Credit Risk Modelling of Mortgage Loans in the Supervisory Stress Test of the Magyar Nemzeti Bank*

András Viktor Szabó

The study aims to develop a model that can estimate potential credit risk losses for housing and home equity loans using both macro and micro data, can be applied uniformly to all banks and takes into account the new accounting standards (IFRS 9). The model is based on a deal-level database for several Hungarian credit institutions, covering an entire business cycle (2004–2018). It uses economic indicators that strengthen risk sensitivity while also including transaction characteristics that mitigate procyclicality. Modelling in a two-step process allows risk groups to be created during forecasting in accordance with various credit characteristics. The results show that the evolution of employment has a stronger effect on riskier groups which potentially have only ad-hoc employment, while net wealth was not even among the explanatory variables for the group containing the best debtors, who presumably rely more on stable earned income.

Journal of Economic Literature (JEL) codes: C320, C530, G210, G280, G510

Keywords: time series modelling, econometric forecasting, bank, stress test, PD, household credit, IFRS 9

1. Introduction and literature review

Stress testing and the corresponding credit risk models gained increasing prominence in the wake of the 2008 global economic crisis. This is attested by the introduction of stress tests by international banking supervision bodies as well as the supervisory function of the *Magyar Nemzeti Bank* (the central bank of Hungary, MNB) since 2017. This study seeks to develop a model that can be applied

* 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.

András Viktor Szabó is a Head of Division at the Magyar Nemzeti Bank. Email: szaboand@mnb.hu

I am indebted to the Applied Research and Stress Testing Department of the MNB for granting me access to the cleansed database used in the modelling, and to all the staff members at the Magyar Nemzeti Bank who contributed to this study with their comments. Responsibility for any remaining mistakes lies exclusively with the author.

The Hungarian manuscript was received on 15 September 2020.

DOI: https://doi.org/10.33893/FER.21.1.56

uniformly to all banks, takes into account the new accounting standards (IFRS 9), and can estimate potential credit risk losses for housing and home equity loans to households using both macro and micro data. The model is based on a transaction-level database for several Hungarian credit institutions, covering an entire business cycle (2004–2018), and uses economic indicators that strengthen risk sensitivity while also allowing deal characteristics that mitigate procyclicality to be included in the modelling. One of its main unique aspects is that modelling is conducted in a two-step process, which allows different macroeconomic variables to exert a varying impact on risk groups created according to different credit characteristics, i.e. debtors in heterogeneous situations and risk buckets. The results confirm this theory because, for example, the evolution of employment has a stronger effect on the riskier groups who potentially have only ad-hoc jobs, while net wealth was not even among the explanatory variables of the group containing the best debtors, who presumably rely more on stable earned income.

Experts at credit institutions and regulatory authorities realised the risks these institutions face to due to their lending activity decades ago. Towards the end of the 20th century, as quantitative methodologies were being developed, portfolio-level modelling of credit risk became increasingly popular, leading to a wealth of studies and referenced models. Among reduced credit risk models, which estimate the parameters of a default through an exogenous jump process rather than through the change in market capitalisation (*Bielecki – Rutkowski 2004*), the literature differentiates between two types. The first comprises intensity models, concerned with the time of default, while the second one, which is more interesting for the purposes of this work, includes models based on credit migration. Since the 1990s, most models focused on estimating three parameters: the probability of default (PD), loss given default (LGD) and the correlation between PD and losses (*Crouhy et al. 2000*). *Das et al.* (*2009*) observed that the models estimating PD based on the concept of default as understood by Merton[1] started to be crowded out by reduced-form models in the early 2000s, as they could include any number of explanatory variables, even client-specific or macroeconomic variables, which further improved the accuracy of the estimate. Of course, this came at the same time as the publication of the Basel II accords in 2004, which underscored the need for banks' internal credit risk assessment, thereby encouraging earlier methods to be renewed and made more precise.

[1] In Merton's approach, a highly stylised model yields the probability of default, and the only explanatory variable is the value of the company's assets.

The idea of credit rating modelling is attributed to *Altman* (*1968*), who used accounting characteristics in his study[2] in an attempt to estimate the PD of various firms. This train of thought was continued and fine-tuned by several others over the following years, for example, by *Martin* (*1977*), *Platt – Platt* (*1991*) and *Sommerville – Taffler* (*1995*), to name but a few. *Lawrence et al.* (*1992*) did not model companies' probability of default; instead they approached the problem from the side of household lending, which makes it partly similar to the methodology used in the present study. The main objection to credit scoring models is that their explanatory variables are pieces of static, accounting information that are unable to immediately capture sudden changes, as they do so only with a delay (*Agarwal – Taffler 2008*). Having recognised this, credit risk experts increasingly turned towards factor models in the early 2000s. These models aimed at information compression usually use two vectors for the estimation. The first typically includes mostly accounting information that deals with rapid effects too rigidly, but captures client quality characteristics well, while the second vector complements it by mainly including macroeconomic indicators supporting dynamism. If it can be considered unchanged over time, the first (e.g. type of loan repayment, time to maturity, educational attainment of the client at the time of application) is mostly used for exploring and quantifying individual risks, while the latter can capture external, systemic risks. Many authors have published articles describing factor or multi-factor models. Among these, *Pederzoli and Torricelli* (*2005*) deserves special mention, in which the authors described the contradictory relationship between risk sensitivity and procyclicality in connection with the Basel II capital requirement calculation, and proposed to mitigate it with a hybrid PD model using both risk group-based (rating) method and a TTC (through-the-cycle) approach. The present study also presents a modelling practice using two types of vectors or set of variables.

The 2008 global economic crisis opened up the eyes of financial market participants to the fact that the modelling methods used until then were unable to provide an accurate picture of the banking sector's credit exposure and fell short in qualitative or quantitative terms in terms of their approach. The unanimous wish of the stakeholders in the financial sector was clear: among other things, they called for the IAS 39 accounting standard applicable back then to be replaced. Several publications, including *Chae et al.* (*2019*) discuss the shortcomings of the earlier, backward-looking standard that only supported provisioning for loss events that had already happened. The authors claimed that the credit risk losses suffered across the banking sector in huge amounts of discrete packages represented a major threat to financial stability, in an especially volatile and tense situation. To eliminate

[2] Altman Z-score: working capital/assets, retained earnings/assets, earnings before taxes and interest/assets, market value of shares/debt, sales revenue/assets.

such procyclical provisioning, the *International Accounting Standards Board* (IASB) introduced the currently applicable accounting rules in January 2018. IFRS 9, which underpins this study, also stipulates a forward-looking provisioning rule set that established three risk stages and is based on expected losses. This allows credit institutions to prepare for potential crises by increasing reserves and opens up new avenues for credit risk modellers in terms of developing various stage-migration models (see *Landini et al. 2019, Gross et al. 2020*, for example).

The literature is quite limited when it comes to presenting the credit risk simulation models used in stress testing, which is partly due to the practice only recently having been introduced and partly to the small number of entities conducting this activity. Stress tests assessing credit institutions' profitability and risk profile are typically run by banks themselves or the bodies supervising them, and thus the publications also come from this small group. The methodology used in practice which is most relevant for the European banking sector is provided by the *European Banking Authority* (EBA) and the *European Central Bank* (ECB), principally intended for internal use by national supervisors. *Gross et al.* (*2015*) offers relevant experience in terms of modelling, with their presentation of the practical application of the Bayesian model averaging. Ideas can also be gained from the practices of other European national competent authorities, among which the Dutch central bank's publication, *Daniels et al.* (*2017*), stands out due to its similar mortgage loan modelling methodology. The present paper adds to the stress testing literature best reflecting the features and risks of the Hungarian banking sector, using Hungarian data from the household sector, in contrast to the majority of publications that have mainly focused on corporate clients until now. With respect to the latter, *Lang and Stancsics* (*2019*) and *Horváth* (*2021*) certainly deserve mention. They address the corporate segment of the credit risk section in the MNB's stress tests. While the former approaches the banking sector from the macroprudential side, establishing stages based on the number of days past due and then estimating transitional probabilities with the model, the latter uses a logit model utilising client and macroeconomic data to augment the framework of the supervisory stress test.[3] Presentation of the literature relevant for this study ends with the comparison to the model in *Banai et al.* (*2014*), which is similar to the methodology described below, inter alia, in terms of the database used and the incorporation into the stress testing framework. The differences are partly due to the fact that while the aforementioned authors include the deal and client characteristics enabling risk categorisation in the same model as the time series variables, in this work it was considered better to include them in two different models on account of the

[3] For a summary on the supervisory stress testing framework, see the MNB's latest methodological handbook, 'The Internal Capital Adequacy Assessment Process (ICAAP), the Internal Liquidity Adequacy Assessment Process (ILAAP) and their supervisory review, and the Business Model Analysis (BMA)'.

methodological features of supervisory stress testing. The other main difference lies in the predictor variable, owing to the introduction of the new accounting practices discussed above. While *Banai et al.* (*2014*) focused their model the probability of default, this paper examines the sensitivity of the transitional probabilities between the stages.

The paper is structured as follows: *Section 2* presents the risk categories based on loans' deal and client characteristics, describing the database used for modelling, the selection of variables as well as the result of clustering. After this, the time series modelling framework of the estimated PDs along the resulting risk categories is illustrated, before shifting the focus to the evaluation and backtesting of the results in *Section 3*. The main thrust of the paper is *Section 4*, describing the steps of converting the modelled PDs into transitional probabilities across stages. *Section 5* summarises the conclusions.

2. The credit risk classification framework

In line with the above summary, the first part of this section mainly focuses on presenting the parameters of the database used and the characteristics of the available variables, to provide a solid basis for establishing the risk clusters described in the second half of the section and the time series models in later sections.

2.1. The database used and data cleansing

The modelling uses a database containing three of the eight largest credit institution groups operating in Hungary based on balance sheet total, using their individual reporting on client and deal characteristics at the time of application, while also incorporating features that change over time. The data were shared by the three institutions with the MNB for research purposes. The dataset includes all loans to households from the three banks in question between December 2004 and December 2018, covering an entire business cycle. Development of the deals over time can be traced at a quarterly frequency. Among the submitted data, mortgage loans include both housing loans and home equity loans, the latter of which comprised a larger share of household loans when the 2008 global economic crisis erupted, but still account for over 10 per cent of the total volume, making their detailed credit risk modelling warranted. However, a large portion of the mortgage-backed loans include housing loans, which also account for a massive share within total household lending, at around 50 per cent. Unsecured household loans and those with collateral other than real estate are not covered by this paper. The database includes more than 9 million observations, with the developments in around 370,000 individual deals over time.

The descriptive statistics of the variables can be found in *Table 4* of the *Annex*, showing that several deal and client characteristics required data cleansing due to missing and/or outlier, probably misreported, values. Data cleansing was conducted using three main strategies. In the case of the variables where the modelling was complicated only by missing values and this only concerned a negligible number of cases (no more than one-thousandth of the observations), the observations were deleted. With the variables where outlier values were also included besides missing information, the appropriate distribution was achieved by rescaling the data points and dragging them back to the percentile yielding realistic values.[4] The third technique, which potentially influenced the modelling the most, was used heavily in the case of the variables where the share of missing values was not large in percentage terms but above the negligible limit. In such cases, following the best-performing method based on *Little and Rubin* (*2002*), which looked at data cleansing techniques, the data missing in the quarters in question were filled in with the average of the variables concerned[5] observed in the relevant period.

According to *Acuña and Rodriguez* (*2004*), the treatment of missing data becomes problematic at over 5 per cent of the total sample and only affects the interpretation of the results from over 15 per cent. Fortunately, no shortcomings in the variables crossed any of these thresholds, and although data cleansing was performed to improve the accuracy of the results, no major effect is attributed to this from here on. For a comprehensive description of the data cleansing techniques, readers are referred to the descriptive statistics.

2.2. Predictor and explanatory variables

In addition to the study's direct credit risk contribution to supervisory stress testing, another aim is to confirm the assumption that the household loans disbursed in Hungary (in this case: mortgage loans) have improved considerably compared to the situation before the last financial crisis in terms of the distribution of risks. This may be driven by various factors, in particular the government and regulatory measures from recent years,[6] but banks' risk appetite has also changed in the meantime. The impact of debt cap regulations[7] on household lending was addressed by, among others, *Fáykiss et al.* (*2018*), pointing out that they did in fact achieve their goal, and thus the structure of lending was preserved, while the riskiest loans were crowded out. The secondary objective of the paper is to approach this phenomenon,

[4] Client age was limited between 18 and 80; the thresholds were 0–100% for the PTI and 0–200% for the LTV; the loan amount was capped at HUF 140 million, while the maturity period was limited to 40 years.
[5] Such missing data were only observed in the case of the PTI, the LTV and the transaction interest rate.
[6] E.g. forced conversion of FX loans, debt cap regulations (loan-to-value ratio, payment-to-income ratio)
[7] MNB Decree No. 32/2014 (IX. 10.) on the regulation of the payment-to-income and loan-ro-value ratios

combined with the change in banks' overall risk appetite, from the perspective of mortgage loans' PD, which is expected to show a decreasing trend outside the PIT (point-in-time) view, meaning a lower PD even in the TTC approach in recent years.

At the time of publication, the stress test used by the MNB's supervisory function calculated credit risk loss under the assumption that the loans amortising over the two-year stress scenario are replaced with the same risk profile. In practical terms, this means that if a 30-year-old mortgagor with a college degree and net income of HUF 300,000 at the time of the loan application repays their loan during this two-year period, they are replaced with a debtor with equivalent characteristics, which guarantees the stability of the risk composition and the loan portfolio. Of course, the size of the loan portfolio can also be tweaked in the stress scenario by adding the appropriate dynamics in line with the macroeconomic environment, just as in the stress test, however, the portfolio's risk profile stays constant in line with the methodology in use at the time of publication. To warrant any change in this, first it has to be known whether such risk consolidation was observable over time, and its extent is also important to develop the appropriate methodology. With this in mind, instead of applying panel regression estimation using macroeconomic and individual deal and client characteristics (e.g. pooled OLS), a two-step modelling approach was used, making the above-mentioned phenomenon better observable and more easily measurable. This represents one of the greatest differences as compared to the household PD model of *Banai et al.* (*2014*).

First, the loans were classified based on their underlying risk and then time series modelling was applied to them. Initially, an accurate definition of default, a variable measuring risk well, had to be given (as the database did not contain such a field), and it had to be quantified, which was eventually calculated with the formula:

$$Default = \begin{cases} 1, if\ DPD_0 < 90\ and\ (DPD_1 \geq 90\ or\ DPD_2 \geq 90\ or\ DPD_3 \geq 90\ or\ DPD_4 \geq 90); \\ 0 \end{cases}$$

where the subscript shows the number of quarters since the starting point, and DPD (days past due) denotes the number of days elapsed since the loan repayment fell due. Thus, the variable can take one of two values (the transaction either defaults or not): it takes 1, meaning a default, if a client which was performing at the start of the period is late on their loan repayment by at least 90 days any time during the next year, irrespective of whether the late payments were made by the end of the one-year period or not. Therefore borrowers cannot exit default during the year, which is consistent with the assumptions in supervisory stress testing. The application of the above rule yielded 69,205 defaults, representing just over one-sixth of the entire database.

Since the modelling aims to enable supervisory stress testing to provide the most accurate possible picture of the credit risk of the credit institutions and credit institution groups operating in Hungary,[8] the projected migration probabilities should reflect the different risk levels in banks' portfolios. Accordingly, the modelling was divided into two: 1) creation of homogenous groups, and 2) modelling the established groups (as portfolios) with time series methods. The first step helps eliminate the procyclicality of the modelling by including various characteristics, which is necessary because of the Pillar 2 guidance, the end product of the supervisory stress test. It can also prevent us from obtaining a less-than-accurate picture of a bank that uses a stricter-than-average credit scoring system or tightens its system in the meantime.

In practice, this micro-level approach can be achieved by creating risk groups other than the IFRS 9 stages, based on deal and client characteristics. Along the default rate (DR),[9] characteristics and explanatory variables need to be chosen that can best separate the loans according to riskiness. To narrow down the group of variables with potentially strong explanatory power in the database, the relationship between the characteristics and the corresponding average through-the-cycle DR independent from time was examined. The charts on these relationships can be seen in *Figure 4* of the *Annex*. The charts were prepared to answer two questions. First, whether the relationship between the explanatory variables and the DR was strong and varied in space, in other words whether the range of a given characteristic was clustered around various average DRs. And second, whether this relationship was linear or not. *Jagric et al.* (*2011*) used bank data from Slovenia to model the relationship between credit risk and explanatory variables, finding that non-linear relationships have a huge impact on model performance. To achieve linearity and ensure easier and more accurate modellability, several continuous variables with broad ranges were transformed into categorical variables.

The next paragraph addresses the variables that required closer inspection not only because of the test statistics and preliminary intuition but also because the features of the database. In the case of the loan-to-value ratio (LTV),[10] it was suggested that its risk segmentation capacity, detected by *Holló* (*2009*) in his work on household mortgage loans, may have been distorted after the crisis, especially due to the introduction of the regulations that took effect around 2015. In theory, this would mean that although during the crisis and directly thereafter those loans which were disbursed with a higher LTV could also have a higher payment-to-income ratio (PTI),[11]

[8] Although the outcome of the stress test is mainly influenced by credit risk costs, in practice, market, counterparty, operational and profitability risks are also quantified during an assumed potential economic downturn.
[9] The proportion of DR weighted for exposure over the entire period
[10] Loan amount / current value of the collateral
[11] (Monthly) loan repayment / verified total net monthly income of the loan applicant and the co-debtors in the loan contract

this may have potentially turned around later. According to the hypothesis, this may have been caused by banks becoming more risk-averse due to regulatory measures but continuing to lend to clients with a lower PTI at a higher LTV. The hypothesis was easy to examine using the data, and it turned out that the theory has no basis that could be detected with the current modelling database, as any given LTV has similar PTI levels in all periods, which rises as the former increases, i.e. deteriorates. This is also confirmed by the fact that the coefficients of the regressions without an LTV do not deviate from those observed in the equation that include the variable. The other interesting variable with a seemingly strong separating power was the time elapsed since the loan application, which could be used quite intuitively: the further along the debtor in repayment, i.e. the more they paid back from the loan, the more likely they are to continue making the repayments until maturity. When the DRs are arranged based on the time elapsed, there was a turning point at around 5–6 years, where the DRs suddenly started to decrease sharply. The use of the variable and the accuracy of the turning point is confirmed by *Balás et al.* (*2015*), who sought to build a cross-sectional model that included the explanatory variables best explaining default risk. Still, the application of the variable was met with some doubt, as even a simple analysis showed that in the 1–5-year range, due to the features of the time series and the dataset, the high DR could also be explained by the fact that in 2008–2010, when most of the defaults occurred, over 90 per cent of the loans in the sample were younger than 5 years. In other words, the DR does not reflect the risk profile of the transactions, but simply the age at which they entered the crisis. Nevertheless, in the end the literature, test statistics, the regressions performed[12] and the inclusion in the modelling of the macroeconomic data varying over time that will be presented below proved to be convincing, and it was decided that the time elapsed would be used.

Eventually, the average annual interest rate and the year of loan disbursement (vintage) were not included among the modelling variables based on the figures and the test statistics. Ultimately, 10 of the 12 potential explanatory variables in the short list shown in *Figure 4* of the *Annex* were retained: (*adjusted*) *loan amount,[13] educational attainment of the client, age of the client, time elapsed since borrowing, time to maturity, existence of a co-debtor, PTI, LTV, loan type, loan currency*. Apart from the time elapsed, all variables are unchanged over time (with static correlations deliberately retained), so each of them reflects their status at the time of disbursement.

[12] It was examined whether the inclusion of the time fixed effects (basically the disbursement period dummies) in the logit model divert the time elapsed parameters, and as both the coefficient and the standard error proved to be stable, the use of the variable was considered justified.

[13] To maintain the time value, the future value of the pre-2018 disbursements was used throughout, which was produced using the growth rate of the cumulative sectoral average wage.

2.3. Establishing the risk categories

Before the separation, i.e. risk classification, the next step was a more detailed statistical analysis of the selected characteristics. Logistic regression (logit), the most widely used method in banks' risk management practices, was employed in further testing explanatory variables.

$$Y(default = 1, performing = 0) = G\left(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i\right)$$

where:

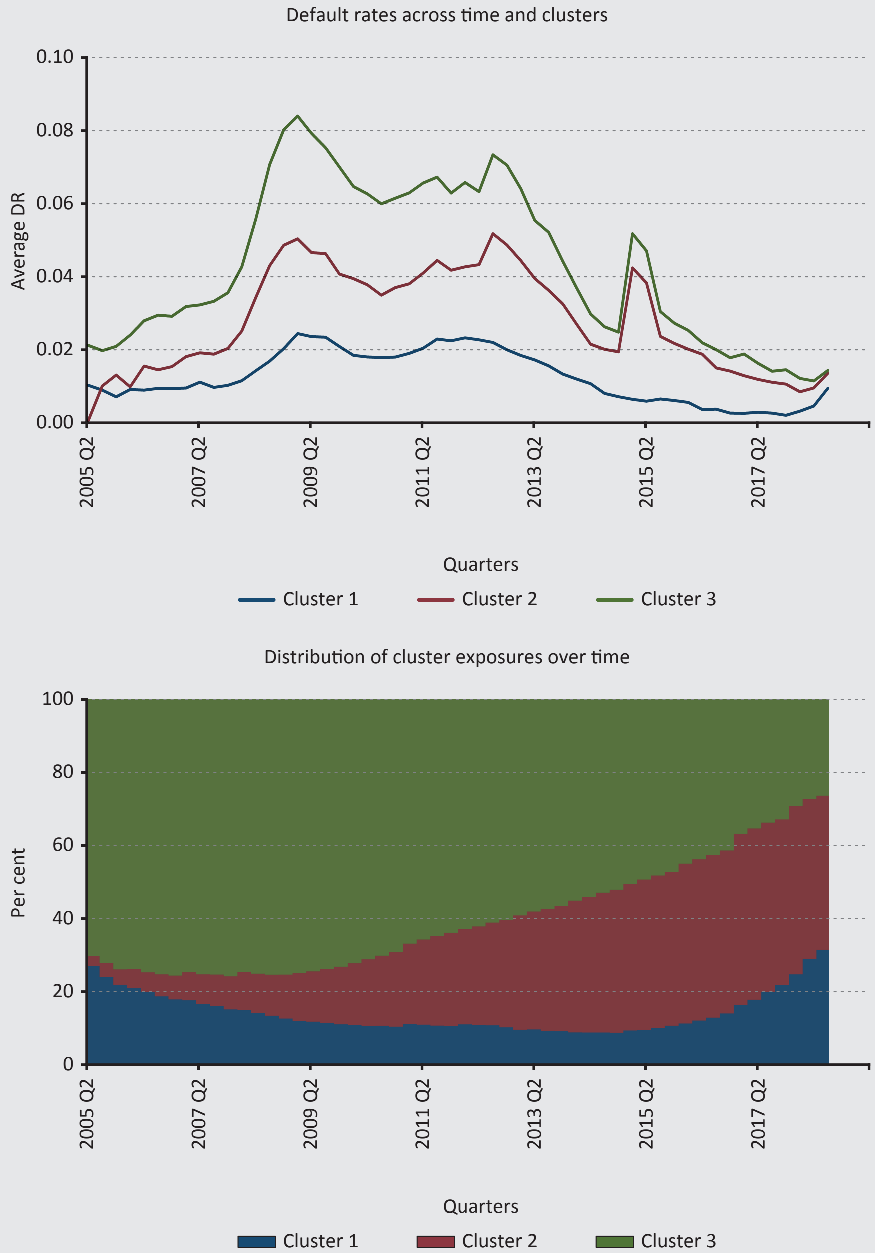
$$G(x) = \frac{e^x}{1 + e^x}$$

All of the explanatory variables in the model proved to be significant together and separately as well, and thus were in line with the earlier intuition that they have a strong separation power, and also with the purpose of the modelling. The results of the model and the statistical tests can be seen in *Table 5* of the *Annex*, while *Figure 5* of the *Annex* shows the ROC (Receiver Operating Characteristic) curve, which backtests the accuracy of the logit model, producing around 70 per cent, which is a relatively good result compared to similar models. Due to the relatively high number of characteristics involved, the extent of multicollinearity, i.e. the correlation among explanatory variables was also examined (see *Table 5* of the *Annex*). The test statistics show that the model has a strong correlation effect. Besides the potential multicollinearity, the variables' other characteristics also pointed towards transformation and dimension reduction. It was observed that the variables used for the separation were heterogeneous in terms of their distribution and range, which tallies with the above multicollinearity issue and the non-linearity problem affecting logistic regression. *McDonald et al.* (*2012*) showed that the non-linear relationship between predictor and explanatory variables that distorts logistic regression can be caused by the correlation between the various risk variables included in the model. To avoid this, the authors recommend the principal component analysis (PCA). This study also used principal component analysis to perform the transformation and dimension reduction. As proposed by *Kovács* (*2014*) for variables of varying standard deviation and measure, the continuous variables were first standardised, which satisfies the normality requirement for the distribution of data. Although the category variables do not have normal distribution and using them in principal component analysis is therefore controversial, several studies, including *Kolenikov and Angeles* (*2004*), have shown that their use does not lead to a great degree of distortion, especially when combined with multiple continuous variables. This principal component analysis produced 4 transformed variables from the 10 existing deal and client characteristics.

After the PCA, the focus was shifted to classifying the loans based on risk, for which a widely used dimension reduction process, cluster analysis, was employed. But first, it was necessary to examine what type of algorithm was permitted by the chosen variables and the dataset. The literature usually ties the choice of methodology to the amount of data and outliers. The sample size in this study makes non-hierarchical cluster analysis an obvious choice, and one of the most popular types, the *k*-means method was picked for creating the groups. This algorithm assigns all data points to the cluster whose centre falls nearest to them. The centre is usually the average of a (random) group of points, and it can typically be applied to points in continuous η-dimensional spaces, so the variables should be collected on the same scale. One feature of the methodology is that during an out-of-sample, idiosyncratic clustering of another bank's portfolio, the same number of groups will be produced as measured in the sample, which may have a distortive effect when supervisory stress testing is performed for relatively homogenous client structures, for example banks lending only to good clients. To reduce or even eliminate any potential distortion, the clusters are calibrated at the banking sector level, which allows individual banks to have very different cluster structures than the average, if warranted by their clientele.

The cluster analysis was performed using the arising principal components, and although the results would have warranted the creation of four different risk group, two clusters produced very similar distributions for the past probabilities of default, so they were merged, and three different clusters were established. The results of this process are summarised in *Table 1*, while the statistics for the PCA and the *k*-means cluster analysis can be seen in *Table 6* and *Figure 6* of the *Annex*. The upper panel of *Figure 1* illustrates that the categories are well-differentiated based on their probabilities of default over time, so the classification produced results in line with expectations.

Figure 1
Variation of the PD (upper panel) and share (lower panel) of the clusters over time



The lower panel of *Figure 1* also shows the share of the different clusters in the period under review. The rapid reduction of Cluster 3 is the first thing that stands out in the chart. The share of the cluster holding the riskiest loans clearly drops from approximately 70 per cent around the 2008 financial crisis to below 30 per cent by the end of the modelling period. This is consistent with what *Bodnár et al.* (*2014*) observed in their paper on the relationship between financial crises and lending. They found that the run-up to crises is usually characterised by the build-up of an excessive amount of bad loans, as seen in Hungary in connection with the household sector's FX loans. However, according to the database used here and the lessons from *Figure 1*, the distribution of loans was worse than today not only in terms of denomination, but also regarding other deal and client characteristics. *Figure 1* also illustrates two interesting events, especially in Clusters 1 and 2, and both of them were tied to government measures. The first appears around 2011, when the share of bad debtors starts to decline fairly steeply, which is attributable to more restrained bank lending. But it can also be seen that the proportion of the best debtors is stable or changes only slowly, which is due to the low loan penetration and the slight dilution of the mortgage loan portfolio in the post-crisis period. This tallies with the claim of *Balás et al.* (*2015*), namely that the early repayment scheme launched in 2011, allowing borrowers to pay off their debt for free, mostly benefitted the best debtors. The second turning point, leading to a larger share of higher-quality transactions, came around 2015–2016. This coincided with the introduction of the requirement to use the LTV and the PTI during credit scoring, whereby applicants who wish to become overindebted are not even admitted to the portfolio, and thus – coupled with the positive effects of an uptick in demand and falling interest rates – the share of good debtors began to rise.

Granular data also show the type of loans that are more likely to be removed from banks' mortgage loan portfolios with this process. While the proportion of those with at most a secondary school diploma is 35 per cent in the least risky Cluster 1 over the entire period, the same figure is 79 per cent in Cluster 3. A similar distribution can be observed for all variables included in clustering. For example, 64 per cent of FX loans and 84 per cent of home equity loans are in the riskiest cluster. The same holds true for continuous variables. The transactions in the best Cluster 1 have a PTI that is 26 percentage points higher on average than in the worst group, while the loan amount (HUF 7.6 million–10.1 million) and time to maturity (187 months–250 months) also exhibit a different distribution. Interestingly, as regards the LTV and the time elapsed, Cluster 2 has the highest average values, but it can also be observed that the other characteristics of these outliers are mostly below-average from a risk perspective, and therefore their placement in the middle category is justified.

3. Time series modelling framework using risk groups

This section describes the other key part of the modelling framework's backbone, namely dynamics, which is represented in the work after a static view of clustering. Similar to the previous section, this section starts with the description of the database helping the modelling, followed by the presentation of the time series models of PDs, before ending with robustness analyses and backtesting of model.

3.1. Data used

Any given portfolio's overall riskiness can be estimated based on deal and client characteristics, yielding a TTC-type measure. This helps differentiate banks from each other according to their vulnerability in line with the prevailing conditions, producing a sort of ranking, but the measurability of the model's response to stress arises from channelling in variables depending on the business cycle. Such procyclical variables can be macroeconomic indicators that change over time and capture the prevailing economic climate of a country well. Yet the macroeconomic variables included in the mortgage default forecasting model should also have another major feature, namely the ability to capture the relationship between household debtors' propensity to repay and business cycles.

Linear regression models were employed to determine this relationship, linking the PDs of the various risk groups to the chosen macroeconomic variables. The choice of macroeconomic variables included in modelling was influenced by two factors. First, the results need to reflect the impact of stress on the PD of household mortgage loans, and even at the expert level, intuitive variables should be included. Second, by virtue of its prediction model nature, the explanatory variables need to be forecastable so that PDs can be estimated for later periods, even 2–3 years ahead. Only variables could be used for which forecasting was available. The latter requirement limited the analysis to 18 macroeconomic measures. The 1–4-quarter lagged values of these variables were also included in the model, to manage protracted effects, potentially drawn out over a year. The variables were first examined on an expert basis and then with a statistical approach; they are shown in *Table 1*.

Table 1
List of macroeconomic variables used in the modelling, their abbreviations and calculation methods

Variable name	Calculation method	Variable name	Calculation method
Households' net financial wealth (wealth)	at 2015 prices	Household's disposable income (hinc)	at 2015 prices
Inflation (infl)	(year-on-year)	EUR/HUF exchange rate (eurhuf)	average
Unemployment rate (unemp)	based on labour force survey	BUBOR-interest (bub3m)	3-month
GDP (gdp)	(year-on-year)	BUX index (bux)	at 2015 prices
Exports (exp)	at 2015 prices	Volatility of the BUX index (buxvola)	quarterly
Imports (imp)	at 2015 prices	EURIBOR interest (eurib3m)	3-month
Employment in the private sector (emp)	thousand people	Benchmark yield curve (gov1y, gov3y, gov5y, gov10y)	government securities market, %
Gross average earnings in the private sector (wage)	at 2015 prices		

Note: GDP and its subitems are seasonally and calendar-adjusted, balanced data.

The predictor variable sought to be estimated is the annual forward-looking PD for all three previously created clusters. To determine this, the defaults that can take [0,1] values calculated from the above formula on the basis of the number of days in default were aggregated for the individual dates and clusters in the database, then, using the share of the defaulting loans' exposures, the average DR was calculated for the period (see *Figure 7* in the *Annex*). In line with the predictor variable, the database of explanatory variables contains past values for 2004–2018.

3.2. The structure of the models, and the results produced

The PDs were forecast along the arising risk categories, by developing three regression equations in total. During this, the includability in the model based on the explanatory power, the interactions and the appropriateness of the time series were all evaluated. As a result of these examinations, the explanatory variables used for the regressions were differentiated once to achieve stationarity.[14] In the case of the predictor variable, it was proposed that, due to its limited nature, it could be stationary on a long time series, unlike standard economic time series, but this was

[14] A stochastic process is broadly speaking stationary if its joint distribution function is independent from time (*Matyasovszky 2002*).

not confirmed by the Dickey–Fuller test, so this time series was also differentiated. The regression equations that emerged were written with the standard structure:

$$\Delta DR_{annual\ average} = \beta_0 + \sum_{i=1}^{n} \beta_i \cdot \Delta x_i$$

where Δx_i is the change in the *i*th explanatory variable in the regression, and n denotes the number of variables.

Before picking the explanatory variables, the autocorrelation of the predictor variable's time series was analysed, in other words the explanatory power of the lagged values and their correlation with the actual period's values were examined. The autocorrelation tests produced a value within the significance level for the lagged values in the first 1–2 quarters in all clusters. But since the predictor variable includes annual PDs, it inevitably overlaps with the values of the next four quarters, and *Table 9* of the *Annex* shows that the Durbin–Watson alternative tests performed on the entire model did not confirm an autocorrelation, so the autocorrelation of the predictor variable within the year did not hamper modelling.

To select the macroeconomic variables necessary for forecasting the change in PD, first simple linear regressions were performed between the change of the predictor variable and of macroeconomic variables or their lagged versions. The variables that turned out to be significant based on the equations and the test statistics were merged, and then a backward[15] elimination method was used to pick the variables that remained significant even when they were together. During this process, special attention was paid not to exclude from the final models the macroeconomic variables that could potentially be eliminated due to the correlation between the explanatory variables (see *Table 7* of the *Annex*); accordingly, after the process was performed, every variable that had been filtered out was attempted to be reinserted into the model one by one, and the exclusion was only considered final after that. In the case of the variables that remained in the model, the correlation could not exceed 0.6, which was set as the critical threshold. The equations with the greatest explanatory power and robustness in the three clusters can be found in *Table 2*, while the corresponding test statistics are shown in *Tables 8, 9* and *10* of the *Annex*. *Table 2* also illustrates what differentiates the methodology presented here, in addition to the duration and current nature of the time series, from the assumptions in the MNB's macroprudential stress test as described in *Banai et al.* (*2014*). While the present study allowed the debtors with heterogeneous situations and riskiness to be affected differently by the various macroeconomic

[15] The selection process involves the following steps: 1) incorporation of all the variables logically related to the explanatory variable into the model; 2) calculation of the partial *t*-test values for the explanatory variables' parameters; 3) if the *t*-value of the variable with the lowest *t*-value is lower than the value at the given significance level, the variable is excluded from the regression; 4) constructing a new model with the remaining explanatory variables; 5) repeating this until only significant variables remain in the model.

variables, *Banai et al.* considered it appropriate to estimate the PD for all borrowers using the same external conditions. The results may support this idea, because from an economic perspective it can be deduced intuitively that for example the development of employment, which has an increasing coefficient along the different clusters, has a greater impact on the workers without a degree and employed in worse, more vulnerable and potentially only ad-hoc jobs than on those graduates who presumably have a more permanent position. It can also be observed that net wealth is important for those in the more vulnerable second and third clusters, while this variable was not even included in the equation for the most reliable debtors based on their repayments. Similarly to the previous observation, this is perhaps explained by the fact that people in more secure jobs earn more (which is also reflected in the PTI levels) and rely more on their earned income than on their existing wealth.

Table 2
Results of multivariate linear regressions determined with the backward selection process

	Clusters		
	1	2	3
Predictor variable / Explanatory variables	*d_DR_y*	*d_DR_y*	*d_DR_y*
d_emp	–0.05367**(0.0260)	–0.14823**(0.0619)	–0.28462***(0.0886)
d_exp	–0.01632**(0.0067)		
l1_d_gdp	–0.00028**(0.0001)		
d_gov10y	0.00044**(0.0002)		
d_bub3m	0.00145***(0.0003)	0.00229**(0.0009)	0.00549***(0.0012)
d_wealth		–0.05383**(0.0226)	–0.08171**(0.0323)
l3_d_gov1y		0.00193***(0.0006)	
l1_d_bux		–0.01027*(0.0055)	
l3_d_hinc			–0.08444**(0.0330)

*Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are shown in brackets. 'd' is the annual change in the variable, 'l' is the quarterly lag, while 'y' indicates the annual nature of the probability of default.*

From a credit risk perspective, the sample covers quite an eventful period, replete with government and regulatory measures that had a major impact on the risk segmentation of household loans and thus also the development of PDs. Perhaps the largest government measure that caused a huge fluctuation in the historical time series of PDs was the forced conversion of FX loans that coincided with the settlement. In the short run, these steps led to a temporary rise in PD, followed by a decline. Several studies have been published on this phenomenon from 2011–2012. *Sepsi* (*2014*) attributes the rise to the fact that between the announcement of the government decree and the actual implementation of this measure debtors may not have cared about paying their upcoming repayment instalments, as they knew that they would be able to make full repayment soon. *Balás et al.* (*2015*) explain the temporary spike in PDs by claiming that the banking system mostly lost mortgagors who performed well, which may have led to a shrinking of the denominator of PDs, without triggering any change in the numerator. The situation was similar with the forced conversion of FX loans, as there the denominator diminished as the value of FX-denominated debt declined. Two solutions were identified for smoothing the spikes in *Figures 1* and *3*: the first is the inclusion of a dummy variable[16] in the model, potentially covering the break in the trend of the time series that could presumably be explained only incorrectly with macroeconomic variables. The other possible solution is the truncation of the time series, whereby the fairly volatile quarters of 2015 are removed from the model estimation. In the end, the truncation of the time series was used, for two reasons: first, the issue only affected one year in the time series spanning 14 years, and second, the case for using the dummy variable was not convincing statistically or from the perspective of forecasting and backtesting, as it lagged behind the second solution in both scenarios.

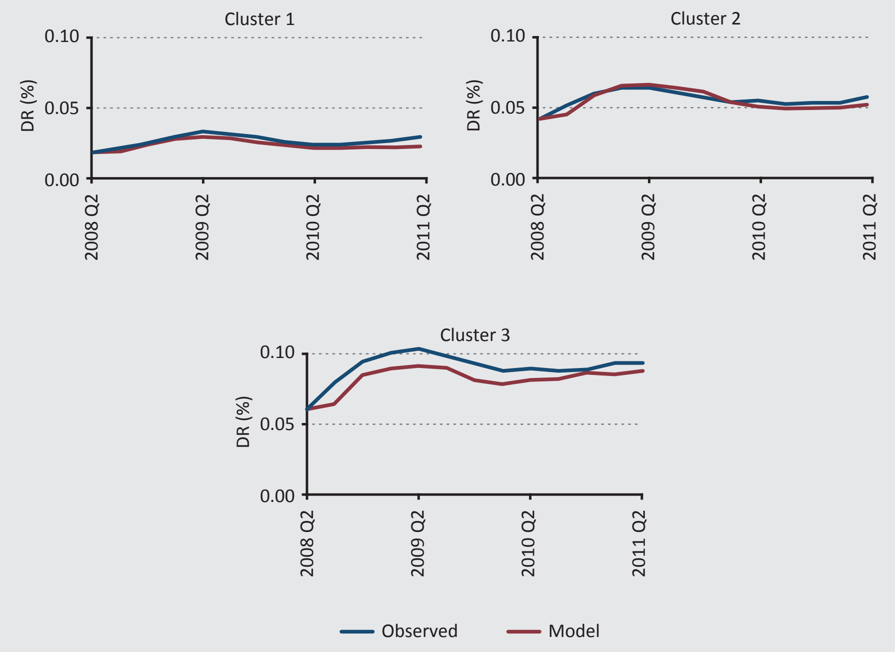
3.3. Cross-validation, robustness analysis

During testing, the results were examined using two approaches, which are presented in this subsection. The first and most important step was to assess the stability of the models by checking the stability of the coefficients and significance levels of the explanatory variables included. The aim was to prove that the model assigns similar coefficients to the variables upon cross-sectional and time series shrinking of the sample, and also when leaving out certain variables from the equation, while preserving the significance of the remaining macroeconomic variables. To aid implementation, random sampling was used before modelling to remove 25 per cent of the existing total dataset. Having performed the cross-validation on this test sample, it was found that the 10 per cent significance level (*p*-value) determined as critical for testing is breached by only a negligible number of variables used in the three equations. After the cross-sectional data truncation, the time series of the modelling database was reduced by half, with similarly positive results. The robustness analysis of the models ended with the exclusion of a few variables. In this case, the significance levels remained within the critical 10 per cent with the exception of one item, and the coefficients of the explanatory variables did not deviate substantially from the values observed in the original models. The details of the analyses can be seen in *Table 11* of the *Annex*.

[16] It takes a value of 1 in the quarters where the break in the trend occurs (2014 Q4–2015 Q4) and 0 in all others.

In the second round of testing, the accuracy of the models was assessed by comparing actual PDs with the values predicted by the model. The sample was chosen to be the first three years of the 2008 financial crisis, for two reasons: first, as this is a stress forecasting model, a volatile period in PDs was necessary, and second, supervisory stress testing usually simulates similar downturns. The backtesting results are shown in *Figure 2*, where the model's estimates move in tandem with the actual data, without any major deviations.

Figure 2
Actual and modelled PDs by clusters, backtested on the data from the 2008 crisis



Therefore, based on the cross-validations and backtests performed that also measure robustness and the goodness of the models, the models can appropriately capture the shifts in macroeconomic variables over time as well as the separating effect of the deal and client characteristics that form the basis of risk segmentation.

4. Transforming PDs into the transitional probabilities of stages

The forecasting of PDs would only allow performing and non-performing transactions to be differentiated, which would have been insufficient as the new accounting standard IFRS 9 became widely used. This section presents how the two groups categorised in terms of the loan repayment performance of debtors are turned into four credit risk stages that better separate clients' solvency.

4.1. Creating stages

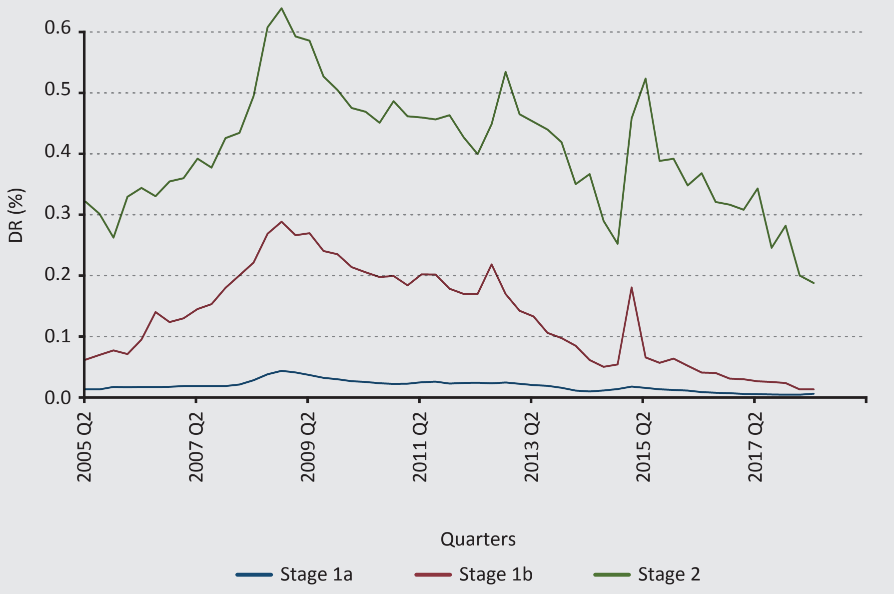
In accordance with the new accounting standard recommendation (*IASB 2013*), three credit risk categories must be established. The first ('Stage 1') comprises the transactions that are not past due or only slightly so (by no more than 30 days). The interim category between non-performing and performing transactions (Stage 2), which is one of the features that sets apart IFRS 9 from the previous accounting standard,[17] includes the loan contracts past due by more than 30 but no more than 90 days. Based on the recommendation of the accounting standard, several transaction and subjective characteristics[18] can be taken into account while creating the groups. This is also confirmed by the experience from ICAAP analysis, namely that there are almost as many rules for defining Stage 2 as there are banks. The present model follows the Stage 1 logic due to its comparability, simplicity and the availability of the data, and only the number of days past due was taken into account while creating the group, with a value of 31–90 days. Finally, the non-performing category ('Stage 3') was created as well. It contained the transactions that were more than 90 days past due at the time of observation.

An additional analysis according to the distribution of the number of days past due highlighted that the riskiness of Stage 1 was too heterogeneous for uniform modelling. *Figure 3* shows that Stage 1 clients who were late with their payments by more than 30 days at least once since the disbursement were more likely to do so again than their always performing peers. It can also be seen that the transactions in the interim category (Stage 2) are much more likely to default on average than any of their Stage 1 peers, and they also respond stronger to economic fluctuations. Consistent with this phenomenon, but in contrast to the IFRS 9 recommendation, the creation of four, rather than three, risk categories was warranted. The first group ('Stage 1a') included the best transactions that were currently past due by no more than 30 days and, in contrast to the usual categorisation in the banking sector, also performed well before the time of observation, so they were never more than 30 days past due since origination. The second Stage 1 subcategory ('Stage 1b') includes the transactions that were not more than 30 days past due at the time of observation, but had been at least once during their lifetime.

[17] IAS 39 only differentiated between two risk categories: performing and non-performing. Directive 2006/48/EC of the European Parliament and of the Council was the first to include an official definition of non-performance (default).

[18] Restructuring, difference between initial and observed PD values, risk rating, expert judgement

Figure 3
Average annual PD of performing mortgage loans of different risk categories



4.2. Forecasting transitional probabilities

As four risk groups were created from the original two, the number of cross-group migration probabilities sought to be forecast also increased considerably. In line with the stress testing methodology of the supervisory authority and the *EBA* (*2021*), there is no exit from default or Stage 3, so that direction was not explored. However, the Stage 1a–Stage 2, Stage 1a–Stage 3, Stage 1b–Stage 2, Stage 1b–Stage 3, Stage 2–Stage 1b, Stage 2–Stage 3 directions are all important for accurately forecasting credit risk losses. The transformation of the PDs is presented in this subsection.

The transformation was performed based on the methodological guidelines prepared by the ECB. The document was drawn up for the EU-wide EBA stress test covering the largest banking groups, strictly for the internal use of national competent authorities, helping to control and compare the internal models of participating banks, such as OTP Group from Hungary, during the quality assurance process performed as part of the exercise. In the ECB's model, the PDs are only used directly for forecasting Stage 1–Stage 3 (TP^{1-3}) and Stage 2–Stage 3 (TP^{2-3}) migrations, based on the below formulas:[19]

[19] Of course, in line with the section's introduction, the Stage 1 loans were divided into two groups, in contrast to the ECB's practice, breaking up the Stage 1 formulas into 'a' and 'b' parts.

$$TP^{1-3}_{T_0+h} = \Phi\left(\Phi^{-1}\left(TP^{1-3}_{T_0}\right) + \Phi^{-1}\left(DR_{T_0+h}\right) - \Phi^{-1}\left(DR_{T_0}\right)\right)$$

$$TP^{2-3}_{T_0+h} = \Phi\left(\Phi^{-1}\left(TP^{2-3}_{T_0}\right) + \Phi^{-1}\left(DR_{T_0+h}\right) - \Phi^{-1}\left(DR_{T_0}\right)\right)$$

where T_0 denotes the values of the starting period, ϕ is the cumulative distribution function of standard normal distribution, while h shows the number of periods since the starting point. The initial (T_0) probabilities are calculated based on the actually observed exposure migrations across stages in the previous year (actual year). The formula determining future, hypothetical $TP^{1–3}$ transitional probabilities is of course duplicated when Stage 1 is broken down into two, so the probability of migration is calculated for both performing and past due transactions. The calculation is performed separately for all three clusters, yielding nine different numbers for the probability of the loans in the given group becoming non-performing over the course of the next year.

Forecasting the probability of migration to the non-performing category is also paramount, as Stage 1–Stage 2 and Stage 2–Stage 1 forecasts are also based on that. This was estimated using a simple linear regression relationship, with the following formulas:

$$\Phi^{-1}(TP^{1-2}_t) = \beta_0 + \beta_1 \cdot \Phi^{-1}(TP^{1-3}_t)$$

$$\Phi^{-1}(TP^{2-1}_t) = \beta_0 + \beta_1 \cdot \Phi^{-1}(TP^{2-3}_t)$$

where first the coefficient of the past covariance of $TP^{1–2}$ and $TP^{1–3}$ (β_1) is estimated, then, assuming that distributions remain stable over time, the value of $TP^{1–2}$ is projected using this coefficient through the previously forecast $TP^{1–3}$. The methodology is similar for $TP^{2–1}$: the only difference is that this migration is complemented by $TP^{2–3}$ as the explanatory variable. When estimating $TP^{1–2}$ and $TP^{2–1}$, unlike in the transitions $TP^{1–3}$ and $TP^{2–3}$, the sample is not broken down into clusters, since the different risks of these groups are incorporated through actual data and the explanatory variables of the regressions.

Table 3

Results of the univariate linear regressions between the transitional probabilities across stages

Explanatory variables	Predictor variables		
	d_invn_s1as2_y	*d_invn_s1bs2_y*	*d_invn_s2s1b_y*
d_invn_s1as3_y	0.52438*** (0.0854)		
d_invn_s1bs3_y		0.23393*** (0.0678)	
d_invn_s2s3_y			–0.69007*** (0.1002)

*Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are shown in brackets. 'd' is the annual change in the variable, 'invn' is short for the inverse of the cumulative standard normal distribution, while 'y' indicates the annual nature of the transitional probabilities.*

In the end, the transitional probabilities were differentiated once to ensure the fulfilment of the stationarity requirement of the variables, which is vital when using time series regressions. The results of the models designed as described above can be seen in *Table 3*, which also shows, among other things, that while the coefficients of TP^{1-3} take positive values in regressions, in line with the preliminary intuition, and thus a positive correlation is assumed with TP^{1-2} values, TP^{2-3} has a negative value, showing that two processes moving in opposite directions are examined there.

5. Summary and conclusions

At the end of the entire modelling process, several conclusions and lessons can be drawn that can help to more accurately assess the risks related to the mortgage loan portfolios of the credit institutions and credit institution groups operating in Hungary as well as facilitate the appropriate stress testing of such. One of these, the reason for which was not this clear from the earlier supervisory stress testing exercises, is that under the current conditions and loan portfolios, and with a macroeconomic shock similar to the 2008 global crisis, the same levels of credit risk losses cannot be estimated for Hungarian banks. There are two reasons for this:

1. prior to the crisis, initial DRs were much higher than what can be observed in the period before the publication of the study; and

2. partly due to the denomination of the loan, but partly due to other factors, a riskier and lower-quality portfolio built up before the turbulent period, which thus entailed a much greater potential PD.

The first reason could be the result of the second to some extent, but this is also partly related to the lower or more limited levels of financial awareness, the psychology of loan repayments and lending controls in earlier times. Another conclusion is that a much more accurate picture can be gained about the credit risk of different banks that originate varying qualities of loans if loan portfolios are stressed using transaction-level models and various characteristics. It has been observed that the difference in PDs between the riskiest and the best client groups can amount to several percentage points. Using the above-mentioned loan characteristics and creating a two-step model could also potentially contribute to a more accurate estimate through the heterogeneous use of macroeconomic variables. The results of the model show that with loan repayments the indicators that shape the PDs of the debtors classified into the different risk categories are mostly influenced by the stability in earned income and the extent of relying on various assets.

The models therefore help in calculating the component with the greatest influence on the stress scenario results, i.e. credit risk costs. The lessons learnt from the separation of the through-the-cycle step using lending characteristics can be used to decide, by touching on the dynamics of the stress test, whether to make the loans disbursed under the stress scenario consistent with the maturing ones, or the portfolio's credit characteristics should be modified, and if so, in what direction and to what extent.

References

Acuña, E. – Rodriguez, C. (2004): *The Treatment of Missing Values and its Effect on Classifier Accuracy*. In: Banks, D. – McMorris, F.R. – Arabie, P. – Gaul, W. (eds.): Classification, Clustering, and Data Mining Applications, pp. 639–647. https://doi.org/10.1007/978-3-642-17103-1_60

Agarwal, V. – Taffler, R. (2008): *Comparing the performance of market-based and accounting-based bankruptcy prediction models*. Journal of Banking and Finance, 32(8): 1541–1551. https://doi.org/10.1016/j.jbankfin.2007.07.014

Altman, E.I. (1968): *Financial ratios discriminant analysis and the prediction of corporate bankruptcy*. Journal of Finance, 23(4): 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Balás, T. – Banai, Á. – Hosszú, Zs. (2015): *A nemteljesítési valószínűség és az optimális PTI-szint modellezése egy háztartási kérdőíves felmérés felhasználásával* (*Modelling Probability of Default and Optimal PTI Level by Using a Household Survey*). MNB Occasional Papers 117, Magyar Nemzeti Bank. https://www.mnb.hu/letoltes/117-mnb-tanulmany-hun.pdf

Banai, Á. – Hosszú, Zs. – Körmendi, Gy. – Sóvágó, S. – Szegedi, R. (2014): *Stress testing at the Magyar Nemzeti Bank.* MNB Occasional Papers 109, Magyar Nemzeti Bank. https://www.mnb.hu/letoltes/op109-final.pdf

Bielecki, T.R. – Rutkowski, M. (2004): *Credit Risk: Modeling, Valuation, and Hedging*. Springer, New York. https://doi.org/10.1007/978-3-662-04821-4

Bodnár, K. – Kovalszky, Zs. – Kreiszné Hudák, E. (2014): *Recovery from crises and lending*. Financial and Economic Review, 13(4): 57–85. https://en-hitelintezetiszemle.mnb.hu/letoltes/3-bodnar-en.pdf

Chae, S. – Sarama, R.F. – Vojtech, C.M. – Wang, J.Z. (2019): *The Effect of the Current Expected Credit Loss Standard (CECL) on the Timing and Comparability of Reserves*. Finance and Economics Discussion Series 2018-020, Board of Governors of the Federal Reserve System (U.S.). https://doi.org/10.17016/FEDS.2018.020r1

Crouhy, M. – Galai, D. – Mark, R. (2000): *A comparative analysis of current credit risk models*. Journal of Banking and Finance, 24(1–2): 59–117. https://doi.org/10.1016/S0378-4266(99)00053-9

Daniels, T. – Duijm, P. – Liedorp, F. – Mokas, D. (2017): *A top-down stress testing framework for the Dutch banking sector*. Occasional Studies, 15-3, De Nederlandsche Bank. https://www.dnb.nl/media/ss1h5zks/201707_nr_3_-2017-_a_top-down_stress_testing_framework_for_the_dutch_banking_sector.pdf

Das, S.R. – Hanouna, P. – Sarin, A. (2009): *Accounting-based versus market-based cross-sectional models of CDS spreads*. Journal of Banking and Finance, 33(4): 719–730. https://doi.org/10.1016/j.jbankfin.2008.11.003

EBA (2021): *2021 EU-Wide Stress Test – Methodological note*. European Banking Authority. https://www.eba.europa.eu/sites/default/documents/files/document_library/Risk%20Analysis%20and%20Data/EU-wide%20Stress%20Testing/2021/Launch%20of%20the%20ST/962559/2021%20EU-wide%20stress%20test%20-%20Methodological%20Note.pdf. Downloaded: 5 February 2021.

Fáykiss, P. – Palicz, A. – Szakács, J. – Zsigó, M. (2018): *Experiences of Debt Cap Regulations in Hungarian Retail Lending*. Financial and Economic Review, 17(1): 34–61. https://doi.org/10.25201/FER.17.1.3461

Gross, M. – Laliotis, D. – Leika, M. – Lukyantsau, P. (2020): *Expected Credit Loss Modeling from a Top-Down Stress Testing Perspective.* IMF Working Paper WP/20/111, International Monetary Fund. https://doi.org/10.5089/9781513549088.001

Gross, M. – Población, J. (2015): *A false sense of security in applying handpicked equations for stress test purposes*. ECB Working Paper 1845, European Central Bank. https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1845.en.pdf

Holló, D. (2009): *Risk developments on the retail mortgage loan market.* MNB Bulletin, 4(3): 14–19. https://www.mnb.hu/letoltes/hollo-2009-okt-en.pdf

Horváth, G. (2021): *Corporate Credit Risk Modelling in the Supervisory Stress Test of the Magyar Nemzeti Bank*. Financial and Economic Review, 20(1): 43–73. https://doi.org/10.33893/FER.20.1.4373

IASB (2013): *Financial Instruments: Expected Credit Losses*. Exposure Draft, ED/2013/3, International Accounting Standards Board. https://www.ifrs.org/content/dam/ifrs/project/fi-impairment/exposure-draft-2013/published-documents/ed-expected-credit-losses.pdf. Downloaded: 4 March 2021.

Jagric, V. – Kracun, D. – Jagric, T. (2011): *Does Non-linearity Matter in Retail Credit Risk Modeling.* Czech Journal of Economics and Finance, 61(4): 384–402.

Kolenikov, S. – Angeles, G. (2004): *The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices.* https://www.measureevaluation.org/resources/publications/wp-04-85.html. Downloaded: 11 March 2021.

Kovács, E. (2014): *Többváltozós adatelemzés* (*Multivariate data analysis*). Typotex Kiadó, Budapest.

Landini, S. – Uberti, M. – Casellina, S. (2019): *Credit risk migration rates modelling as open systems II: A simulation model and IFRS9-baseline principles*. Structural Change and Economic Dynamics, 50(September): 175–189. https://doi.org/10.1016/j.strueco.2019.06.013

Lang, P. – Stancsics, M. (2019): *Modeling loan loss provisions under IFRS 9 in the top-down solvency stress test of the Central Bank of Hungary*. EBA Policy Research Workshop "The future of stress tests in the banking sector – approaches, governance and methodologies", Paris. https://www.eba.europa.eu/sites/default/documents/files/document_library/Calendar/Conference-Workshop/2019/8th%20annual%20workshop%20documents/10%20Modeling%20loan%20loss%20provisions%20under%20IFRS%209%20in%20the%20top-down%20solvency%20stress%20test%20of%20the%20Central%20Bank%20of%20Hungary.pdf?retry=1. Downloaded: 14 April 2021.

Lawrence, E.L. – Smith, S. – Rhoades, M. (1992): *An analysis of default risk in mobile home credit*. Journal of Banking and Finance, 16(2): 299–312. https://doi.org/10.1016/0378-4266(92)90016-S

Little, R.J.A. – Rubin, D.B. (2002): *Statistical Analysis with Missing Data*. Wiley, New Jersey. https://doi.org/10.1002/9781119013563

Martin, D. (1977): *Early warning of bank failure: A logit regression approach*. Journal of Banking and Finance, 1(3): 249–276. https://doi.org/10.1016/0378-4266(77)90022-X

McDonald, R.A. – Sturgess, M. – Smith, K. – Hawkins, M.S. – Xiao-Ming Huang, E. (2012): *Non-linearity of scorecard log-odds*. International Journal of Forecasting, 28(1): 239–247. https://doi.org/10.1016/j.ijforecast.2011.01.001

Matyasovszky, I. (2002): *Statistical Climatology* (*Time Series Analysis*). ELTE Eötvös Kiadó, Budapest.

Pederzoli, C. – Torricelli, C. (2005): *Capital requirements and business cycle regimes: Forward-looking modelling of default probabilities*. Journal of Banking and Finance, 29(12): 3121–3140. https://doi.org/10.1016/j.jbankfin.2005.01.004

Platt, H.D. – Platt, M.B. (1991): *A note on the use of industry-relative ratios in bankruptcy prediction*. Journal of Banking and Finance, 15(6): 1183–1194. https://doi.org/10.1016/0378-4266(91)90057-S

Sepsi, B. (2014): *Így változott a 90 napon túli késedelmes devizaalapú hitelek helyzete Magyarországon* (*That is how the situation of the FX loans more than 90 days past due changed in Hungary*). Közgazdász Fórum (Forum on Economics and Business) 17(119/120), pp. 168–181.

Sommerville, R.A. – Taffler, R.J. (1995): *Banker judgement versus formal forecasting models: The case of country risk assessment*. Journal of Banking and Finance, 19(2): 281–297. https://doi.org/10.1016/0378-4266(94)00051-4

Appendix

Table 4
Descriptive statistics of the explanatory variables used from the database

Variable	Observations	Average	Standard deviation	Min	Max	Data cleansing
termek (loan type)	9,350,798	1.5	0.5	1.0	2.0	
kesnap (days past due)	9,350,798	110.2	405.1	0.0	4,794.0	
szla_deviza (currency)	9,350,798	2.0	0.5	1.0	3.0	
szla_futamido (repayment period)	9,350,798	226.5	82.0	8.0	480.0	
szla_ltv (ltv at disbursement)	9,350,385	0.5	0.2	0.0	9.5	deletion, scaling
ugyf_rendjov (disposable income)	9,350,798	135,017.7	194,403.8	0.0	82,000,000.0	
ugyf_kor (age)	9,350,798	37.5	9.8	1.0	152.0	scaling
ugyf_nem (gender)	7,689,811	1.4	0.5	1.0	2.0	
ugyf_eltartott (dependent)	9,331,701	0.8	1.0	0.0	32.0	
ugyf_kereso (number of earners in family)	9,350,798	1.6	0.6	0.0	25.0	
ugyf_torlkiad (repayment)	9,308,471	29,192.2	283,734.4	0.0	269,000,000.0	
ugyf_iskveg (education)	9,343,639	2.3	0.5	1.0	3.0	deletion
ugyf_csalallapot (marital status)	9,346,990	1.7	0.7	1.0	3.0	deletion
adostars (co-debtor)	9,350,798	0.6	0.5	0.0	1.0	
jaras (district)	9,326,903	91.8	57.4	1.0	198.0	
szla_arfolyam (exchange rate)	9,350,798	147.7	75.3	1.0	316.0	
szla_kamat (interest rate)	9,350,798	5.9	2.2	2.0	19.5	
pti_felv (pti at disbursement)	8,717,907	7.2	741.6	0.0	230,674.0	average, scaling
felv_hitelossz (loan amount)	9,350,798	11.9	9.3	0.5	345.0	
vintage	9,350,798	575.8	27.9	524.0	707.0	
eltelt_ido (time from disbursement)	9,350,798	52.2	37.7	0.0	183.0	

Note: The table shows the missing data in the different variables as well as the data cleansing method used.

Figure 4
The relationship between the chosen deal and client characteristics as well as the corresponding PDs

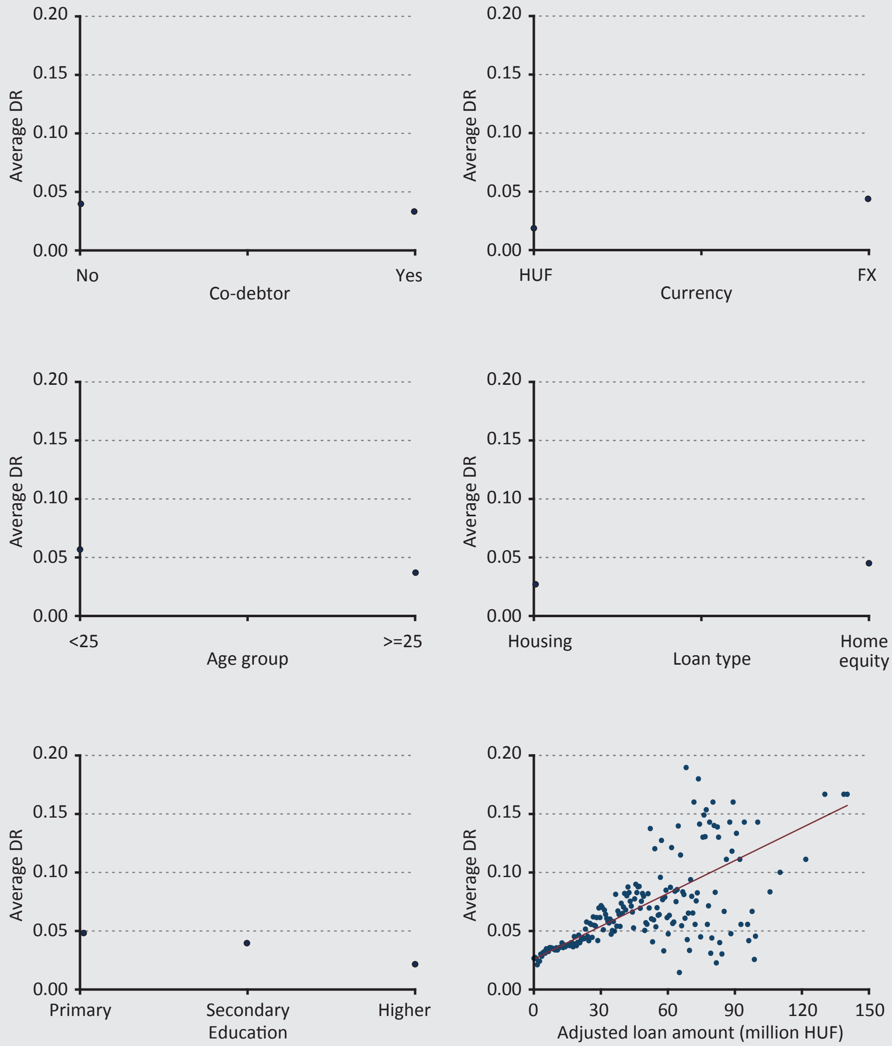
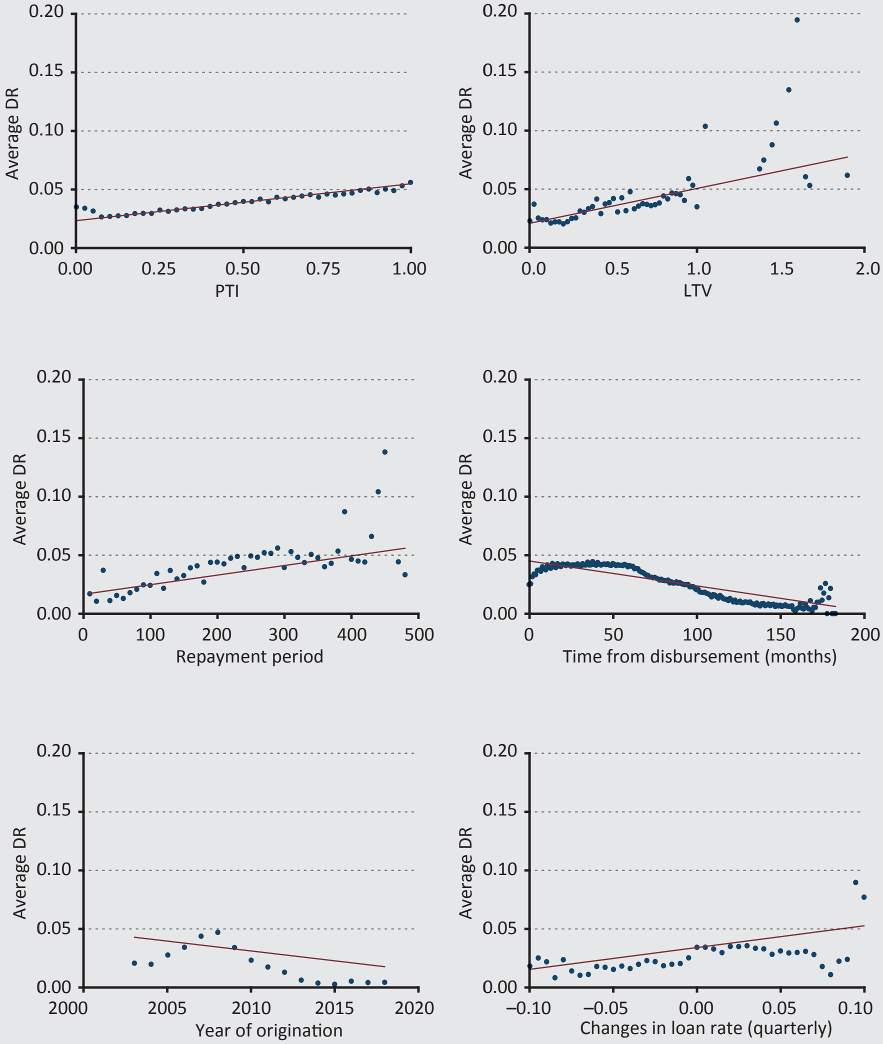


Figure 4
The relationship between the chosen deal and client characteristics as well as the corresponding PDs (part II)



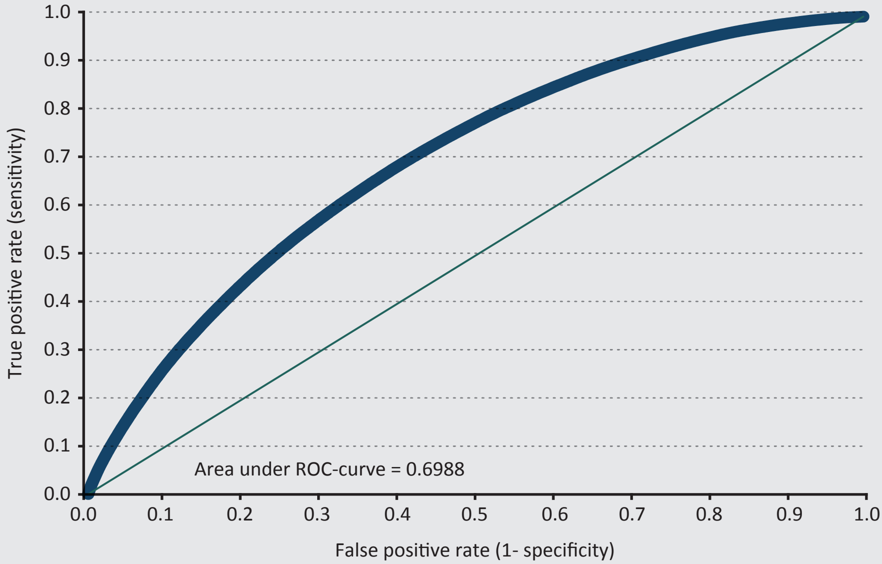
Note: In the case of continuous variables, the figure also shows the line fitted based on the simple linear regression of the given variable and the average PD.

Table 5
Results of the logit regressions used for picking the client and deal characteristics and other test statistics

	Univariate logit	Multivariate logit
Predictor variable / Explanatory variables	*Default (1=default; 0=performing)*	*Default (1=default; 0=performing)*
szla_ltv (ltv at disbursement)	0.85152*** (0.0106)	1.20208*** (0.0136)
ugyf_iskveg (education)	–0.51747*** (0.0041)	0*** –0.20131 (0.0098) –0.74801 (0.0108)
korosztaly (age cohort)	–0.47459*** (0.0076)	0*** –0.28976 (0.0078)
eltelt_ido (time from disbursement)	–0.00705*** (0.0001)	–0.00870*** (0.0001)
adostars (co-debtor)	–0.21042*** (0.0042)	0*** –0.22223 (0.0043)
szla_futamido (repayment period)	0.00240*** (0.0000)	0.00305*** (0.0000)
pti_felv (pti at disbursement)	0.82692*** (0.0077)	0.55789*** (0.0074)
termek (loan type)	0.60353*** (0.0042)	0*** 0.90221 (0.0052)
szla_deviza (currency)	1.12028*** (0.0085)	0*** 0.79321 (0.0087)
hitelossz_kereset (inflation adjusted loan amount)	0.02188*** (0.0002)	0.01036*** (0.0003)
vintage	–0.05582*** (0.0010)	
d_szla_kamat (interest rate)	5.48997*** (0.3872)	
	VIF	**1/VIF**
szla_ltv (ltv at disbursement)	9.57	0.1045
ugyf_iskveg (education)	0 11.82 6.33	0 0.0846 0.1581
korosztaly (age cohort)	0 13.63	0 0.0734
eltelt_ido (time from disbursement)	3.17	0.3151
adostars (co-debtor)	0 2.66	0 0.3755
szla_futamido (repayment period)	11.80	0.0848
pti_felv (pti at disbursement)	4.84	0.2065
termek (loan type)	0 2.42	0 0.4133
szla_deviza (currency)	0 6.56	0 0.1525
hitelossz_kereset (inflation adjusted loan amount)	4.60	0.2174
Average VIF	7.04	

*Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are shown in brackets. In the case of the category variables, the one taking the value of 1 is the basis, and the PD of the rest is compared to that.*

Figure 5
The ROC curve of the logit regression



Note: In a perfect model, the area under the ROC curve is 1. In a simple random model, the area under the ROC curve is 0.5.

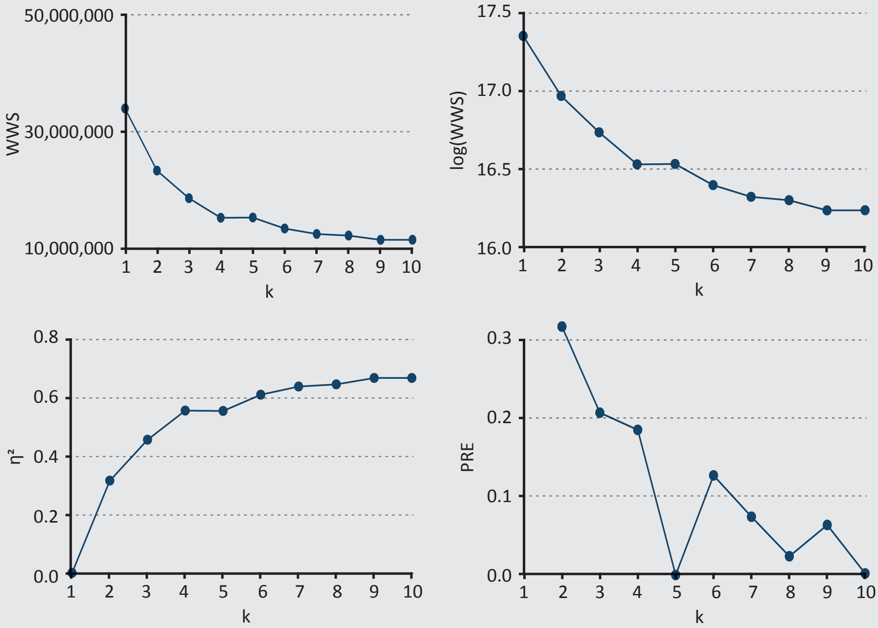
Table 6
Statistics and figures of the PCA and *k*-means cluster analysis

	PC1	PC2	PC3	PC4
szla_ltv (ltv at disbursement)	0.5100	–0.0577	0.0110	0.0918
ugyf_iskveg (education)	0.1646	–0.4606	0.1207	0.0835
korosztaly (age cohort)	–0.1102	–0.0608	0.4031	0.6107
eltelt_ido (time from disbursement)	0.0504	0.3255	–0.4959	0.4856
adostars (co-debtor)	–0.0778	0.1885	0.4448	0.4671
szla_futamido (repayment period)	0.4971	0.0594	–0.1819	0.1111
pti_felv (pti at disbursement)	0.1601	0.4773	0.4086	–0.3275
termek (loan type)	–0.4230	0.3105	0.1090	–0.1410
szla_deviza (currency)	0.1164	0.5443	–0.1918	0.0771
hitelossz_kereset (inflation adjusted loan amount)	0.4764	0.1236	0.3611	–0.1109

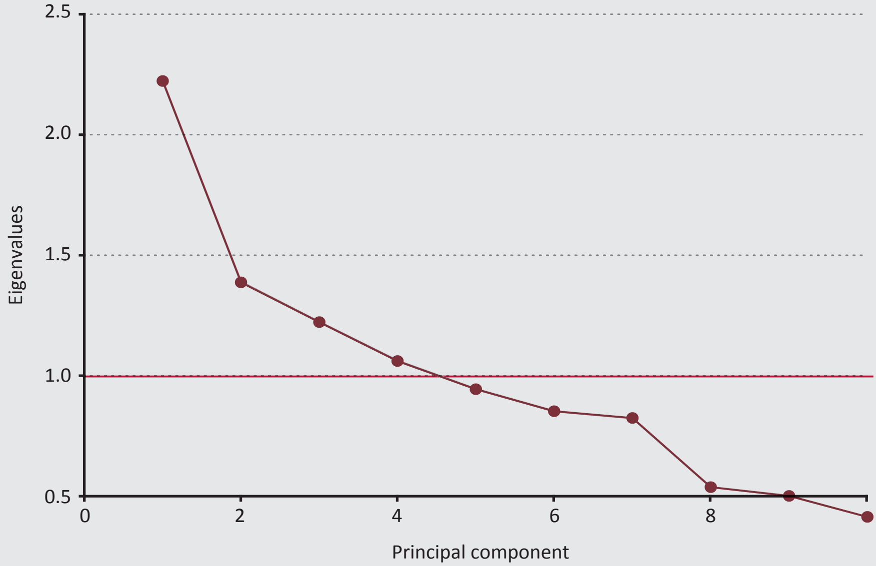
Note: 'PC' denotes the established principal components, while the values show the explanatory variables' coefficients in the different principal components.

Figure 6
Figures of the PCA and the *k*-means cluster analysis

a) The cohesive power of cluster structures



b) Eigenvalues of the principal components arising along the PCA



Note: The cohesive power of the different cluster structures is measured along WSS (within sum of squares), η^2 and PRE (the percentage drop in the WSS).

Table 7
Correlation matrix of the explanatory variables used for the time series analysis

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1)	d_infl	1.0000														
(2)	d_buxvola	–0.0004	1.0000													
(3)	d_bux	0.2952	–0.0877	1.0000												
(4)	d_bub3m	0.0810	0.0635	–0.4726	1.0000											
(5)	d_eurib3m	0.0734	0.0720	0.2457	0.0293	1.0000										
(6)	d_gov10y	–0.0410	0.0000	–0.4501	–0.0152	–0.2682	1.0000									
(7)	d_wealth	0.3231	–0.1070	0.6185	–0.1611	0.2650	–0.6026	1.0000								
(8)	d_eurhuf	–0.3043	0.0029	–0.5902	0.1138	–0.4678	0.6116	–0.4498	1.0000							
(9)	d_hinc	–0.2171	–0.0803	–0.1870	–0.0234	0.1020	–0.0901	0.0308	0.1205	1.0000						
(10)	d_wage	0.1661	–0.0769	0.1080	0.3364	0.1063	–0.1876	0.2527	–0.1072	–0.1297	1.0000					
(11)	d_emp	–0.0409	0.2646	–0.0656	0.0123	0.3273	0.0349	0.1307	–0.0076	0.3078	0.0996	1.0000				
(12)	d_imp	0.2304	–0.1627	0.4284	–0.2343	0.5885	–0.2115	0.3278	–0.3333	–0.0223	0.2057	0.1302	1.0000			
(13)	d_exp	0.2054	–0.0292	0.3170	–0.1755	0.5531	–0.1249	0.2677	–0.2142	0.0060	0.1950	0.1990	0.8694	1.0000		
(14)	d_gdp	0.0658	–0.1337	0.2483	–0.3437	0.4642	–0.2040	0.2963	–0.2195	0.1080	0.1499	0.1923	0.6172	0.6355	1.0000	
(15)	d_unemp	0.0057	–0.0594	–0.0524	0.0495	–0.3801	0.1158	–0.0462	0.2026	–0.3743	0.3026	–0.4708	–0.2308	–0.1605	–0.0572	1.0000

Note: The table includes the correlation coefficients between the (potential) time series explanatory variables.

Figure 7
Clusters' time series PDs and averages

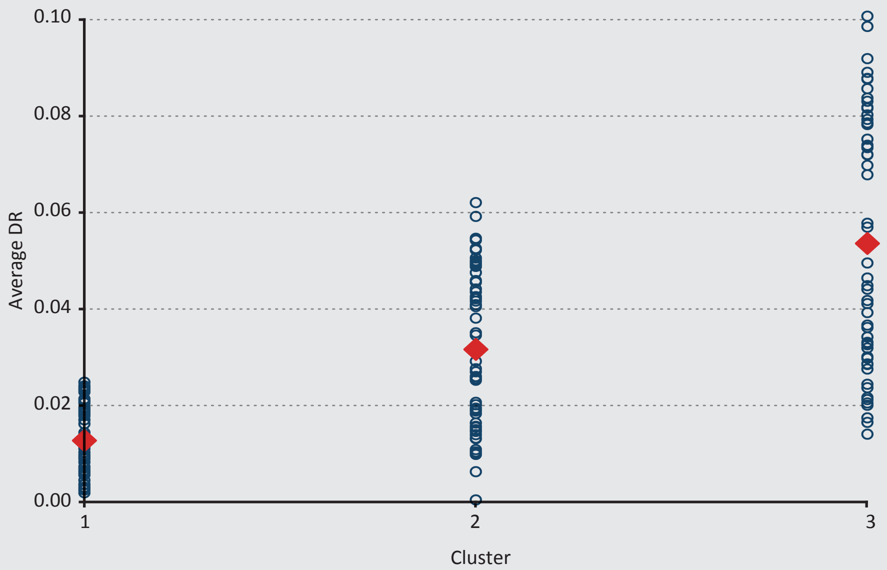


Table 8
Results of the stationarity tests

d_DR_y	–3.328	d_eurhuf	–6.332
d_infl	–6.140	d_hinc	–8.101
d_buxvola	–10.81	d_wage	–8.637
d_bux	–4.816	d_emp	–4.938
d_bub3m	–4.965	d_imp	–5.363
d_eurib3m	–3.554	d_exp	–5.412
d_gov10y	–8.130	d_gdp	–5.207
d_wealth	–6.066	d_unemp	–7.739

Note: The critical value is –3.576 at 1 per cent, –2.928 at 5 per cent, and –2.599 at 10 per cent.

Table 9
Results of the time series regression tests by clusters

	Clusters		
	1	2	3
Predictor variable / Tests	*d_DR_y*	*d_DR_y*	*d_DR_y*
Breusch-Pagan / Cook-Weisberg test (heteroscedasticity)	2.13 (0.1443)	3.72 (0.0537)	0.01 (0.9137)
Durbin-Watson alternative test (autocorrelation)	0.92 (0.3383)	0.22 (0.6377)	0.12 (0.7341)
Ramsey RESET test (excluded variable)	0.97 (0.4167)	2.49 (0.0764)	1.34 (0.2756)
Wu-Hausman F-test (endogeneity)		0.46 (0.4994)	

Note: The p-value is shown in brackets.

Table 10
Results of the multicollinearity time series regression tests by clusters

	Clusters		
	1	2	3
Predictor variable / Explanatory variables	**VIF**	**VIF**	**VIF**
d_emp	1.14	1.02	1.02
d_exp	1.09		
l1_d_gdp	1.10		
d_gov10y	1.02		
d_bub3m	1.03	1.33	1.04
d_wealth		1.06	1.06
l3_d_gov1y		1.06	
l1_d_bux		1.27	
l3_d_hinc			1.02
Average VIF	**1.08**	**1.15**	**1.03**

Table 11
Robustness analysis with three approaches (cross-sectional, time series truncation, exclusion of variables)

	Clusters		
	1	2	3
Predictor variable / Explanatory variables	*d_DR_y*	*d_DR_y*	*d_DR_y*
cross-sectional validation (25%)			
d_emp	–0.09887** (0.0413)	–0.17727*** (0.0651)	–0.25181*** (0.0886)
d_exp	–0.01866* (0.0107)		
l1_d_gdp	–0.00029 (0.0002)		
d_gov10y	0.00051* (0.0003)		
d_bub3m	0.00128** (0.0005)	0.00223** (0.0010)	0.00472*** (0.0012)
d_wealth		–0.04133** (0.0237)	–0.09045** (0.0323)
l3_d_gov1y		0.00101 (0.0007)	
l1_d_bux		–0.01005* (0.0058)	
l3_d_hinc			–0.09309** (0.0330)
time series validation (2007Q2–2014Q1)			
d_emp	–0.03989 (0.0340)	–0.17727* (0.0796)	–0.27969** (0.1078)
d_exp	–0.01786* (0.0107)		
l1_d_gdp	–0.00041** (0.0002)		
d_gov10y	0.00044* (0.0003)		
d_bub3m	0.00136** (0.0004)	0.00223** (0.0008)	0.00546*** (0.0015)
d_wealth		–0.04133* (0.0296)	–0.08887** (0.0398)
l3_d_gov1y		0.00101** (0.0013)	
l1_d_bux		–0.01005 (0.0071)	
l3_d_hinc			–0.12337** (0.0477)
exclusion of variables (1: d_emp, 2: d_wealth, 3: d_wealth)			
d_emp		—0.16524** (0.0650)	–0.31230*** (0.0935)
d_exp	–0.01906*** (0.0068)		
l1_d_gdp	–0.00035*** (0.0001)		
d_gov10y	0.00042** (0.0002)		
d_bub3m	0.00141*** (0.0003)	0.00251** (0.0010)	0.00602*** (0.0012)
d_wealth			
l3_d_gov1y		0.00193*** (0.0007)	
l1_d_bux		–0.01192** (0.0058)	
l3_d_hinc			–0.09476*** (0.0348)

Table 11
Robustness analysis with three approaches (cross-sectional, time series truncation, exclusion of variables)

	Clusters		
	1	2	3
Predictor variable / Explanatory variables	*d_DR_y*	*d_DR_y*	*d_DR_y*
exclusion of variables (1: d_exp, 2: d_emp, 3: d_emp)			
d_emp	–0.0661** (0.0267)		
d_exp			
l1_d_gdp	–0.0003** (0.0001)		
d_gov10y	0.0005** (0.0002)		
d_bub3m	0.0016*** (0.0003)	0.0021** (0.0010)	0.0053*** (0.0013)
d_wealth		–0.0600** (0.0237)	–0.0946** (0.0355)
l3_d_gov1y		0.0020*** (0.0007)	
l1_d_bux		–0.0109* (0.0058)	
l3_d_hinc			–0.0825** (0.0365)
exclusion of variables (1: l1_d_gdp, 2: l3_d_gov1y, 3: d_bub3m)			
d_emp	–0.0672*** (0.0262)	–0.1582** (0.0690)	–0.2646** (0.1086)
d_exp	–0.0190** (0.0069)		
l1_d_gdp			
d_gov10y	0.0005*** (0.0002)		
d_bub3m	0.0016** (0.0003)	0.0028*** (0.0010)	
d_wealth		–0.0546** (0.0252)	–0.1086*** (0.0391)
l3_d_gov1y			
l1_d_bux		–0.0073 (0.0058)	
l3_d_hinc			–0.0707* (0.0403)
exclusion of variables (1: d_gov10y, 2: d_bub3m, 3: l3_d_rendjov)			
d_emp	–0.0497* (0.0270)	–0.1374** (0.0656)	–0.2872*** (0.0925)
d_exp	–0.0182** (0.0069)		
l1_d_gdp	–0.0003** (0.0001)		
d_gov10y			
d_bub3m	0.0014*** (0.0003)		0.0054** (0.0012)
d_wealth		–0.0592** (0.0238)	–0.0891*** (0.0334)
l3_d_gov1y		0.0023*** (0.0007)	
l1_d_bux		–0.0161*** (0.0053)	
l3_d_hinc			

Table 11
Robustness analysis with three approaches (cross-sectional, time series truncation, exclusion of variables)

	Clusters		
	1	2	3
Predictor variable / Explanatory variables	*d_DR_y*	*d_DR_y*	*d_DR_y*
exclusion of variables (1: d_bub3m, 2: l_d_bux, 3: -)			
d_emp	–0.0461 (0.0307)	–0.1540** (0.0637)	
d_exp	–0.0205** (0.0079)		
l1_d_gdp	–0.0003** (0.0001)		
d_gov10y	0.0004* (0.0002)		
d_bub3m		0.0030*** (0.0009)	
d_wealth		–0.0591** (0.0231)	
l3_d_gov1y		0.0018*** (0.0007)	
l1_d_bux			
l3_d_hinc			

*Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are shown in brackets. 'd' is the annual change in the variable, 'l' is the quarterly lag, while 'y' indicates the annual nature of the probability of default.*