

Estimation technique for deriving the Basel LGD estimate on a retail bank mortgage portfolio

Credit Scoring and Credit Control XV conference, 2017,
University of Edinburgh, Edinburgh, Scotland.

Morne Joubert

Helgard Raubenheimer & Tanja Verster
Centre for Business Mathematics and Information

Indirect LGD model development

- LGD is calculated on accounts that are in default where default is assigned according to the Basel default definition.
- The loss given default is calculated as $LGD_{i,d} = \frac{E(L_{i,d}/default)}{EAD_{i,0}}$, where d is the number of months account i is in default. The expected loss amount for a defaulted account is $E(L_{i,d}/default)$ and the exposure at the default is $EAD_{i,0}$.
- The expected value of the loss amount can be written as the sum product of the loss amount components and the probability components and loss given default as

$$LGD_{i,d} = \frac{E(L_{i,d}/default)}{EAD_{i,0}} = \frac{L_{i,d}|W \times P_{i,d}(W) + L_{i,d}|C \times P_{i,d}(C) + L_{i,d}|IC \times P_{i,d}(IC)}{EAD_{i,0}}.$$

- The loss severity component and the probability components are modelled separately, this is known as the indirect approach.



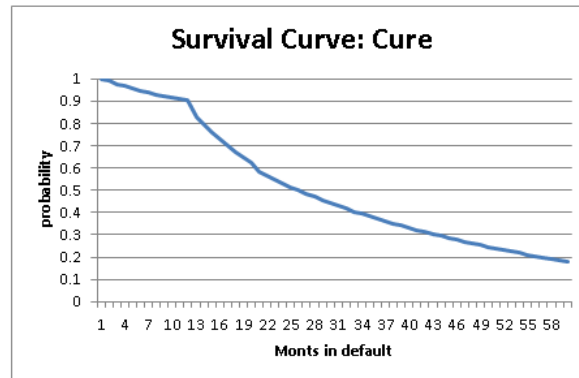
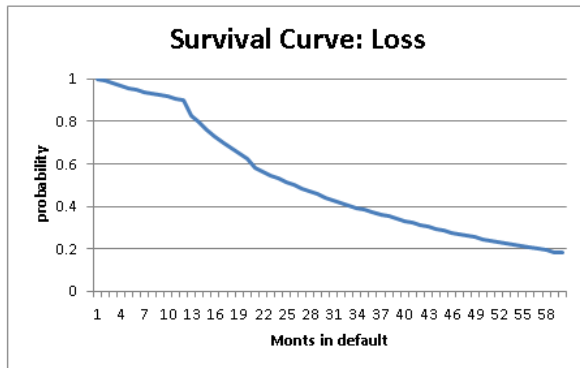
Indirect LGD model development: Probability model

- The aim of the probability model is to estimate $P_{i,d}(W_e)$, $P_{i,d}(C_e)$ and $P_{i,d}(IC_e)$.
- Leow & Mues (2012:193) made use of logistic regression to model probabilities. Survival analysis instead of logistic regression will be described and used to predict the probability component.
- A Survival curve $S(d)$, is defined as the probability that an account that is in default at time d remains in default until the end of the workout period, T .
- An account can exit the default state by either writing-off, curing or by remaining incomplete.
 - **Cure** (eg.. by paying off outstanding arrears and returning to the non-default book) or
 - **Write-off**
 - Remain **incomplete** at the end of the workout period (did not exit the default state after a set period of time i.e. cure or loss event did not occur).



Indirect LGD model development: Probability model

- The two survival curves $S^w(d)$ and $S^c(d)$ above can be estimated for the entire population or can be estimated at an individual level. The cox proportional hazards model is used to model these survival curves for individuals.
- The Kaplan–Meier estimate is the empirical survival curve estimated from the data.
- Kaplan Meier estimator of the survival curve and is represented graphically:

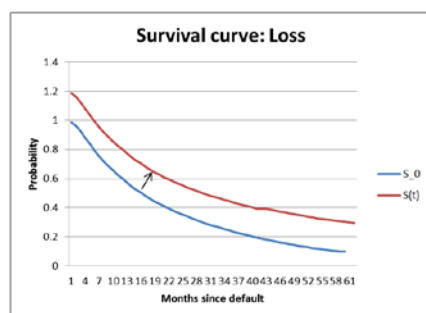
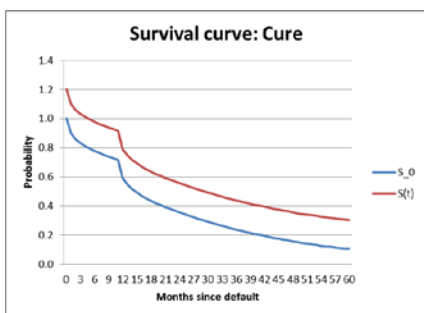


Indirect LGD model development: Probability model

- The cox proportional hazards model is used to model these survival curves for individuals.
- The general form of the Cox model can be written as,

$$S(d) = [S_0(d)]e^{\sum \beta_i X_i}$$

- $S_0(d)$ is called the Baseline survival function whilst the second of these is the exponential expression to the linear sum of the covariates.
- Dummy variables are produced for each of the covariates, X_i , where i is the number of covariates in the model.
- The baseline survival curve, $S_0(d)$, is estimate by selecting the population where all the dummy variables are equal to their baseline groups and calculating the Kaplan–Meier estimate for this population.
- The Kaplan–Meier estimate is the empirical survival curve estimated from the data.
- If the values of an individual's covariates value falls outside the dummy variables baseline group, the baseline survival curve, $S_0(d)$, will be adjusted, $[S_0(d)]e^{\sum \beta_i 1}$, to yield a survival curve, $S(d)$, that is the Kaplan–Meier estimate for the population equal to the individuals new covariate values.



Variable	Estimate	Variable	Estimate
C_DC_SEG_TRADE_PRICE_C		L_DC_REGION_C	
[0,80)	0	Personal Sales Dealers and FORD FINANCIAL SERVICE	-0.1346
(80,100)	-0.5201	Personal Sales Internal and SMALL	-0.2391
>=100 and Missing	-1.009	CENTRALISED SALES, Risk Management and Others	0
C_DC_BKF_INST_METHOD_CODE_C		L_DC_BKF_INST_METHOD_CODE_C	
DEBIT ORDER	0.3629	DEBIT ORDER	-0.2724
CASH	0	CASH	0
Others	0.5621	Others	-1.0246
C_DC_MANUFACTURER_CODE_C		L_DC_MANUFACTURER_CODE_C	
TOYOTA, ISUZU, OPEL, VOLKSWAGEN, SUZUKI, HYUNDAI and FIAT	0	CHEVROLET, PEUGEOT, TOYOTA, HYUNDAI, AUDI, TATA, B M W, HONDA, MERCEDES-BENZ, ALFA ROMEO, OPEL, MITSUBISHI, RENAULT, VOLKSWAGEN, CHRYSLER, CITROEN, SUZUKI	0
CHEVROLET and FORD	-0.0609	HAFEI, GONOW, CMC, CHANA, C.A.M, IVECO, DAEWOOD and Others	0.0872
LAND ROVER, MERCEDES-BENZ, RENAULT, KIA, CITROEN, PEUGEOT, CHRYSLER, VOLVO, SUBARU, DODGE, JAGUAR, MINI, LEXUS and PORSCHE	-0.2151	LEXUS, JAGUAR, SUBARU, MINI, PORSCHE, KIA, FORD, JEEP, LAND ROVER, MAZDA, FIAT, ISUZU, VOLVO, NISSAN, GWM	0.11
GEELY, DAIHATSU, TATA, JEEP, GWM, SAHRA, MAHINDRA, CHANA, C.A.M, CMC, 0, GOLDEN JOURNEY, JINBEI	-0.3882		
Others	0.1243		



Centre for BMI
Senteramo ya DII
Sentrum vir BWI



NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
POTCHEFSTROOMKAMPUS

Indirect LGD model development: Probability model

- The two survival curves $S^w(d)$ and $S^c(d)$ is combined by applying the cumulative incidence function to produce the three probabilities $P_{i,d}(W_e)$, $P_{i,d}(C_e)$ and $P_{i,d}(IC_e)$.
- Rodriguez (2012:2) states that the cumulative incidence for a failure of type j and distinct failure times $0 < d_{(1)} < \dots < d_{(T)} <$ is estimated as

$$\hat{I}_j(d) = \sum_{i:d_{(i)} \leq d} \hat{S}^j(d_{(i)}) \frac{m_{ij}}{n_i}$$

where m_{ij} is the number of events of type j at time $d_{(i)}$, n_i is the total number of observations at risk at time $d_{(i)}$, $\hat{S}(d_{(i)})$ is the Kaplan–Meier estimator of survival to time $d_{(i)}$. The Kaplan–Meier estimate is empirical value of the survival curve calculated from the data. The sum of all cumulative incidences has the feature that it equals the complement of the overall Kaplan–Meier estimate of survival considering failures of any kind,

$$\sum_j \hat{I}_j(d) = 1 - \hat{S}(d).$$

The cumulative incidence function is applied to $S^w(d)$ and $S^c(d)$ and the sum taken to produce the three probabilities $P_{i,d}(W_e)$, $P_{i,d}(C_e)$ and $P_{i,d}(IC_e)$.



Indirect LGD model development: Probability model

- A loss will be incurred when the Haircut, \hat{h}_d , exceeds the loan to value at default point, LTV_0

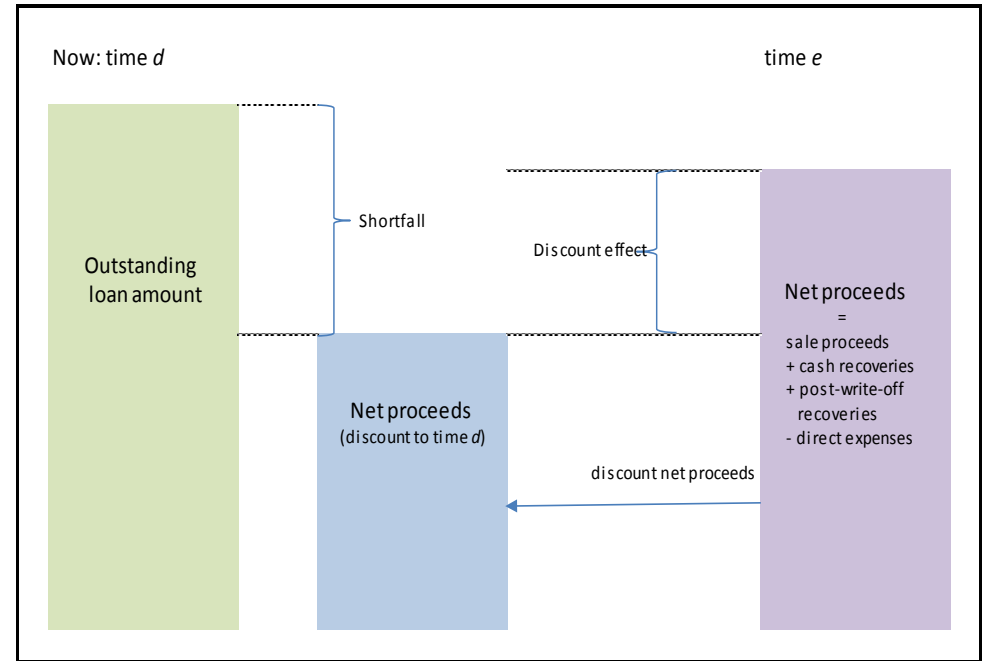
$$LTV_0 = \frac{\text{outstanding loan amount}_0}{V_0}$$

$$\hat{h}_d = \frac{\text{net proceeds}}{\text{valuation of the property}}$$

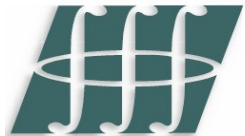
- Each cashflow component of the net proceeds calculation is summarised in the schematic below:

$$\begin{aligned} L_{i,d}|W &= E(\text{shortfall percentage}|W) \times V_d \\ &= E\left(\frac{\text{outstanding loan amount}_0}{V_0} - \frac{P_d}{V_d} |W\right) \times V_d \\ &= E(LTV_0 - \hat{h}_d |W) \times V_d \\ &= \int_{-\infty}^{DLTV} p(h)(LTV_0 - \hat{h}_d) dh \end{aligned}$$

where $p(\cdot)$ denotes the probability density function of the distribution for h .



Questions?



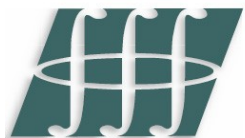
Centre for BMI
Senteramo ya DII
Sentrum vir BWI



NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
POTCHEFSTROOMKAMPUS

References

- Leow, M. & Mues, C. (2012). Predicting loss given default (LGD) for residential mortgage loans: A two-stage model and empirical evidence for UK bank data. *International Journal of forecasting* (2012) 28, 183-195.
- Rodriguez, G. (2007). *Generalized Linear Models*. Princeton university.



Centre for BMI
Senteramo ya DII
Sentrum vir BWI



NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
POTCHEFSTROOMKAMPUS