A Questionnaire to Address Subjectivity in LGD Modeling

Chandrakant Maheshwari
Sanghamitra Mukherjee

Abstract

LGD is one of the three critical components to estimate expected loss and there are various subjective aspects of LGD estimation which should be taken care of. Many of these aspects cannot be captured easily and requires a deeper understanding about business, recovery process and underlying collaterals. From the practical modeling experience and studies we arrived at the conclusion that those subjective concepts are critical in influencing the behavior of loss, and thus crucial for LGD estimations but difficult to address just by analyzing the data in a statistical manner. This is because the Loss behavior is not only limited to the obligor's behavior or transaction history, it's also linked to the time frame between the occurrence default event and the period of recovery. So modeling the variability of LGD is complex than predicting the occurrence of default (PD) or estimating the exposure at the time of default (EAD).

In this paper we aim to put together all those issues that should be addressed during LGD model development. We presented it in a form of simple set of questions, addressing which will help to develop a robust model capturing the subjectivity.

Key Words:

Credit Risk, Loss Given Default, Subjective concepts, Recovery, Collateral

Introduction

With the advent technology and data availability more and more financial institutions are opting Advanced Internal Rating Based (A-IRB) approach. This enables them to determine their own estimates of Probability of Default (PD), Loss Given Default (LGD) and Exposure at Default (EAD).

For this approach risk model development team should have a clear understanding of the business requirements. They should have complete understanding of input data and application of output. Their methodologies should be accurate and validated by the validation teams.

In the case of Loss Given Default (LGD) this issue becomes critical. It is because LGD is dependent on recovery and data on recovery is quite fragmented and hence unreliable.

It has been observed that recovery rate (and hence LGD) of commercial obligors depends on

- The bank's relationship with their customers, this impacts terms of debt renegotiation with debtors, compromise and settlements which are country specific.
- Type of loans, its seniority in the firm's capital structure and quality of collateral attached
- Business tangibility: The value of liquidated assets dependent on the industry of the borrower.
- Macro-economic factors during the time of default:

There are other operational factors on which recovery is dependent and which are difficult to model:

- Financial Institution's understanding about the business of the obligor
- Recovery agent and relationship history of the obligor with the financial institution
- Negotiation of restructuring of loans
- Optimal combination of "time of resolution" and price received by selling collateral

- Current microeconomic factors which are impacting both the parties
- Indirect costs which include cost of maintaining the recovery staff etc.
- Country specific rules and regulations of recovery

Imagine two default cases of two similar shopkeepers (here their shop is collateral) on a same lane. Recovery of one shop may be better (in terms of time to resolution and recovered amount) just because more parking spaces are available for customers around it in comparison to other.

Data associated with such specific information is thin hence development of a robust model is a challenging task. It is never guaranteed that two deals with same collateral (also obligor, industry, macroeconomic situation) will result in same recovery. Also no complete variable list is available on which LGD depends. This leaves scope of subjective judgments in the right measurement and prediction of LGD.

Because of the subjectivity in LGD measurements, monitoring framework for LGD models need to be taken in a different paradigm. LGD Monitoring has to focus on portfolio dynamics, quality of the data and use of the model. One can monitor changes at the data level (population stability), loss of model performance, or calibration issues (ie, by comparing estimated against observed LGD at the pool level).

We provide a list of questions which analyst should be exploring so that he can logically document the assumptions he makes and articulate logical explanations why he is choosing certain real world characteristics in his modeling exercise and ignoring others. Finally this information will help him in setting up a robust model and model monitoring framework.

Business related questions

1. What is the industry segment of the portfolio for which LGD model has to be developed?
2. As per the business team what are the factors and scenarios under which they can face losses?
3. Who are the target customers and their ratings? What are their sizes?
4. What is the seniority of the loan?

5. Is the business directly lending the end customer or is a third party vendor? Who and at what proportion guarantees the loans? How reliable is the guarantor? What are the legal clauses of the loan agreement?
6. What is the credit worthiness of the Guarantor? What if the guarantor defaults?
7. Is there any risk sharing agreement with any business partner/dealer etc.?
8. What are the underwriting procedures/criteria?

Collateral related questions

1. What are the collateral's asset classes?
2. How liquid is the market of these assets?
3. What is the loan to value (of collateral) ratio followed for various products in the portfolio?
4. How efficient is the appraisal process of the collateral? How often this process is back tested? What is the frequency of the appraisal process?

Recovery related questions

1. What is the average recovery cycle for the products in the portfolio and explanation around whether there has been any policy change which has resulted in longer/shorter collection cycle with passage of time?
2. How bank is managing unsecured direct exposure? For secured loan, the recovery pattern primarily depends on underlying assets. But for unsecured loans the recovery primarily relies on guarantor.
3. If the loan is securitized using risk mitigation instrument such as insurance/bonds etc. how it's going to affect the recovery?
4. What is the default history of the portfolio? In some portfolios where there is no default in the history, the recovery mechanism might not have developed hence may have high LGD when there is a default.
5. What is the discount rate used to for the recovery cash flows? What is the rationale to use that rate?

6. How is the business understanding of the sales team helps in the recovery in case of default? How is this understanding reflected in the recovery process?
7. What are types of default? This is important because
    1. For payment defaults the recovery process depends on the bank's internal rule
    2. If the default is due to bankruptcy, a legal procedure is involved there in following the bankruptcy filed by the obligor. Therefore, recovery may take longer time given the legal procedures involved.
    3. In case of debt restructuring it is more or less in control of bank to initiate the recovery process post default event.
8. Whether the obligor Business is Single Ownership/Partnership/Public Company
    1. If the business is single ownership/partnership the recovery primarily depends on the financial health of the sole owner/co-owner of the business.
    2. If it is public company the recovery is not solely dependent on the financial health of co-owners, rather more on shares/bonds and other company assets.

Data Quality related issue/questions

1. How can the available modeling data can be triangulated so that its correctness is verified?
2. How frequently the database is updated?
3. What is the distribution of historical LGD of whole portfolio and segment wise?
4. At what level default data is captured in the system? As per Basel II requirement the data should be at transaction level but there are certain scenarios where cash flow is not calculated at transaction level rather it is calculated at obligor level. In such cases transaction level LGD might not be available, therefore, either model should be developed at obligor level or there is a need to define a methodology to break-down obligor level LGD at transaction level.
5. Among the total default cases how many are still open? LGD modeling is done on closed default cases because they are the ones which have definite LGD. But if majority of the total default cases are still open, this implies we might be missing important insights in the data.
6. What are the default scenarios that have already started but not yet been resolved?

7. How are TDR (Troubled Debt Restructuring) loan/customers treated? Are the included in the LGD modeling? Is a loss recorded in these transactions? How is it measured?
8. Is periodic watch on the ratings (or CDS) of the obligor maintained? It is important because they become volatile just before default.
9. Are there any portfolio specific external/internal variables which act as good predictors of LGD? What is the proof that this behavior is genuine?
10. Outlier/Exclusion/Noise: What data should be used for modeling purpose is a major question.
    1. The final economic LGD is calculated after a certain time period when it is assumed that no more recovery is expected from the transaction. But an interim economic LGD is also calculated for the transactions for which the recovery process is yet not complete. The modeler needs to take the call whether these transactions should be included in the modeling process depending on at which proportion of such open transactions are present in the data, whether the interim LGD is in line with portfolio LGD etc.
    2. Outlier: The LGD model generally produces the predicted LGD ranging between 0%-100%. But in reality the observed economic LGD can be <0% or >100%. The modeler needs to take a call whether model should be developed on the observed LGD as it is or LGD should be floored and capped at 0% and 100% respectively.
11. Some of the data might be impacting the whole portfolio LGD because of unexpected recovery costs, so they have to be excluded. What business processes established to ensure that such cases may not happen in future? If none, then these data points should be included such that they do not impact the whole model and also their impact is captured.

Estimation technique related questions

LGD can be estimated using various statistical techniques and there is no thumb rule on the best technique. The selection of modeling technique largely depends on the below questions:

1. Whether the data is sufficient to develop a statistically robust model? If the underlying portfolio is a small one with large exposure, it is often difficult to develop a robust statistical estimation.
2. What is the distribution of observed LGD? The distribution may follow a uniform pattern, bi-modal or uni-modal distribution. Estimation process is largely determined by the distribution.
3. Whether to apply a single step model or a multi-step model? In single step model, the LGD is estimated directly from one model equation. A very common practice for single step model is decision tree, linear regression, multinomial logit or beta regression. But in multi-step model, the estimated LGD is obtained from a combination of model. This is very typical practice if bimodal or unimodal distribution is observed. The model can be developed through combination of logistic and linear approach, ordered logit and linear etc.
4. How to select the appropriate weight during modeling? One can choose EAD as the weight. It captures the variation in exposure influencing LGD. But if any outlier is present, the model will be impacted. Generally, the EAD weighted approach can be used for small ticket size portfolios. Alternative option is count weighted model. This approach entirely relies upon the frequency distribution of the observed LGD. Since no importance is given to EAD, LGD of a transaction with lower EAD can be higher which may prove to be costly for small ticket size portfolio. But this can be an option for large ticket portfolio where EAD is not too diverse.
5. What is the incremental benefit of the multi-step model as the implementation of a complicated model costs more? If the model performance is improved marginally by using a multi-step model, it is often discouraged to implement the complex one. So unless there is a noticeable improvement in the model performance, one can propose a simpler solution.
6. How well macro- economic factors are able to explain the estimation of the LGD along with other explanatory variables? Macro-economic parameter value is same for all transactions for a given time point, so how it can discriminate the LGD pattern of one transaction to other?
7. How to use obligor level explanatory variables as they do not discriminate at transaction level, although may be very critical to determine the LGD pattern?

8. What type of statistical transformation should be applied to explanatory variables? How to explain the relationship between the transformed variable and observed LGD?

Bench marking related questions

1. How the LGDs are benchmarked? What are the benchmarking criteria?

1. Benchmarking is a challenging task, because it should ensure right seniority, rating of the debt, portfolio nature

1. Is there any industry study about LGD or related concept about similar portfolio?
2. What are the LGD numbers of similar portfolio around the other businesses in the organization?

Outcome analysis level questions

1. LGDs are used for underwriting, Allowance of Loan and Lease Losses (ALLL), Capital allocation, stress testing. Is the data used for modeling and output conceptually compliant with these requirements? It is important to note that
    1. For Economic Capital, LGD has to be a function
    2. For Stress testing, stressed LGD should take into account the change in relationship between LGD numbers and independent variables which it is dependent. This relationship will not be same in stressed times.
    3. For underwriting and ALLL modeling we need expected LGD if an obligor defaults.
    4. LGD estimate for an obligor which has just defaulted will be a separate number.

1. Whether the model is using data from at least one downturn cycle

1. As per Basel II requirement the LGD model should be developed using 5- 7 years of data so that a full business cycle is covered including one downturn period. If downturn cycle is not used in modeling, there is a possibility that the model may under predict the loss when model is applied during economic distress.

3. What are the key metrics to judge the model performance and it's threshold level of model acceptance?
4. What will be the impact on the live portfolio if the model is implemented?
5. What will be the impact on the output LGD on the output of other models which are going to use these LGDs as an input?

Industry study on LGD behavior

As discussed LGD modeling is majorly about business understanding, for completeness and holistic overview we would like to include some industry wide study and opinion about LGD and recoveries. We suggest that all the findings below should be taken as pinch of salt because we have observed that the results of one research team at times are refuted by other research team. We feel that the results of these studies are extremely data dependent, but we would still like to share the perspectives.

Business understanding

1. Using data on observed prices of defaulted securities in the United States over the period 1982-1999, Acharya, Bharath and Srinivasan (2007) find that seniority and security are important determinants of recovery rates. While this result is not surprising and in line with previous empirical studies on recoveries, their second main result is rather striking and concerns the effect of industry-specific and macroeconomic conditions in the default year. Indeed, industry conditions at the time of default are found to be robust and important determinants of recovery rates. They show that creditors of defaulted firms recover significantly lower amounts in present-value terms when the industry of defaulted firms is in distress and also when non-defaulted firms are rather illiquid and if their debt is collateralized by specific assets that are not easily redeployable into other sectors. Also, they find that there is little effect of macroeconomic conditions over and above the industry conditions and the latter is robust even with the inclusion of macroeconomic factors.
2. Amihud, Garbade and Kahan (2000) point out that "private loans better control the agency costs of debt through tighter covenants, renegotiation, and closer monitoring" (p. 116). Bankers are able to exploit their lending relationship to firm up their position at top of capital structure in anticipation of bankruptcy thereby raising

expected recovery. The more fluid and dispersed nature of bond ownership makes it impractical for bondholders to renegotiate the core terms and conditions of the bond contract as the firm's condition changes.16 Bankers are not so constrained

3. Asarnow and Edwards (1995) look at Citibank's middle market and large corporate lending from 1970–93 and find no relation between loss given default and size of loan. Carty and Lieberman (1996), using Moody's data on syndicated lending, arrive at a similar negative result. Thornburn (2000), in her study of Swedish small business bankruptcies, also found that firm size doesn't matter in determining LGD. Eales and Bosworth (1998) look at Australian small business and larger consumer loans such as home loans and investment property loans and conclude that size does matter, at least a little. They report an average severity of 30% with a median of 20% (their distribution too is bimodal). Interestingly they find that loss recovery is U-shaped29 with the trough of around A$100-500k. They note that business bankruptcy almost always results in higher severity than consumer bankruptcies.
4. Duffie, Eckner, Horel, and Saita (2009), Tang and Yan (2010), and Qi, Zhang, and Zhao (2009), find that defaults are mainly driven by firm-level risk factors. Together, these findings suggest that the joint distribution of default and recovery is more likely due to idiosyncratic risk than to systematic risk.

Independent Variables

1. Böttger et al. (2008) conclude that corporate debt is mainly driven by the six factors: seniority, securitization, jurisdiction, industry, economic cycle, and expected liquidity of the secondary market for the debt type.
2. Qui/Yang (2009) analyze residential mortgage loans that had insurance protection by one of the six private mortgage insurance companies operating in the US and that were settled between 1990 and 2003. Since a complete housing market cycle is covered, Qui/Yang (2009) claim to be the first paper that empirically models an economic downturn LGD for residential mortgages. They find that current LTV is the single most important determinant of LGD and also a much better predictor for LGD than original LTV.
3. As per Acharya, Bharath, and Srinivasan [2007]; Altman and Kishore [1996], a utility industry dummy variable is also included because the utility industry has a significantly higher recovery rate than other industries in our sample.

4. There is strong evidence that recoveries in recessions are lower, often much lower, than during expansions. Altman, Brady, Resti and Sironi (2003) show that when aggregate default rates are high, recovery rates are low.

Modeling

1. The forecasting of LGD for retail credit using macroeconomic variables is relatively unexplored by literature. The growth rate of GDP is significant in calculating the loss rate for Altman et al. (2005) on US bonds and for Figlewski et al. (2007) also on US bonds. The same variable is not significant for Bruche and Gonzalez-Aguado (2008) ´ and Acharya et al. (2007). The results agree on the relevance of the unemployment rate to explain the LGD (Acharya et al., 2007; Bellotti and Crook, 2012; Bruche and Gonzalez-Aguado, 2008, Caselli et al., 2008). Other macroeconomic covariates chosen in the literature to predict the recovery rates are the interest rate (Bellotti and Crook, 2012; Figlewski et al., 2007), stock market return (Acharya et al., 2007; Figlewski et al., 2007), investment growth (Bruche and Gonzalez-Aguado, 2008; Caselli et al., 2008) and inflation (Figlewski et al., 2007).

BENCHMARKING

1. Rauh and Sufi (2010) find that different types of firms tend to have different debt structure— firms with high credit quality mainly rely on senior unsecured debt, whereas firms with lower credit quality usually use multiple tiers of debt, including bank loans, secured, senior unsecured, and subordinated debts. As a result, unsecured debts from different firms may not be directly comparable, and recovery from unsecured debts may not necessarily be lower than secured debts.

Concluding thoughts

LGD is one concept in risk management which is dependent not only on objective variables but on subjectivities because recovery varies from business to business, firm to firm, time to time etc. There is no end to this list.

Due to this variation in recoveries we find most varieties in modeling prepositions in LGD modeling. On a given data set one analyst may propose a simple linear regression, another may propose Decision trees, and others may propose Neural Networks, genetic algorithm or any other exotic fuzzy logic based models. Interestingly these analysts may not agree on the choice of variables as well.

In fact LGD modeling is all about business and data understanding. This understanding is important for the model developer to develop a right framework of model and model monitoring.

To develop that understanding he should be prepared to ask questions to gather information so that he can develop right monitoring framework to evaluate an existing model. We provided some list of questions which can help him to clarity around the business requirements and hence develop right risk solution.

We tried to keep our questions as unbiased as possible by refraining any questions related to model. Any model is basically an opinion of an analyst which is dependent in his perception about data. Any such question would have directed the whole questionnaire towards an un-warranted analytical direction.

References

1. Acharya, V. V., S. T. Bharath, and A. Srinivasan. 2003. Understanding the Recovery Rates on Defaulted Securities. Working Paper, London Business School. ———. 2007. "Does Industry-wide Distress Affect Defaulted Firms? Evidence from Creditor Recoveries." Journal of Financial Economics 85, no. 3:787–821. http://jairajgupta.weebly.com/uploads/6/3/7/0/6370607/crm_2.pdf
2. Amihud, Yakov, Kenneth Garbade and Marcel Kahan (2000), "An Institutional Innovation to Reduce Agency Costs of Public Corporate Bonds", Journal of Applied Corporate Finance, Spring, 114-121. http://fic.wharton.upenn.edu/fic/papers/04/0401.pdf
3. Asarnow, Elliot and David Edwards, (1995), "Measuring Loss on Defaulted Bank Loans: A 24- Year Study", Journal of Commercial Lending, 77(7), 11-23

4. Carty, L. and D. Lieberman (1996), "Defaulted Bank Loan Recoveries", Moody's Special Comment, November.
5. Thorburn, K. (2000), "Bankruptcy auctions: costs, debt recovery and firm survival", Journal of Financial Economics, 58, 337-368.
6. Eales, R. and E. Bosworth (1998), "Severity of Loss in the Event of Default in Small Business and Larger Consumer Loans", Journal of Lending & Credit Risk Management, May, 58-65.
7. Duffie, Darrell Andreas Eckner, Guillaume Horel, and Leandro Saita. "Frailty Correlated Default." Journal of Finance 64(5) (2009): 2089–123.
8. Qi, Min, Xiaofei Zhang, and Xinlei Zhao. "Unobservable Systematic Risk Factor and Default Prediction." Working paper, Office of the Comptroller of the Currency (2009). http://www.occ.gov/publications/publications-by-type/occ-working-papers/2013-2009/wp2011-2.pdf
9. Tang, Dragon and Hong Yan. "Market Conditions, Default Risk and Credit Spreads." Journal of Banking and Finance 34 (2010): 743–753
10. Bellotti, T. and J. Crook (2011). Loss given default models incorporating macroeconomic variables for credit cards. International Journal of Forecasting (forthcoming) https://statistik.econ.kit.edu/download/Michigan2011.pdf
11. Qi, M. and X. Yang (2009, May). Loss given default of high loan-to-value residential mortgages. Journal of Banking and Finance Vol. 33(5), 788–799.
12. Acharya, Viral.V., Sreedhar T. Bharath, and Anand Srinivasan. "Does Industry-Wide Distress Affect Defaulted Loans? Evidence from Creditor Recoveries." Journal of Financial Economics 85 (2007): 787–821. http://www.occ.gov/publications/publications-by-type/occ-working-papers/2013-2009/wp2011-2.pdf
13. Altman, Edward I. and Vellore Kishore. "Almost Everything You Wanted to Know About Recoveries on Defaulted Bonds." Financial Analyst Journal November/December (1996): 57-64.
14. Altman, E., B. Brady, A. Resti and A. Sironi (2003), "The Link between Default and Recovery Rates: Theory, Empirical Evidence and Implications," forthcoming, Journal of Business http://fic.wharton.upenn.edu/fic/papers/04/0401.pdf

15. Altman E. I., Brady B., Resti A., Sironi A.: The link between default and recovery rates. Theory, empirical evidence and inplications. Journal of Business 78, 2203–2228 (2005)
16. Figlewski S., Frydman H., Liang W.: Modeling the Effect of Macroeconomic Factors pn Corporate Default and Credit Rating Transitions. NYU Stern Finance Working Paper, November (2007)
17. Acharya Viral V., Bharath Sreedhar T., Srinivasan A.: Does industry-wide distress affect defaulted firms? Evodence from creditor recoveries. Journal of Financial Economics. 85(3), 787–821 (2007)
18. Bellotti T. and Crook J.: Loss Given Default models for UK retail credit cards. Credit Research Centre, working paper (2009)
19. Bruche, M., Gonzalez-Aguado C.: Recovery Rates, Default Probabilities and the Credit Cy- ´ cle. CEMFI, working paper (2008)
20. Caselli, S., Gatti, S. Querci, F.: The sensitivity of the loss given default rate to systematic risk: New empirical evidence on bank loans. Journal of Financial Services Research. 34, 1– 34 (2008) https://boa.unimib.it/retrieve/handle/10281/20170/25613/Predicting_Bank_Loan_Recovery_Rates_in_a_mixed_continuous.pdf
21. **http://businessperspectives.org/journals_free/imfi/2013/imfi_en_2013_04_Hochstotter.pdf**
22. **https://www.capitaliq.com/media/162422-LGD_%20Decision_Tree_Case_Study.pdf**