

IMPACT ASSESSMENT OF LOSS GIVEN DEFAULT (LGD) MODELS' RISK ON REGULATORY CAPITAL: A BAYESIAN APPROACH

YANG LIU
HSBC Bank plc, UK

ABSTRACT

"The model is wrong!" so it is determined. All of the estimated output using the model becomes un-reliable immediately. And so is every other result calculated using the unreliable output. So what is the impact of the model being "wrong" in the later calculations? To address this question, this paper presents a Bayesian approach that provides a quantitative assessment for the impact on downstream results calculated using the unreliable estimates. Section 1 details the practical challenge in the financial industry and discusses why this is important. Section 2 starts the discussion with a description of the overall framework for this Bayesian approach, introducing and defining each individual component. Then Sections 3 and 4 carry on discussing the prior and likelihood distributions, respectively. Section 5 then obtains the target posterior distribution by applying the Bayesian posterior update using obtained prior and likelihood results. Then conditioning on value of the unreliable estimate already in place in the portfolio, the density distribution obtained can be used to update the output of the "wrong" model and assess the impact in further calculations. This approach bridges the practitioners' initial expectations with the model performance and provides an intuitive quantitative assessment for the impact in the follow-up calculations which are largely affected by the unreliable estimate. The presented approach is the first in literature to raise the concern of uncertain impact caused by "wrong" models and propose a solution. The pioneer demonstration using uncertainty in the loss given default (LGD) models as an example and assessing the impact on the then calculated regulatory capital provides a timely assessment tool for model risk management in the current banking industry. Note that the abuse of the word *wrong* in quotation marks is an exaggeration of the uncertainty involved, in practice, impact analysis could be requested at any level of uncertainty.

Keywords: model risk management, impact analysis, expected loss ratio, observed loss rate, Bayesian LGD estimate, loss given default (LGD), risk weighted assets, post-observation, best effort loss estimate.

1 INTRODUCTION

Modelling is a heated topic in the finance industry, especially in the recent post-crisis period due to 2 main reasons: First, it has a large impact for everyone in the market in monetary terms. Second, uncertainty in models cannot be completely avoided. Therefore, impact analysis for both model output and any further calculation involving the output is vital for risk management in the current market.

Defined as the percentage ratio of the total underlying exposure given default, the loss rate estimate comes in several different forms, market and traded estimates often assume a fixed loss rate of 40% [1], [2]. In the case of regulatory internal rating-based (IRB) risk models, the empirical loss given default (LGD) is modelled directly from bank's internal loss experience [3], [4].

Empirical studies show that, the observed loss ratio at portfolio level is known to follow the "L", "J" or "U" shaped distribution which could vary over different time and economic conditions [5]. Therefore, it is difficult to obtain a well formulated theoretical distribution to describe the distribution of LGD in general terms.



As a result, it is not uncommon to see uncertainty in LGD models raised under the regulatory IRB framework. While the reason behind the uncertainty can be complex, the requested impact analysis regarding these concerns are highly desirable.

We focus on the LGD model and the regulatory risk weighted assets (RWA), especially for the Banking Book, in this paper as an example because: 1) the Banking Book is unlikely to be traded or hedged, therefore the “wrong” estimate stays with the bank; 2) The LGD model could cover hundreds of billions of exposure, the impact of a “wrong” estimate could be extreme; 3) The RWA is a vital component in calculation of the regulatory Tier 1 Capital Ratio which is published to public and reported to both regulatory authority and the shareholders, it is not desired to report a result based on “wrong” estimates, and the consequences could be disastrous.

Another practical challenge banks face in reality is: re-development or re-calibrating the Banking Book models and obtain regulatory approval to implement the model can be very time consuming, in the interim, live and new deals that is measured in hundreds of billions of exposure is again rated by the model in question. Therefore it is unrealistic to quickly develop a new model to mitigate the risk without a few years of risk bearing.

Together with probability of default, exposure at default, and the maturity, the regulatory RWA can be calculated using paragraph 272 as shown in [3], here we denote the function by F and express the RWA as:

$$RWA = F(PD, LGD, EAD, m). \quad (1)$$

Now, considering the uncertainty already raised, the situation is that: 1) the LGD model is “wrong”; 2) The estimated LGDs are un-reliable; 3) So is the RWA calculated using the LGD. The elephant in the room is: *what is the true RWA for the current portfolio?*

2 BAYESIAN FRAMEWORK AND ANALYTICAL TARGET

To know what is the true RWA for the portfolio is (too) hard or impossible, given the analyst is left with a “wrong” model and a portfolio of un-reliable LGD. However, inspired by the classical Bayesian analysis, this paper propose a solution to answer an alternative question: *Given the un-reliable LGD value is x , which value, θ , has the highest probability to be the realized LGD?*

Denote the realized LGD by θ , $\theta \in \Theta$, where Θ is all possible outcomes of the realized LGD. Further define the modelled LGD by x , $x \in X$, X is the set of modelled LGDs. The probability of observing θ , conditioning on estimated x , is defined as: $q(\theta|x)$. Next, we follow the Bayesian approach to evaluate this probability for all $\theta \in \Theta$, and $x \in X$.

Assume that θ is randomly distributed with some distribution defined by $q(\theta)$, this unconditional probability distribution where our initial uncertainty of θ is embedded is our prior distribution.

The distribution of estimated LGDs conditioning on the realised LGD being θ , $p(x|\theta)$, is the likelihood obtained from observed data. This likelihood is sometimes used to describe the predictive power or model performance in practice, as the distribution of $p(x|\theta)$ describes both how many x values was estimated and the density in each estimate, given that the outcome is θ .

The joint distribution of the realised LGDs Θ and the estimated LGDs X is given by:

$$p(\theta, x) = q(\theta) \cdot p(x|\theta), \theta \in \Theta, x \in X. \quad (2)$$

The marginal distribution of the modelled LGDs can be obtained by integrating through the space of the possible outcomes:



$$p(x) = \int_{\theta \in \Theta} q(\theta) \cdot p(x|\theta) d\theta. \quad (3)$$

According to the Bayes's Theorem, this is the posterior distribution (see more in [6]):

$$q(\theta|x) = \frac{q(\theta) \cdot p(x|\theta)}{p(x)} = \frac{q(\theta) \cdot p(x|\theta)}{\int_{\theta \in \Theta} q(\theta) \cdot p(x|\theta) d\theta}. \quad (4)$$

In the sections below, we discuss each of the individual components in detail.

3 PRIOR DISTRIBUTION: $q(\theta)$

Prior selection is an important topic in the application of Bayesian frameworks. This distribution express the initial belief on the unknown random variable, the LGD θ in this case, in the meantime, the choice of this distribution could have large influence on the later obtained posterior distribution, so it is important to understand the goal of this selection beforehand.

It is common in practice to use a so called “vague” or “non-informative” uniform prior to limit the influence of making specific prior choices. In plain language, this can be expressed as: we do not know anything about the unknown variable of interest, so we equally accept any outcome and the chance of observing any outcome is the same.

We acknowledge that one of the disadvantage of the Bayesian approach is the subjectivity in the choice of prior, and the fact that using a non-informative prior well limit impact of the subjectivity.

However, in case for the internal experience based Bayesian framework, the prior distribution $q(\theta)$ is the probability distribution of observing the realised LGD of θ . In reality it is unlikely for a bank, especially an IRB bank, to state that there is no knowledge with regards to the distribution of observed LGDs, after all the regulatory approval on the use of loss estimate models assume that the bank have empirical evidence and understand the internal loss experience.

On the other hand, owning such knowledge and yet purposely choose to ignore it by choosing non-informative prior distribution will not only purposely mislead the algorithm with false information, it will worsen the other widely known weakness of the Bayesian approach: heavy computational cost.

To best represent the empirical loss experience, our choice of prior is loss data based. But it is unlikely in practice for the internal loss data to be large enough to illuminate the impact on the posterior, even though it is known that the realised LGDs can be heavily bi-modal or skewed towards boundary values [7]. This leads to the general sampling problem between sample distribution and the true population distribution.

Suppose the sample distribution is $\tilde{q}(\theta)$, and the link between the sample distribution and the population distribution is a constant normalization parameter Z_q where:

$$q(\theta) = \frac{1}{Z_q} \cdot \tilde{q}(\theta). \quad (5)$$

The issue here is that the sample distribution can be easy to obtain, but the constant Z_q is impossible or too hard to compute.

Depending on size of the available data there are several approaches to help overcome the sampling problem and estimate the true population distribution, widely known and considered methodology include: Markov Chain Monte-Carlo (MCMC), Kernel Density Estimation methods (KDE) and Bootstrap algorithms, see [8]–[10] for more details.

In this paper, we employ the MCMC approach and use a Metropolis–Hastings sampler to create the chain of target distribution. We will not discuss the details of MCMC nor the

M–H sampler here. More discussion of the algorithm can be found in [8]. We only followed the vanilla approach using Uniform sampling distribution and “independent” constant probability for the proposal distribution.

Denote the internal observed historical data set as Θ_h , the observed LGD values as θ_h . The observed sample distribution is $\tilde{q}(\theta_h)$ while the population distribution does not change.

Fig. 1 demonstrate the density plot of the data and the outcome of the MCMC process. Fig. 1(a) is the density distribution of the observed LGDs, the y-axis mark the density, the x-axis is the observed LGD. Fig. 1(b) is the estimated population distribution of the LGD.

4 THE LIKELIHOOD: $p(x|\theta)$

As mentioned in Section 2, the likelihood is the conditional probability of the modelled LGD x given a realised LGD θ . The likelihood is often reviewed in a bucketed way in case for regulatory IRB LGD models because of conservatism or downturn assumptions embedded, also, the bucketed view well adopt the error tolerance in model predictions.

In this paper, we follow the practice and classify the realised LGD θ in five equal buckets between percentage values of 0 and 100%.

Note that bucketing is not compulsory for the Bayesian analysis, we adopt this practice in order to reflect the reflect the analytical approach in the industry.

The details of this bucketing scheme is defined as shown in Table 1.

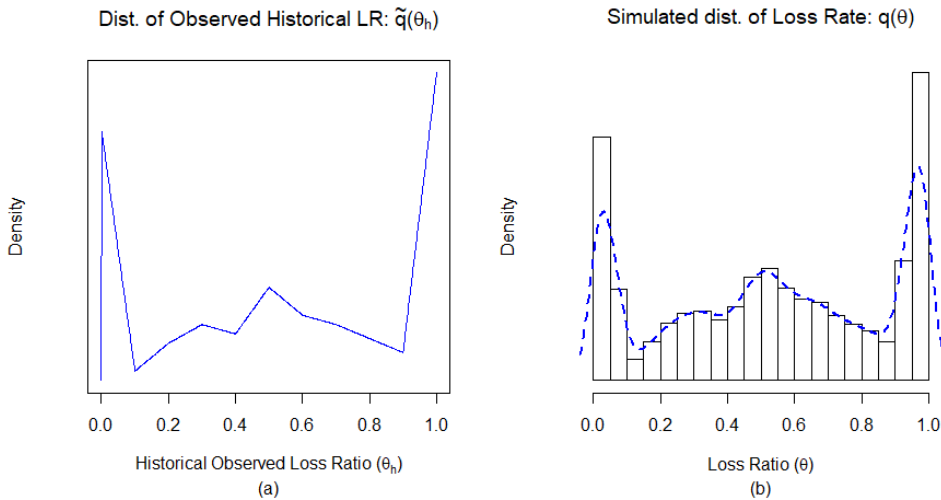


Figure 1: Historical observed LGD distribution vs. simulated LGD distribution.

Table 1: Example of a 5-bucket LGD scale.

Bucket name	Lower LR bound	Expected LR	Higher LR bound
θ_1	0.00%	10.00%	20.00%
θ_2	20.00%	30.00%	40.00%
θ_3	40.00%	50.00%	60.00%
θ_4	60.00%	70.00%	80.00%
θ_5	80.00%	90.00%	100.00%

There are many ways to derive the distribution for $p(x|\theta)$. Empirically summarising the mass probability for each x_1 conditioning on realised observation of θ_1 is straightforward for data rich loss experience. However, as the sensitivity test and distributional analysis of the underlying statistical model are probably available as part of model development or review, one can follow the model assumptions and construct the theoretical likelihood distribution.

To demonstrate the modelling and impact in details, we assume that the likelihood distribution follows a mixture beta distribution, see [11] and [12] for details. The choice of mixture distribution here is just to demonstrate the boundary densities and the different modality properties for different LGD buckets respectively.

Mathematically, the mixture distribution is as following:

$$p(x|\theta_i; \alpha_i^j, \beta_i^j, \rho_i^j) = \sum_{j=1}^{j=2} \frac{\rho_i^j}{B(\alpha_i^j, \beta_i^j)} \cdot x^{\alpha_i^j-1} \cdot (1-x)^{\beta_i^j-1} \text{ for } i \in [1, \dots, 5]. \quad (6)$$

For simplicity we set the weight of the mixture distribution $\rho_i^j = 0.5$ for all LGD buckets $i = 1, \dots, 5$. In practice, the distribution properties should reflect the underlying model in use and there for is less flexible in terms of parameter setting.

Here the mixture distribution in each bucket is centred on the mean of the bucket but given the model development knowledge of the boundary properties, we assume the modelled likelihood reflects higher chance to observe extreme boundary LGDs.

Fig. 2 demonstrates the distribution based initial expectation according to the loss rating bucketing scheme as defined in Table 1. It is important to highlight that the likelihood of obtaining boundary losses are purposely inflated to mimic the modelling effort towards the more frequently observed extreme losses. For example, in the first bucket, when the outcome is a lower end loss the model is expected to predict lower loss more often. While in the 4th bucket, when the observed loss is between 60–80%, more higher-end prediction is expected, even in case of conservative estimates between 80–100% in the right tail.

Assume that we have a data set X^d , recording the recent experience of realized LGD θ^d for each of the model estimated LGD x^d . One then first bucket the realized LGD according to the bucketing scheme shown in Table 1 to obtain θ_i^d , for $i \in (1, 5)$. Then the observed conditional likelihood distribution of $p(x^d|\theta_i^d)$ for all buckets can be obtained.

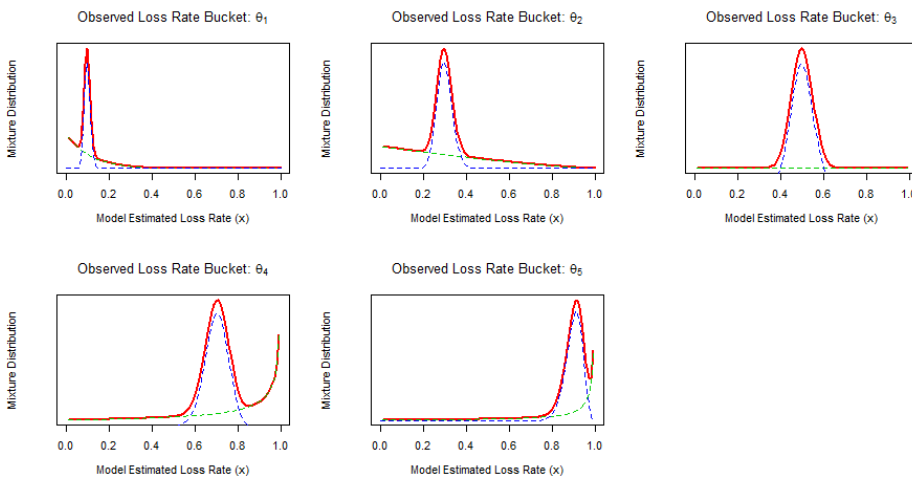


Figure 2: Distribution of modelled estimates for each of the observed loss rate buckets.

Fig. 3 is the consolidated likelihood plot over the observed loss rates obtained directly from the data set X^d , one can see that the issue with distribution obtained from the data is that it is incomplete, in θ_1^d for example. Fortunately, there is no need to employ the numerical techniques to overcome the sampling problem in order to get the conditional likelihood distribution of $p(x|\theta)$ in case the model is built in house and the distribution is known to the analyst. Therefore, obtaining the conditional likelihood is preferred if model specification is available. The complete likelihood distribution curves is shown in Fig. 4.

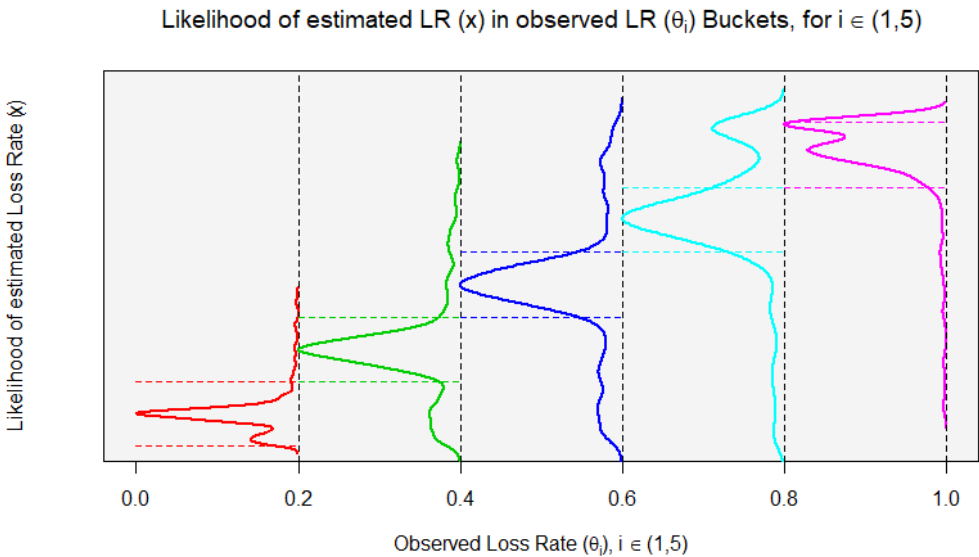


Figure 3: Distribution of modelled estimates for realized loss rate θ .

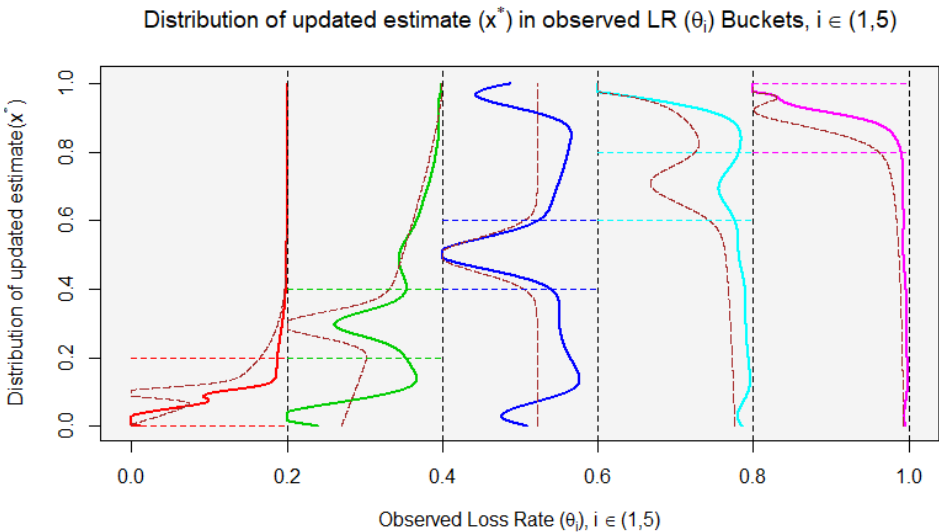


Figure 4: Distribution of updated modelled estimates x^* for realized loss rate θ .



Table 2 summarise the likelihood distribution of the mixture double beta distribution calibrated according to the modelled LGDs as discussed earlier. Conditioning on realised LGD buckets, cells in Table 2 present the likelihood of obtaining a modelled outcome in different buckets. For example, cell $[x_1, \theta_1]$, 93.28% as shown, means that the model is highly likely to produce an estimate of x_1 which later resolves into θ_1 , the likelihood of in-bucket estimation helps to measure the model performance in this bucket.

5 POSTERIOR DISTRIBUTION: $q(\theta|x)$

The posterior distribution can now be obtained using the Bayes's Theorem:

$$q(\theta|x) = \frac{q(\theta) \cdot p(x|\theta)}{p(x)} = \frac{q(\theta) \cdot p(x|\theta)}{\int_{\theta \in \Theta} q(\theta) \cdot p(x|\theta) d\theta}. \quad (7)$$

From the above sections we see that the both the prior distribution and the likelihood function are directly derived from the real observed data: while the likelihood bridge the model performance in accordance to the realised LGD, the prior distribution could be purely data based.

It is important to understand that the conditional probability obtained as posterior is a full distribution curve based primarily on the observed data and the underlying model, and therefore this probability describes the chance of observing a LGD θ given the modelled LGD x , based on the internal experiences of model estimate and observation.

This is different from the conditional probability obtained directly from the observed data, i.e. $\tilde{q}(\theta^d|x^d)$, especially the periodical performance review data. The distribution of the observed data set is unlikely to provide complete information of the observed LGDs.

One may wonder if she could apply the same technique that is used to derive the prior distribution, such as MCMC, on the periodic model review data to generate the conditional probability $q(\theta|x)$ from the observed data directly. The constraint here is the condition on modelled LGD, as both realised and modelled LGDs follow different distributions, simulation over a condition with unknown distribution can be difficult and inaccurate.

Once the posterior is obtained, the expected LGD can then be calculated conditioning on the estimate in place.

$$x^* = E(\theta = x^*|x) = \int_{-\infty}^{\theta} \theta \cdot q(\theta|x) d\theta = \sum_{\theta} \theta \cdot q(\theta|x). \quad (8)$$

Now, recall the recently observed data set X^d . One can test and plot the distribution of the updated LGD estimates to visually illustrate the impact of posterior update on the observed data set.

Fig. 4 is the updated posterior distribution plotted in comparison with the modelled likelihood distribution, we still plot and demonstrate the distribution in the same conditional setting as for the likelihood to show the impact of posterior update, under the same bucketing scheme.

Table 2: Likelihood for LGD buckets.

Bucket name	θ_1	θ_2	θ_3	θ_4	θ_5
x_1	93.28%	20.34%	9.99%	4.59%	1.46%
x_2	6.79%	64.29%	11.09%	5.30%	1.84%
x_3	0.49%	9.46%	57.80%	8.32%	2.49%
x_4	0.01%	4.72%	11.09%	55.98%	5.03%
x_5	0.00%	1.17%	10.00%	25.79%	89.16%

Similar to Table 2, Table 3 summarizes the probability x^* for each bucket of observed LGD θ to obtain the post-observation updated distribution.

Browsing through Table 3 by columns, one may draw some initial findings from the table itself, for example, for observed LGD in bucket θ_1 , the likelihood of observing an in-bucket model estimate, x_1^* , is 92.41% which continues to support the expected concentration of observations near 0%, although this is lower than the likelihood of 93.28% according to the initial model specification. However, while for model estimates x_5^* in bucket θ_4 , the 51.13% likelihood indicates that: for all of the losses that resolves into bucket θ_4 , the model is most likely to produce an out-of-bucket over-estimate of x_5^* . Likewise, the likelihood of 37.60% in cell $[x_1^*, \theta_2]$, comparing to the likelihood of 34.00% in cell $[x_2^*, \theta_2]$ indicates that the model is slightly more likely to offer an estimate of x_1^* for losses later resolve in to bucket θ_2 , an under-estimate in this case.

6 UPDATED LGD AND CAPITAL ADEQUACY

Fig. 5 illustrates the prior, likelihood and posterior distribution of the LGD, it is clear that because of the model in place, the post-observation updated LGDs could be observed on different tendencies.

Before proceeding to apply the RWA formula using the updated estimate x^* , we'd like to stress that the aim of this paper is to quantify the impact of uncertainty instead of updated and replace the regulatory parameters used under approval from authority. It is recommended that the x^* reviewed internally for risk management purposes. If gap is found between the $RWA|x$ and the $RWA^*|x^*$, the risk mitigation should be considered as a buffer in addendum rather than a direct update of the regulatory reported RWA itself.

Table 3: Updated likelihood for LGD buckets.

Bucket name	θ_1	θ_2	θ_3	θ_4	θ_5
x_1^*	92.41%	37.60%	15.76%	6.78%	2.64%
x_2^*	6.79%	34.00%	12.36%	6.15%	2.63%
x_3^*	0.77%	18.83%	36.04%	11.86%	5.36%
x_4^*	0.01%	7.68%	14.27%	24.05%	6.72%
x_5^*	0.00%	1.86%	21.56%	51.13%	82.62%

Prior, Likelihood and Posterior distribution of the Loss Rate (θ)

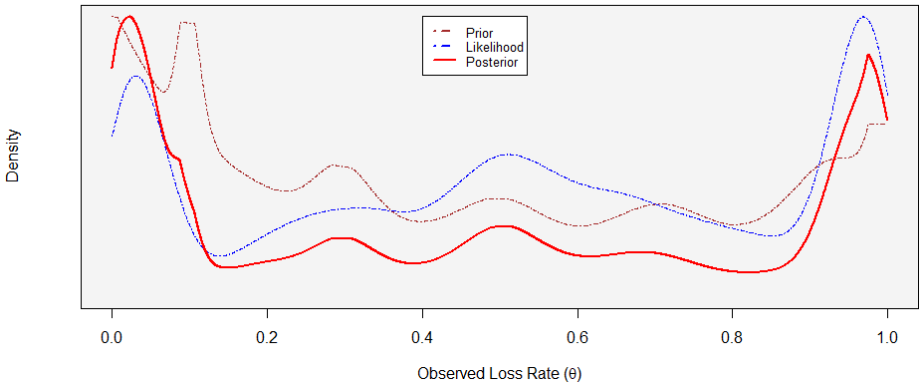


Figure 5: Overall OLR distribution.

The regulatory Tier 1 capital ratio compares the equity capital to the RWA, where the latter is calculated directly using the model estimated LGDs for advanced IRB banks. Therefore, the uncertainty in the underlying LGD estimate could result doubts in the bank's capital ratio, hence applying the updated LGD estimation to update the calculation of the RWA is a direct assessment of adequacy conditioning on the model in place.

Recall the simple expression of RWA from Section 1:

$$RWA = F(PD, LGD, EAD, m). \quad (9)$$

All forms of the updated RWAs relates to updated LGD can be mathematically expressed as:

$$RWA^{L^*} = F(LGD^*, \cdot). \quad (10)$$

The updated RWA can then be calculated by using the updated LGD, to help quantify and address the impact, in case uncertainty is found in the quality of modelled estimates.

7 CONCLUSION

In this paper, we start the discussion with a practical problem in the finance and banking industry where model performance is questioned and the impact in downstream calculations are difficult to assess. We then propose a Bayesian framework to estimate the conditional probability of observed LGD conditioning according to the model estimated LGD.

This approach allows the practitioner to timely assess current positions according to latest observed real life experience and understand the adequacy of the calculated RWA exposure-by-exposure. This allows assessment of impact of uncertainty from the model which results in the follow-up calculation, e.g. the RWA, quantitatively at penny level accuracy even with dynamic portfolio changes.

The Bayesian approach can be easily extended to review model performance, for example, we have presented a mixture double beta distribution based on our assumptions for the model in Section 4 for likelihood estimation, for review tests, this assumption could be an initial theoretical assumption that is different from the model. This way, if we define the theoretical likelihood as $p^t(x|\theta)$ and actual modelled likelihood as $p^m(x|\theta)$, the impact of assumptions can then be quantified easily by comparing the different likelihoods.

Similarly, the approach is applicable in case for model selection as well, assume that two models m_1 and m_2 , where the likelihoods and conditional posterior distribution is p^{m_1} , p^{m_2} , q^{m_1} and q^{m_2} respectively, one can easily test for the better performing model between the two candidates.

This approach is the first in literature specifically aiming to address the impact when uncertainty arose from a model and results further uncertainty in follow-up calculations, we demonstrate the detailed thought process and step-by-step demonstrate the procedures using an example in which uncertainty in the RWA resulted from an "ill" LGD estimate model.

Lastly, it is important to point out that aside from the widely accepted Bayesian Theorem, our approach also benefit from the fact that there is no additional data requirement aside from the historical and recently observed data, moreover, there no additional assumption or scenario change required on the underlying model.

ACKNOWLEDGEMENT

Special thanks to Qining Liu, loving wife and caring mother, who smiles sweetly at my swearing, on each failed attempt to solve the target problem.



REFERENCES

- [1] Viral, V., Bharath, S.T. & Srinivasan, A., Understanding the Recovery Rates on Defaulted Securities, Acharya, Online. <https://ideas.repec.org/p/cpr/ceprdp/4098.html>. Accessed on: 16 Jan. 2018.
- [2] Madan, D.B., Bakshi, G. & Zhang, F.X., Understanding the Role of Recovery in Default Risk Models: Empirical Comparisons and Implied Recovery Rates, Online. <https://ssrn.com/abstract=285940>. Accessed on: 31 Jan. 2018.
- [3] Bank for International Settlements, Basel II: International Convergence of Capital Measurement and Capital Standards, Online. www.bis.org/publ/bcbs128.htm. Accessed on: 31 Jan. 2018.
- [4] Bank for International Settlements, Basel III: Finalising Post-crisis Reforms, Online. www.bis.org/bcbs/publ/d424.htm. Accessed on: 31 Jan. 2018.
- [5] Altman, E.I., Brooks, B., Resti, A. & Sironi, A., The link between default and recovery rates: theory, empirical evidence, and implications. *The Journal of Business*, **78**(6), pp. 2203–2228, 2005.
- [6] Berger, J.O., *Statistical Decision Theory and Bayesian Analysis*, Springer Science+Business Media: New York, pp. 126–132, 1985.
- [7] Schuermann, T., What Do We Know about Loss Given Default, Online. <https://ssrn.com/abstract=525702>. Accessed on: 31 Jan. 2018.
- [8] Stone, C.A. & Zhu, X., *Bayesian Analysis of Item Response Theory Models Using SAS*, SAS Institute Inc.: Cary, NC, pp. 45–48, 2015.
- [9] Zhang, X., King, M.L. & Hyndman, R.J., A Bayesian approach to bandwidth selection for multivariate kernel density estimation. *Computational Statistics & Data Analysis*, **50**(11), pp. 3009–3031, 2006.
- [10] Efron, B., The bootstrap and Markov-Chain Monte Carlo. *Journal of Biopharmaceutical Statistics*, **21**(6), pp. 1052–1062, 2011.
- [11] Kieschnick, R. & McCullough, B.D., Regression analysis of variates observed on (0, 1): percentages, proportions, and fractions. *Statistical Modelling*, **3**(3), pp. 1–21, 2003.
- [12] Ferrari, S.L.P. & Cribari-Neta, F., Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, **31**, pp. 799–815, 2010.