

THE CREDIT RISK MODEL VALIDATION PROCESS UNDER MULTIREGIONAL PORTFOLIO STRUCTURE

Paweł Siarka¹

¹ ORCID 0000-0001-7488-0782

Wrocław University of Economic and Business
Faculty of Management, Computer Science and Finance
Komandorska 118/120, 53-345 Wrocław, Poland
E-mail: pawel.siarka@ue.wroc.pl

Abstract: *A comprehensive credit risk management process should cover a back testing procedure regarding risk projections. Model validation is especially important for global banks operating in multinational environment and having multiregional portfolio structure. Territory segmentation implies various analytical solutions for various regions. However comprehensive models quality assessment process should be performed with respect all identified regions.*

In the paper the author presents an approach that can be leveraged under the back-testing analysis. TLA (Traffic Light Approach) method has been presented and discussed. The author defined three disjointed intervals for selected statistic which allows to assess the quality of risk projections. Furthermore, the author referred to vintage analysis methodology and its usefulness under probability of default estimation process.

The paper presents an example that illustrates the use of the TLA method and vintage analysis. These methods were leveraged on the ground of internal credit scoring system and used to assess the quality of probability of default estimates.

Key words: *credit risk, probability of default, model validation*

JEL codes: *C51, C53, E58, G2, G33*

1. Introduction

In 1996 the Basel Committee for the first time referred to the problem of quantitative assessment of the risk management models (Basel 1996). The recommendation issued at that time concerned market risk and was included in the document supplementing the first capital accord established in 1988 (Basel 1988). On this basis, the Traffic Light Approach (TLA) was

developed and implemented in many financial institutions. This method allows to verify the quality of risk projections by setting three disjointed intervals for risk measures, where each interval is marked with various colour i.e. green, yellow, red. When the actual risk value is in the green area, then the prediction is deemed as reliable. When the risk value is in the yellow area, it suggests that the forecasting model underestimates the actual risk, however the model is still acceptable. It is then necessary to identify and verify the reasons of forecast inaccuracy. The red area means that the model accuracy is not acceptable, therefore the model should be thoroughly verified and withdrawn from risk system immediately.

The model validation process cannot be conducted in isolation from regions that are geographically, culturally and economically coherent. The specificity of the regions arise from the fact that factors affecting borrower's creditworthiness vary within regions. This applies to both macroeconomic and microeconomic factors. In addition, the strength of their impact on the client's ability to repay the loans can also vary significantly. Hence, the model validation procedure should take into account regional diversification and other clusters identified within segmentation analysis.

The process of credit risk model validation focuses mostly on probability of default (PD) estimates which are being compared with actual values. Such an analysis is called backtesting. The Basel Committee on Banking Supervision emphasizes its importance within the credit risk management process. The PD value is one of the key parameter of the minimum capital requirements calculation process for credit risk. For the retail banking the definition of default event recommended by the Basel Committee (Basel 2006) is based on days past due indicator (DPD), i.e. default is defined as DPD over 90 days. Forecast horizon of PD covers one year under New Basel Capital Accord.

Under the backtesting process one of the most powerful tool is a vintage analysis which comes to calculation of the fraction of defaulted loans within the total amount of originated loans in specific period of time. So, unlike the PD estimates it has the same denominator value over the whole life of the portfolio. Based on the above data, a comparative analysis i.e. forecasts versus actual risk can be performed. Therefore vintage analysis fits the comprehensive credit risk management systems.

The TLA approach has also been applied for credit risk analysis. It appeared to be useful for PD models assessment. Some rules regarding constructing of 'decision areas' for PD were presented by Blochwitz (2004). He leveraged normal distribution properties to get disjointed intervals. The Basel Committee suggested in its studies, that the minimum number of 'decision

areas' (marked with various colours) should not be less than three. Moreover the idea of using a third colour which separates 'decision areas' such as 'model is perfect' and 'model is bad', was also widely accepted by many professionals. It results from the complexity of the risk nature due to multiple factors affecting its value. Substantial simplification by using only two 'decision areas' may lead to fateful decision. Under the backtesting procedure there are two types of decision errors. First consists in consideration the good forecasting model as bad and the second refers to situation where bad model is considered as good. Both errors are undesirable, however the second one seems to be much more dangerous for financial institutions. Hence, the implementation of another 'decision area' marked in yellow gives more flexibility under the model validation process. Further researchers developed the TLA approach by implementing wider colour palette. Castermans (2007) in his paper presented an approach based on five 'decision areas'.

The European Banking Authority (EBA) has published draft mapping report (European Banking Authority 2015) for credit rating agencies that are recognised by the European Securities and Markets Authority. The mapping allows to identify the borrowers with the same credit quality and hence to perform backtesting procedure for homogeneous groups. It is also worth to notice that credit rating agencies use published IRBs (Internal Ratings Based Systems) for the assessment of asset-backed securities (Moody's 2015). Finally, passing backtesting implies that a bank's internal models are able to conservatively capture the future potential risk with a certain reliability.

The implementation of TLA approach under the credit risk management is challenging due scarce data regarding defaults. A possible solution is to extend the forecast horizon from 1 year to e.g. 5 years. Such an approach is used within the vintage analysis. The advantages of this approach are discussed by Burns and Stanley (2001). Some benefits of vintage analysis were also highlighted by Breden (2004) who pointed out that it could be leveraged for stress testing analysis. Anderson (2007) showed that credit risk forecast accuracy analysis can be performed on vintage outcomes. He examined the differences between forecasted defaults and the actually observed bankruptcies.

A comprehensive review of multiple backtesting methods was provided by Zhang and Nadarajah (2018). The authors presented application of statistical testing theory for VaR models backtesting. Also Pelletier and Wei (2015) proposed the geometric VaR test comprising of three individual hypotheses. Since VaR approach has been successfully applied in the banking

industry, Nadarajah and Chan (2016) presented comprehensive review of the mathematical properties and estimation techniques.

The purpose of this paper is to address the problem of PD models accuracy assessment in the context of the TLA approach. This approach was originally used for the market risk models assessment, however the author wants to examine whether it can be leveraged for credit risk. Therefore, an example of implementation of TLA method will be presented on the ground of recommended by Basel Committee definition of default. Additionally, the concept of vintage analysis will be leveraged to assess the credit risk model. Based on the results, the hypothesis regarding high quality of examined model will be verified.

The paper consists of an introduction followed by the concept of vintage analysis and the TLA approach. Next, the author examines the sensitivity of estimated 'decision areas' to changes in the assumed probability of exceeding the forecasted PD by the actual value. Then, an example of TLA approach is presented. The paper ends with conclusions and discussion regarding conducted research.

2. Accuracy assessment of PD forecasts

To assess the accuracy of the PD forecast, the BDR (Backtest Default Rate) statistic can be used. This exercise is performed ex-post based on historical data observed for a specific loan portfolio. The BDR is derived according to the following formula (Svec 2009):

$$BDR = \frac{DF}{AF}$$

where DF is the number of loans that defaulted within the specified 12-months period. AF is the number of non-default loans at the beginning of considered period. Therefore, BDR is a measure which is compared with previously estimated forecast of PD. Both values DF and AF are usually collected from vintage analysis performed by analysts.

The term *vintage* was taken by bankers directly from the wine world. Wine connoisseurs for long time have been assessing wines focusing on the grapes strain as well as the year they were harvested. This analogy was leveraged by bankers for credit risk assessment. It is widely known that one of the key factor determining wine's quality is the year when grapes ripened. This results from the fact that poor sunlight during the growing season causes the low sugar content, which impacts final wine flavour. This is why, wine experts create vintage tables where each wine gets a mark depending on the year of production. So, these marks shows whether a given wine should be preserved in order to wait for the optimal taste, or should be consumed.

Figure 1 presents an example of a vintage table prepared for French wines. Individual wines are evaluated based on the scale ranging from 1 to 100, where more points represent better taste. Furthermore, various colours in the table allow to plan the optimal strategy e.g. wait with the wine, drink it now, can drink but it is better to wait, etc.

Fig. 1 French wines vintage table

| EUROPE | | 2004 | 2003 | 2002 | 2001 | 2000 | 1999 | 1998 | 1997 | 1996 | 1995 | 1994 | 1993 | 1992 | 1991 | 1990 | |
|-----------------------------|--------------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|--|
| FRANCE | Bordeaux | | | | | | | | | | | | | | | | |
| | <i>Pomerol/St. Emilion</i> | 91 | 86 | 87 | 89 | 97 | 88 | 92 | 85 | 88 | 90 | 89 | 86 | 80 | 81 | 98 | |
| | <i>Médoc</i> | 90 | 90 | 86 | 89 | 98 | 89 | 89 | 84 | 93 | 91 | 88 | 85 | NR | 82 | 98 | |
| | <i>Graves (red)</i> | 89 | 89 | 86 | 88 | 96 | 89 | 89 | 84 | 91 | 91 | 88 | 85 | NR | 81 | 97 | |
| | <i>Graves (white)</i> | 91 | 86 | 89 | 89 | 94 | 90 | 89 | 89 | 90 | 89 | 84 | 90 | 87 | 84 | 92 | |
| | <i>Sauternes/Barsac</i> | 87 | 90 | 91 | 94 | 81 | 90 | 88 | 90 | 88 | 88 | 80 | NR | NR | NR | 96 | |
| | Burgundy | | | | | | | | | | | | | | | | |
| | <i>Côtes de Nuits (red)</i> | 93 | 89 | 92 | 88 | 84 | 89 | 88 | 89 | 92 | 90 | 85 | 88 | 80 | 86 | 91 | |
| | <i>Côtes de Beaune (red)</i> | 92 | 89 | 91 | 87 | 84 | 88 | 87 | 90 | 91 | 91 | 83 | 87 | NR | 84 | 92 | |
| | <i>Chablis</i> | 92 | 87 | 95 | 87 | 89 | 87 | 89 | 89 | 93 | 90 | 88 | 84 | 91 | NR | 90 | |
| | <i>Côtes de Beaune (white)</i> | 93 | 86 | 93 | 90 | 88 | 88 | 90 | 89 | 92 | 91 | 87 | 83 | 91 | NR | 92 | |
| | <i>Maconnais</i> | 89 | 84 | 90 | 90 | 88 | 87 | 89 | 90 | 92 | 91 | 87 | 84 | 90 | NR | 91 | |
| | Beaujolais | 91 | 90 | 90 | 87 | 84 | 88 | 88 | 90 | 90 | 90 | 84 | 87 | 80 | 84 | 90 | |
| | Northern Rhône | | | | | | | | | | | | | | | | |
| | <i>Reds</i> | 86 | 92 | 86 | 91 | 90 | 93 | 92 | 91 | 87 | 91 | 89 | NR | NR | 93 | 93 | |
| | <i>Whites</i> | 90 | 86 | 85 | 90 | 87 | 92 | 92 | 90 | 89 | 92 | 87 | NR | 88 | 90 | 90 | |
| | Southern Rhône | | | | | | | | | | | | | | | | |
| | <i>Reds</i> | 89 | 89 | 84 | 90 | 93 | 89 | 97 | 85 | 85 | 93 | 88 | 87 | NR | NR | 96 | |
| | <i>Whites</i> | 87 | 87 | 85 | 87 | 87 | 90 | 93 | 87 | 86 | 91 | 86 | 89 | 87 | NR | 91 | |
| | Loire | | | | | | | | | | | | | | | | |
| <i>Dry Whites</i> | 91 | 89 | 90 | 88 | 84 | 89 | 86 | 85 | 91 | 90 | 85 | 90 | 88 | 87 | 90 | | |
| <i>Sweet Whites</i> | 84 | 90 | 87 | 88 | 83 | 89 | 87 | 91 | 92 | 91 | 82 | 89 | NR | NR | 91 | | |
| <i>Reds</i> | 83 | 91 | 88 | 89 | 93 | 90 | 87 | 87 | 91 | 90 | 87 | 86 | 84 | 80 | 84 | 92 | |
| Alsace | 93 | 89 | 88 | 94 | 86 | 84 | 89 | 96 | 90 | 88 | 86 | 84 | 80 | NR | 93 | | |
| Champagne | 94 | 86 | 93 | NV | 85 | 88 | 87 | 85 | 92 | 89 | NV | 87 | NV | NV | 98 | | |
| Languedoc-Roussillon | 88 | 90 | 84 | 90 | 90 | 89 | 95 | 85 | 84 | 90 | 86 | 86 | 81 | 80 | 90 | | |
| Provence | 87 | 90 | 85 | 90 | 91 | 88 | 94 | 86 | 85 | 91 | 85 | 85 | 80 | 81 | 91 | | |

Source: <http://www.winemag.com>

The analogy noticed by bankers between wine maturation and credit risk profile observed within a given period of time resulted in incorporation of vintage approach in banking industry. Currently, vintage analysis is a popular tool for credit risk management. Its simplicity and functionality allows ongoing monitoring of credit risk. It is also a source of data for other analyses including backtesting.

3. Vintage analysis as an input of PD backtesting process

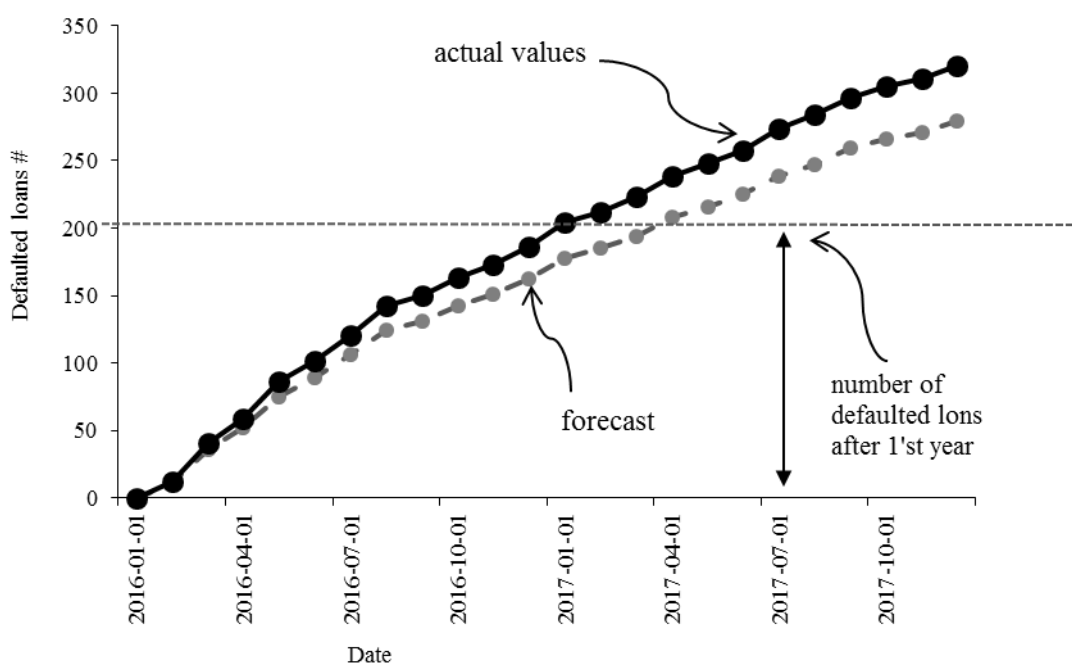
The idea of vintage approach comes to credit risk profile analysis, where the portfolio is presented in two dimensional space of characteristics. First characteristic, analogously to the wine methodology, is the moment of loans origination. It is usually represented by the month or quarter of origination. The second characteristic represents DPD (Days Past Due) measure, which is grouped in thirty days intervals. For each individual pair of DPD and origination period there risk measure is calculated. These measures can be derived in two alternative ways. First

focuses on the amount of loans in specified bucket in relation to original number of originated loans, while the seconds refers to its actual outstanding value and its original value.

Vintage method allows credit risk profile analysis over the time. The trend can be identified and used for further projections. So, the analysis provides historical data which can be leveraged for time series models. Furthermore, it can be used as a comparative analysis for two various portfolios. Hence, the impact of credit risk policy change on the final risk level can be assessed by risk profile comparison.

Figure 2 presents vintage analysis outcomes for +90DPD loans originated in January 2016. The data were collected from a consumer finance bank operating in Poland where loans were collateralized with cars. The grey line represents risk forecasts and the black line shows actual values. The differences between these lines can be used for forecast accuracy assessment.

Fig. 2 vintage analysis outcomes



Source: own study

The figure above shows cumulated number of defaulted loans over the observed period of time. It can be noticed that actual values are above forecasted and therefore the lines diverge. It clearly shows that statistical model used for forecasting underestimates the actual credit risk.

Thus, vintage methodology allows to collect data for further credit risk analysis. The outcomes can be used for comparison of risk profiles of various portfolios. Furthermore, any changes in credit risk policy or macroeconomic variables can be analysed in terms of its impact

on bank's losses. An unexpected spike observed on the chart can be a result of credit policy liberalisation or the beginning of economic downturn.

4. Missed forecasts – the number of breaches

Vintage analysis allows to identify the trend of the number of defaulted loans in the portfolio since the origination moment. These values are used for BDRs estimation process according to the equation presented above. When the BDR value is higher than the predicted PD, then an event of forecast breach occurs. It means that PD projection underestimated its actual value. Such an event can be observed even for robust projection models, however should appear relatively rarely. This is why the model validators assume the probability of risk underestimation q at a conservative level e.g. 1%. Hence, when the events are independent, the probability of k breaches with N tries can be calculated using following formula:

$$p(k) = \frac{N!}{k!(N-k)!} q^k (1-q)^{(N-k)}$$

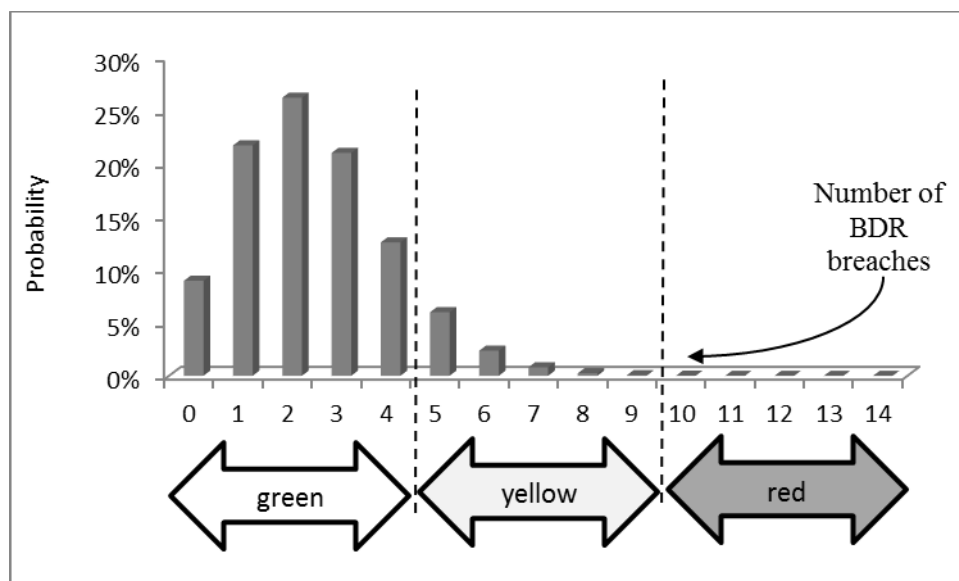
The key parameter is the probability of breaching which refers to risk underestimation. It can be interpreted as a measure of conservatism implemented by a bank under the process of model validation. The smaller is the assumed probability, the more difficult it is to deem the model as accurate. And vice versa, once the probability of breaching q gets larger, the PD projection model is likely to be identified as accurate. According to Basel Committee recommendations the probability of breaching should not be higher than 1%.

Under the TLA approach, it is necessary to determine the number of tries N . This number is usually not high as it represents the number of forecasts. This is why it is convenient to perform portfolio segmentation and derive forecast for each segment. Typically, the credit scoring system provides natural segmentation where each identified group of borrowers is associated with various PD. Furthermore, a twelve months forecast horizon can be applied for loans originated in various periods (months / quarters) which also leads to more tries.

According to the recommendations of the Basel Committee, individual areas associated with different colors should correspond to various strategies. The example analysis (figure 3) was performed for 240 tries, where twelve forecasts for twelve groups were analysed. The green area covers the number of breaches where the probability of higher number of breaches then the maximum given for green area is not higher than 5%. In other words, the actual number of breaches observed in the green area means that the projection model can be deemed as reliable.

The area mark with yellow colour is associated with the probability of 1%, so the observations within this interval should initiate immediate model audit procedure .

Fig. 3 TLA approach for 240 tries and probability of breaching equal to 1%

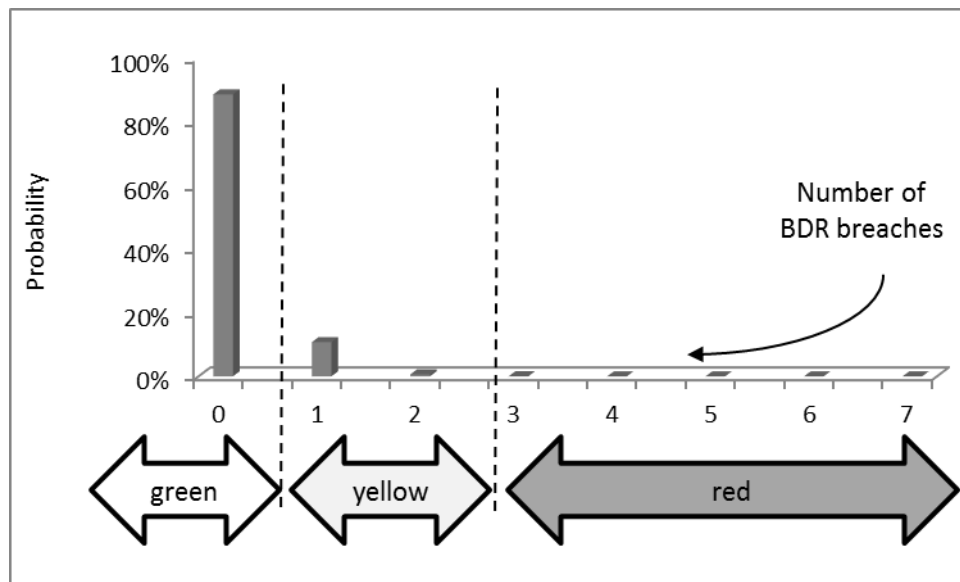


Source: own study

The TLA approach can also be leveraged independently for each individual group of loans identified under the scoring system. However, when the number of tries decreases, the outcomes are more difficult to interpret. Figure 4 shows probability distribution function for breach events where only twelve forecasts were analysed and probability of individual breach was assumed to be equal 1%.

The three areas calculated for 12 tries appeared to be relatively narrow comparing to previous example with 240 tries. The interval marked with green colour covers only non-breaching event where zero events are accepted for reliable projection model. The yellow area assuring conditional model approval includes one and two possible breaches. The number of events where BDR is higher than PD cannot be above 3 for robust forecasting model. If this is the case, the model should be considered as inaccurate and should be thoroughly examined.

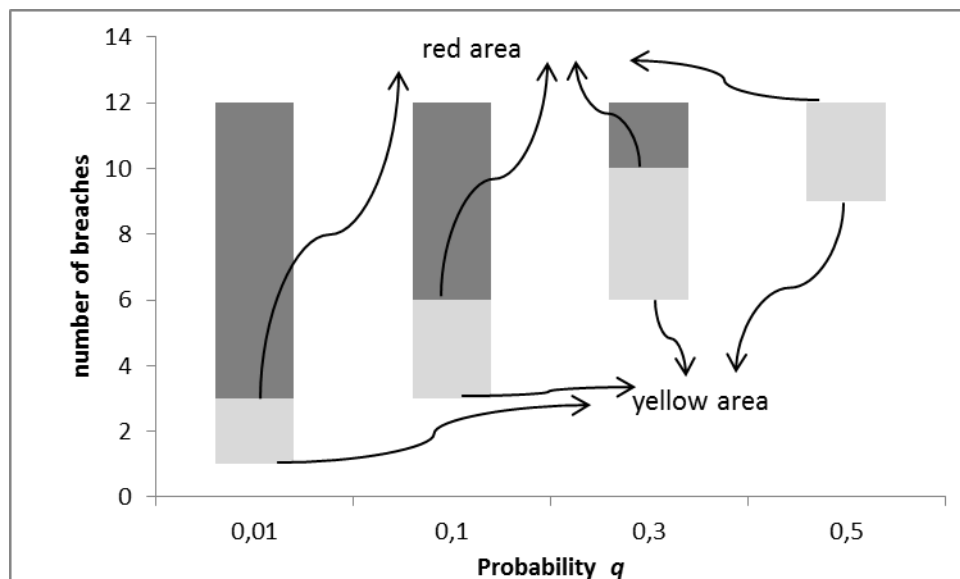
Fig. 4 TLA approach for 12 tries and probability of breaching equal to 1%



Source: own study

According to the presented above formula for probability of k breaches calculation, the TLA intervals depend substantially on the value of the arbitrarily assumed probability of individual breach q . The probability reflects the measure of conservatism under the model validation process. Once q gets higher, also the bounds of intervals are being shifted upwards. If this is the case, it is easier to deem the model as reliable. Figure 5 presents the outcomes of sensitivity analysis, where intervals were calculated for various q values.

Fig. 5 sensitivity analysis of TLA intervals (12 tries)



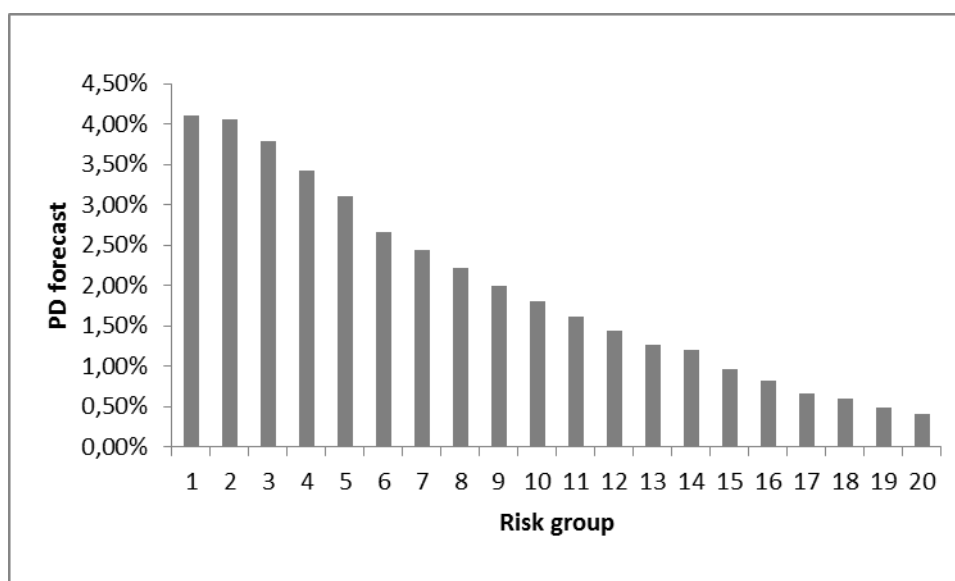
Source: own study

Under the sensitivity analysis there were considered various breach probabilities q i.e. 0.01, 0.1, 0.3 and 0.5. For q equal to 0.01, the yellow area starts with one breach, and the red area covers three and above breaches. When the probability of individual breach was assumed to be 0.1, then the yellow area starts at three breaches, while the red area starts at six. This tendency continuous with rising the probability q . So, higher q values impact interval boundaries by shifting it upwards. Therefore, high probabilities q increase the risk of considering a faulty forecasting model as reliable. Hence, the identified relation underlines the importance of probability q under the model validation process. Its inadequate value can significantly impact the final validation conclusions and lead the bank to severe losses.

5. TLA implementation example

To illustrate the TLA approach, the data from a financial institution was acquired. The financial institution is a leading consumer finance bank in Poland. Based on the credit scoring model, 20 groups of loans with similar credit risk (PD) were identified. The scoring model was developed using 31 000 loans originated between year 2012 and 2015. The loans were originated to finance car purchases by individual persons. For each of the identified group the probability of default was estimated.

Fig. 6 PD forecasts for identified risk groups



Source: own study

It can be noticed (Figure 6) that the higher group number (ranging from 1 to 20) the lower is PD which corresponds with higher score value. The scoring model was developed based on

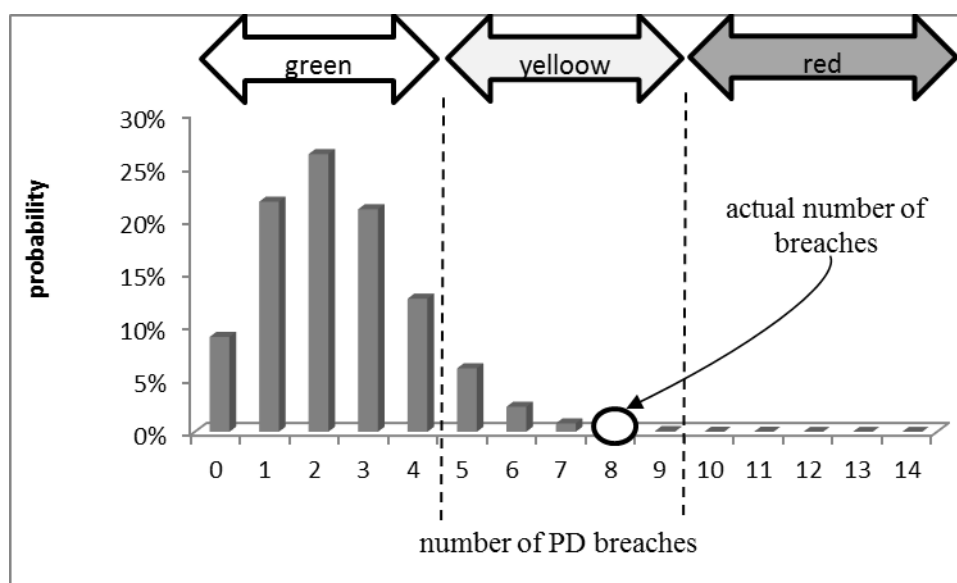
the following characteristics: the borrower's net income, the ratio of the loan value and the value of the purchased car, the number of years of employment, the age of the borrower, the number of loan installments, the number of family members.

The vintage analysis outcomes were leveraged to calculate the number of defaulted loans in examined portfolio. Then, the ratios of defaulted loans and non-defaulted loans were calculated over the period of 12 months to estimate BDR statistics. Finally for year 2016 a total of 240 (12 months* 20 scoring groups) BDR estimates were derived.

The BDR estimates were used in line with previously projected PD forecasts to verify the hypothesis regarding the high quality of PD projection model. It was assumed that the model provides robust risk projections and therefore can be used as a reliable tool. Each individual BDR was compared with its original PD forecast adequate for risk group. According to the performed analysis, the number of cases where BDR was above adequate PD was equal to eight.

Figure 7 presents empirically observed number of PD breaches and TLA intervals derived in line with previously presented methodology. Eight breaches placed its value within the yellow interval. This means that the projection model can be deemed as reliable and can be used for further forecasting process. However, the yellow area implies that it is necessary to examine the model in order to identify reasons of relatively large number of breaches. So, the model can be conditionally used, but thorough model testing is recommended.

Fig. 7 Actual number of breaches under the TLA approach



Source: own study

Further analysis of the financial institution policy revealed key factors which led to the unexpected increase of credit risk. It was noticed that during the years 2014-2016 the financial institution realized aggressive market strategy causing dynamic growth of car loan sales. At that time the sales network has been substantially expanded impacting the sales volume. On the other hand dynamic growth was accompanied by the credit risk policy loosening. It was noticed that the cut-off point of the scoring system has not been changing over time, however some credit policy rules so called 'boundary conditions' were revised. The identified liberalization concerned wider range of possible collaterals as well as wider range of acceptable legal forms of employment. These changes may impact the portfolio composition and lead to portfolio quality deterioration. The scoring model used to estimate the probability of default was developed based on historical population where other credit risk rules were applied. So, this could be the reason of relatively high level of PD breaches. Therefore, the management team of the financial institution was recommended to recalibrate the scoring model, as well as to perform the back-testing analysis more frequently.

6. Conclusions

In contemporary banking industry, probability of default models constitute an integrated parts of the credit risk management systems. With standardized definition of default, implemented by Basel Committee, it is easier to compare the analysis outcomes presented by various financial institutions. Therefore risk management process became more transparent. However estimation process of risk measures requires consideration of specific regional macro variables. It results directly from the regulations like US CCAR (Comprehensive Capital Analysis and Review), EU-wide stress tests or International Financial Reporting Standard (IFRS9). All these regulations encourage financial institutions to project credit losses according to macro factors reflecting regional economic conditions. Furthermore risk forecasts are expected to be validated by independent experts.

The paper shows that TLA approach can be leveraged for PD model validation purposes. It is especially useful for global banks with regionally diversified portfolios, where the projections can be assumed to be independent. One of the key advantage of the method results from its relatively straightforward implementation and outcomes interpretation. It was shown that the necessary data can be collected from traditional vintage analysis.

The paper shows empirical implementation of TLA approach which illustrates the process of PD model quality assessment. The examined model used for PD projections was deemed as

reliable, however it was noticed that the number of underestimated forecasts reached relatively high level. Thus, the model still can be leveraged for further forecasting process in line with null hypothesis.

Further complementary analysis revealed the reasons of observed PD breaches. The dynamic sales growth was accompanied with changes in bank's credit risk policy. It was noticed that this new strategy may impact portfolio composition and overall average PD. Furthermore, macroeconomic changes observed over the past years could also impact the overall credit risk losses.

References

- Anderson R.(2007) *The Credit Scoring Toolkit*, Oxford University Press.
- Basel Committee on Banking Supervision (1996). *Supervisory Framework for the Use of "Backtesting" in Conjunction with the Internal Models Approach to Market Risk Capital Requirements*.
- Basel Committee on Banking Supervision (2006), *International Convergence on Capital Measurement and Capital Standards*, Bank for International Settlements, Basel.
- Basel Committee on Banking Supervision. (2006) Annex 10a: *Supervisory Framework for the Use of "Backtesting" in Conjunction with the Internal Models Approach to Market Risk Capital Requirements*. In: *International Convergence of Capital Measurement and Capital Standards. A Revised Framework Comprehensive Version*. Bank for International Settlements,.
- Basle Committee on Banking Supervision (1988), *International Convergence of Capital Measurement and Capital Standards*, Basel.
- Blochwitz, S. Hohl, Tasche, D., Wehn, C.S. (2005), *Reconsidering Ratings*. Deutsche Bundesbank, Working Paper, May.
- Blochwitz, S., Hohl, S, Tasche, D., Wehn, C.S. (2004) *Validating Default Probabilities on Short Time Series*. Deutsche Bundesbank, Working Paper.
- Braun A., Rymaszewski P., Schmieser H.(2010), *A Traffic Light Approach to Solvency Measurement of Swiss Occupational Pension Funds*, Working Papers on Risk Management and Insurance No.74, University of St. Gallen.
- Breeden L.(2004), *Stress Testing 2004-2005 Retail Origination,s* The RMA Journal.
- Burns P., Stanley A. (2001),*Managing Consumer Credit Risk*, Discussion Paper, Federal Reserve Bank of Philadelphia.
- Castermans, G., Martens, D., Van Gestel, T., Hamers, B., Baesens, B. (2007), *An Overview and Framework for PD Backtesting and Benchmarking*. In *Credit Scoring and Credit Control*, Edinburgh (U.K.).
- Coppens F., Gonzalez F., Winkler G.(2007), *The performance of credit ratings systems in the assessment of collateral used in eurosystem monetary policy operations*, Occasional Paper Series, European Central Bank.
- Dungey N., Dwyer J., Flavin T., *Vintage and Credit Rating: What matters in the ABX data during the credit crunch?*, <http://www.frbsf.org/economics/conferences/0901/Dungey-Dwyer-Flavin.pdf> 2008.

- Emmer S., Tasche D.(2005), Calculating credit risk capital charges with the one-factor model <http://arxiv.org/abs/cond-mat/0302402v5>.
- Engelman B., The Basel II Risk Parameters: Estimation, Validation, and Stress Testing, Springer 2006.
- European Banking Authority (EBA) (2015), Implementing technical standards on the mapping of ECAIs credit assessments. <https://www.eba.europa.eu/regulation-and-policy/external-credit-assessment-institutions-ecai/draft-implementing-technical-standards-on-the-mapping-of-ecais-credit-assessments>.
- Hamerle A., Liebig T., Scheule H. (2004), Forecasting Credit Portfolio Risk, Deutsche Bundesbank, Working Paper, May.
- Jackson P., Perraudin W. (2000), Regulatory implications of credit risk modelling Journal of Banking & Finance 24.
- Merton R. C. (1974), On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, Journal of Finance, No. 29, s. 449-470.
- Moody's (2015), Moody's global approach to rating SME balance sheet securitizations. https://www.moody's.com/researchdocumentcontentpage.aspx?docid=PBS_SF393152.,
- Nadarajah, S. and Chan, S. (2016). Estimation methods for value at risk. Chapter 12 of Extreme Events in Finance: A Handbook of Extreme Value Theory and Its Applications (edited by F. Longin), pp. 303-373. John Wiley and Sons, Chichester.
- Pelletier, D. and Wei, W. (2015). The geometric-VaR backtesting method. Journal of Financial Econometrics, 2015, 1-21.
- Rösch D., Scheule H. (2004), Forecasting retail portfolio credit risk. Journal of Risk Finance, Winter/Spring, s. 16–32.
- Schuermann T., Hanson S.(2004), Estimating Probability of Default, Federal Reserve Bank of New York Staff Reports.
- Svec M.(2009) PD Backtesting Empirical Study on Credit Retail Portfolio, CSOB Bank.
- Tasche, D.(2003) A Traffic Lights Approach to PD Validation. Deutsche Bundesbank, Working Paper.
- Uyemura D.g., Wan Deventer D.R. Zarządzanie ryzykiem finansowym w bankach Związek Banków Polskich Warszawa 1997.
- Vasicek O.(2002), The Distribution of Loan Portfolio Value, Risk..
- Y. Zhang & S. Nadarajah (2018) A review of backtesting for value at risk, Communications in Statistics - Theory and Methods, 47:15, 3616-3639, DOI: 10.1080/03610926.2017.1361984