we also relied on the default databases provided for the MNB in the course of the supervisory review process, in the context of banks' data supply.

Focusing on the stability of the model and the estimated PDs, we paid attention to making sure that the risk segment (micro, small, medium-sized, large enterprise) of a given customer is fixed and that changes in the performance of the company, particularly its decline before default results in no change in the segment to which it is assigned; therefore, we fixed the companies' segments on the basis of the historical maximums of their sales revenue, balance sheet total and headcount data. The quantity of data also made it possible to prepare a model for the large

[3] Projects – e.g. commercial real estate financing exposures – tend to have profoundly different risk profiles than financial enterprises, and these segments are modelled separately by banks as well, therefore our model does not cover these types of exposures.
[4] Typically real estate

corporate segment, which is the smallest segment in terms of the number of entities it contains, but at the same time is the most important one in terms of the magnitude of risks.

The model to be presented links the negative event (default) and the explanatory variables based on the balance sheet and the profit and loss statement via logistic regression; however, the range of the variables taken into account was expanded significantly in comparison to the range of variables used in the model of Inzelt et al. Another change is that while our model uses the same regression coefficients across all segments, the PDs associated with the score were calibrated separately for each segment, making it possible to estimate corporate PDs that adequately reflect the long-term default rate and that can be used for capital requirement calculations. The main differences and similarities between the two models are presented in detail in *Table 1.*

Table 1
Comparison of the two models

	Inzelt et al. (2016)	**Adjusted model**
Corporate data used	All non-financial enterprises registered in Hungary, using double-entry book-keeping	Only normal companies financed by domestic large banks (excluding project loans, retail micro enterprises, micro enterprises with product-based financing, and financial enterprises)
Negative event	Negative legal events (liquidation proceedings, bankruptcy proceedings, court deregistration proceedings, completed liquidation, compulsory winding-up)	Bank default events
Model development sample period	1999–2013	2006–2017
Segmentation	Based on current sales revenue	Fixed segmentation on the basis of historical maximums of their sales revenue, balance sheet total and headcount data
Modelled segments	Micro, small and medium-sized enterprises	Micro, small and medium-sized enterprises and large corporates
Negative event explanatory variants	In the case of micro enterprises: four indicators; in the case of small and medium-sized enterprises: two indicators, from the following: debt burden, long- and short-term liquidity position, productivity indicator	Six indicators in each segment: long- and short-term liquidity position, profitability indicator, leverage, debt coverage, size
Model	Logistic regression	
Calibration	Separate logistic regression by segment	One logistic regression for each segment, but separate PD calibration for each
Application	Supervisory monitoring: comparing risks, analysing changes	IRB-based capital requirement calculation, Pillar 2

Source: Inzelt et al. (2016) with our own supplements

4. Data used

The corporate PD model was based on corporate default databases collected by major domestic banks and banking groups with advanced risk measurement methodologies (consequently, with long, reliable time series). Not only banks using the IRB approach under Pillar 1, but also most large banks collect credit risk loss and default data (for the calculation of the Pillar 2 capital requirement (ICAAP)) which they use in their rating systems.

For the development of the corporate PD model, we performed a variety of data filtering and cleansing routines:

- In the first step, we only retained normal corporate customers and aimed at separating all segments with radically different risk profiles, particularly projects and financial enterprises, which necessitate a profoundly different modelling approach and for which even banks themselves develop separate rating systems. Micro portfolios with product-based financing under retail management were not integrated into the model, because we found that on the one hand they have significantly higher default rates than other similar-sized micro enterprises under corporate management, and on the other hand their risks may also be affected by product attributes which we do not wish to take into account in a general corporate model.
- An annual customer-level database was prepared, in which each company appears in the modelling database only once a year (provided it had a period of performing status during the given calendar year), regardless of whether it was financed by more than one bank.
- A company was regarded as having defaulted if it was in default with at least half of its financing banks. Default events – as the target variable of our corporate PD model – were registered in the year in which they occurred. In the case of customers with multiple banks, we checked whether this choice causes any significant distortion because in the case of customers financed by multiple banks each of the banks concerned tended to register default events; in general, differences appeared in the timing of the default. In the case of multiple default events when customers kept shifting between performing and non-performing status over the years, the default events were combined into a single default event and assigned to the date of the first default event.
- In the case of large corporates, manual data cleansing was performed for the defaulting entities, by also checking the appropriateness of assignment to the default category based on publicly accessible data.

Figure 1
Default rate time series calculated from the large bank default databases used

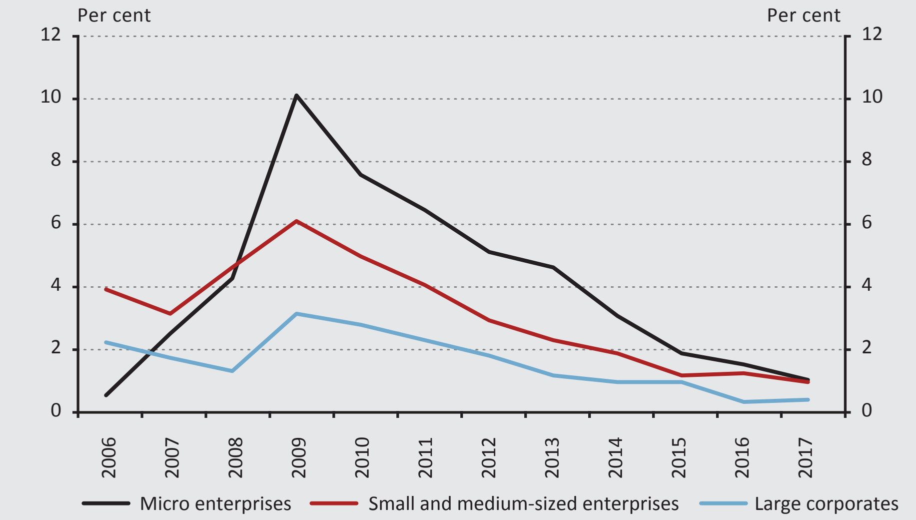


Figure 1 shows the calculated default rate time series for the different company size ranges as a result of the filtering and data cleansing processes.

As the default database is not complete for the year 2006, with full data not available for the micro segment, only the data from after 2007 were used for model building (*Table 2*). The final database includes 286,000 observations per year and per customer, including some 10,000 default events. The annual averages of the default rates were 4.33, 3.00 and 1.49 per cent for the micro, small/medium-sized and large corporate segments, indicating that model calibration was necessary for each segment.

Table 2
Composition of the default database used in modelling

Segment	Number of customers per year	Number of defaults	Average default rate
Micro enterprises	98,727	4,385	4.33%
Small and medium-sized enterprises	174,318	5,386	3.00%
Large corporates	13,400	211	1.49%
Total	286,445	9,982	3.38%

Source: Calculated on the basis of banks' default databases

The explanatory variables required for predicting the probability of negative events – default – over the long-term were generated from the balance sheet and profit and loss statement data of the company information database and the required segmentation was also devised on the basis of headcount, sales revenue and balance sheet total figures taken from the same source. *Act XXXIV of 2004 on Small and Medium-sized Enterprises, and the Promotion of Their Development* (SME Act) was taken into account as a basis for the definition of the micro, small and medium-sized enterprises, as well as the large corporate segments, however the 'and' relationship was not required in the conditions regarding headcount. The HUF amounts we calculated as equivalent to the amounts in euros to be found in *SMA Act* are presented in *Table 3*.

Table 3
Definition of size-based segmentation

SME category	Headcount (no. of employees)		Annual net sales revenue (HUF million)		Balance sheet total (HUF million)
Micro enterprises	< 10	or	≤ 300	or	≤ 300
Small enterprises	< 50		≤ 2,000		≤ 2,000
Medium-sized enterprises	< 250		≤ 15,000		≤ 15,000
Large corporates	≥ 250		≥ 15,000		≥ 15,000

Source: 2004. évi XXXIV. törvény a kis- és középvállalkozásokról, fejlődésük támogatásáról (Act XXXIV of 2004 on Small and Medium-sized Enterprises, and the Promotion of Their Development)

Segmentation based on the current headcount, sales revenue and balance sheet total may result in significant migration between segments. Switches from segment to segment may be particularly problematic when it is a consequence of the declining economic performance (falling balance sheet total, profit, headcount) of the company facing problems before defaulting, because in this case the default would be shown in a size category smaller than that of the customer's original segment, resulting in underestimation of the larger segments' default rates. Therefore, in defining the modelling segments we used the maximum of the headcount, sales revenue and balance sheet total figures from 2000 on, so that where the given customer belonged to a larger size category according to any one of the indicators on the basis of which segmentation is determined, such higher category was regarded as the customer's final segment (i.e. the indicators are in an "or" relationship with one another). This method enabled the customers' segment to be fixed for the entire modelling time horizon.

5. Rating system – PD model

The purpose of the benchmark model is to assign to each company the particular PD value which best reflects the long-term average default rate of companies with similar risk profiles, across successive cycles. Moreover, a model is expected to clearly separate exposures based on risk exposure and risk profile, distinguish good companies from poorly performing ones and enable the monitoring of changes in portfolio quality driven by non-systemic factors (i.e. changes independent of economic cycles). Estimating PD parameters as independently as possible from cycles (TTC-type PD parameters) is important for a variety of reasons. On the one hand, the IRB capital requirement calculation requires an unconditional PD value as an input parameter, and on the other hand both the European Central Bank's guideline for the evaluation of internal models (*ECB 2019*) and the European Banking Authority's guideline on PD and LGD estimation (*EBA 2017a*) require TTC-type calibration of the PD parameters, i.e. the estimate must reflect the long-term average default rate. Moreover, from a supervisory aspect it is necessary to assess risks independently of cycles and thus set up a stable capital requirement that is not sensitive to economic cycles (to avoid underestimation during an upswing or overestimation during a downturn).

Rating systems based on logistic regression have been widely adopted by banks in practice for distinguishing by risk profile; therefore just like Inzelt et al., the authors of this study also opt for this approach. By linking historical default events and the explanatory variables characterising the customer's risk profile by a function in the model, it is through regression that we determine the weight and coefficients of the explanatory variables as detailed below (with x_i as explanatory variables and β_i as weights/coefficients):

$$\text{Default probability} = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i)}} \quad (1)$$

In selecting the explanatory variables, we primarily used the variables found in the study published by *Inzelt et al.* (*2016*), and we made our selection on the basis of objective financial indicators found in large banks' corporate models that can be generated from balance sheets and profit and loss statements. We aimed to select variables with high explanatory power and also ensure that we use simple, economically meaningful variables from each major group of variables. The indicators were selected from the following main indicator groups, taking into account their correlations as well: indebtedness/capital leverage, liquidity position, balance sheet structure, debt coverage, profitability and size.

It should be emphasised that banks' experts and analysts usually have much more information concerning companies' credit quality, as compared to the information that can be extracted solely from companies' financial data. Banks' corporate models usually have an expert module as well (in addition to the financial module), containing the above mentioned expert factors. The management's/owner's expertise and commitment, which may be reflected even through the involvement of private guarantees, the company's market position and the industry's outlooks may all contribute to the model's explanatory power. Selecting such factors is, however, clearly complicated by their heterogeneity across banks and their subjectivity, but their prospective integration may be the target of future development of the model to be presented.

The initial list of large banks' variables consisted of about 50 different financial variables; this was reduced by correlational analysis to a total of 6 variables, including 1 of each of the above groups of variables. One of the key considerations in the selection of variables was the aim to dampen the model's Point-in-Time (PiT) nature as far as possible, and therefore in the case of profitability type indicators we avoided the use of profit (loss) before taxation, while a negative profit (loss) before taxation figure is one of the strongest indicators of default. For the most part, highly similar indicators were defined in the main indicator groups, e.g. in the case of the capital leverage-type indicators either the shareholders' equity or the balance sheet total was typically adjusted (e.g. for intangible assets). In these cases we chose the simpler options. The financial indicators used in the model were defined as follows:

$$\text{Long-term liquidity} = \frac{\text{If(long-term liabilities=0, }-1\text{, long-term liabilities)}}{\text{tangible assets + financial investments + intangible assets}}$$

$$\text{Short-term liquidity} = \frac{\text{cash and liquid assets + securities}}{\text{short-term liabilities}}$$

$$\text{Profitability} = \frac{\text{material + personnel + other expenditures}}{\text{sales revenue}}$$

$$\text{Leverage} = \frac{\text{shareholders' equity}}{\text{balance sheet total}}$$

$$\text{Debt coverage} = \frac{\text{operating profit + depreciation}}{\text{long-term + short-term liabilities}}$$

$$\text{Size} = \text{sales revenue}$$

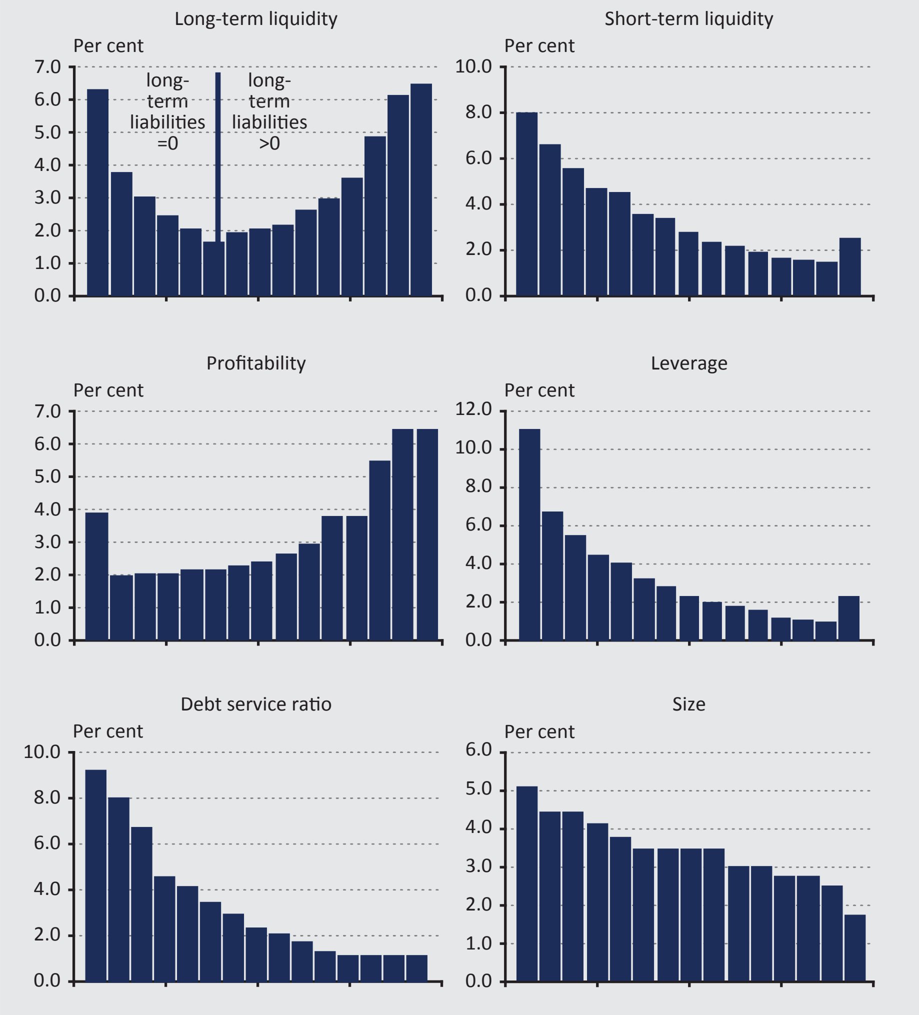
The long-term liquidity position indicator had to be split because the balance sheets of a significant proportion of obligors included only short-term liabilities.

Size, however, was taken into account not only through the segments in the model but also as a variable, by fixing the historical maximum for each company over the period starting from 2000. The use of the maximum value enables avoidance of excessive cyclic patterns by ensuring that the customer's quality does not deteriorate more in the case of a decrease in its sales revenue than the deterioration caused by the current sales revenue decrease already reflected in the profitability ratio itself, so in this case again, the goal of achieving a TTC model was given priority over increasing the explanatory power.

The explanatory power of each indicator was analysed during the selection of variables. Explanatory power means the extent to which it is possible to separate good (non-defaulting) customers from bad (defaulting) customers on the basis of the given indicator. A continuous indicator separates customers effectively when the observed default rate is monotonous for the indicator and there is a large difference between the default rates of the customers with the best and the customers with the worst financial indicators.

The model's discriminatory power was assessed in two different ways: customers were first assigned on the basis of the indicators to 15 categories, with the same number of customers assigned to each category. The default rate within each category was checked and where the relationship was not monotonous – this could be observed only in the case of some neighbouring categories – the categories

Figure 2
Financial indicators of the model divided into 15 categories and the average default rates calculated within each category



concerned were combined. The missing values were assigned to a separate "missing" category. The default rates calculated on the basis of the final categories of variables are presented in *Figure 2*, while the missing values were assigned to the first or last categories.

Even at first glance, *Figure 2* shows that debt coverage and leverage are the most powerful variables, with an 8–10 time difference between the default rates in the lowest and the highest variable categories. Size appears to have the weakest explanatory power where this difference is less than threefold. Size, however, will be of relevance particularly in the large corporate segment. The way this is taken into account is specifically discussed during the model's calibration.

We then also measured the Gini indices of the various variables, the metric most often used by banks for measuring explanatory power. Instead of raw variables, however, the model uses the WoE values, which are widely used for automatically dealing with non-linearities, extreme values and missing values and can be calculated for the categories of variables. The weight-of-evidence was calculated for each category as described below:

$$\mathrm{WoE}_i = \ln\left(\frac{\text{ratio of non-defaulting customers assigned to category i to all non-defaulting customers}}{\text{ratio of defaulting customers assigned to category i to all defaulting customers}}\right)$$

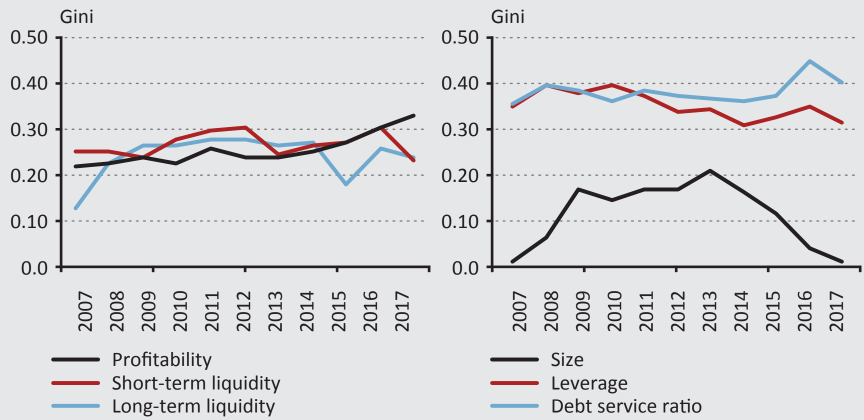
Since in our model we used the WoE values calculated for the above 15 categories, the Gini indices were also calculated on the basis of the same WoE variables. The explanatory power values in each segment, as characterised by the Gini index, calculated for the financial variables used in the model, are presented in *Table 4*.

Table 4
Gini indices characterising the explanatory power of each financial indicator used in the model, by segment

Segment	Gini index					
	Long-term liquidity	Short-term liquidity	Profitability	Leverage	Debt coverage	Size
Micro enterprises	0.23	0.24	0.24	0.33	0.33	0.06
Small and medium-sized enterprises	0.24	0.38	0.29	0.45	0.46	0.08
Large corporates	0.17	0.31	0.38	0.31	0.55	0.13
Total corporate portfolio	0.26	0.30	0.26	0.40	0.40	0.14

One important question is how stable over time each variable can be regarded, i.e. whether they have reliable explanatory power in the long-term. This is of particular relevance to periods during which large numbers of defaults occurred. The "crisis resistance" of the model is shown by the Gini value of each variable during years with high default rates, i.e. 2009–2013 according to the modelling database. Explanatory power is less relevant to years with low default rates, because the small numbers of defaults are caused by factors (we regard as idiosyncratic factors) that have only little negative impact on the model's long-term performance. Based on *Figure 3* therefore we can declare also that size – the variable with the smallest explanatory power – has the highest Gini values during the years between 2009 and 2013. The larger the size category, the smaller the number of companies it holds, and the greater the explanatory power of size is; and since the bulk of bank exposures is associated with medium-sized and large enterprises, the inclusion of this variable is all the more important if we are to build up a rating system that is well aligned to actual observations.

Figure 3
Changes over time in the explanatory power of each financial indicator



The indicators were primarily chosen with a view to minimising overlaps between the balance sheet and profit/loss data used to establish them. We also carried out correlational analyses to establish the extent to which each variable can be expected to add to the discriminatory power. If the ranking order set up on the basis of one variable is very similar to the ranking order established using another variable, the integration of the two variables cannot be expected to add much to the model's explanatory power in comparison to just using only one. This type of relationship is measured by rank correlation, the result of which is presented in *Table 5*. The largest overlap is found between the ranking set up on the basis of debt coverage and the one based on leverage and profitability. The 0.55 correlation between debt coverage and leverage can be regarded as adequate, considering the explanatory power of each. An even stronger (0.59) correlation was found between debt coverage and profitability; profitability has low explanatory power on the whole, and therefore only a little extra value could be expected for the whole model. It should also be taken into account, however, that this indicator is one of the best variables in the large corporate segment, and therefore we decided to retain it.

Table 5
Spearman's rank correlation coefficient among financial indicators used in the model

	Long-term liquidity	Short-term liquidity	Profitability	Leverage	Debt coverage	Size
Long-term liquidity	1.00	–0.11	–0.01	–0.24	–0.10	0.16
Short-term liquidity	–0.11	1.00	–0.20	0.46	0.39	–0.14
Profitability	–0.01	–0.20	1.00	–0.25	–0.59	0.11
Leverage	–0.24	0.46	–0.25	1.00	0.55	0.00
Debt coverage	–0.10	0.39	–0.59	0.55	1.00	0.00
Size	0.16	–0.14	0.11	0.00	0.00	1.00

After selecting and examining the explanatory power and correlations for the variables, we carried out the model's logistic regression alignment. As we have shown, instead of the financial indicators themselves, their WoE values, allocated to 15 categories, were assigned as explanatory variables to the default indicators. The logistic regression was carried out using the SAS software; the Wald test results show that each variable is highly significant. In terms of the Gini coefficient, the model has an explanatory power of 0.507, a very good result for a model using purely financial indicators and covering the complete range of companies in terms of size. Our supervisory experience shows that Gini indices over 0.6 are produced only by models which also use some behavioural variables or other variables based on more recent financial indicators than those based on annual reports.

Although the incorporation of behavioural variables and current information in the model would have increased the explanatory power, this would also have made the model pro-cyclical which we wished to avoid since our model is to be used for capital calculations. In addition to capital calculation, however, some relevant risk management and risk monitoring considerations require banks to monitor current information. Early intervention is one of the most effective means for mitigating risks and minimising losses; and impairment also has to reflect the current prospects.

The coefficients and significance of the explanatory variables are presented in *Table 6*.

Table 6
Results of the logistic regression in the SAS software and Wald test of the coefficients

	Coefficient's value	Standard error	Estimate's significance (p-value)
Intercept	–3.3181	0.0112	<0.0001
Long-term liquidity	–0.5012	0.0247	<0.0001
Short-term liquidity	–0.6093	0.0208	<0.0001
Profitability	–0.1654	0.0276	<0.0001
Leverage	–0.3739	0.0179	<0.0001
Debt coverage	–0.5125	0.0216	<0.0001
Size	–0.5018	0.0411	<0.0001
Gini value		0.507	

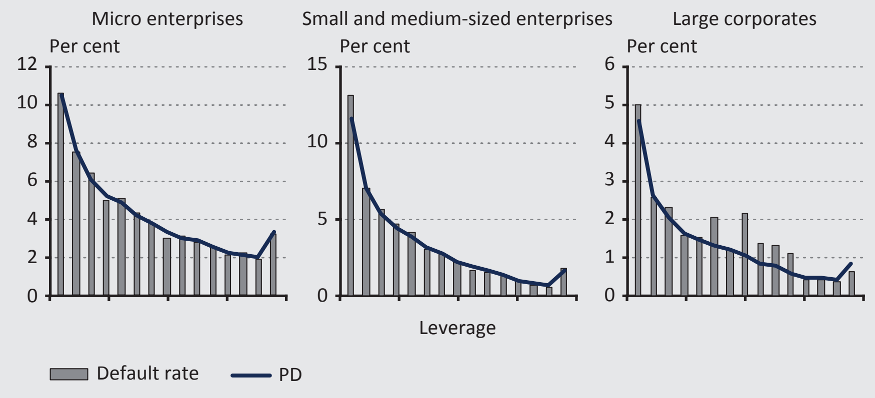
The larger the company, the more important size is as a variable. However, when establishing the 15 size categories with the same number of companies in each rank, all of the companies with HUF 15 billion in sales revenue are added to the largest size category, because of the small number of large enterprises. In response, we supplemented the above model in the case of the large corporates with a continuous size variable, defined as the historical maximum of the sales revenue and the balance sheet total. Thereafter, we produced the natural logarithm for this value. The "score" $\left(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i\right)$ value as per the above model and the size variable as explanatory variable were used in the large corporate logistic regression which resulted in a Gini of 0.603 (see *Table 7*).

Table 7
Result in the large corporate segment of supplementing the model with size; the value and significance of each coefficient

Large corporate size calibration	Coefficient's value	Standard error	Estimate's significance (p value)
Intercept	5.2099	0.8157	< 0.0001
Score (from the corporate model)	1.1727	0.0886	< 0.0001
ln (max (sales revenue, balance sheet total))	–0.3144	0.0492	< 0.0001
Gini value		0.603	

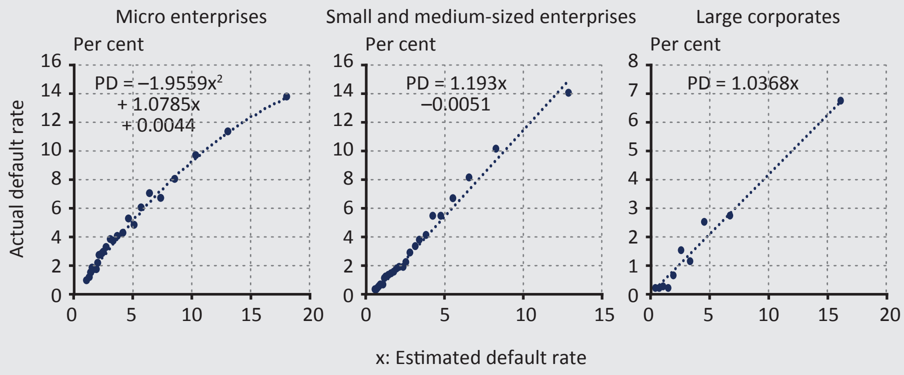
As we have seen, the different explanatory variables for each corporate size segment explain the default event probability to varying extents, while each financial indicator can be regarded as adequate for each size range. Two different solutions can be used to tackle this: on the one hand, the segments could already have been separated along with the logistic regression process, which would have resulted in different sets of coefficients for each segment, or, on the other hand, the single model can be calibrated separately for each segment. Having reviewed both options, we chose PD calibration by segment, because in this case explanatory power and alignment corresponding to those produced by separate logistic regression can be achieved, while the single logistic regression provides a simpler and more robust model. At this point, we only mention the fact that using the same set of variables and coefficients we built up an effectively aligned model during supervisory reviews in the retail micro segment with product-based financing, merely by recalibrating the model. *Figure 4* illustrates – using the leverage indicator as an example – that even by mere PD calibration we can achieve excellent alignment by segment and there is no need for separate logistic regression models. Another possible direction for continued model development is examining whether different financial indicators may be the best explanatory factors in the different size segments, in which case even specific models could be developed for each segment. There are also examples in banks' practices for separate modelling for large enterprises and medium-sized enterprises using different sets of variables.

Figure 4
Alignment of calibrated PD to the default rates by calibration segment, subject to leverage



PD is calibrated by plotting the actual default rate subject to the default rate generated by logistic regression, modelled in accordance with function (1). The PD calibration function was defined with the adequately chosen regression function between the modelled and the actual default rates; its result yielded the final PD parameter values that can be used for IRB capital requirement calculation. The alignment and the PD calibration functions are presented in the charts in *Figure 5*.

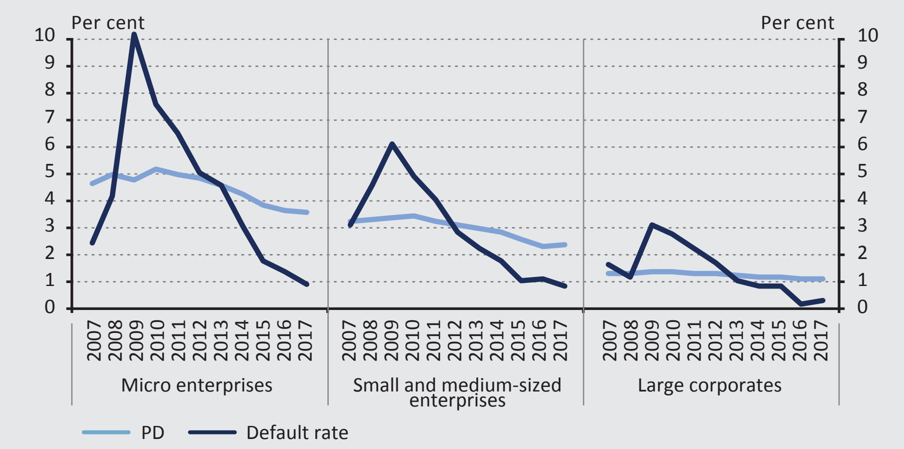
Figure 5
Alignment of PD calibration functions on each corporate segment



6. Results

From the PDs calculated with our corporate benchmark model, we can conclude that the model produces a stable result, closely aligned to historical default rates (*Figure 6*). The time series of the estimated PDs show no such cyclical pattern as does a Point-in-Time model which includes behaviour variables as well, and which cannot capture companies long-term credit quality by following fluctuations in the annual default rates. Significant improvement is evident, however, in the PD time series; this may stem not only from the improved composition of the financed portfolios and from idiosyncratic effects, but also from favourable – cyclical – effects of an economic upswing on financial indicators.

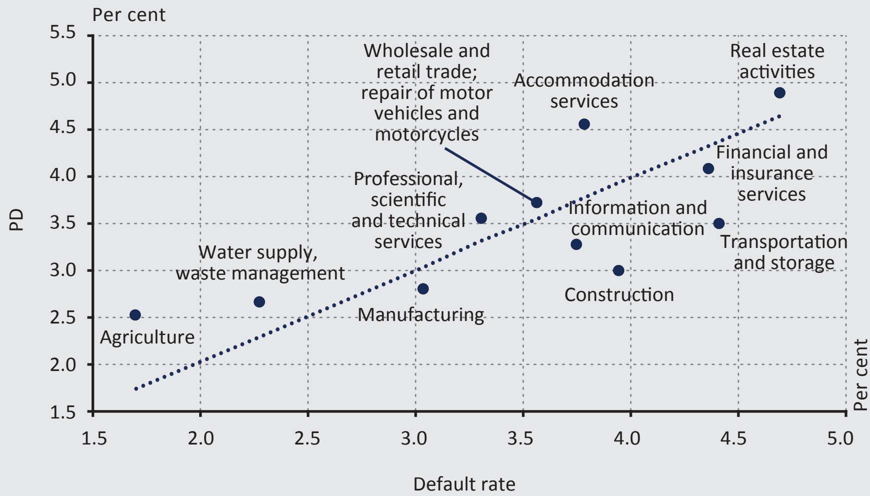
Figure 6
Historical default rates and PDs calculated with the benchmark model, by segment



Another important decision point in the development of the model was whether industry should be included in the set of explanatory variables. Our considerations were: 1) if the average default rates of the various industries are captured sufficiently closely by the model based on purely financial indicators during backtesting, and 2) if the model reflects the risk ranking order by industry, then we do not incorporate it in the PD model, because in this case we can say that industry specifics are already reflected by the selected financial indicators. *Figure 7* shows how effectively the results of the PD model reflect industry specifics. Agriculture carries the lowest risks even in the PD as per the model, while the highest PD is attached to the real estate activities which is the segment with the highest default rate. Although the PDs calculated for construction and for transportation and storage, are below the relevant actual default rates, the size of the difference and the number of such cases did not necessitate the integration of industry as a variable in this model. Based on just one cycle one cannot expect the long-term default rate to be reflected in the case of every single industry, and therefore we wished to avoid 'over-fitting' the model, that is, modelling relationships that may not actually exist.

It is also possible however, that different factors have different influences on credit quality in the various basic sectors – e.g. production, service provision, trade – but this would take additional, even deeper, analyses that would go beyond the limits of this article.

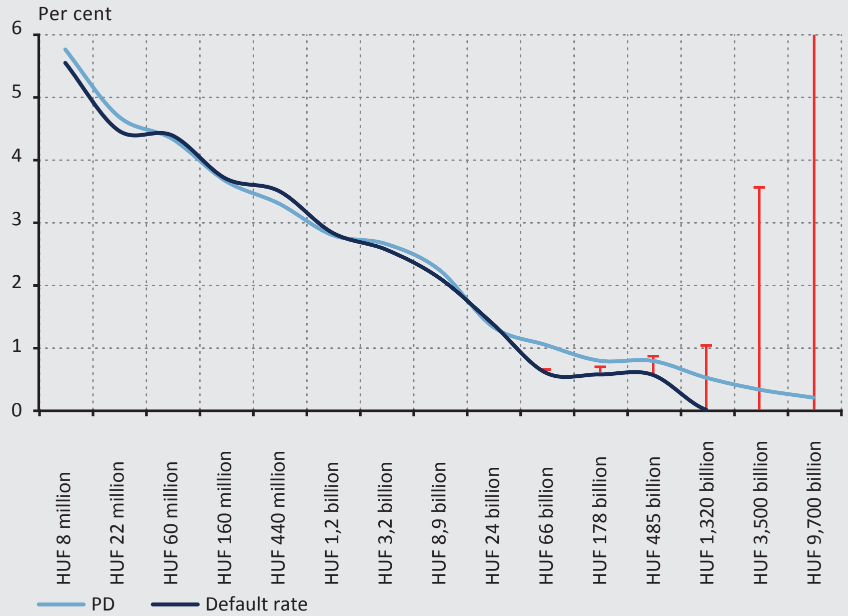
Figure 7
Estimated sectorial average PDs versus observed default rates



In our view, one of the key merits of our benchmark model is that it can be effectively applied to an extremely wide size range. We have managed to develop a robust model with remarkable explanatory power not only for the micro segment which comprises a large number of businesses and which can therefore be efficiently modelled, but also for the large corporate segment. The supervisory authority has a difficult time assessing a PD model with a small number of defaults, one that may even rely on data of foreign banking group members via the parent bank, but by combining domestic banking system data used in our model and with the PD model calibrated on it, it became possible to carry out quantitative assessments of large corporate PDs as well.

Figure 8 shows how accurate the alignment between the actual default rate and the modelled PD is across the entire size category. There are so few companies and defaults in the largest size category (> HUF 50 billion) that even a single default can cause a significant shift in the default ratios. For this very reason, the calculated PD values are considered to be adequately conservative in this category.

Figure 8
Alignment of the corporate benchmark PD to the actual average default rate across the entire size range



Note: Size was logarithmically shown, the band of error was plotted (in red) with a single additional default.

The results, backtested by banks and portfolio segments, show that the historical default rates of the various portfolio segments are in close correlation with the calibrated PDs in the case of individual banks as well, i.e. banks' specific risk management practices and qualitative elements do not systematically and materially deflect the credit quality level from the level that would be implied by financial data alone.

Capital requirement calculation is the single most important use of our corporate PD benchmark model. A supervisory benchmark PD model is an important tool for checking, and, where necessary, revising the PD estimates of institutions already using internal models; or for directly establishing the Pillar 2 capital requirement in the case of institutions that do not have their own internal models. Our benchmark model was also tested on large banks' analytical credit data collected during the supervisory review process in 2019 also by comparing benchmark PDs with banks' own PD estimates and we examined the differences between the IRB capital requirements calculated with banks' PDs and the IRB capital requirements calculated with the benchmark PDs.

Our results (*Table 8*) show that in 2019 banks' corporate PD estimates were closely aligned with the benchmark PDs presented in this article, and there were, on the whole, negligible differences between the IRB capital requirements calculated using them. In 80 per cent of the various banks' portfolios (in terms of size), the differences between the banks' PDs and the benchmark PDs were within 10 per cent, with the maximum difference falling in the 20–25 per cent band in both the negative and the positive domains.

Table 8
Comparison of large banks' own PDs and the benchmark PDs, as well as the IRB capital requirements (HUF billion and %)

Segment	Exposure (HUF billion)	Institution's PD	Benchmark PD	IRB: capital requirement calculated with institution's PD (HUF billion)	IRB: capital requirement calculated with benchmark PD (HUF billion)
Micro enterprises	209	3.89%	4.02%	14	15
Small and medium-sized enterprises	2,250	2.35%	2.72%	144	152
Large corporates	2,881	1.29%	1.10%	198	189
Total	5,339			356	357

7. Summary

In this study, we describe the process of building a corporate benchmark PD model – based on the model developed by *Inzelt et al.* (*2016*) – that can be reliably used for capital requirement calculations as part of the supervisory review process. The model has excellent explanatory power and is well aligned to the observed default rates in each size category and industry: we are convinced that in the large corporate segment it yields PDs more reliable than those estimated by banks' models. Our model enables the homogeneous, consistent measurement of all corporate portfolios at all banks. This benchmark model makes it possible for the MNB to calculate the capital requirement in a risk-sensitive manner even for institutions without advanced rating systems, or whose rating systems are not reliable. Accordingly, the authors of this article managed to successfully apply the benchmark PDs even in calculating small banks' capital requirements.

Use revealed certain shortcomings of the model which can only be resolved on a case-by-case basis. Regarding groups of companies, integration of the group/parent company PD into the final PD with an adequately conservative weight may be contemplated. The financial indicators of holding companies and companies established for the purpose of acquiring shareholdings do not always adequately reflect the given company's risks. In such cases, expert judgement may be exercised and may result in a revised PD. In the case of large companies with an international background, EBA benchmark PD values – available for MNB – may also be taken into account.

Based on lessons drawn from supervisory reviews and model validation procedures, we are aware of the main differences between the PD benchmark model presented herein and banks' PD models. Banks' corporate models always include qualitative ("soft") elements in addition to objective financial indicators, including management's experience, customer track record, companies' market positions, etc. Group/parent company influence may also be taken into account in banks' models, and deflection by experts ("overruling") may also play a substantial role. Without disputing the value added by such expert elements, we emphasise that our benchmark model provides an adequate risk level on average, which may be revised in individual cases. Supplementing the model with qualitative perspectives may significantly increase the model's explanatory power, and therefore this may be an important development direction.

Finally, a note on PD models' PiT/TTC aspects. Current financial indicators are bound to add a cyclical element to PD estimates. To be able to measure risks regardless of cycles and to avoid the customary underestimation and overestimation of risks during upswings and downturns, respectively, the model definitely needs improvement towards the TTC direction. 'More TTC' may be added to the model by a variety of solutions, whether by calculating averages of financial indicators or based on the relations between the variables during the given year, but these might be examined in another study.

References

Banai, Á. – Körmendi, Gy. – Lang, P. – Vágó, N. (2016): *Modelling the credit risk of the Hungarian SME sector*. MNB Occasional Papers 123, Magyar Nemzeti Bank.

Bauer, P. – Endrész, M. (2016): *Modelling bankruptcy using Hungarian firm-level data*. MNB Occasional Papers 122, Magyar Nemzeti Bank.

BCBS (2006): Basel Committee on Banking Supervision: *International Convergence of Capital Measurement and Capital Standards*. http://www.bis.org/publ/bcbs128.pdf. Downloaded: 2 May 2020.

BCBS (2010): *Basel III: A global regulatory framework for more resilient banks and banking systems*. Basel Committee on Banking Supervision. https://www.bis.org/publ/bcbs189_dec2010.pdf. Downloaded: 2 May 2020.

BCBS (2013): *The regulatory framework: balancing risk sensitivity, simplicity and comparability*. Discussion paper, Basel Committee on Banking Supervision. https://www.bis.org/publ/bcbs258.pdf. Download date: 2 May 2020.

BCBS (2017): *High-level summary of Basel III reforms*. Basel Committee on Banking Supervision. https://www.bis.org/bcbs/publ/d424_hlsummary.pdf. Downloaded: 2 May 2020.

EBA (2013): *Report on the comparability of supervisory rules and practices*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/documents/10180/15947/aab59482-8c84-44b7-8d5f-8a588eb7276e/20131217%20Report%20on%20the%20comparability%20of%20supervisory%20rules%20and%20practices.pdf?retry=1. Downloaded: 2 May 2020.

EBA (2015): *EBA Report Results from the 2014 Low Default Portfolio (LDP) Exercise*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/documents/10180/950548/b3adc4f7-653d-408d-b950-da2e4229e294/EBA%20results%20from%20the%202014%20Low%20Default%20portfolio%20%28LDP%29%20exercise.pdf?retry=1. Downloaded: 6 March 2020.

EBA (2016): *Final Draft Regulatory Technical Standards: on the specification of the assessment methodology for competent authorities regarding compliance of an institution with the requirements to use the IRB Approach in accordance with Articles 144(2), 173(3) and 180(3)(b) of Regulation (EU) No 575/2013*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/documents/10180/1525916/e8373cbc-cc4b-4dd9-83b5-93c9657a39f0/Final%20Draft%20RTS%20on%20Assessment%20Methodology%20for%20IRB.pdf. Downloaded: 6 March 2020.

EBA (2017a): *Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/documents/10180/2033363/6b062012-45d6-4655-af04-801d26493ed0/Guidelines%20on%20PD%20and%20LGD%20estimation%20(EBA-GL-2017-16).pdf. Downloaded: 6 March 2020.

EBA (2017b): *EBA Report Results from the 2016 High Default Portfolios (HDP) Exercise*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/documents/10180/15947/7d246d06-4fca-4662-8a8e-71106008c4c8/EBA%20Report%20results%20from%20the%202016%20high%20default%20portfolio%20exercise%20-%20March%202017.pdf?retry=1. Downloaded: 6 March 2020.

EBA (2017c): *EBA Report Results from the 2017 Low Default Portfolios (LDP) Exercise*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/documents/10180/15947/8dd027a7-08d9-490a-87d3-b5cd0ea9c9bf/EBA%20Report%20results%20from%20the%202017%20LDP%20Credit%20Risk%20benchmarking.pdf?retry=1 Downloaded: 6 March 2020.

EBA (2019): *EBA Report Results from the 2018* Low and High *Default Portfolios Exercise*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/documents/10180/2087449/431db2b7-559a-4a64-b3e8-0451edaebfef/EBA%20Report%20results%20from%20the%202018%20Credit%20Risk%20Benchmarking%20Report.pdf?retry=1. Downloaded: 6 March 2020.

EBA (2020): *EBA Report Results from the 2019* Low-*Default* and High-*Default Portfolios Exercise*. European Banking Authority. https://eba.europa.eu/sites/default/documents/files/document_library/Publications/Reports/2020/EBA%20Report%20-%20Results%20from%20the%202019%20Credit%20Risk%20Benchmarking%20Exercise.pdf Downloaded: 6 March 2020.

ECB (2017): *Guide for Targeted Review of Internal Models*. European Central Bank. https://www.bankingsupervision.europa.eu/ecb/pub/pdf/trim_guide.en.pdf. Downloaded: 2 May 2020.

ECB (2019): *ECB guide to internal models*. European Central Bank. https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.guidetointernalmodels_consolidated_201910~97fd49fb08.en.pdf. Downloaded: 2 May 2020.

Imre, B. (2008): *Bázel II definíciókon alapuló nemfizetés-előrejelzési modellek magyarországi vállalati mintán* (*2002–2006*) (*Default-prediction models based on Basel II definitions, for a sample of Hungarian companies* (*2002–2006*)). PhD thesis, Miskolc University.

Inzelt, Gy. – Szappanos, G. – Armai, Zs. (2016): *Supervision by robust risk monitoring – a cycle-independent Hungarian corporate credit rating system*. Financial and Economic Review, 15(3): 51–78. https://en-hitelintezetiszemle.mnb.hu/letoltes/gyorgy-inzelt-gabor-szappanos-zsolt-armai.pdf

Kristóf, T. – Virág, M. (2019): *A csődelőrejelzés fejlődéstörténete Magyarországon* (*Historical development of Hungarian bankruptcy prediction*). Vezetéstudomány (Budapest Management Review), 50(12): 62–73. https://doi.org/10.14267/VEZTUD.2019.12.06

Madar, L. (2014): *Scoring rendszerek hatásai a gazdasági tőkeszámítás során alkalmazott portfóliómodellek eredményeire* (*The effects of scoring systems on the results of portfolio models used in economic capital calculations*). PhD thesis. University of Kaposvár.

MNB (2020): *A tőkemegfelelés belső értékelési folyamata* (*ICAAP*)*, a likviditás megfelelőségének belső értékelési folyamata* (*ILAAP*) *és felügyeleti felülvizsgálatuk, valamint az üzleti modell elemzés* (*BMA*) (*The Internal Capital Adequacy Assessment Process* (*ICAAP*)*, the Internal Liquidity Adequacy Assessment Process* (*ILAAP*) *and their Supervisory Review, and the Business Model Analysis* (*BMA*)). Magyar Nemzeti Bank. https://www.mnb.hu/letoltes/icaap-ilaap-bma-kezikonyv-2019-final.docx. Downloaded: 6 March 2020.