Acta Oeconomica
DOI: 10.1556/032.2024.00015

AKADÉMIAI KIADÓ

A stochastic loss given default model of Hungarian residential mortgages

BÁLINT VÁRGEDŐ[1,2]*

[1] Magyar Nemzeti Bank (MNB, the Central Bank of Hungary), Krisztina krt. 55, Budapest, H-1013 Hungary

[2] Corvinus University of Budapest, Hungary

Received: August 4, 2023 • Revised manuscript received: February 21, 2024 • Accepted: June 20, 2024

ABSTRACT

The paper analyses the Loss Given Default (LGD) rates of residential mortgages, using a model based on stochastic collateral value. The implementation of the model is based on exponential Ornstein-Uhlenbeck processes fitted to the Hungarian regions' house price indices. According to the model results, in case of a mortgage with a 80% loan-to-value ratio at origination, the expected LGD is around 30–40%, depending on the region. The highest LGD rates are estimated for villages, while the lowest rates are expected in Budapest and cities in the middle of the country. The range of the regional differences can reach 7 percentage points. According to the LGD Risk index based on the aggregated model, the LGD risk profile of recently issued mortgages has improved significantly since 2009 in Hungary.

Due to the strong negative relation between the house prices and mortgage default rates, the expected return on defaulted collateral value tends to be low. The results could be relevant for credit institutions in their mortgage origination decisions and enhance analysis of lending processes and the associated risks.

KEYWORDS

mortgage loans, recovery rates, stochastic processes, Hungary

JEL CLASSIFICATION INDICES

G21, G33

* Corresponding author. E-mail: vargedob@mnb.hu

1. INTRODUCTION

The determination of credit risk losses is influenced not only by the probability of default (PD) but also by the extent of loss given default (LGD). While there are several advanced solutions for PD modeling of mortgages in Hungary (Banai et al. 2013; Pollák – Popper 2021; Szabó 2022; Burger 2022), the methodologies for estimating LGD are not as well-developed. This deficiency is not limited to a specific country, as it is also evident in the international literature. However, accurate estimation of LGD is crucial for risk management processes within financial institutions, as it directly impacts lending decisions, capital adequacy assessment, and expected credit losses. Moreover, financial supervisors have an interest in developing LGD models as well, to ensure that lending institutions meet capital requirements and to assess the resilience of the financial system to potential distress. The 2008 financial crisis emphasized the importance of accurate modeling of credit risk components for prudent risk management, as oversimplified models can jeopardize financial stability (ECB 2007). Therefore, the development of more advanced tools for LGD modeling not only enhances risk management practices but also promotes financial stability by replacing commonly used constant LGD models in the industry (James 1991; Jokivuolle – Peura 2003).

In the case of PD models, numerous well-known risk drivers exist, including the loan-to-value (LTV) ratio, borrower income, credit history, behavioral patterns, employment stability and fraud events (Deng et al. 1996; Calem – LaCour-Little 2004; Ambrose et al. 2004; Keresztúri et al. 2023). Conversely, LGD ratios are primarily influenced by the collateral quality, LTV ratios (Leow – Mues 2012), and the foreclosure process (Pelizza – Schenk-Hoppé 2020), with borrower-specific information playing a lesser role. Many studies have highlighted a strong correlation between PD and LGD components (Altman et al. 2005; Acharya et al. 2007) due to underlying macroeconomic factors such as house price dynamics. Others have highlighted that fire sales in times of financial distress can play a significant role in determining liquidation values (Shleifer – Vishny 1992), and as a result, increasing the co-movement of PD and LGD. This correlation further underscores the importance of comprehending the drivers of LGD and accurately calculating the severity of losses.

This paper builds upon the research conducted by Frontczak and Rostek (2015) and aims to investigate an LGD model that incorporates a stochastic collateral value. The model's parameterization involves fitting exponential stochastic Ornstein-Uhlenbeck processes to Hungarian house price indices while aligning other parameters with those found in the international literature. The resulting model outcomes are compared to the figures generated by a benchmark model employed by the ECB (2018).

My contributions to the literature are threefold:

- The utilization of this model offers the advantage of effectively distinguishing between various factors that influence the LGD rate, such as the expected appreciation of collateral, expected return, and liquidation costs. The primary focus of the study centers around examining the impact of changes in house prices. Although the modeling framework allows for the quantification of stress scenarios, the main objective is to establish expected LGD rates. While the modeling framework is based on Frontczak – Rostek (2015), the implementation consists of several novelties.

- Due to the granularity of house price indices disaggregated by region and settlement size (capital, city, municipality), it becomes possible to analyze the LGD rate at different geographical levels. This enables a comparison of the LGD risk across diverse regions and settlement types.
- As the resulting expected LGD function solely relies on the loan-to-value (LTV) ratio at origination, it becomes feasible to construct a through-the-cycle LGD risk index using the LTV distribution of new loans. This index provides an estimation of the expected LGD risk within the banking sector based on the quality of the loan portfolio. By capturing the emerging risks associated with relaxed lending conditions in the banking sector, this theoretically grounded measure offers a more comprehensive assessment compared to simpler measures like average LTV value dynamics. Mortgage's LTV developments are important due to their primary role in macroprudential regulation (Mérő 2017; Kim 2021). Moreover, the examination of the results obtained from stochastic processes fitted to Hungarian house price indices contributes a novel perspective to the existing literature. By analyzing the process parameters segmented by region and settlement type, it becomes possible to formulate statements concerning return and risk in property investments.

The subsequent Section 2 provides a summary of the relevant literature, while Section 3 outlines the methodology employed. Section 4 presents the data utilized in the study, Section 5 describe the empirical results and the modeled value of the expected LGD. Finally, Section 6 concludes the paper by providing a summary of the findings.

2. LITERATURE REVIEW

The available literature extensively covers studies on the probability of default (PD) parameter of credit risk, whereas fewer papers focus on modeling loss given default (LGD). The LGD literature is characterized by a division based on the granularity of the data employed. Some studies utilize granular transaction-level data and statistical models to predict transaction- and portfolio-level LGD indicators. Others predominantly rely on aggregate and publicly available data for modeling purposes. Granular models offer better fit and more accurate predictions, while aggregate models are suitable for analyzing mortgage markets lacking uniform, granular datasets (Palmroos 2016a). Aggregate models also require fewer inputs, enabling the definition of less detailed stress scenarios for stress tests.

Several studies concentrate on modeling LGD in the corporate segment proposing models based on granular data. For instance, Gupton and Stein (2002) and Chalupka and Kopecsni (2009) examined LGD in the corporate loans and bonds context. One of the main findings for corporate exposures that holds across different asset classes is that LGDs tend to have bi-modal distributions (Qi – Zhao 2011). In the realm of retail loans, Bellotti and Crook (2012) model LGD rates of credit cards using United Kingdom (UK) data from 1999 to 2005. Their findings revealed that macroeconomic variables, alongside customer data variables, exhibit significant explanatory power. Regarding residential mortgages, Leow and Mues (2012) conducted an empirical analysis using data from a UK bank. They employed a two-stage regression model, which outperforms a single-stage regression approach. According to their findings, the discount rate on the sale price of a repossessed property is an important factor of LGD. Qi and Yang (2009) investigated LGD rates of high loan-to-value (LTV) mortgages. Their study reveals that the LGD is better explained by the current LTV (CLTV) rather than the LTV at origination, and that mortgage loss severity is notably higher in distressed housing markets. Calabrese (2014) proposed a methodology for calculating expected downturn LGD following the Basel regulation, incorporating a mixture of expansion and recession distribution. The results, based on an Italian dataset, suggest that downturn values can exceed average LGDs by up to 17 percentage points.

Frye's (2000) model, based on aggregate data and theoretical foundations, highlights the correlation between collateral value reduction and increasing default rates. He established a link between the change in collateral value and PD through a systematic risk factor. Jokivuolle and Peura (2003) investigated LGD in corporate loans and bonds, considering the relationship between stochastic collateral value and business valuation. Their findings demonstrate that expected LGD decreases as the drift of collateral value increases but rises with increased volatility of collateral value. Van Damme (2011) presented a general framework for incorporating stochastic LGD models into credit risk management and derivative pricing processes. The proposed framework is also suitable for estimating downturn LGDs.

Frontczak and Rostek (2015) introduced the modeling of residential mortgages' LGD with stochastic collateral value. Recognizing that LGD resembles the payoff of a long-put option, they employed derivative pricing methods. Building upon Fabozzi et al. (2012), Frontczak and Rostek utilized an exponential Ornstein-Uhlenbeck process to model residential property values. Pelizza and Schenk-Hoppé (2020) demonstrated that defaulted loans in the Italian mortgage market may be overvalued due to higher-than-expected LGD rates. Examining LGD models used in stress testing provides a more comprehensive understanding of LGD modeling's practical applications. Siemsen and Vilsmeier (2017) employed a CLTV-LGD function to quantify LGD in the Bundesbank's mortgage stress test, by deriving a meta-dependency between CLTV and LGD using the results of Qi and Yang (2009) and Palmroos (2016b). A polynomial regression is applied to combine the results into a continuous function. Holló (2009) calculated LGD during a stress test of Hungarian retail mortgages based on the portion of LTV exceeding 100% at the time of default, updating the value of collateral homes using the house price index scenario. Tajti (2011) analyzed an empirical interbank LGD database and found that the average LGD for home equity loans in Hungary is lower (16.26%) than for residential mortgages (27.11%), albeit with considerable standard deviation in both cases. Tajti also identified a bimodal distribution of registered LGD ratios for Hungarian banks, potentially resulting from technical defaults recorded in the database with zero losses on the instruments.

The proposed model requires less data than the model employed by Tajti (2011), making it suitable for analyses targeting the entire banking system or situations with data shortages. The model parameters are easily interpretable, and the solid theoretical foundation adequately reflects the non-linear relationship of various parameters or variables. This is particularly relevant for the relationship between LTV and LGD rates, as demonstrated by the presented results. A comparison with other relevant models is shown in Table 1.

Berki and Szendrei (2017) analyzed house price dynamics in Hungary using the MNB House Price Index to examine the long-run equilibrium level of house prices determined by fundamentals and the difference between observed prices. Their VECM methodology results indicate a cointegration relationship between the house price index, average new housing loans, household disposable income, and the stock of dwellings. The authors find that house prices take time to return to equilibrium levels and observe a house price gap. These results are in line with the related international literature (Glaeser – Nathanson 2017), where the role of idiosyncratic movements of individual house prices is also emphasized (Case – Schiller 1989).

Table 1. Model comparison

	Tajti (2011)	Holló (2009)	Siemsen – Vilsmeier (2017)	ECB (2018)	Proposed model
Data requirements	Granular	Aggregate	Aggregate	Aggregate	Aggregate
Model type	Reduced	Structural	Reduced	Structural	Structural
Methodology	Regression	Excess exposure	Metadependency	Formula	Stochastic modeling
Source of data	Hungary	Hungary	USA, Finland	Unknown	Hungary

3. METHODOLOGY

3.1. Stochastic LGD model

When mortgages default, banks can mitigate their losses using the collateral. If the market value of the real estate serving as collateral was higher than the bank's exposure at default (EAD), and the bank would be able to sell it at the market price immediately, realising zero loss on the transaction. But if the cash flow from selling the collateral (including the time value of money and workout costs) is less than the outstanding debt, the loss equals the difference between the EAD and the collateral value. In formulaic terms:

$$Loss = \left\{ \begin{array}{cc} EAD - \tilde{C}_{T_D}, & \tilde{C}_{T_D} < EAD \\ 0, & if EAD \leq \tilde{C}_{T_D} \end{array} \right\} \tag{1}$$

where *Loss* is the loss suffered by the bank, T_D is the time of default, and $\tilde{C}_{T_D}$ is the present value of cash flows from selling the collateral, at the time of default. $\tilde{C}_{T_D}$ is equal to the value of collateral at the time of liquidation (C_{T_L} , where T_L is the time of liquidation), discounted at an expected return of r, less the k cost ratio:

$$\tilde{C}_{T_D} = e^{-r(T_L - T_D)}(1-k)C_{T_L} \tag{2}$$

From a modelling perspective, the workout cost ratio (k) can also be seen as the discount factor for the defaulted debtors' properties. It can also be interpreted as a discount factor of foreclosures, which is a popular way for liquidating collateral. The bank's expected return from the collateral's cash flow (r) is the sum of the risk-free return and the required risk premium, which compensates for the uncertainty of the idiosyncratic real estate price.

Building on the above-mentioned notions (Equations (1) and (2)), the LGD can be expressed as a function of the collateral value as follows:

$$\mathrm{LGD}_{T_D} = \max\left(0, \frac{EAD - e^{-r(T_L - T_D)}(1-k)C_{T_L}}{EAD}\right) \tag{3}$$

The parallels can be seen here between *LGD* and the payoff of a long-put option (for the $\frac{\tilde{C}_{T_D}}{EAD}$ process, with a strike of 1).

Accordingly, the expected LGD is:

$$E[LGD_{T_D}] = E\left[\max\left(0, \frac{EAD - e^{-r(T_L - T_D)}(1-k)C_{T_L}}{EAD}\right)\right] \tag{4}$$

The expected value can be derived by defining an appropriate C_T process, using integral functions similar to pricing the option. Unlike in option pricing, estimating the LGD requires a "real world" probability distribution, therefore the parameters of house price developments can only be defined using historical data.[1]

3.2. Stochastic house price processes

After defining an appropriate stochastic process, the distribution of the collateral value at the time of liquidation can be derived. The collateral value process was modelled as the sum of a general real estate market process and a process describing idiosyncratic price movements. The cumulative logarithmic return of the collateral value (Y_t) is thus derived as the sum of the market-based cumulative log return (Y_t^{market}) and the idiosyncratic cumulative log return (Y_t^{idio}):

$$Y_t = Y_t^{market} + Y_t^{idio}. \tag{5}$$

The collateral value can be expressed with the following equation:

$$C_t = C_0 e^{\left(Y_t^{market} + Y_t^{idio}\right)} = C_0 e^{Y_t}. \tag{6}$$

Fabozzi et al. (2012) recommend using an exponential Ornstein–Uhlenbeck (OU) process for modelling the price level of the real estate market. The proposed process satisfies three phenomena observed in international residential property markets:

- property prices fluctuate in the short term but reverse to a long-run trend (mean reversion),
- property market returns are autocorrelated,
- property prices are guaranteed to be positive at all times.

If real estate prices follow an exponential OU process, the log of the prices (Y_t^{market}), which can be seen as the cumulative log return of house price indices, follows an OU process in all $0<t$ periods. The stochastic differential equation defining the OU process is:

$$dY_t^{market} = \left(\frac{d\varphi_t}{dt} + \kappa\left(\varphi_t - Y_t^{market}\right)\right)dt + \sigma^{market} dW_t^{market} \tag{7}$$

where φ_t is the long-run trend of logarithmised house price indices, and it can be approximated based on a simple linear model. κ denotes the parameter for the strength of mean reversion, and σ^{market} shows the volatility of the process. W_t^{market} is the standard Brownian motion, which is responsible for the uncertainty of the process.[2] The three terms on the right side of the equation are responsible for trend growth, mean reversion and uncertainty, respectively.

[1] No replicating portfolio has to be created during this process. It would only be required when looking at how banks can hedge their exposure to property market price changes through the LGD by replicas built using property market assets.

[2] The stochastic differential equation is an extension of the Vasicek Interest Rate Model, using a trend instead of the long-run average.

The solution of the stochastic differential equation is the following Gaussian process for all $t<T$ times:

$$Y_T^{market} = \varphi_T - \left(\varphi_t - Y_t^{market}\right)e^{-\kappa(T-t)} + \sigma^{market}\int_t^T e^{-\kappa(T-s)}dW_s^{market} \tag{8}$$

Based on the solution of the stochastic differential equation, the conditional growth of the Y_t process (μ_Y) can be derived,

$$\mu_Y^{market} = E\left[Y_T^{market}|Y_t^{market}\right] = \varphi_T - \left(\varphi_t - Y_t^{market}\right)e^{-\kappa(T-t)} \tag{9}$$

along with its conditional standard deviation (σ_Y^{market}),

$$\sigma_Y^{market} = \sqrt{Var\left[Y_T^{market}|Y_t^{market}\right]} = \sigma^{market}\sqrt{\frac{1-e^{-2\kappa(T-t)}}{2\kappa}}. \tag{10}$$

The term on the right side of Equation (10) shows the volatility-reducing effect of mean reversion. It also has to be noted that if κ is close to zero, i.e. when mean reversion is very low or non-existent, the OU process is reduced to a Brownian motion with a drift. Therefore, the conditional standard deviation approximates σ^{market} too, if the mean reversion cannot mitigate the uncertainty of the process. The results obtained so far for the real estate market process are practically identical to Frontczak and Rostek (2015), the contribution to the theoretical model framework is the addition of the idiosyncratic price process described next.

Besides the general property market developments, the collateral value is also influenced by the idiosyncratic price movements of the given collateral. These idiosyncratic price movements cannot be captured by house price indices, which are designed to filter these out. The following assumptions were applied to the idiosyncratic price movements measured in log returns:

- expected value of zero,
- standard deviation reflects the idiosyncratic standard deviation of properties,
- independent from overall real estate market trends,
- independent from earlier and later idiosyncratic price movements,
- follow normal distribution.

Accordingly, the cumulative idiosyncratic log returns (Y_t^{idio}) can be written as follows:

$$dY_t^{idio} = \sigma^{idio}dW_t^{idio}, \tag{11}$$

where σ^{idio} is the standard deviation of the idiosyncratic log returns, W_t^{idio} is the standard Brownian motion, which is independent from the W_t^{market} process shaping market returns. The solution shows that its conditional variance is as follows:

$$Var\left[Y_T^{idio}|Y_t^{idio}\right] = \left(\sigma^{idio}\right)^2 \cdot (T-t) \tag{12}$$

The results of the market and idiosyncratic processes can be used to determine the distribution of collateral value. Substituting the processes into Equation (5), one arrives at the following equation, according to which this process is also a Gaussian process:

$$Y_T = \varphi_T - \left(\varphi_t - Y_t^{market}\right)e^{-\kappa(T-t)} + \sigma^{market}\int_t^T e^{-\kappa(T-s)}dW_s^{market} + \sigma^{idio}\cdot W_T^{idio}, \tag{13}$$

the expected value and variance of which is as follows, taking into account that the two processes are independent:

$$\mu_Y = \mu_Y^{market} + 0 = \varphi_T - \left(\varphi_t - Y_t^{market}\right)e^{-\kappa(T-t)} \tag{14}$$

$$\sigma_Y^2 = Var\left[Y_T^{index}|Y_t^{index}\right] + Var\left[Y_T^{idio}|Y_t^{idio}\right] = \left(\sigma^{market}\right)^2 \frac{1-e^{-2\kappa(T-t)}}{2\kappa} + \left(\sigma^{idio}\right)^2 \cdot (T-t) \tag{15}$$

By looking at the log of house prices, it can be seen about house price levels that the $C_{T_L}|C_t$ probabilistic variable follows a lognormal distribution in all $t<T$ periods, as its log $(Y_{T_L}|Y_t)$ follows a $N(\mu_Y, \sigma_Y^2)$ distribution. If the *LGD* function from the previous section is substituted as a function of C_{T_L} (Equation (3)), along with the previously determined $C_{T_L}|C_t$ distribution, the integral equation of the expected *LGD* is derived, which can be solved by substituting Equation (6). The integral equation can be interpreted as the expected value of the transformed property price, where the first term is the real estate price transformed with the LGD function, and the second is the distribution of the property price.

$$E[LGD_{T_D}] = \int \max\left(0, \frac{EAD - e^{-r(T_L-T_D)}(1-k)C_{T_L}}{EAD}\right) \frac{1}{C_{T_L}\sqrt{2\pi\sigma_Y^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln C_{T_L} - (\ln C_t + \mu_Y)}{\sigma_Y}\right)^2\right] \tag{16}$$

The expected *LGD* formula can be further developed by solving the integral,

$$E[LGD_{T_D}] = \Phi(-d) - (1-k)e^{-r(T_L-T_D)}\frac{C_t}{EAD}e^{\mu_Y+\frac{1}{2}\sigma_Y^2}\Phi\left(-(d+\sigma_Y)\right) \tag{17}$$

where $d = \frac{\ln\frac{C_t}{X}+\mu_Y}{\sigma_Y}$, and $X = \frac{EAD}{1-k}e^{r(T_L-T_D)}$.

In line with intuition, the produced expected *LGD* formula decreases in the C_t value, as a higher initial collateral entails lower LGD values. Similarly, as the $\frac{C_t}{EAD}$ ratio, or the LTV, grows, loss declines (Frontczak – Rostek 2015). The formula in Equation (17) was validated with Monte Carlo simulations (see Appendix).

3.3. Calibrating the collateral value parameters

The collateral value parameters (μ_Y, σ_Y) were determined using market parameters and idiosyncratic standard deviation, as shown in Equations (14) and (15). The market parameters were estimated based on the house price index process, while the idiosyncratic standard deviation was determined using results from the international literature. The house price trend was not estimated with the trend of the raw house price index, but instead a trend was calculated using the house price index and a default rate time series. This method controlled for the covariance of mortgage defaults and housing market developments. Applying this method was necessary because low real estate market returns are overrepresented among defaulted loans due to the composition effect caused by the defaults.[3]

[3]It is necessary to consider that default events are not independent of the return, when estimating the average return on a defaulted collateral. This phenomenon implies a composition effect, as there are more observations when the default rates are high.

The following sections detail and justify the calibration steps. First, the data cleaning performed on the raw house price index process and the estimations of the OU parameters are described, as recommended by Fabozzi et al. (2012). Then the modified trend function is defined. Finally, the idiosyncratic standard deviation is incorporated.

Choosing an appropriate house price index is critical for modeling property price developments. In Hungary, three well-known indices are published quarterly: the MNB House Price Index, the Hungarian Central Statistical Office (HCSO) index, and the Takarékbank index. The MNB House Price Index, created by Banai et al. (2017) using hedonic regression and period pair estimates, offers an aggregate index for the entire country, along with a detailed breakdown by region and settlement size. Since this breakdown is essential to produce regional results, the MNB index was used during the analysis. The estimation of the parameters of the exponential OU process fitted on the house price index was performed in two stages, as recommended by Fabozzi et al. (2012). Fabozzi et al. (2012) advise that as a first step, a linear trend model – $\varphi_t = a + b \cdot t$ – should be estimated for a logarithmised price index (Y_t^{index}).

$$Y_t^{index} = \varphi_t + \tilde{Y}_t = a + b \cdot t + \tilde{Y}_t \tag{18}$$

The derived $\tilde{Y}_t$ is the detrended transformation of the log house price process. Therefore, the detrended part is derived as the error term of the trend model, with an expected value of zero, and it follows an OU process according to the model. Then the parameters of the derived OU process with a constant mean, are estimated with the least square error (LSE) method outlined by Chaiyapo and Phewchean (2017).

Substituting $T = t + \Delta t$ into Equation (8), using the results of Equations (9) and (10) and the substitution $Y_t^{index} = \varphi_t + \tilde{Y}_t$, the following discretised equation is derived, where ΔW_t^{market} is the change in the Brownian motion with an independent and identical distribution, following a $N(0, 1)$ distribution:

$$\tilde{Y}_{t+\Delta t} = \tilde{Y}_t \cdot e^{-k\Delta t} + \sigma^{market} \sqrt{\frac{1 - \exp(-2k\Delta t)}{2k}} \Delta W_t^{market} \tag{19}$$

The parameters of the equation can be estimated with the following model:

$$\tilde{Y}_{t+\Delta t} = \beta \cdot \tilde{Y}_t + \varepsilon_{t+\Delta t} \tag{20}$$

where $\beta = e^{-\kappa \cdot \Delta t}$, $\varepsilon_{t+\Delta t} \sim N\left(0, \sigma^{market} \sqrt{\frac{1 - \exp(-2\kappa\Delta t)}{2\kappa}}\right)$ is independent and identically distributed. By transforming the equations, one arrives at the following results:

$$\kappa = \frac{-\log(\beta)}{\Delta t} \tag{21}$$

$$\sigma^{market} = sd(\varepsilon_t) \cdot \sqrt{\frac{2\kappa}{1 - \exp(-2\kappa\Delta t)}} \tag{22}$$

During the estimation, the value of Δt was set as one quarter of a year, based on the data frequency in the house price index.

In distressed economic times, default rates often spike all across the economy, including the mortgage market (Koopman 2009; Ali 2010). Moreover, housing market returns typically underperform in times of stress (Rünstler 2018). Accordingly, low returns are overrepresented among defaulted loans, just like defaults occurring in times of stress are overrepresented among non-performing loans in the whole period. Therefore, the historical average of house price returns does not equal the average price of the homes used as collateral for the non-performing loans, simply because of the timing of the returns.[4]

Assuming that the collateral is sold in a similar housing market environment as observed at the time of default, the historical average return on the collateral ($\bar{c}_t^{coll}$) can be approximated with the weighted house price index returns (c_t^{index}) by using default rates from the given year (DR_t), as the number of collateral items covering the non-performing loans is proportional to the default rate. Introducing $\overline{DR}$ for the average of the DR_t time series, the estimation of the historical average is formulaically done as follows:

$$\bar{c}_t^{coll} = \sum_{t=1}^{n} c_t^{index} \cdot \frac{DR_t}{\overline{DR}} \cdot \frac{1}{n} \tag{23}$$

The $\bar{c}_t^{coll}$ thus derived is used in the trend model of the collateral, instead of the b coefficient estimated for the unmodified house price index. The other market OU parameters (σ^{market}, κ) used during modelling still come from the detrended and logarithmised house price index ($\tilde{Y}_t$).

No published papers have been attempting to determine the idiosyncratic standard deviation in the Hungarian housing market, therefore the results of international studies were used. Miller and Pandher (2008) employed a two-factor model to examine the price changes in American sub-housing markets distinguished by ZIP codes. The two explanatory factors were equity market returns and the changes in the national house price index, which helped the authors adjust for the systematic economy-wide risks and the overall risk of the housing market. They estimated that the representative idiosyncratic risk was around 10%. Giacoletti (2021) analysed individual house price transactions in California, adjusting for the price changes in the ZIP code-level housing market to determine abnormal returns related to individual homes. According to the curve of the idiosyncratic standard deviation that they derived, idiosyncratic volatility is 19% in the case of the homes sold within 2 years, 12% for those sold within 5 years, and only 7.5% for those sold within 15 years. Based on the results from the international literature, the price changes of individual houses may experience an annualised idiosyncratic standard deviation (σ^{idio}) of around 10% in excess of the movements in the regional house price index.

In connection with idiosyncratic standard deviation, the country-level aggregate house price index of villages has to be addressed. As seen in the case of cities, the country-level aggregate index conceals some of the standard deviation in regional indices, meaning that the standard deviation of the aggregate indicator is lower. To allow this effect to take hold in the case of villages, where this could be even more pronounced than in the case of cities due to the large number of settlements and their potentially higher heterogeneity, synthetic village volatility parameters were created for each region. The LGD rates estimated using these can be employed to compare the LGD levels across different regions. The derived volatility parameter reflects the overall risk of villages and the risk of the region relative to the country, which was quantified based on cities' volatilities. Formally, these were derived based on the city index ($\sigma_{C,i}^{market}$) for the

[4]The collateral behind the defaulted loans may also underperform the total stock of properties in a cross-sectional analysis, but this is controlled for in the size of the auction discount rather than in the parameters of the house price process.

given region (i), by scaling up the volatility parameter ratios of the municipality ($\sigma_{V,aggr}^{market}$) and city ($\sigma_{C,aggr}^{market}$) aggregate indices:

$$\sigma_{V,i}^{market} = \sigma_{V,aggr}^{market} \cdot \frac{\sigma_{C,i}^{market}}{\sigma_{C,aggr}^{market}}. \tag{24}$$

This yields the following equations describing the collateral value process parameters of the different regions, for every region and settlement type (i). The expected value (based on Equation (14)) is the sum of the log return of the expected collateral and the difference between the given region or settlement type and the national trend, multiplied by the time until liquidation:

$$\mu_Y(i) = \left(\bar{c}_t^{coll} + b_i - b_{aggr}\right) \cdot T_L \tag{25}$$

Conditional variance (σ_Y^2, based on Equation (15)) is the sum of the market variance of the region or settlement type and the idiosyncratic variance:

$$\sigma_Y^2(i) = \left(\sigma_i^{market}\right)^2 \frac{1 - e^{-2\kappa_i \cdot T_L}}{2\kappa_i} + \left(\sigma^{idio}\right)^2 \cdot T_L. \tag{26}$$

3.4. Other LGD parameters

Besides the house price process parameters, several other factors need to be determined in the outlined stochastic LGD model, see Equation (17). These parameters were set in line with the related literature. The EAD was set in a conservative approach, as the exposure at default equals the loan amount at origination. In other words, the loan amount does not amortise. This is in line with the assumption of the LGD model used in the stress test of Siemsen and Vilsmeier (2017). This simplifying assumption also ensures the static balance sheet assumption at the portfolio level, because in the period until the time of default new loans can be expected along with potential repayments, and these two effects cancel out each other across the portfolio as a whole. The assumption that *EAD* remains constant can be considered conservative, but has a limited effect on account of the short, 1-year timeframe until the default. The cash recovery rate outlined in Frontczak and Rostek (2015) was not considered either, in other words no further repayments are expected from defaulted debtors.

The parameter k was used as the discount factor for the collateral covering non-performing loans. The fact that a debtor default may offer additional information about the value of their property. This was confirmed by Campbell et al. (2011), who showed that the market applied a 27% cross-sectional discount for foreclosed residential properties in their sample from Massachusetts. The authors attributed this to the condition of the properties and illiquidity problems, among other things. In the light of their study, the k parameter was set at 30%, so that it reflected the workout costs incurred by the bank.

The value of the cash flows received after the default date needs to be discounted. In accordance with the industry standard established by European Banking Authority (2017), this is the sum of a 3-month reference rate (e.g. BUBOR) and a 5-percentage point risk premium. In the interest rate environment of 2021, this means a discount rate of 10%. The time between the starting date and the default date, T_D, is 1 year, the time until selling, T_L, is 4 years, in line with the selling time of 3 years following the default date as recommended in the model documentation of the LGD used as a benchmark. This means that house prices change for 4 years, up until the time of sale. The Financial Stability Report of the MNB from December 2019 also suggests that defaulting debtors are usually in arrears for several years before coming to an agreement.

4. DATA

The country-wide aggregate time of the MNB's House Price Index is available from 1990 Q1, while the time series broken down by region and settlement size starts in 2001 Q1, with quarterly frequency. The last observation in the sample is from 2021 Q3. The analysis uses the nominal house price index, as the nominal appreciation or depreciation of the collateral is what matters to banks from a credit risk perspective.

As the residential property market is heterogeneous from a geographical standpoint, the price changes in the different regions and settlement types may differ considerably from the country-wide average. To avoid losing any information, more granular subindices were used instead of an aggregate country-wide index. This yields different models for the different settlement types and regions, which can provide interesting results from an analytical perspective. Furthermore, this ensures that the idiosyncratic standard deviation of the regions does not disappear during aggregation, which can also be important as the LGD grows in house price volatility. The advantages arising from using subindices outweigh the drawbacks of having a shorter time series. To produce LGD results that can be interpreted as a country-wide average, the subindex weights are used to produce an aggregated model.

The default rate time series used for revealing the relationship between house price returns and the frequency of default was prepared using residential mortgage data from 5 Hungarian banks. The period covered stretches between 2004 and 2017, therefore it includes a full credit cycle. According to a uniform definition of default, loans become non-performing if they are persistently 90 days past due[5] or if the period of delinquency is over 365 days. Since loans must be multiple times 90 days past due to be considered defaulted, the probability of a non-performing loan becoming performing is much lower than with the popular 90-day default concept. The liquidation of the collateral and the selling of the loan is relevant in the case of the debtors who have been delinquent for a longer period, and this stricter definition of default fits the model well.

LTV distributions were necessary for determining the portfolio-level LGD values presented as the LGD risk index later. These distributions were produced using the data about typical LTV ratios in Hungary as published in the December 2017 Macroprudential Report of the MNB, and its May 2020 and June 2021 Financial Stability Report. The distributions were derived from the LTV values of the new mortgages disbursed in the two years prior to the year in question. The distribution time series thus produced has a frequency of two years.

[5]Persistently, in this context it means that borrowers need to be in 90 days latency multiple times to be defined as defaulted, a one-time 90 days past due is not sufficient.

5. RESULTS

5.1. Exponential Ornstein-Uhlenbeck process fitted to house price indices

The development of the (aggregate and granular) house price indices covered in the analysis between 2001 Q1 and 2021 Q3 is shown in Figure 1. It is clear that the nominal prices of residential properties increased between 2001 and 2008, before falling until 2013–2014 as a result of the economic crisis, which started on the American mortgage market (Marer 2010). Up until the end of the period under review, real estate prices were steadily characterised by strong growth, even though this period includes the Covid crisis that erupted in early 2020. The figure indicates that property prices moved similarly across the various regions of the country in the same periods.

The quantitative analysis is based on the log house price index, but it is also worth looking at the difference between the log house prices, the log returns, to gain a better understanding of the data. Table 2 contains the descriptive statistics of the quarterly log house price returns produced from the MNB House Price Index. The lowest average log returns were observed in the case of the property prices in villages, lagging well behind the other indices in the covered period. The greatest average growth was recorded in cities in Central Hungary, the Northern Great Plain and Central Transdanubia, followed by Budapest. The standard deviation of the log returns was the

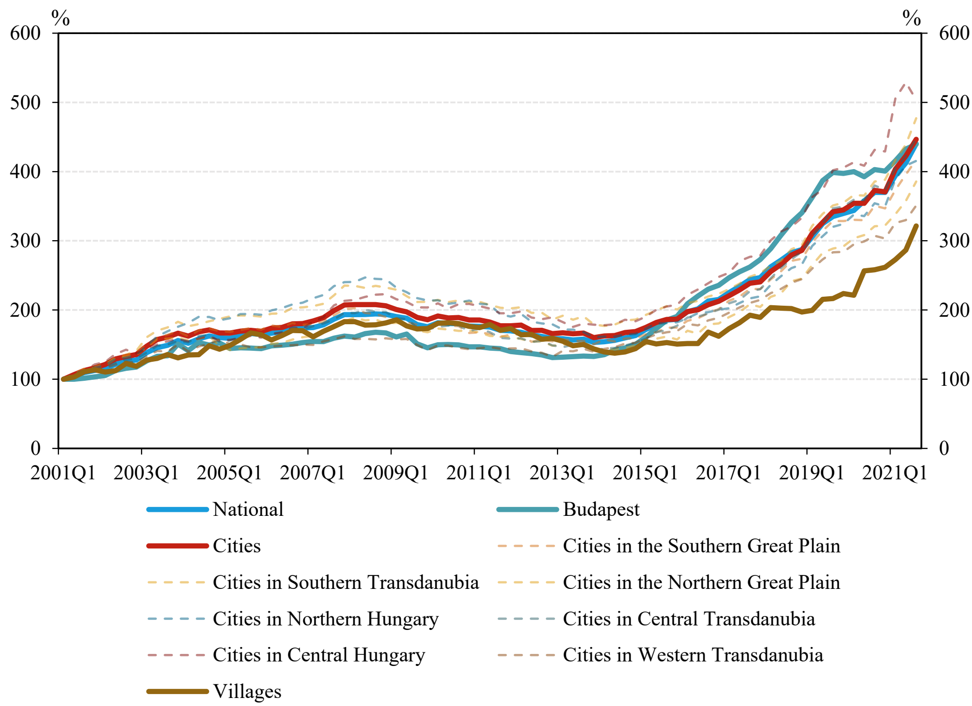

Fig. 1. MNB House Price Index by regions, 2001 Q1 – 2021 Q3
Source: MNB.

Table 2. Descriptive statistics of the quarterly log returns derived from the MNB House Price Index, 2001 Q1 – 2021 Q3 (%)

	Mean	Standard deviation	Minimum	Maximum	Median
National	1.81	2.71	−4.94	8.00	1.73
Budapest	1.81	3.51	−8.92	11.86	1.74
Cities	1.82	2.96	−4.45	8.50	1.87
Cities in the Southern Great Plain	1.74	3.35	−6.73	8.37	1.80
Cities in Southern Transdanubia	1.65	3.57	−9.67	8.93	1.35
Cities in the Northern Great Plain	1.91	3.36	−6.49	13.03	1.87
Cities in Northern Hungary	1.74	3.46	−7.20	10.56	1.65
Cities in Central Transdanubia	1.83	3.37	−6.80	8.01	2.15
Cities in Central Hungary	1.97	3.70	−4.91	16.42	2.10
Cities in Western Transdanubia	1.53	3.07	−6.83	7.20	1.79
Villages	1.42	4.22	−5.37	14.78	1.03

largest in villages, meaning that the lowest returns were coupled with the greatest risk in such settlements based on standard deviation. This underperformance may be explained by the urbanisation processes of the past two decades. In the case of Central Hungary's index, the higher return is coupled with greater standard deviation, the largest behind villages. This is partly attributable to log returns from 2021 Q1, which was exceptionally high at around 20%, making it the highest in the whole sample. The low standard deviation of the aggregate city index shows why it is worth employing subindices so that as little of the uncertainty of real estate prices is lost during modelling as possible.

In accordance with the Methodology Section, the trend fitted on logarithmised house prices was determined with a simple linear model. During the calibration of this parameter, the dependent variable in the linear model used for trend-adjustment is the log house price, and the explanatory variable is time (measured in years), besides the constant, see Equation (18). From a modelling perspective, the trend thus established can be considered the deterministic part of house prices, while the deviations from the trend can be seen as the stochastic part. The results of the regressions for the individual regional indices are shown in Table 3.

The regression coefficients were clearly significant (even at 1%) in all regions, meaning that nominal property prices exhibited a growing trend across the country during the period under review. The results of the trend models are consistent with the geographical heterogeneity observed with the average log returns of the subindices, as well as with the relative differences between the regions.

The process for establishing the expected return from the collateral derived from the covariance of house price log returns and default rates can be found in the Methodology section. The DR_t time series has an annual frequency, so annual log returns were used for the modelling. The default rate time series is shown in Figure 2. The average rate was 1.54% in the period under

Table 3. The results of the linear trend model

		α	β	sd(β)	R^2
National	Estimate	4.768	0.045	0.003	0.713
	P-value	0.000***	0.000***		
Budapest	Estimate	4.587	0.057	0.004	0.708
	P-value	0.000***	0.000***		
Cities	Estimate	4.833	0.042	0.003	0.678
	P-value	0.000***	0.000***		
Cities in the Southern Great Plain	Estimate	4.832	0.039	0.003	0.608
	P-value	0.000***	0.000***		
Cities in Southern Transdanubia	Estimate	4.789	0.036	0.003	0.607
	P-value	0.000***	0.000***		
Cities in the Northern Great Plain	Estimate	4.933	0.041	0.004	0.659
	P-value	0.000***	0.000***		
Cities in Northern Hungary	Estimate	4.964	0.034	0.004	0.491
	P-value	0.000***	0.000***		
Cities in Central Transdanubia	Estimate	4.768	0.043	0.004	0.625
	P-value	0.000***	0.000***		
Cities in Central Hungary	Estimate	4.789	0.053	0.003	0.780
	P-value	0.000***	0.000***		
Cities in Western Transdanubia	Estimate	4.707	0.039	0.003	0.697
	P-value	0.000***	0.000***		
Villages	Estimate	4.816	0.028	0.003	0.576
	P-value	0.000***	0.000***		

Note: Estimated for the regional log house price indices. For each region, the first row contains the estimated coefficients, the second row the related *P*-values. The third column represents the standard errors of the estimated trend coefficients. The last column contains the R^2 indicators of the trend models.

review. The figure indicates that the house price index and default rates changed for the worse in 2008 and started recovering in 2013–2014. Therefore, there is a qualitative relationship between the two indicators. This relationship is statistically significant, with a Pearson correlation of −0.86. The negative sign of the correlation coefficient is consistent with economic logic. After calculating the weighted average (Equation (23)), it can be derived that the annual expected return of collateral is −0.16%. In other words, an average growth rate of close to zero is typical for collateral value in the various regions, in contrast to the roughly 4% rates seen in the trend models.

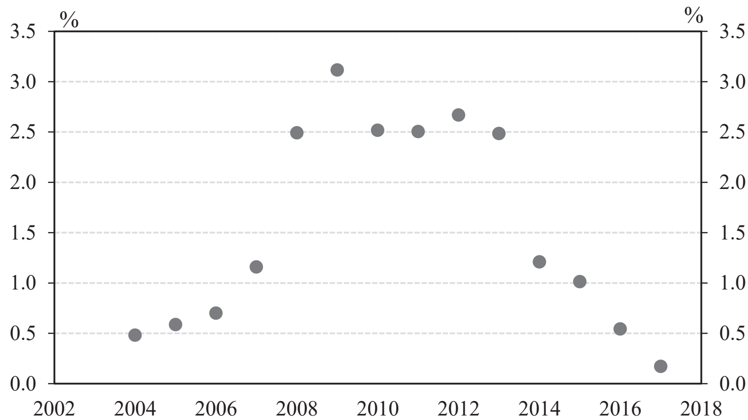

Fig. 2. Default rates of residential mortgages
Note: The rates were determined employing a uniform definition of default between 2004 and 2017, using data from 5 Hungarian banks.
Source: MNB.

The detrended log house price index is derived by taking the log house price index of a given period less the value estimated based on the trend model. The resulting detrended time series oscillates around zero, as shown in Figure 3. The phases in the fluctuation of the housing market processes described at the section on the house price index can also be seen. The transformed time series also indicates that the deviation from the trend between 2004 and 2008 stagnated for a long time. By contrast, this stagnation cannot be observed in the real estate market cycle that began in 2015, even though the deviation from the trend is already well over the 2008 value in all regions.

The mean reversion κ and the index volatility σ^{market} parameters were determined using the detrended time series and the estimation process described in the Methodology section.[6] The estimated values of κ and σ^{market} can be seen in Table 4, with a time unit of one year. It is interesting to note that for example in the case of the country-wide aggregate house price index, the estimated parameter of the mean reversion is negative and close to zero. Additionally, the values that are very close to zero are typical for all subindices except for the indices of villages and the cities from a few regions. This proves that the nominal property prices in Hungary are not characterised by a strong mean reversion, or at least the period under review definitely did not reflect this. This is consistent with the results of Berki and Szendrei (2017) who found that the house price gap is persistent in Hungary. However, this finding runs counter to what is observed in international indices. Fabozzi et al. (2012) estimated a κ parameter of 0.11 for the house price index of the UK, while Pelizza and Schenk-Hoppé (2020) found a value between 0.06 and 0.15 for Italian regional indices. Besides the special features of the Hungarian housing

[6]Results of the estimated autoregressive models are not published due to lack of space but can be obtained directly from the author.

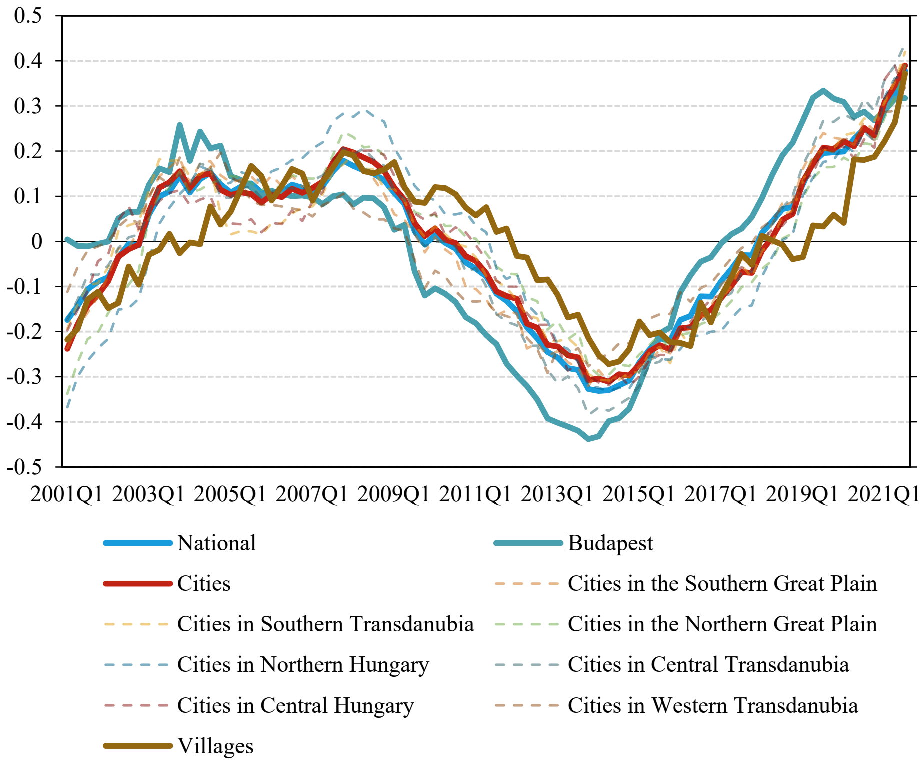

Fig. 3. Detrended logarithmised house price indices, 2001 Q1 – 2021 Q3
Note: The detrended value is the difference between the log house price index and the value estimated based on the trend model.

market, one possible explanation could be the duration and cyclical nature of the sample period. The period under review is not long, and while it covers an entire cycle, it only contains the growth phase of the other cycle that began in 2014. It can be assumed that if the model covered four or five whole real estate market cycles, it would be less sensitive to this growth phase.

Nevertheless, these estimated parameters were used for two reasons. First, from a modelling perspective the threshold value of the OU process approximates the Brownian motion (with a drift) when the κ^{index} values are around zero, which can thus be considered a special case of the general model. Second, from the perspective of LGD modelling, lower κ^{index} values entail greater standard deviation, leading to somewhat more conservative LGD rates, which can be considered a prudent estimate.

According to the model's assumption, the residuals follow normal distribution, as they are derived as the growth in the Brownian motion. The normality of the residuals can be confirmed with the Jarque-Bera test. The results of the test show that the hypothesis of the residuals' normality cannot be rejected at a 5% level, only in the case of a subindex for cities in Central Hungary. And even in this case the low P value is due to the outlier value from the

Table 4. Estimated parameters of the Ornstein–Uhlenbeck processes

	κ^{index}	σ^{index}
National	−0.042	0.054
Budapest	0.000	0.070
Cities	−0.016	0.059
Cities in the Southern Great Plain	−0.007	0.067
Cities in Southern Transdanubia	−0.027	0.071
Cities in the Northern Great Plain	0.050	0.067
Cities in Northern Hungary	0.062	0.070
Cities in Central Transdanubia	−0.041	0.067
Cities in Central Hungary	0.040	0.074
Cities in Western Transdanubia	−0.027	0.061
Villages	0.072	0.085

Note: Estimated for log house price indices adjusted for the regional trend.

above-mentioned 2021 Q1. After removing this observation, the normality hypothesis can be confirmed for this subindex, too.

5.2. Volatility of individual property prices

Using the estimated volatility of the index, the mean reversion parameter and the idiosyncratic standard deviation, the volatility of the cumulative log returns of individual properties can be derived based on the conditional variance equation (Equation (26)). These can differ in all region and settlement pairs due to their different risk profile, which is reflected in the estimated parameters of the Ornstein-Uhlenbeck process. In the case of the village subindex, the volatility parameter was derived synthetically as described in the Methodology section.

The volatility of the cumulative log returns of individual properties is an important component in the LGD model, and it can serve as a useful indicator for property owners when measuring the riskiness of their investments. A common feature of real estate exposures is that investors hold a low number of assets, mostly only one, therefore they are exposed to the associated idiosyncratic risks. Hence a substantial part of real estate investors do not enjoy the benefits of diversification since the real estate part of their portfolio is relatively large compared to their total wealth. Figure 4 shows the standard deviation of the cumulative returns as a function of the number of years elapsed, along with the simple estimate for the same calculated from the volatility of the country-wide index, which does not take into account the idiosyncratic standard deviation. It can be seen that the "naïve" estimate very much underestimates the riskiness of the investment if the magnitude of the idiosyncratic standard deviation applied is accepted. Since the estimated volatility parameter of the regional index pertaining to villages is greater, a higher risk is coupled with these indices in the first few years. However, it is

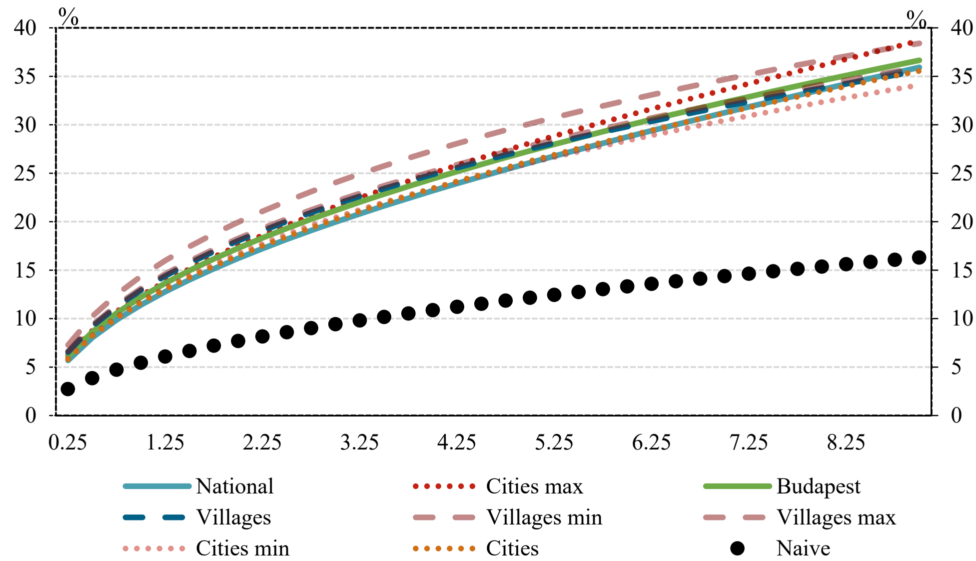

Fig. 4. Standard deviation of individual property investments' cumulative log returns'
Note: The figure shows the standard deviation of the cumulative log returns of individual property investments in various locations, as a function of the investment horizon.

striking that over time, around the 6th or 7th year, the volatility of other indices becomes higher than the value for the villages. This is due to the relatively strong mean reversion parameter of the villages, which reduces risk in the long run. The detailed volatilities broken down by regions and settlement types can be found in Table 6 of Appendix.

5.3. Regional breakdown of the resulting LGD values

The parameters of the estimated housing market processes and the other parameters relevant from a credit risk perspective discussed in the Methodology section can be used to create an LGD model based on Equation (17). The return expected until the time of liquidation and volatility could be determined using Equations (25) and (26), broken down by regions and settlement types. The expected return (μ_Y) was between -7.5% and 4%, while volatility (σ_Y) was between 23.4% and 27.3%, depending on the region, see Table 7. All the terms of Equation (17) were determined for all regions, except for $\frac{C_t}{EAD}$, which is the LTV of the mortgage. The LTV is a widely used measure of risk in the industry, so the LGD rates were established using that.

Figure 5 shows the regional LGD rates as a function of the LTV, with the country-wide LGD produced as the weighted average of the regional models depicted with black dots. The weights are proportional to the number of property transactions in a region. Assuming that loan origination is also proportional to the number of transactions, the weighted indicator can be interpreted as the LGD of a bank's loan representative of the whole country. As expected, the weighted regional LGD is higher than the one calculated from the country-wide index for all

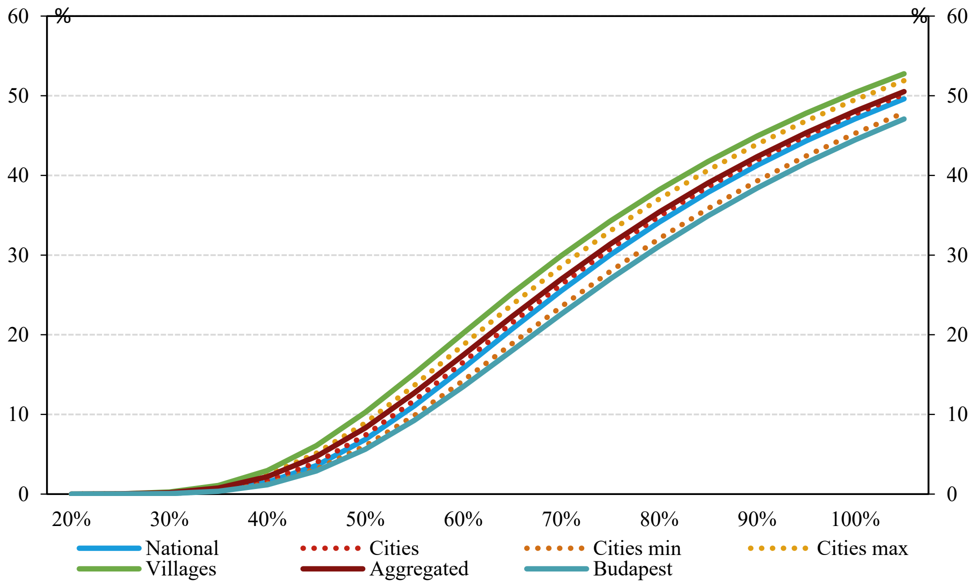

Fig. 5. **LGD rates as a function of initial LTV by settlement types**

LTV values, and the maximum difference is 1.6 percentage points. The weighted LGD function can be approximated accurately using MSE optimalisation, if $\mu_Y = -2.9\%$ and $\sigma_Y = 25.1\%$ are substituted in Equation (17). The advantage of these closed-form solutions is that the low calculation requirement can be maintained for the weighted model.

In the case of non-performance, the loss given default is the highest in villages and the lowest in Budapest and the cities of Central Hungary. This suggests that the expected return had the greatest effect on the LGD, in contrast to volatility and the reversion parameter that influences it. With an 80% LTV, the LGD is 38% for villages in the Northern Great Plain and 31% for Budapest, so the difference between the regions is around 7 percentage points. The average country-wide LGD of 35%, estimated at an 80% LTV, could be relevant for the currently effective debt cap rules, too. The detailed LGD rates broken down by region and settlement type can be found in Table 8 of Appendix.

Figure 5 clearly shows the non-linear nature of the LGD, as the slope of the function is small with lower LTVs, but it increases with LTVs of over 50%. Accordingly, debtors with a higher LTV entail a much higher LGD for banks in the case of non-performance than those with a low LTV. With extremely high LTVs, the slope starts to decline, but this is less relevant due to the current debt cap rules. The figure indicates that the 80% debt cap limit can be found on the steep section of the LGD functions, and even a small change would have a relatively major effect on the LGD.

5.4. Comparison to a benchmark model

To examine how realistic the produced results are, they should be compared to the numbers generated by a challenger model. The chosen benchmark model is the residential mortgage LGL

(Loss Given Loss) model used by the ECB during its 2018 Asset Quality Review (ECB 2018). The model also employs the LTV to determine the loss given default with the help of a continuous function. The initial LTV is modified by the model through the change in real estate prices and the cost ratio, thereby producing an adjusted LTV. Then using a continuous sigmoid function taking into account the haircuts used for selling the collateral and the volatility of sales, the LGD rate is assigned to the adjusted LTV. In order to use the model, the change in property prices, the sales ratio and standard deviation of collateral, the time elapsed until the default and the sale and the appropriate discount rate need to be established. The manual proposes values for these so that a prudent "downturn" or distressed LGD function can be determined. In this case, a 10% fall in real estate prices makes the function distressed.

Figure 6 compares the weighted country-wide LGD function to the downturn benchmark model and the results of a non-downturn benchmark function derived without any contraction in house prices. The discount rate was set at 10% for both benchmark models, similar to the stochastic model. The figure indicates that if the LTV is greater than 40%, the weighted country-wide LGD function is lower than the downturn benchmark model but higher than the non-downturn function. Basically, the two non-downturn LGD functions are similar.

5.5. Hungarian residential mortgage portfolio

The country-wide LGD model can be used to estimate the expected LGD rate of a loan portfolio, if the LTV distribution is known. The estimation about the portfolio, just like the LGD function, is about an expected rate, without taking into account the macroeconomic outlook in the given period. In this sense, it better reflects the quality of the loan portfolio, derived from the LTV distribution, which is the key factor influencing the LGD. Therefore, it can be considered a through-the-cycle estimate, since the cyclical components or expectations of real estate prices do not influence the index, only the loan a characteristic at origination, the LTV. A point-in-time index could be produced by incorporating a forecast about real estate prices. Due to the

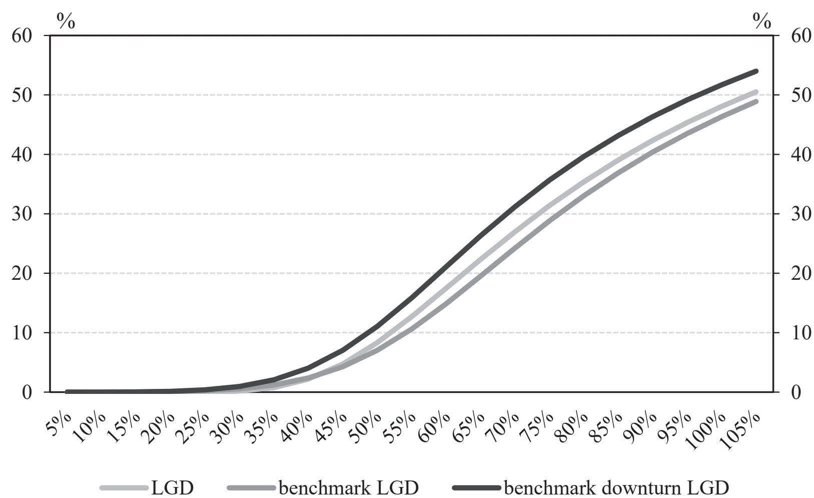

Fig. 6. LGD rates of aggregate country-wide and benchmark models

non-linearity of LGD functions, portfolio-level expected LGD rates are mostly determined by the share of new loans with a high LTV.

The time series describing the expected LGD of the portfolio captures an important component of credit risk and summarises the overall LTV distribution of new Hungarian mortgages in an indicator that is relevant from a risk and business perspective. Therefore, the LGD can be used as a risk index, too. Figure 7 shows the development of the index between 2005 and 2021. It can be seen that the LGD risk profile of new loans reached a high level in the two years prior to 2009, with an expected LGD of 25%. In the crisis period after 2009, probably due to banks' stringent lending conditions, this risk indicator improved considerably. The index continued to decline after 2017, presumably affected by the family home creation allowance and the baby support government programme, which often complemented housing loans from the market. Benchmark models show a similar process. For the loans disbursed before 2021, the model estimates an expected LGD of around 10% based on portfolio quality.

5.6. Limitations of the model

From the theoretical side, the proposed estimated LGD and the PD's correlation is not modelled in detail, which limits the usability of the model. The through-the-cycle nature of the model also implies that in cases where point-in-time estimates are needed, some modifications are necessary to the estimation process. Additionally, some parameters of the model were taken from the relevant literature, due to the non-availability of Hungary-specific information. This also true for cash recovery for defaulted loans that is not considered currently.

Regarding the empirical part of the model, the estimation was based on a real estate price index, which inherently contains some uncertainty about the underlying price development. While the more than 20 year-long-price index time series contains a sufficient number of observations, due to the long real estate cycles, additional data points could increase the model's accuracy.

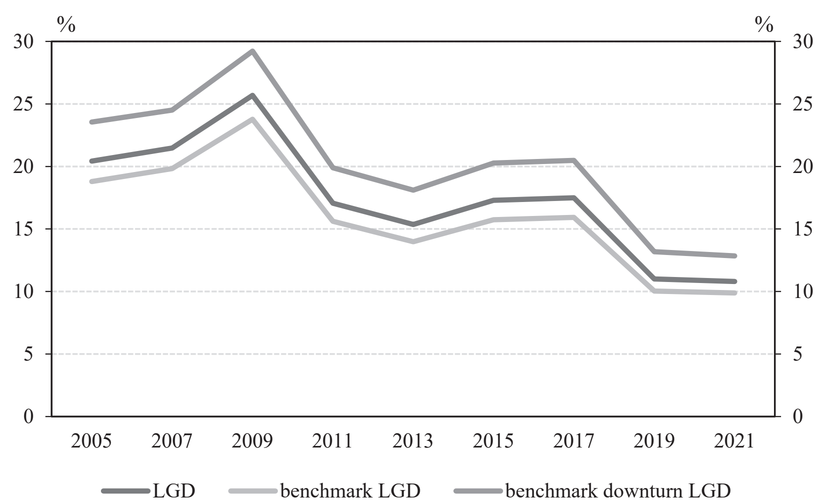

Fig. 7. Expected LGD rates of new mortgage loans based on the LTV distribution, 2005 – 2021
Note: LTV distributions are volume-based and refer to loans disbursed in Hungary in a two-year period.

6. SUMMARY

The paper aimed to calibrate a model based on Frontczak and Rostek (2015), focusing on the expected LGD of residential mortgages with a stochastic collateral value, for the features of the Hungarian residential mortgage market. The specifically Hungarian model was created by fitting a stochastic process, the exponential Ornstein-Uhlenbeck (OU) process on the MNB House Price Index. To lose as little information as possible, this was done at the level of indices broken down by region and settlement type. Moreover, the model considers the additional information about the expected return of the collateral value arising from the default of the loan, as well as the uncertainty of the price change of individual homes. The other parameters related to non-performance were established in line with the international literature. By way of a robustness analysis, the LGD values of the constructed model were compared to the results of a benchmark model.

While performing the stochastic fitting of the Hungarian house price processes, several new results were obtained. Regional house price indices are heterogeneous with respect to the growth trend of the prices and the standard deviation of returns. The growth trend of the prices was high in the capital, Budapest and the cities of the Central Hungary and Central Transdanubia regions, while the growth rate of the villages was low. The villages index was characterised by high volatility. As regards the logarithmised house price indices, in the OU model the value of the kappa parameter measuring mean reversion is almost always close to zero, meaning that property prices can persistently deviate from the trend. In other countries this reversion is usually significant, while in Hungary it was only observed in the villages' index, and only weakly so. In the time series, the deviation from the trend was the greatest in 2021 Q3, which may offer information about the position of the housing market cycle at that time. The standard deviation derived from the OU model fitted on the house price index and the cumulation of idiosyncratic standard deviation far exceeds the volatility of the naïve model fitted only on the house price index, at all points on the time horizon. The results show that the expected return on non-performing collateral value is low due to the strong correlation between the real estate market and defaults. This conditional annual expected return is 1% at most, and the modelling suggests that it can go as low as -2% depending on the region.

The derived LGD models can be used to identify the regions and settlement types that are riskier from the perspective of expected LGD rates. According to the model, the rates for a given level of leverage are the highest in villages and the lowest in Budapest and the cities of Central Hungary. The difference can be as large as 7 percentage points. The difference between the country-wide LGD model prepared by aggregating the models broken down by region and the one calculated based on the country-wide house price index is around 1–1.5 percentage points. The results of the implemented stochastic model show that the LGD function is highly non-linear as a function of the origination LTV. The expected LGD estimation based on the country-wide model can be performed for loan portfolios, too, and as such it can be used as a through-the-cycle LGD risk index. This shows that because of the reduction in lending with high LTV, the LGD profile of new mortgages has improved significantly in Hungary since 2009.

The results could be relevant for credit institutions in their mortgage origination decisions and risk management processes. Banks can use the LTV at origination – expected LGD mapping to assess and measure the expected loss of their existing portfolios. Furthermore, the function can serve as an input to risk-based pricing of mortgages. Since it only depends on the LTV at origination, it is easier to use than models with more granular data requirements. Supervisory authorities may also use the proposed model as a benchmark in the Supervisory Review and Evaluation Process (SREP) for expected and unexpected loss calculations and with some modifications for stress testing purposes.

The results may also confirm banks' practices regarding smaller settlements, namely that they determine more stringent credit insurance values for the properties in such locations. Similarly, the higher expected LGDs of smaller settlements may also support the fact that the conditions of statistical valuation are determined based on settlement size. The LGD risk index may provide input to analysts about lending processes and the associated risks or inform the decisions about the countercyclical capital buffer targeted at mortgage loans. Higher index values indicate more risk-taking from the banks side, which could be countered by increasing the capital buffer. The through-the-cycle nature of the index however limits this application as expectations about future house price developments do not affect on the index.

The results of the housing market processes may offer information to real estate investors. Investment in residential buildings is considered safe among Hungarian retail investors, and studies about the risks of such investments are not common. Future research could investigate the financial risk of Hungarian real estate investments in more detail to provide more information to the public.

ACKNOWLEDGMENT

I would like to Gábor Fukker, Nikolett Tőrös-Barczel , Donát Kim and Csaba Burger for their useful comments.

REFERENCES

Acharya, V. V. – Bharath, S. T. – Srinivasan, A. (2007): Does Industry-Wide Distress Affect Defaulted Firms? Evidence from Creditor Recoveries. *Journal of Financial Economics*, 85(3): 787–821.

Ali, A. – Daly, K. (2010): Macroeconomic Determinants of Credit Risk: Recent Evidence from a Cross-Country Study. *International Review of Financial Analysis,* 19(3): 165–171.

Altman, E. I. – Brady, B. – Resti, A. – Sironi, A. (2005): The Link Between Default and Recovery Rates: Theory, Empirical Evidence and Implications. *The Journal of Business*, 78(6): 2203–2228.

Ambrose, B. W. – Lacour-Little, M. – Sanders, A. B. (2004): The Effect of Conforming Loan Status on Mortgage Yield Spreads: A Loan Level Analysis. *Real Estate Economics*, 32(4): 541–569.

Banai, Á. – Hosszú, Zs. – Körmendi, Gy. – Sóvágó, S. – Szegedi, R. (2013): Stress Testing at the Magyar Nemzeti Bank. *MNB Occasional Papers,* No. 109.

Banai, Á. – Vágó, N. – Winkler, S. (2017): The MNB's House Price Index Methodology. *MNB Occasional Papers,* No. 127.

Bellotti, T. – Crook, J. (2012): Loss Given Default Models Incorporating Macroeconomic Variables for Credit Cards. *International Journal of Forecasting*, 28(1): 171–182.

Berki, T. – Szendrei, T. (2017): The Cyclical Position of Housing Prices – A VECM Approach for Hungary. *MNB Occasional Papers*, No. 126.

Burger, Cs. (2022): Defaulting Alone: The Geography of SME Owner Numbers and Credit Risk in Hungary. *MNB Occasional Papers,* No. 144.

Calabrese, R. (2014): Downturn Loss Given Default: Mixture Distribution Estimation. *European Journal of Operational Research*, 237(1): 271–277.

Calem, P. S. – Lacour-Little, M. (2004): Risk-Based Capital Requirements for Mortgage Loans. *Journal of Banking & Finance*, 28(3): 647–672.

Campbell, J. Y. – Giglio, S. – Pathak, P. (2011): Forced Sales and House Prices. *American Economic Review*, 101(5): 2108–31.

Case, K. E. – Shiller, R. J. (1989): The Efficiency of the Market for Single-Family Homes. *The American Economic Review*, 79(1): 125–137.

Chaiyapo, N. – Phewchean, N. (2017): An Application of Ornstein-Uhlenbeck Process to Commodity Pricing in Thailand. *Advances in Difference Equations*, 2017(1): 179.

Chalupka, R. – Kopecsni, J. (2009): Modelling Bank Loan LGD Of Corporate and SME Segments: A Case Study. *Czech Journal of Economics and Finance*, 59(4): 360–382.

Deng, Y. – Quigley, J. M. – Van Order, R. – Mac, F. (1996): Mortgage Default and Low Downpayment Loans: The Costs of Public Subsidy. *Regional Science and Urban Economics*, 26(3–4): 263–285.

European Banking Authority (2017): Guidelines on PD Estimation, LGD Estimation and Treatment of Defaulted Assets. Https://Www.Eba.Europa.Eu/Regulation-And-Policy/Model-Validation/Guidelines-On-Pd-Lgd-Estimation-And-Treatment-Of-Defaulted-Assets.

European Central Bank (2007): ECB Financial Stability Review December 2007. Https://Www.Ecb.Europa.Eu/Pub/Pdf/Fsr/Financialstabilityreview200712en.Pdf.

European Central Bank (2018): Asset Quality Review Phase 2 Manual. Https://Www.Bankingsupervision.Europa.Eu/Ecb/Pub/Pdf/Ssm.Assetqualityreviewmanual201806.En.Pdf.

Fabozzi, F. J. – Shiller, R. J. – Tunaru, R. S. (2012): A Pricing Framework for Real Estate Derivatives. *European Financial Management*, 18(5): 762–789.

Frontczak, R. – Rostek, S. (2015): Modeling Loss Given Default with Stochastic Collateral. *Economic Modelling*, 44: 162–170.

Frye, J. (2000): Collateral Damage. *Risk*, April: 91–94.

Glaeser, E. L. – Nathanson, C. G. (2017): An Extrapolative Model of House Price Dynamics. *Journal of Financial Economics*, 126(1): 147–170.

Giacoletti, M. (2021): Idiosyncratic Risk in Housing Markets. *The Review of Financial Studies*, 34(8): 3695–3741.

Gupton, G. M. – Stein, R. M. (2002): Losscalc: Model for Predicting Loss Given Default (LGD). *Moody's Investors Service*.

Holló, D. (2009): Risk Developments on the Retail Mortgage Loan Market. *MNB Bulletin*, October 2009.

James, C. (1991): The Losses Realized in Bank Failures. *Journal of Finance*, 46(4): 1223–1242.

Jokivuolle, E. – Peura, S. (2003): Incorporating Collateral Value Uncertainty in Loss Given Default Estimates and Loan-To-Value Ratios. *European Financial Management,* 9(3): 299–314.

Keresztúri, J. L. – Berlinger, E. – Lublóy, Á. (2023): Blowing the Whistle on Corporate Fraud: The Role of Regulators and Journalists in the Financial vs Non-Financial Sectors. *Applied Economics*, 55(11): 1273–1284.

Kim, D. (2021): Debt Cap Regulations in Mortgage Lending. *Economy and Finance: English-Language edition of the journal Gazdaság és Pénzügy*, 8(2): 110–145.

Koopman, S. J. – Kräussl, R. – Lucas, A. – Monteiro, A. B. (2009): Credit Cycles and Macro Fundamentals. *Journal of Empirical Finance*, 16(1): 42–54.

Leow, M. – Mues, C. (2012): Predicting Loss Given Default (LGD) for Residential Mortgage Loans: A Two-Stage Model and Empirical Evidence for UK Bank Data. *International Journal of Forecasting*, 28(1): 183–195.

Marer, P. (2010): The Global Economic Crises: Impacts on Eastern Europe. *Acta Oeconomica*, 60(1): 3–33.

Mérő, K. (2017): The Emergence of Macroprudential Bank Regulation: A Review. *Acta Oeconomica*, 67(3): 289–309.

Miller, N. G. – Pandher, G. (2008): Idiosyncratic Volatility and the Housing Market. *Journal of Housing Research*, 17(1): 13–32.

Palmroos, P. (2016a): Modeling the Current Loan-To-Value Structure of Mortgage Pools without Loan-Specific Data. *Journal of Credit Risk*, 12(4): 1–15.

Palmroos, P. (2016b): Correlation Between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools. *Aalto University Working Paper*, No. 3/2009.

Pelizza, M. – Schenk-Hoppé, K. R. (2020): Pricing Defaulted Italian Mortgages. *Journal of Risk and Financial Management*, 13(2): 31.

Pollák, Z. – Popper, D. (2021): Stress Tests in Hungarian Banking After 2008. *Acta Oeconomica*, 71(3): 451–463.

Qi, M. – Yang, X. (2009): Loss Given Default of High Loan-To-Value Residential Mortgages. *Journal of Banking & Finance*, 33(5): 788–799.

Qi, M. – Zhao, X. (2011): Comparison of Modeling Methods for Loss Given Default. *Journal of Banking & Finance*, 35(11): 2842–2855.

Rünstler, G. – Vlekke, M. (2018): Business, Housing, and Credit Cycles. *Journal of Applied Econometrics*, 33(2): 212–226.

Shleifer, A. – Vishny, R. W. (1992): Liquidation Values and Debt Capacity: A Market Equilibrium Approach. *The Journal of Finance*, 47(4): 1343–1366.

Siemsen, T. – Vilsmeier, J. (2017): A Stress Test Framework for the German Residential Mortgage Market: Methodology and Application. *Bundesbank Discussion Paper*, No. 37/2017.

Szabó, A. V. (2022): Credit Risk Modelling of Mortgage Loans in the Supervisory Stress Test of The Magyar Nemzeti Bank. *Financial and Economic Review*, 21(1): 56–94.

Tajti, Zs. (2011): The Methodological Opportunities of Quantifying the Retail Mortgage Loan's. LGD in Hungary. *Ph.D. Dissertation*.

Van Damme, G. (2011): A Generic Framework for Stochastic Loss-Given-Default. *Journal of Computational and Applied Mathematics*, 235(8): 2523–2550. https://doi.org/10.1016/j.cam.2010.11.006.

Appendix

Results of Monte Carlo simulations

The expected LGD rates calculated with the help of Monte Carlo simulations and the analytical solution (Equation (17)) were compared as a function of the initial collateral value. The parameters used were based on the parameters in the paper:

$$T_D = 1\ year; T_L = 4\ ;\ k = 30\%;\ r = 10\%;\ EAD = 1;\ \mu_Y = 0\%;\ \sigma_Y = 25\%$$

The results shown in the figure were produced with 10,000 simulations.[7]

[7]Figure not published due to lack of space but can be obtained directly from the author.

Standard deviation of cumulative log returns in the individual regions (%)

Table 6. The standard deviation of the cumulative log returns of individual property investments, as a function of the investment horizon

Year	Southern Great Plain		Southern Transdanubia		Northern Great Plain		Northern Hungary		Central Transdanubia		Central Hungary		Western Transdanubia	
	C	V	C	V	C	V	C	V	C	V	C	V	C	V
1	12.1	13.7	12.3	14.1	12.0	13.7	12.1	13.9	12.1	13.6	12.4	14.4	11.8	13.1
2	17.1	19.0	17.5	19.5	16.8	19.1	16.9	19.3	17.2	19.0	17.4	20.0	16.7	18.3
3	20.9	22.9	21.6	23.6	20.4	23.0	20.5	23.3	21.3	22.9	21.1	24.0	20.5	22.1
4	24.2	26.1	25.0	26.8	23.4	26.2	23.5	26.5	24.7	26.1	24.3	27.3	23.8	25.2
5	27.1	28.8	28.1	29.5	26.1	28.9	26.1	29.3	27.9	28.8	27.0	30.1	26.7	27.8
6	29.7	31.2	31.0	31.9	28.4	31.3	28.4	31.6	30.8	31.2	29.4	32.5	29.4	30.2
7	32.1	33.3	33.6	34.1	30.5	33.4	30.4	33.8	33.5	33.3	31.5	34.7	31.9	32.2
8	34.3	35.2	36.1	36.0	32.4	35.3	32.3	35.7	36.1	35.2	33.5	36.6	34.2	34.1
9	36.5	37.0	38.5	37.8	34.2	37.1	34.1	37.5	38.7	37.0	35.4	38.4	36.5	35.9

Note: The first line of the table shows the region, the letter in the second shows whether the column contains the volatility of cities (C) or villages (V), in percentages. The first column of the table shows the duration of the investment, in years.

Expected return and standard deviation until the time of liquidation of the collateral

Table 7. Estimated effect of the cumulative expected return (μ_Y) and standard deviation (σ_Y) on collateral value until the time of liquidation (4 years), in a regional breakdown, %

	μ_Y	σ_Y
National	−0.66%	23.19%
Budapest	3.97%	24.43%
Cities	−1.86%	23.43%
Cities in the Southern Great Plain	−3.38%	24.18%
Cities in Southern Transdanubia	−4.43%	25.01%
Cities in the Northern Great Plain	−2.51%	23.45%
Cities in Northern Hungary	−5.39%	23.51%
Cities in Central Transdanubia	−1.48%	24.73%
Cities in Central Hungary	2.50%	24.26%
Cities in Western Transdanubia	−3.15%	23.80%
Villages	−7.53%	24.88%
Villages in the Southern Great Plain	−7.53%	26.12%
Villages in Southern Transdanubia	−7.53%	26.80%
Villages in the Northern Great Plain	−7.53%	26.20%
Villages in Northern Hungary	−7.53%	26.53%
Villages in the Central Transdanubia	−7.53%	26.10%
Villages in Central Hungary	−7.53%	27.33%
Villages in Western Transdanubia	−7.53%	25.19%

Note: The regional parameters of the villages were created synthetically, as described in the body text.

LGD rates in a detailed regional breakdown

Table 8. The resulting LGD rates as percentages, as a function of the LTV by regions and settlement types

	LTV	20%	30%	40%	50%	60%	70%	80%	90%	100%
	National	0.0	0.1	1.5	6.9	15.8	25.5	34.1	41.3	47.1
	Budapest	0.0	0.1	1.2	5.6	13.5	22.6	31.1	38.4	44.5
	Cities	0.0	0.1	1.7	7.4	16.6	26.3	34.9	41.9	47.7
	Villages	0.0	0.3	3.0	10.3	20.2	29.9	38.2	44.9	50.4
	Aggregate	0.0	0.2	2.2	8.3	17.5	27.0	35.4	42.4	48.1
Cities	Southern Great Plain	0.0	0.1	2.0	8.2	17.6	27.3	35.7	42.7	48.4
	Southern Transdanubia	0.0	0.2	2.4	8.9	18.4	28.0	36.3	43.2	48.8
	Northern Great Plain	0.0	0.1	1.8	7.7	17.0	26.7	35.3	42.3	48.0
	Northern Hungary	0.0	0.1	2.2	8.9	18.7	28.6	37.0	43.9	49.5
	Central Transdanubia	0.0	0.1	1.9	7.6	16.6	26.1	34.5	41.6	47.3
	Central Hungary	0.0	0.1	1.3	6.1	14.2	23.5	32.1	39.3	45.3
	Western Transdanubia	0.0	0.1	1.9	8.0	17.4	27.1	35.6	42.6	48.3
Villages	Southern Great Plain	0.0	0.3	3.3	10.6	20.3	29.9	38.1	44.8	50.3
	Southern Transdanubia	0.0	0.4	3.4	10.8	20.4	29.9	38.0	44.7	50.2
	Northern Great Plain	0.0	0.3	3.3	10.6	20.4	29.9	38.1	44.8	50.2
	Northern Hungary	0.0	0.4	3.4	10.7	20.4	29.9	38.0	44.7	50.2
	Central Transdanubia	0.0	0.3	3.3	10.6	20.3	29.9	38.1	44.8	50.3
	Central Hungary	0.0	0.4	3.6	10.9	20.5	29.9	38.0	44.6	50.1
	Western Transdanubia	0.0	0.3	3.0	10.3	20.2	29.9	38.2	44.9	50.4