

https://dspace.org/

The impact of deterioration in ratingmodel discriminatory power on expected losses

S Zhou; Gary Van Vuuren

INCISIVE MEDIA
https://hdl.handle.net/10539/47109
WOS

Downloaded from a DSpace repository. https://dspace.org/

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/383685379

The impact of deterioration in rating-model discriminatory power on expected losses

Article *in* The Journal of Risk Model Validation · January 2024

DOI: 10.21314/JRMV.2024.004

CITATIONS
0

READS
122

2 authors:



Gary van Vuuren
North-West University
144 PUBLICATIONS **1,193** CITATIONS
SEE PROFILE



Siyi Zhou
Emirates NBD Bank
3 PUBLICATIONS **0** CITATIONS
SEE PROFILE

All content following this page was uploaded by Gary van Vuuren on 02 October 2024.

The user has requested enhancement of the downloaded file.

Journal of Risk Model Validation **18**(2), 1–13
DOI: 10.21314/JRMV.2024.004

Copyright Infopro Digital Limited 2024. All rights reserved. You may share using our article tools. This article may be printed for the sole use of the Authorised User (named subscriber), as outlined in our terms and conditions. https://www.infopro-insight.com/termsconditions/insight-subscriptions

Research Paper

The impact of deterioration in rating-model discriminatory power on expected losses

Siyi Zhou[1] and Gary van Vuuren[2]

[1]Group Model Validation, Emirates NBD PJSC, Baniyas Road, Deira, PO Box 777, Dubai, United Arab Emirates; email: siyiz@emiratesnbd.com

[2]Centre for Business Mathematics and Informatics, North-West University, 11 Hoffman Street, Potchefstroom 2531, South Africa; email: vvgary@hotmail.com

(Received November 6, 2023; revised March 21, 2024; accepted March 22, 2024)

ABSTRACT

This paper sets out a methodology for estimating the impact on a portfolio's expected credit loss (ECL) caused by potential model risks in underwriting, such as those arising from scoring models (for retail banking) and rating models (for wholesale banking). Such models are used not only for underwriting assessments but also for the determination of through-the-cycle probabilities of default (for regulatory capital purposes) and point-in-time probabilities of default (for International Financial Reporting Standard 9 ECL reporting). An important metric in credit risk management and assessment – required by regulators and auditors – is the impact on ECL if the model's Gini coefficient deteriorates. We establish the quantum of Gini weakening that breaches a bank's internal validation standards (as assessed by a consequently large impact on the ECL), and we formulate an analytical approach, caveated with realistic assumptions, that allows for tractable, computationally simple measurement of ECL as the Gini coefficient worsens. The efficiency gained by the simplifying assumptions comes at the cost of losing granularity and some degree of accuracy.

Keywords: model risk; Gini index; expected credit losses; point in time (PIT); through the cycle (TTC); validation.

Corresponding author: G. van Vuuren

Print ISSN 1753-9579 | Online ISSN 1753-9587
© 2024 Infopro Digital Risk (IP) Limited

1 INTRODUCTION

In this paper we propose a methodology for estimating the impact of using potentially weak-performing underwriting models, such as scoring models in retail banking and rating models in wholesale banking, to compute a portfolio's expected credit loss (ECL). In addition to underwriting purposes, such models are used extensively for computing through-the-cycle (TTC) probabilities of default (PDs), transformed into point-in-time (PIT) PDs for International Financial Reporting Standard 9 (IFRS 9) ECL reporting. While these models aim to provide quick solutions to determine ECL without undue cost, they raise the following question (often asked by regulators and auditors): what is the impact on ECL if the model's Gini coefficient and/or calibration test statistics breach internal validation standards, and how may this impact be mitigated? In response, the UK Prudential Regulation Authority, for example, introduced the concept of post-model overlays to temporarily mitigate model risk (Prudential Regulation Authority 2023). Our analysis could assist banks and regulatory authorities in determining ECL model-risk mitigation buffers.

Inevitably, solutions are highly analytical and come with several assumptions to allow for better tractability and computational ease. Efficiencies are gained at the cost of losing granularity and accuracy. This paper only focuses on the underperformance of scoring/rating models due to poor discriminatory power, as measured, for example, by the Gini coefficient (Dorfman 1979), also known as the accuracy ratio. Another important performance metric for rating/PD models is calibration accuracy, which has been more extensively studied (see, for example, Prorokowski *et al* 2021; Ptak-Chmielewska and Kopciuszewski 2022). The impact on ECL due to miscalibration is conceptually straightforward and hence beyond the scope of this paper. Our analysis assumes a reasonably well-calibrated rating system but a deterioration in discriminatory power.

The remainder of the paper proceeds as follows. Section 2 sets out a survey of the relevant literature. Section 3 establishes the mathematical development of the methodology for large homogeneous portfolios, while Section 4 does the same for large inhomogeneous portfolios. Section 5 states our conclusions and presents suggestions for future work.

2 LITERATURE REVIEW

Scoring models can only be used effectively once an assessment of their prediction accuracy is known. Models are selected according to quality measures determined at the time of development and continuously after deployment into the business, when real conditions prevail. Credit scoring model methodologies, along with quality assessments, have been adequately surveyed by Hand and Henley (1997), Thomas

(2000), Engelmann *et al* (2003) and Crook *et al* (2007). A comprehensive survey of credit scoring approaches and the pitfalls that plague them is provided by Vojtek and Kočenda (2006). Izzi *et al* (2011) set out an applied guide for quantitative and qualitative regulatory credit risk models, while Onay and Öztürk (2018) provide a review of credit scoring research in an age dominated and defined by the availability of big data. Finally, Hayashi (2022) explores emerging deep-learning trends applied directly to credit scoring, while Markov *et al* (2022) examine the trends emerging in (and propose possible trajectories of) the field of credit scoring. Breed *et al* (2022) propose a novel approach for deriving PIT PDs by first deriving a PIT term structure with accompanying segmented term structures and then calibrating credit scores using a Gini-index approach to create account-specific PD term structures.

Little work has been devoted to the assessment of credit scoring model quality. While there are several metrics dedicated to evaluating this quality, there is little consensus among academics and practitioners on the best approach for the task. Řezáč and Řezáč (2011) provide a comprehensive summary of metrics used to evaluate credit scoring model quality, including the Gini index (widely employed in Europe), Kolmogorov–Smirnov (KS) statistics (which enjoy popularity in the United States) and Mahalanobis distances. Gini coefficients and KS statistics may not be optimal or reliable metrics: the Gini coefficient is a global measure, so it is difficult to employ it to assess the quality of local outputs, while the KS statistic is only ideal if the expected cutoff values are near the point where the statistic is realized, which is rarely the case in credit portfolios. Although the Gini coefficient and similar metrics can be used as model-performance indicators, their secondary impact on an institution's profitability and/or capital adequacy is rarely understood by academia and the industry. This paper is, to the best of our knowledge, the first attempt in the literature to study the impact on the ECL of deteriorated Gini coefficients in credit scoring models.

3 LARGE HOMOGENEOUS PORTFOLIOS

Credit scorecards are often used to produce customer rank orderings, helping the business to sell products to qualified customers. In IFRS 9 modeling, credit scorecards are also used to estimate TTC PDs, which are subsequently transformed into PIT PDs and driven by selected macroeconomic variables.

This section provides approaches to estimate the impact on the ECL of a large homogeneous portfolio. If the exposure is the same for each obligor, this portfolio is "size-homogeneous". If the number of portfolio obligors tends to infinity, the portfolio is "large". Most retail banking portfolios consist of small-ticket loans, so they meet the criteria of large, homogeneous portfolios. For a comprehensive discussion

of whether the Gini coefficient is an appropriate metric for evaluating low-default portfolios, see Frunza (2013).

Suppose there is a retail portfolio consisting of n obligors, where each obligor has the same exposure size and loss given default.[1] Denote by x_i the score produced by a scorecard. Then, the probability of default for obligor i is $p(x_i)$, and the portfolio expected credit loss, $\mathrm{ECL_p}$, is

$$\mathrm{ECL_p} = \mathrm{EAD} \times \mathrm{LGD} \sum_{i=1}^{n} p(x_i),$$

where EAD is the exposure size and LGD is the loss given default.

Calculating ECL is equivalent to a randomized process. Assume there is a pool of an infinite number of customers (the population). Select n customers independently from this population (the sample). Before a customer is selected, the PD that should be used is unknown, so x_i is a random variable (the same applies for $p(x_i)$). Theoretically, x_i can follow any distribution, but for illustration we assume this to be a normal distribution, $N(\mu_x, \sigma_x)$;[2] note the subscript i has been removed, because x_i is the ith realization in this distribution.

The distribution for $p(x)$ depends on the function $p(\cdot)$. In many scorecards the PD is a calibrated logistic function:

$$p(x) = \frac{1}{1 + \mathrm{e}^{-s(x)}},$$

where $s(x) = \alpha + \beta x$. Because s is a linear transformation of x, it also follows a normal distribution, so $s \sim N(\mu_s, \sigma_s)$, where $\mu_s = \alpha + \beta\mu_x$ and $\sigma_s = \beta\sigma_x$. Also, $p(s)$ follows a logit-normal distribution:

$$p(s) \sim P[N(\mu_s, \sigma_s)].$$

The probability density function of this logit-normal distribution is

$$f_{\mathrm{p}}(p) = \frac{1}{\sigma_s\sqrt{2\pi}} \exp\left[-\frac{(\mathrm{logit}(p) - \mu_s)^2}{2\sigma_s^2}\right] \frac{1}{p(1-p)}, \quad 0 < p < 1,$$

where

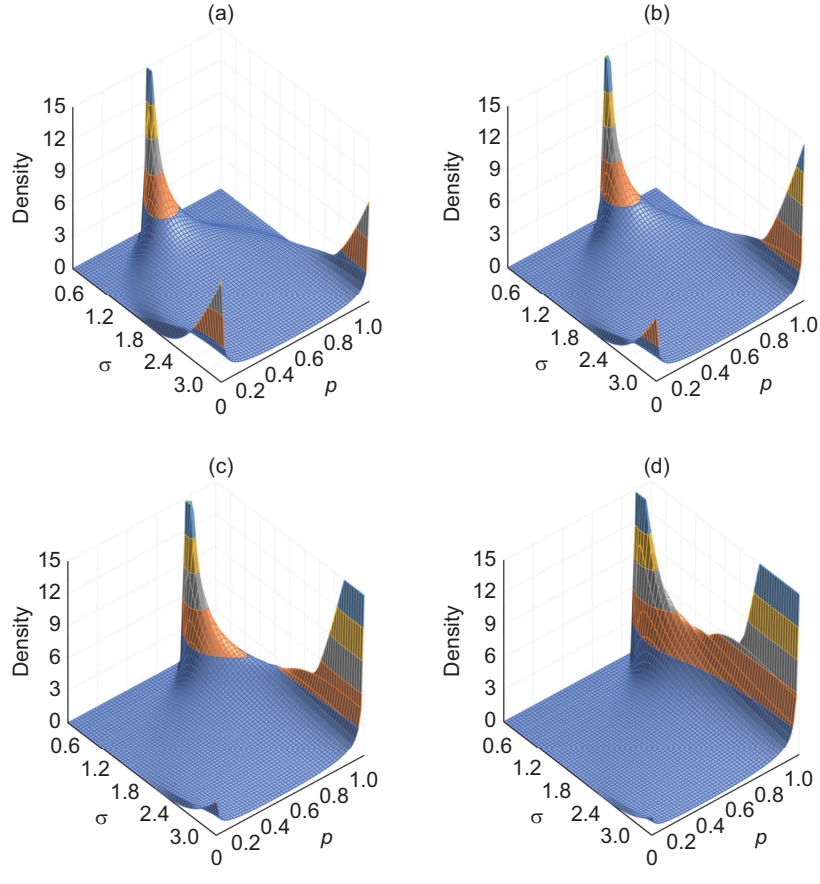
$$\mathrm{logit}(p) = \log\left(\frac{p}{1-p}\right).$$

Figure 1 shows that the probability density function has the following properties:

[1] This assumption may be unnecessary. It is only made for ease of computation and will be assessed separately.

[2] It is not uncommon to assume, incorrectly, that the histogram of credit scores is bell-shaped, especially for the types of models that employ latent variables to model obligor defaults. Relative to other distributions, it is frequently employed in the industry and by academics as it can lead to more tractable solutions.

FIGURE 1 Logit-normal density function for increasing values of μ and σ.



(a) $\mu = 0$. (b) $\mu = 0.6$. (c) $\mu = 1.3$. (c) $\mu = 2.0$.

- it is bounded by 0 and 1;
- when $\mu_s = 0$, the distribution is symmetric;
- when $\mu_s > 0$, it is left-skewed;
- increasing σ_s increases the distribution variance; and
- large σ_s results in a bimodal distribution.

Because there is no analytical solution for the expected value (mean) of the logit-normal distribution, it must be solved numerically (eg, by numerical integration).

However, if the scorecard is calibrated to a portfolio central tendency, denoted by $\bar{p}$, then $E[p(s)] = \bar{p}$.

The calculation of the portfolio ECL can also be expressed as

$$\mathrm{ECL_p} = n\mathrm{EAD} \times \mathrm{LGD}\left[\frac{1}{n}\sum_{i=1}^{n} p(x_i)\right].$$

Therefore,

$$\mathrm{ECL_p} \approx n\mathrm{EAD} \times \mathrm{LGD}\bar{p}. \tag{3.1}$$

From the above result, we can conclude that the ECL in a large homogeneous portfolio is ultimately determined by the central tendency toward which the scorecard is targeted.[3]

How can we reduce the impact of poor Gini values on ECL? The straightforward approach to mitigating a poor Gini coefficient is to develop a model with a better coefficient (to comply with validation standards) and recompute the expect loss. However, provided we do not alter the target central tendency $\bar{p}$, then no matter how we adjust the model to improve its Gini coefficient, the portfolio ECL should not change.

In the context of IFRS 9, banks often recalibrate the scorecard's TTC PD to a forecasted portfolio-level PIT PD driven by predetermined macroeconomic variables. There is no unique way to develop this kind of satellite model to forecast PIT PDs. For example, a linear model may be developed by regressing a portfolio-level (transformed) observed default rate (ODR) on (transformed) macroeconomic variables. Given the macroeconomic scenarios, the model outputs project portfolio-level ODR for future time periods. This projected ODR essentially plays the role of the updated portfolio central tendency to which the scorecard must be calibrated.

In many banks the final scorecard PIT PD calibration involves a parallel shift of the existing TTC PD, achieved by calculating a scalar such as the following:

$$\frac{a_T}{n}\sum_{i=1}^{n} p(x_i) = \bar{p}_T,$$

where $\bar{p}_T$ is the projected portfolio-level ODR for future year T, $p(x_i)$ is the scorecard PD and a_T is the scalar for future year T.

In the above exercise the portfolio average PD is recalibrated to $\bar{p}_T$. The final obligor-level PIT PD can be derived as the product of the scalar and its original TTC PD. Note that this recalibration, resembling a parallel shift, does not alter the rank

[3] We caution against the insinuation that model Gini coefficients are irrelevant to portfolio ECL. If the portfolio central tendency is kept constant, only those changes to the Gini coefficient made after portfolio selection do not impact the ECL.

order of the original PD: the discriminatory power of the PIT PD persists as it is inherited from the underlying scorecard model.

Using similar arguments to those for (3.1), the ECL of this portfolio for year T is

$$\mathrm{ECL_p}(T) \approx n\mathrm{EAD} \times \mathrm{LGD}\bar{p}_T.$$

The logic flow for large homogeneous portfolios is illustrated in Figure 3 (see Appendix A, available online).

4 LARGE INHOMOGENEOUS PORTFOLIOS

In Section 3 we analyzed the impact of different scorecard models on the ECL of a large homogeneous portfolio. In this section we extend the analysis to large inhomogeneous portfolios.

Denote by e_i the exposure at default of obligor i, so that the ECL of the portfolio can be expressed as

$$\mathrm{ECL_p} = \mathrm{LGD}\sum_{i=1}^{n} p(x_i)e_i.$$

As in the case of large homogeneous portfolios, calculating the above ECL is equivalent to a randomized process in which we randomly select n customers independently from an infinite pool of customers. Before the customer is selected, nobody knows what the PD and exposure at default should be. Upon selection, the pair (x_i, e_i) will be realized from a joint distribution $g(x, e)$.

Any theoretical distribution can be used for $g(\cdot)$, but for illustration we assume both x and e follow normal distributions. Then, $g(\cdot)$ is a bivariate normal distribution:

$$\begin{pmatrix} x \\ e \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_x \\ \mu_e \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \rho_{xe}\sigma_x\sigma_e \\ \rho_{ex}\sigma_e\sigma_x & \sigma_e^2 \end{bmatrix}\right).$$

The portfolio ECL is a random variable from a certain distribution,[4] and its expectation is

$$E(\mathrm{ECL_p}) = \mathrm{LGD}\sum_{i=1}^{n} E(p(x_i)e_i).$$

Now, the expectation inside the summation can be dealt with by the law of iterated expectation (the subscript is omitted for simplicity):

$$E(p(x)e) = E(E(p(x)e \mid x)) = E(p(x)E(e \mid x)).$$

[4] The ECL is the sum of the product of two bivariate normal distributions. Unfortunately, there is no closed-form solution for the distribution of this dual product except for independent pairs.

The conditional mean for e is easy to calculate given that e and x are bivariate normally distributed:

$$E(e \mid x) = \mu_e + \rho_{ex}\frac{\sigma_e}{\sigma_x}(x - \mu_x).$$

We plug the conditional mean into the expectation to obtain

$$E(p(x)e) = E\left(p(x)\left(\mu_e + \rho_{ex}\frac{\sigma_e}{\sigma_x}(x - \mu_x)\right)\right).$$

Simple algebra shows that

$$E(p(x)e) = \bar{p}\mu_e + \rho_{ex}\sigma_e E\left(p(x)\left(\frac{x - \mu}{\sigma_x}\right)\right).$$

Define $\tilde{x}$ to be the standardized score, ie,

$$\tilde{x} = \frac{x - \mu_x}{\sigma_x}.$$

As in Section 3, we assume the PD is a logistic function of score x_i, so

$$p(x) = \frac{1}{1 + \mathrm{e}^{-s(x)}},$$

where $s(x) = \alpha + \beta x$. In turn, $s(x)$ can be written in terms of the standardized variable:

$$s(x) = \alpha + \beta(\mu_x + \sigma_x\phi) \quad \forall\phi \sim N(0, 1).$$

Since s is a linear transformation of $\tilde{x}$, it is also a normal distribution, $s \sim N(\mu_s, \sigma_s)$, where $\mu_s = \alpha + \beta\mu_x$ and $\sigma_s = \beta\sigma_x$. Returning to the derivation of the expectation, we now change variable from x to its standardized value, $\tilde{x}$:

$$E(p(x)e) = \bar{p}\mu_e + \rho_{ex}\sigma_e E(p(\tilde{x})\tilde{x}).$$

The expectation of the total portfolio ECL is then

$$E(\mathrm{ECL_p}) = \underbrace{n\mathrm{LGD}\bar{p}\mu_e}_{(3.1)} + n\mathrm{LGD}\rho_{ex}\sigma_e E(p(\tilde{x})\tilde{x}). \tag{4.1}$$

Because the covariance of two random variables is linked to its product expectation, we have

$$E(p(\tilde{x})\tilde{x}) = \mathrm{cov}(p(\tilde{x}), \tilde{x}) - E(p(\tilde{x}))E(\tilde{x}) = \mathrm{cov}(p(\tilde{x}), \tilde{x}),$$

since $E(p(\tilde{x}))E(\tilde{x}) = 0$ because $E(\tilde{x}) = 0$.

Normally we will impose on the scorecard model the condition that high scores imply lower PDs, so it is a necessary condition that $\text{cov}(p(\tilde{x}), \tilde{x}) < 0$, since all other terms are strictly greater than or equal to 0.

Equation (4.1) may also be expressed in terms of covariance:

$$E(\text{ECL}_\text{p}) = n\mu_e \text{LGD}\bar{p} + n\text{LGD}\rho_{ex}\sigma_e \,\text{cov}(p(\tilde{x}), \tilde{x}). \tag{4.2}$$

Keeping n, $\bar{p}$, μ_e, σ_e and LGD constant, the following results are obtained.

- The first part of (4.1) is (3.1) (that is, the ECL from the homogeneous case), which captures only the means of the PD and exposure at default.
- The second term represents an adjustment or add-on.
- Because the covariance term is less than 0, a positive correlation ρ_{ex} reduces the ECL. Very large exposures give very low PDs and vice versa, resulting in lower ECLs.
- If $\rho = 0$, the second term equals 0.
- Higher variance in the exposure at default results in lower (higher) ECLs when $\rho_{ex} > 0$ ($\rho_{ex} < 0$).
- Higher absolute covariance values result in lower (higher) ECLs when $\rho_{ex} > 0$ ($\rho_{ex} < 0$).

If we plot the pair $(\tilde{x}, p(\tilde{x}))$, we can see a PD curve decreasing with respect to the corresponding score. The steepness of the curve is positively correlated with the absolute value of covariance implied by the model. In fact they are equal, as is the case in the linear regression $p(\tilde{x}) = b\tilde{x} + \varepsilon$, where $\hat{b} = \text{cov}(p(\tilde{x}), \tilde{x})$ for all $\sigma_{\tilde{x}} = 1$.

More importantly, the steepness of the PD curve or the covariance is linked to the Gini coefficient implied by the model through the logistic calibration function. In particular, the higher the absolute value of the covariance or the steeper the curve, the larger the implied Gini coefficient.[5] During model development, the calibration function (specifically the PD curve slope) is dependent on the Gini coefficient observed in the development sample. In applications such as the computation of ECL, this Gini coefficient will persevere, since the calibration function is fixed. Given the calibration function, there is a model-implied Gini coefficient, and this affects the ECL computation through the covariance term (an observation that can be proved mathematically; see Appendix B (online) for more details).

[5] The Gini coefficient implied by a PD curve is a model-implied value. That is, it is not the same as the empirical Gini coefficient, which is computed using default data.

TABLE 1 Simulated PD curves: parameters used and output statistics.

Parameter	Scorecard 1	Scorecard 2
$\mu(s)$	0	0
$\sigma(s)$	3	1
$E[p(\tilde{x})]$ (%)	50	50
max($\tilde{x}$)	1.000	0.968
min($\tilde{x}$)	0.000	0.017
cov($p(\tilde{x}), \tilde{x}$)	−0.347	−0.208

FIGURE 2 Simulated PD curves.

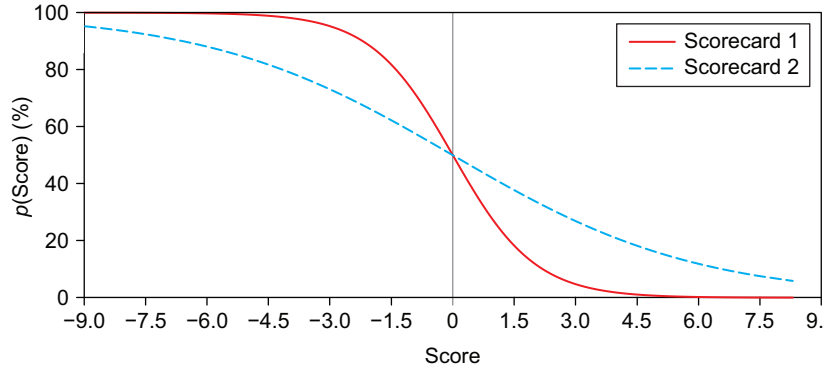


Table 1 lists the parameters (mean and standard deviation) employed when simulating the two PD curves plotted in Figure 2. The PD curve for scorecard 1 is steeper than the PD curve for scorecard 2 (hence the higher absolute value of covariance for scorecard 1 than scorecard 2). More importantly, scorecard 1 implies a higher Gini coefficient than scorecard 2 because it converges to a very low (or very high) PD faster due to its steepness.

What would be the impact on portfolio ECL if the model's discriminatory power (as measured, for example, by the Gini coefficient) deteriorated? First, no matter the model, parameters such as n, $\bar{p}$, μ_e, σ_e and LGD would not change. So the first term in the equation for EL_p, (3.1), would be unaffected. For most retail portfolios, we expect very low correlation between exposure sizes and PDs;[6] hence, the impact on ECL would be negligible. Second, for other portfolios, such as small and medium-sized enterprise (SME) or corporate portfolios, the exposure size may be positively

[6] Empirical evidence is required to confirm this.

correlated with an obligor's turnover, and obligors may tend to receive better ratings if their turnover is high. As a result, positive correlation between exposure sizes and scores is expected.[7]

When $\rho_{ex} > 0$, despite the deterioration in the model's empirical Gini coefficient, the model still calculates portfolio ECL based on the Gini coefficient implied by the PD curve (which is calibrated at the time of model development). In other words, it is calculated as if the Gini coefficient had not deteriorated. In this case, portfolio ECLs are underestimated.

To correct this underestimation, the following steps can be taken.

STEP 1 Calculate the covariance $\mathrm{cov}(p(\tilde{x}), \tilde{x})$ using the existing scorecard – that is, the one implying a higher Gini coefficient (and steeper PD curve) assumed at the time of model development.

STEP 2 Use the deteriorated Gini coefficient, estimated based on recent empirical data, to updated the covariance value.

STEP 3 Holding other parameters constant, calculate $E(\mathrm{ECL_p})$ using the covariance estimated in the previous two steps.

STEP 4 Take the difference calculated in the previous step:

$$\Delta E(\mathrm{ECL_p}) = n\mathrm{LGD}\rho_{ex}\sigma_e \Delta\,\mathrm{cov}(p(\tilde{x}), \tilde{x}).$$

To compute the covariance term, we propose two approaches. One employs empirical data, while the other is a model-based approach combined with numerical solutions such as Monte Carlo simulation.

APPROACH 1 In the first step, calculate the covariance directly by using customer credit scores and the corresponding PDs, or, in other words, by regressing scorecard PDs on standardized scores.[8] In the second step, form the standardized score buckets (or rating categories) and calculate each bucket's PD, approximated by the ODR of that bucket. Finally, regress the obligors' PD on its bucketed scores.

APPROACH 2 In the first step, determine the functional form of $p(\tilde{x})$ and calculate the covariance by Monte Carlo simulation. In the second step, adjust the functional form from the first step to reflect the deteriorated Gini coefficient and then recalculate the covariance.

[7] To be confirmed by empirical evidence.

[8] These data are easily procured from any bank's risk system. For example, a rating system takes a customer's financial statements as inputs to generate a credit score. The score is then transformed into a PD via a calibrated logistic function before being mapped onto the bank's master scale for rating.

Appendix B (online) provides an example of how the functional form of $p(\tilde{x})$ may be determined and its connection to the model-implied Gini coefficient in the framework of the quasi-moment-matching method advocated by Tasche (2013).

5 CONCLUSION AND SUGGESTIONS

This paper proposed an analytical approach for estimating the impact of weak-performing underwriting models on portfolio ECLs in retail and wholesale banking. This approach, though simplified and accompanied by various assumptions, allows for a computationally simple measurement of ECL as the Gini test statistics of scoring and rating models deteriorate. This methodology could be important for banks as it would help them determine the point at which internal validation standards are breached and enable them to take the necessary steps to improve their underwriting and provisioning practices. By quantifying model risk due to poor discriminatory power, it could also help banks, regulatory authorities and auditors to determine post-model overlays as a possible option for model-risk mitigation. While our approach is efficient, such efficiency may come at the cost of losing some granularity and accuracy.

Future studies could be conducted to confirm that a positive correlation between exposure sizes and scores can be expected for other portfolios, such as SME or corporate portfolios.

The performance of other predictive models, as measured, for example, by the mean squared error (which measures the average squared difference between actual and predicted values in regression problems), the Hosmer–Lemeshow test (which assesses the goodness-of-fit of logistic regression models) and the Brier score (which assesses probabilistic prediction accuracy in classification problems), could also be explored in future work.

Finally, a detailed examination of the impact on internal ratings-based risk-weighted assets when a rating model's Gini coefficient deteriorates could provide interesting results with considerable consequences for banks.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES

Breed, D. G., van Jaarsveld, N., Gerken, C., Verster, T., and Raubenheimer, H. (2022). Development of an impairment point in time probability of default model for revolving

retail credit products: South African case study. *Risks* **9**(11), Paper 208 (https://doi.org/10.3390/risks9110208).

Crook, J. N., Edelman, D. B., and Thomas, L. C. (2007). Recent developments in consumer credit risk assessment. *European Journal of Operational Research* **183**(3), 1447–1465 (https://doi.org/10.1016/j.ejor.2006.09.100).

Dorfman, R. (1979). A formula for the Gini coefficient. *Review of Economics and Statistics* **61**(1), 146–149 (https://doi.org/10.2307/1924845).

Engelmann, B., Hayden, E., and Tasche, D. (2003). Measuring the discriminative power of rating systems. Discussion Paper 01/2003, Series 2: Banking and Financial Supervision, Deutsche Bundesbank. URL: https://bit.ly/3AebkDw.

Frunza, M. (2013). Is the Gini coefficient appropriate for low default portfolios? Working Paper, Social Science Research Network (https://doi.org/10.2139/ssrn.2203582).

Hand, D. J., and Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: a review. *Journal of the Royal Statistical Society* A **160**(3), 523–541 (https://doi.org/10.1111/j.1467-985X.1997.00078.x).

Hayashi, Y. (2022). Emerging trends in deep learning for credit scoring: a review. *Electronics* **11**(19), Paper 3181 (https://doi.org/10.3390/electronics11193181).

Izzi, L., Oricchio, G., and Vitale, L. (2011). *Basel III Credit Rating Systems: An Applied Guide to Quantitative and Qualitative Models*. Springer (https://doi.org/nb4q).

Markov, A., Seleznyova, Z., and Lapshin, A. (2022). Credit scoring methods: latest trends and points to consider. *Journal of Finance and Data Science* **8**(4), 180–120 (https://doi.org/10.1016/j.jfds.2022.07.002).

Onay, C., and Öztürk, E. (2018). A review of credit scoring research in the age of Big Data. *Journal of Financial Regulation and Compliance* **26**(3), 382–405 (https://doi.org/10.1108/JFRC-06-2017-0054).

Prorokowski, L., Deev, O., and Guigou, J. D. (2021). Validation nightmare: the slotting approach under International Financial Reporting Standard 9. *The Journal of Risk Model Validation* **15**(2), 63–100 (https://doi.org/10.21314/JRMV.2021.003).

Prudential Regulation Authority (2023). Model risk management principles for banks. Policy Statement 6/23, Bank of England. URL: www.bankofengland.co.uk/prudential-regulation/publication/2023/may/model-risk-management-principles-for-banks.

Ptak-Chmielewska, A., and Kopciuszewski, P. (2022). New definition of default: recalibration of credit risk models using Bayesian approach. *Risks* **10**(1), Paper 16 (https://doi.org/10.3390/risks10010016).

Řezáč, M., and Řezáč, F. (2011). How to measure the quality of credit scoring models. *Czech Journal of Economics and Finance* **61**(5), 486–507. URL: https://is.muni.cz/do/econ/soubory/konference/vasicek/20667044/Rezac.pdf.

Tasche, D. (2013). The art of probability-of-default curve calibration. *The Journal of Credit Risk* **9**(4), 63–103 (https://doi.org/10.21314/JCR.2013.169).

Thomas, L. C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. *International Journal of Forecasting* **16**(2), 149–172 (https://doi.org/10.1016/S0169-2070(00)00034-0).

Vojtek, M., and Kočenda, E. (2006). Credit scoring methods. *Czech Journal of Economics and Finance* **56**(3–4), 152–167. URL: https://journal.fsv.cuni.cz/storage/1050_s_152_167.pdf.

View publication stats