STRESS TEST MODELLING OF PD RISK PARAMETER UNDER ADVANCED IRB

Zoltán Pollák
Department of Finance
Corvinus University of Budapest
1093, Budapest, Hungary
E-mail: ZPollak@bankarkepzo.hu

Dávid Popper
International Training Center
for Bankers (ITCB)
1011, Budapest, Hungary
E-mail: DPopper@bankarkepzo.hu

KEYWORDS

Stress test, Default rate modelling, Migration matrix, Probability of Default, IRB.

ABSTRACT

The 2008 crisis highlighted the importance of using stress tests in banking practice. The role of these stress tests is to identify and precisely estimate the effect of possible future changes in economic conditions on the capital adequacy and profitability of banks. This paper seeks to show a possible methodology to calculate the stressed point-in-time PD parameter. The presented approach contains a linear autoregressive distributed lag model to determine the connection between the logit of default rates and the relevant macroeconomic factors, and uses migration matrices to calculate PDs from the forecasted default rates. The authors illustrate the applications of this methodology using real credit portfolio data.

INTRODUCTION

The European Banking Authority (EBA) requires banks under Advanced Internal Rating-Based (AIRB) approach to prudently measure their risk profile and apply more accurate risk management tools.
According to Basel Capital Requirements Regulation (CRR), institutions "shall regularly perform a credit risk stress test to assess the effect of certain specific conditions on its total capital requirements for credit risk" (575/2013/EU, Article 177(2)).
Banks should have a comprehensive stress testing framework, which is based on the forecasting of the capital adequacy, balance sheet and the P&L statement along different stress scenarios. As a part of this framework, credit institutions should estimate stressed PD parameters that serve as important input factors for the calculation of the bank's performance under the defined scenarios.
One of the ways to forecast PD parameters are econometric models with default rate time series as dependent variable. In this methodology, the estimated stressed default rates are transformed into stressed segment PDs using migration matrices.
The main object of this paper is to reveal connections between various macroeconomic variables and the default rate using econometric methodology on real data, and estimate the stressed PD parameters based on the results. First, we present the used database and the chosen methodology, then we show the results of the default rate model. In the last part, we will transform the stressed default rates into stressed PDs.

INPUT DATA

Default rates

In our presented case, stress test modelling of default rates was carried out based on quarterly default rate time series. The portfolio segment used for modelling is the micro segment (micro-enterprises of a commercial bank's credit portfolio).
For the stress test modelling, yearly default rates were calculated for each quarter during the observation period (from Q1 2007 to Q1 2016).

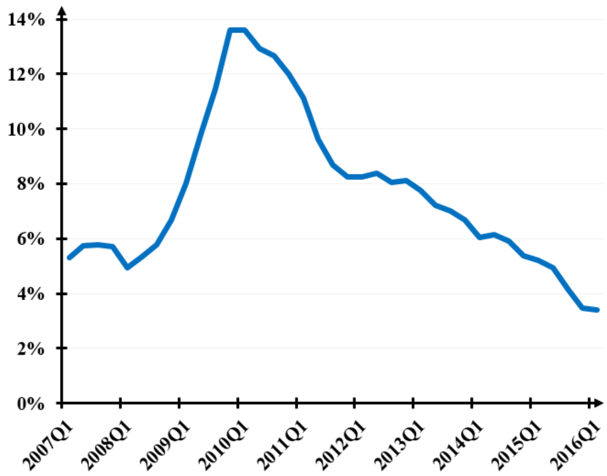

Figure 1: Default rate time series

Figure 1 illustrates the default rate time series for the mentioned micro segment.
To forecast default rates along the stress scenarios, we used different macroeconomic variables. Most of the selected macro variables are available on a quarterly basis, therefore we chose the quarterly frequency for the calculations.
Instead of modelling the default rates directly, we used the logit of them in the equations to avoid estimations outside the [0,1] interval.
Following the Box-Jenkins time series analysis philosophy (Box and Jenkins 1976), as a first step, we checked the stationarity of the dependent variable.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

The results of the unit root test (Augmented Dickey-Fuller and Phillips-Perron) for logits of the default rates are as follows:

Table 1: Unit root tests of the default rate's logit

Level	ADF tau	ADF p-value	PP tau	PP p-value
Not differenced	-0.0794	0.6505	-0.1324	0.6342
First difference	-2.3364	0.0205	-2.4976	0.0133

As the second row of Table 1 shows, p-values for the logits of the default rates (not differenced) are high, so we cannot reject the null hypothesis of the unit root both in the case of Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. Therefore the logits are not stationary at the usually used significance levels (1%, 5% or 10%).

According to the final row of Table 1, the first differences are already stationary on 5% significance level, so variables are first order integrated according to the tests. As a result, we used the first differences of the logits for modelling.

Explanatory macro variables

We chose the following key macroeconomic variables and used during our stress test:

Table 2: Macroeconomic variables used for modelling (the large list)

Variable name	Description	Source
Real GDP (SA) volume growth	GDP volume index, basis period = same period in previous year, seasonally and calendar adjusted, expressed in percent	KSH
Real GDP (SA) gap (percent)	Computed using HP filter on seasonally adjusted real GDP (cyclical component, percentage deviation from „equilibrium" path)	own calculation
Consumer Price Index	Consumer Price Index (3 month average as quarterly data)	MNB
Employment rate (SA)	Employment rate, seasonally adjusted	KSH
Investments (national)	Investments, YoY change	KSH
Investments (companies)	Investments, YoY change	KSH
Industrial Production Index	Industrial Production index, same period in previous year = 100	KSH
BUX index	BUX index	BÉT
DAX index	DAX index	YahooFinance
HUF base rate	HUF base rate	MNB
BUBOR 3M	BUBOR, 3 month	MNB
CHF base rate	CHF base rate, call money rate	SNB
EUR base rate (lending)	EUR base rate, Marginal lending facility	ECB
EUR/HUF exchange rate	EUR/HUF exchange rate	MNB
CHF/HUF exchange rate	CHF/HUF exchange rate	MNB
Housing price index (FHB)	Housing price index (FHB), 2000 =100 (nominal)	FHB
Housing price index (Eurostat)	Housing price index (YoY change)	Eurostat
Retail deposits	Retail deposits (index, December 2001 = 100)	MNB
Leverage ratio	(Domestic corporate loans / Nominal GDP)·100	own calculation, MNB

For time series analysis the explanatory variables are required to be stationary (Hamilton 1999). If stationarity is not satisfied, then the estimated models are not considered reliable, and the risk of spurious regression incurs. This is the reason for why the stationarity of the explanatory variables is inspected. Augmented Dickey-Fuller and Phillips-Perron unit root tests were carried out first for the level of the variables. Regarding the level, only the GDP growth rate and the GDP gap proved to be stationary. For the other explanatory variables, the differences were also examined.

We examined the results for both the first ("d") and the yearly ("d4") difference of the variables. The choice between "d" and "d4" depends on their explanatory power on the default rates. (A preselection of the explanatory variables and their differenced forms is based on the analysis of the cross-correlation matrices.)

According to the unit root tests, all differenced variables can be regarded as stationary.

METHODOLOGY

As we previously mentioned, instead of modelling the default rates directly, we used the logit of them by the estimations. The logit of the default rate is calculated by the following transformation:

$$logit\ of\ the\ default\ rate = log\left(\frac{defrate}{1-defrate}\right) \quad (1)$$

The connection between the explanatory variables and the dependent variable is estimated using linear autoregressive distributed lag model (using Maximum Likelihood estimation method).

The dependent variable is the first difference of the logit of the default rate. The explanatory variables are the previously described (already stationary transformed) macroeconomic variables and their lags of 1-4 quarter.

Furthermore, the autoregressive component of the dependent variable is also included in the regression (only the first lag proved to be correlated with the current value).

The applied autoregressive linear regression can be described by the following formula:

$$Y_t = \alpha_t + \beta_0 \cdot Y_{t-1} + \sum_{i=1}^{k} \boldsymbol{\beta_i} \cdot \boldsymbol{X}_{t-i} + \varepsilon_t \qquad (2)$$

Where

Y_t	differenced logit of the default rate of given segment in time period t
Y_{t-1}	autoregressive component
$\boldsymbol{X_{t-i}}$	macro variables and their lagged values (where k = 4 quarters)
ε_t	random error term
α_t and $\boldsymbol{\beta_i}$	coefficients estimated by the linear regression

In case of the dependent variable, the autocorrelation function can provide information about the numbers of lags to be included in the model. Generally, for the explanatory variables and their lags, the statistical significances of the estimated coefficients determine their inclusion or exclusion in the final model equation.

We made the estimation of the regressions with the help of SAS 9.4 software. The stepwise procedure is an iteration process where we start by including a wider range of explanatory variables into the model. During every step, the explanatory variable with the highest p-value is omitted if this p-value is above the pre-given threshold. If there are no more explanatory variables with p-values higher than the threshold, then the iteration is finished. The final model at the end of the procedure is used to carry out the estimation for the period 2016Q2-2019Q4.

The initial (wide) circle of the explanatory variables was determined with the help of the correlation matrices. The correlation matrices help to choose which differences (either "d1" or "d4") and lags of the possible explanatory variables should be included in the initial regression.

If we set the p-threshold at a relatively low level, then the estimation of the coefficients will be highly reliable as they are significantly different from zero even at a strict confidence level. However, the number of the explanatory variables will be strongly limited. The larger the threshold is, the more macro variables the estimation is based on. So we have more possibility to capture the link between the macro environment and the default rates. The coefficient estimations are however less reliable as less significant variables are also kept in the model.

RESULTS

We started the procedure with a wider range of explanatory variables; then the statistically insignificant variables were omitted during the steps. The variables in the initial large model were chosen based on the cross-correlation matrices. The following variables were included („D" indicates the order of difference, „L" indicates which lag is used):

Table 3: The initial model

Dependent variable	Logit_default rate (D1, L0)
Explanatory variables	Logit_default rate (D1, L1)
	Real GDP (SA) growth (D0, L0)
	Leverage ratio (D4, L1)
	Employment rate (D, L0)
	BUBOR (D4, L1)

After running the stepwise procedure, we got a final model (*** denotes that the variable is significantly different from 0 at 1% significance level; ** indicates 5% and * indicates 10% significance level):

Table 4: The final model

Explanatory variables	**Coefficient**	***t* value**	***p* value**
Intercept	0.0137	1.13	0.3270
Logit_default rate (D1, L1)	0.2543	1.82	0.1657
Real GDP (SA) growth (D0, L0)	-0.0151	-3.01	0.0514^{*}
Leverage ratio (D4, L1)	3.4005	2.48	0.0157^{**}

R^2	**Adjusted R^2**	**Number of obs.**	**Applied *p* threshold**
0.6884	0.6689	36	0.2

As presented in Table 4, the p threshold of 0.2 was applied, and two of the initial explanatory variables were omitted. Real GDP growth is significant at 10%, while Leverage ratio is significant even at 5% significance level.

Residual correlation diagnostics

In this part, we examine whether the final model is correctly specified by checking the residuals of the estimation. In the case of a correct specification, no autocorrelation remains in the residuals. The autocorrelation can be checked with the graphs of ACF and PACF functions, furthermore by the Durbin-Watson test.

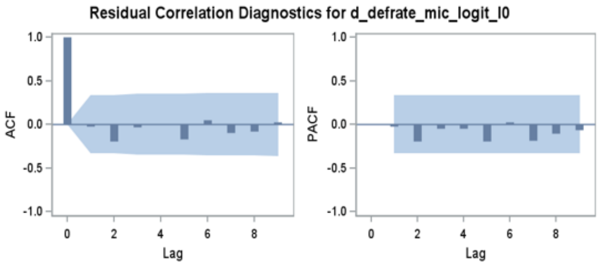

Figure 2: ACF and PACF functions of the residuals

The ACF and PACF functions show that the residual is not strongly correlated with its own lagged values. The Durbin-Watson test with its value close to 2 confirms the conclusion that the residual contains no autocorrelation (DW = 2.1584).

The forecasted default rates

Figure 3 plots the forecast results for the four estimated macroeconomic scenarios (baseline, adverse, severely adverse and the crisis; Crisis(7) in Figure 3 means that the crisis scenario was calculated based on the worst 7 quarters of the 2008 financial crisis period):

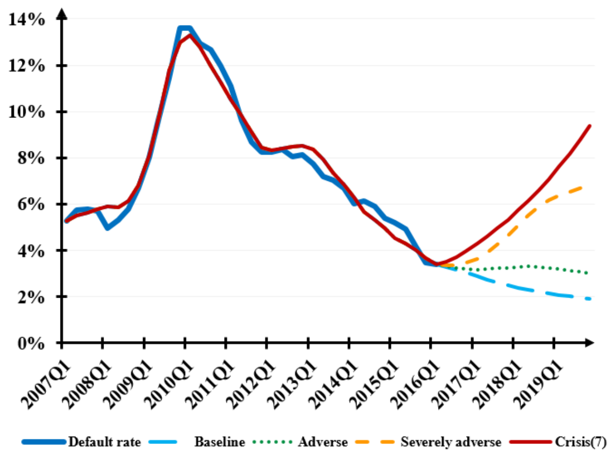

Figure 3: Default rate forecast along the scenarios

Table 5 summarizes the results numerically. Beyond the average yearly default rates (calculated as a simple average of the 4 quarters) it also presents the default multipliers. The multiplier is a ratio comparing the default rate of the adverse, severely adverse and crisis scenario to the respective default rate of the baseline scenario. The multiplier is expressed in percent. It tells how severe the certain scenario is, compared to the baseline forecast.

Table 5: Summary table of the model forecasts

	2017	**2018**	**2019**
Baseline default rate	2.50%	2.14%	1.90%
Adverse default rate	3.26%	3.26%	3.02%
Adverse multiplier	**130.56%**	**152.62%**	**159.15%**
Severely adverse default rate	4.65%	6.18%	6.86%
Severely adverse multiplier	**186.06%**	**289.13%**	**361.04%**
Crisis default rate	5.33%	7.09%	9.37%
Crisis multiplier	**213.49%**	**331.97%**	**493.38%**

TRANSLATING DEFAULT RATES TO PD

For further estimations within the stress testing framework, we need segment-level stressed PD values. That means that we have to relate the forecasted default rates with the probability of default. This is done using migration matrices that tell us how clients' ratings change over time.
We took into account the recommendation of the European Banking Authority (EBA) about stress testing framework and migration matrices (EBA 2016). According to this suggestion, institutions need to calculate point-in-time transition matrices, and these matrices should meet at least the following two criteria:

- The PD for each grade should calculate in line with the scenarios, and
- The probabilities of moving between grades are adjusted according to the scenarios.

Considering the above, we created the one-year observed migration matrix. The matrix show how individual ratings changed between 2015 and 2016. As a next step, we stress these matrices using our estimated default rate multipliers. The methodology can be described as followed:

1. We define a common stress parameter φ that serves as a factor in which we shift the distribution of the ratings in the matrix (φ is different for every scenario).
2. We calculate the yearly default rates for the years 2016-2019 based on the observed migration matrix. These will be the "baseline" default rates.
3. We calibrate φ in a way that the ratio of the stressed and the baseline default rate at the end of the forecasting horizon (2019) would be the same as the corresponding default rate multiplier estimated by the linear regressions.
4. With these stressed migration matrices we calculate the total exposure in every rating category for all four years.
5. Using the fixed PD of the rating categories and the total stressed exposures we can calculate the segment-level average PD for all four years.

The observed one-year migration matrix ("base" migration matrix) is the following:

Table 6: Base migration matrix, number of clients

	C1	C2	C3	C4	C5	C6	C7	C8	D	Total
C1	24	6	1							**31**
C2	312	641	79	2	1	5	5	6	8	**1 059**
C3	89	1 121	601	49	30	33	9	17	39	**1 988**
C4	2	121	229	37	20	10	5	11	23	**458**
C5	1	53	143	39	230	15	3	14	11	**509**
C6		18	63	26	13	10	5	5	20	**160**
C7	2	10	38	14	23	30	8	6	19	**150**
C8		15	22	18	42	28	47	63	54	**289**
D					1			1	531	**533**
Total	**430**	**1 985**	**1 176**	**185**	**360**	**131**	**82**	**123**	**705**	**5 177**

Expressed in percent:

Table 7: Base migration matrix, percentage of clients

	C1	C2	C3	C4	C5	C6	C7	C8	D
C1	77.42%	19.35%	3.23%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C2	29.46%	60.53%	7.46%	0.19%	0.09%	0.47%	0.47%	0.57%	0.76%
C3	4.48%	56.39%	30.23%	2.46%	1.51%	1.66%	0.45%	0.86%	1.96%
C4	0.44%	26.42%	50.00%	8.08%	4.37%	2.18%	1.09%	2.40%	5.02%
C5	0.20%	10.41%	28.09%	7.66%	45.19%	2.95%	0.59%	2.75%	2.16%
C6	0.00%	11.25%	39.38%	16.25%	8.13%	6.25%	3.13%	3.13%	12.50%
C7	1.33%	6.67%	25.33%	9.33%	15.33%	20.00%	5.33%	4.00%	12.67%
C8	0.00%	5.19%	7.61%	6.23%	14.53%	9.69%	16.26%	21.80%	18.69%
D	0.00%	0.00%	0.00%	0.00%	0.19%	0.00%	0.00%	0.19%	99.62%

Using the migration matrix and the sum of the clients in every rating category we can easily calculate the default rates for the forecasted years, that is the percentage of non-defaulted (C1-C8) clients that migrate to D (defaulted) rating category.

Then, we stress the matrix by the factor φ. "Stressing" means that φ% of the clients in a cell is shifted to the next cell to the right, that is they migrate to a worse category. For example, if φ = *10%* then the C1-C1 cell will be *90%·77,42%* = *69,68%*, while C1-C2 cell will be *90%·19,35%* + *10%·77,42%* = *25,16%*. With this methodology, we can stress the whole segment depending only on one factor that can easily be calibrated to be in line with our regression estimates.

As already mentioned, the stress factor φ is calibrated in a way to get the same ratio of the stressed and baseline default rates as estimated with the regressions (marked with blue in Table 5).

Following the same steps as before we calculated the average segment PD. The results are summarized in Table 8:

Table 8: Stressed PD results for micro segment

Scenario		2017	2018	2019	Stress factor (φ)
Baseline	Default rate based on observed migration matrix	2.35%	1.63%	1.24%	
	Average segment PD	**4.12%**	**3.17%**	**2.63%**	
Adverse	Default rate based on observed migration matrix	3.05%	2.35%	1.97%	**20.39%**
	Default rate multiplier	130.01%	144.42%	**159.14%**	
	Average segment PD	**5.13%**	**4.35%**	**3.91%**	
Severely adverse	Default rate based on observed migration matrix	4.99%	4.58%	4.47%	**71.41%**
	Default rate multiplier	212.74%	281.93%	**361.04%**	
	Average segment PD	**8.02%**	**7.90%**	**7.86%**	
Crisis	Default rate based on observed migration matrix	6.03%	5.93%	6.11%	**96.16%**
	Default rate multiplier	256.91%	364.37%	**493.38%**	
	Average segment PD	**9.62%**	**9.95%**	**10.18%**	

As a summary, Figure 4 presents the stressed segment PDs along the 4 scenarios:

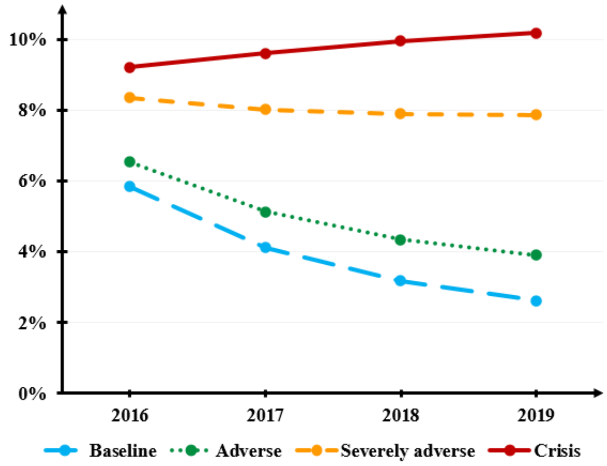

Figure 4: Stressed segment PDs

And finally, with the help of these PDs, we can estimate the effect of the different stress scenarios to the P&L and the capital adequacy of the given financial institution.

CONCLUSION

In this paper, we presented a possible methodology to calculate the stressed point-in-time PD parameter. This estimation is crucial for a sound stress testing, but the final methodology that a financial institution chooses has to be in line with the scope and complexity of the bank.

The regression method we applied requires long enough time series which is not always available. Therefore alternative methods and expert-based considerations should also be taken into account. And even if econometric modelling is possible, the results should be evaluated and sometimes modified by experts who have deep knowledge of the bank's portfolio and risk profile.

REFERENCES

Box G.E.P. and G.M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control.* Holden-Day. San Francisco.

Budapest Stock Exchange Statistics https://bet.hu/ (downloaded on the 27th of October 2016)

European Banking Authority (EBA). 2016 EU-Wide Stress Test Methodological Note. 24 February 2016.

European Central Bank (ECB) Statistics https://www.ecb.europa.eu (downloaded on the 27th of October 2016)

Eurostat Statistics http://ec.europa.eu/eurostat (downloaded on the 27th of October 2016)

Hamilton J.D. 1994. *Time Series Analysis.* Princeton University Press. New Jersey.

Hungarian Central Bank Statistics https://www.mnb.hu/ (downloaded on the 27th of October 2016)

Hungarian Central Statistical Office Statistics http://www.ksh.hu/ (downloaded on the 27th of October 2016)

Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms

Swiss National Bank Statistics https://www.snb.ch/en/ (downloaded on the 27th of October 2016)

YahooFinance Statistics https://finance.yahoo.com/ (downloaded on the 27th of October 2016)

AUTHOR BIOGRAPHIES

ZOLTÁN POLLÁK completed his MSc degree summa cum laude in Finance at Corvinus University of Budapest. He is currently doing a Ph.D. at the Department of Finances. He is lecturing financial courses such as Corporate Finance and Financial Calculations. Beside Ph.D. he works as a partner consultant for the International Training Center for Bankers (ITCB), where he also teaches on banking and investment courses.
His e-mail address is: `ZPollak@bankarkepzo.hu`

DÁVID POPPER graduated from Economics MA program of the Central European University. His main fields of interest are financial economics and economic growth (especially different aspects of economic convergence). He currently works as a junior consultant at ITCB where he is responsible for developing various credit risk models including stress testing methodologies.
His e-mail address is: `DPopper@bankarkepzo.hu`