Estimating loss given default based on time of default

Jamil J. Jaber*
The University of Jordan
Department of Risk Management and Insurance
Jordan
Aqaba Branch
and
Department of Risk Management & Insurance
The University of Jordan
Jordan
j.jaber@ju.edu.jo
Jameljaber2011@hotmail.com

Noriszura Ismail
National University of Malaysia
School of Mathematical Sciences
Malaysia

Siti Norafidah Mohd Ramli
National University of Malaysia
School of Mathematical Sciences
Malaysia

Baker Albadareen
National University of Malaysia
School of Mathematical Sciences
Malaysia

Abstract. The Basel II capital structure requires a minimum capital to cover the exposures of credit, market, and operational risks in banks. The Basel Committee gives three methodologies to estimate the required capital; standardized approach, Internal Ratings-Based (IRB) approach, and Advanced IRB approach. The IRB approach is typically favored contrasted with the standard approach because of its higher accuracy and lower capital charges. The loss given default (LGD) is a key parameter in credit risk management. The models are fit to a sample data of credit portfolio obtained from a bank in Jordan for the period of January 2010 until December 2014. The best parametric models are selected using several goodness-of-fit criteria. The results show that LGD fitted with Gamma distribution. After that, the financial variables as a covariate that affect on two parameters in Gamma regression will find.

Keywords: Credit risk, LGD, Survival model, Gamma regression

*. Corresponding author

1. Introduction

Survival analysis is a statistical method whose outcome variable of interest is the time to the occurrence of an event which is often referred to as failure time, survival time, or event time. Survival data can be divided into three categories; complete, censored and truncated. Complete data is the ideal data that contains the beginning and end dates of which the event time is determined. Censored and truncated data is also called missing data due to the unavailability of information on the beginning or end dates (Klugman et al. 2012). Right censored data can be divided into three types: type I, type II and type III. Type I and type II are also called the singly censored data, while type III is also called the progressively censored data (Cohen 1965). Another commonly used name for the type III censoring is random censoring.

The credit risk is a crucial issue for financial institutions as well as for non-financial companies. Banks are allowed to estimate their own credit risk parameters dependent on the Basel II regulation and under the internal rate approach (IRB) and thus more accurate for the capital requirement with the underlying risk in the credit portfolio. In the past, Jordanian banks had to use standardized approaches to estimate minimum capital requirement under Basel I. Indeed, Basel II leads to a better differentiation of risks and takes into account the diversification of the bank's portfolio to estimate minimum capital requirement for credit risk (Kollar & Kliestik 2014).

In the related literature, there is numerous growths in various academic research and publications focused on credit risk models namely; probability default (PD) and loss given default (LGD) that can be formulated by using survival analysis techniques. It was initiated with the research paper by (Altman 1989) who used actuarial analysis to investigate the mortality rates of US corporate bonds. This was followed by various empirical studies on PD and LGD. Narain in 1992 defined the PD as the complement of the conditional survival function evaluated at the forecast horizon. Later, this technique was developed by (Carling et al. 1998). They used a semi-parametric duration model (Cox's proportional hazards model) with a data set consisting of 4733 individuals who were granted credit by a Swedish lending institution between 1993 and1995 to analyze the factors that determine the time to maturity on a loan and to evaluate loan applicants by their expected duration and profits. In addition,(Stepanova & Thomas 2002) and (Malik & Thomas 2010) modeled the lifetime of individual credit with Cox's proportional hazards model. Furthermore,(Glennon & Nigro 2005) measured the default risk of small business loans by survival analysis techniques with Cox's proportional hazards models. They found that the default behavior of the loans is time sensitive; the likelihood of default increases initially, peaks in the second year, and declines thereafter. However, (Beran & Djaïdja 2007) proposed statistical modelling of credit risk for retail clients based on survival analysis under extreme censoring for the time-to-default variable. (Cao et al. 2009) proposed three different mechanisms to estimate PD by using survival analysis techniques.

However, the survival analysis for predicting LGD has been presented by many researchers. (Dermine & De Carvalho 2006) used mortality approach to measure the percentage of the bad and doubtful loan of corporate bonds that are recovered n-months after the default date. The actuarial-based mortality approach is appropriate because the population sample changes over time. The dataset of this study obtained from micro-data on defaulted bank loans of a private bank in Portugal, Banco Commercial Portugues (BCP). It consists of 10000 short –term loans granted to small and medium-sized companies over the period from June 1995 to December 2000 (66 months). They identify the LGD by the following: $\mathrm{LGD} = \prod_{\mathrm{t}=1}^{\mathrm{T}} \mathrm{SPULB_t}$, Where, $\mathrm{SPULB_t}$ is a sample (weighted) percentage unpaid loan balance at period t, $\mathrm{SPULB_t} = 1-\mathrm{SMRR_t}$. Furthermore, $\mathrm{SMRR_t}$ is a sample (weighted) marginal recovery rate at time t,

$$\mathrm{SMRR_t} = \frac{\sum_{i=1}^{\mathrm{m}} \text{Cashflow received}_{\mathrm{it}}}{\sum_{i=1}^{\mathrm{m}} \text{Loan outstanding}_{\mathrm{it}}},$$

i stands for each of the m loan balances outstanding in the sample, t periods after default. The empirical results show that marginal recovery rate is high in the first five months and heavy tail in the last months. Also, the cumulative recovery rate n-months after default increase gradually from 20% to 70% over the time. Furthermore,(Chen 2018) used the right censored survival techniques for predicting bank loan LGD. The data contains 2644 defaulted loan over eight years (quarterly) with start dates between January 2007 and December 2014 for real estate and non-real estate for a U.S financial services company. The datasets tested by five parametric models namely, gamma, lognormal, Weibull, exponential, and log-logistic. In the figure we can see the LGD is high in the first months then decreases over time to become heavy tail in the last months. Moreover, (Tanoue et al. 2017) study forecasts loss given default of bank loans with multi-stage model. And, (Krüger & Rösch 2017) estimated downturn LGD modeling using quantile regression. Finally, (Thompson & Brandenburger 2019) constructed a model to predict the risk of a cardholder for the lifetime of the account and the survival analysis methodologies applied to a case study from capital card services.

The empirical studies on credit risk have depended mostly on the corporate bond market to gauge losses in the case of default. The purpose behind this is that, as bank loans are private instruments, little information on loan losses are freely accessible. The researchers use parametric, non-parametric, semi-parametric and transformation regression models to estimate LGD. In the other words, the ordinal least squares regression (OLS), Ridge regression (RiR), Fractional response regression, Tobit model, Decision trees model, Beta distribution and Normal distribution for parametric models. However, regression tree, neural networks, Multivariate adaptive regression spline, Least squares support vector machine for non-parametric models. In the semi-parametric models applied the joint beta additive model. However, inverse Gaussian regression, inverse Gaussian regression with beta transformation, Box-Cox transformation, Beta transformation, Fractional logit transformation, and Log transform for Transformation regression see for instance, (Bellotti & Crook 2008, 2012; Bruche & González-Aguado 2010; Calabrese 2012; Giese 2005; Huang & Oosterlee 2011; Loterman et al. 2012; Pykhtin 2003; Qi & Zhao 2011).

Theoretically, this study is of crucial importance for several reasons. First, this study used hybrid model for estimating LGD with new data which get from corporate credit portfolio from Jordanian bank. Second, it uses the actuarial analysis of the progressive right censored data for estimating the loss given default S (t) and recovery rate F(t) of a sample of corporate loans obtained from a bank in Jordan. Third, this study are to fit LGD data to several parametric distributions related with time such as Gama distribution, in order to estimate LGD each month which reduce capital requirement to meet default obligation each month. Finally, the financial variables which significantly affects on two parameters for Gamma regression can be explored. Indeed, the LGD is fixed under Basel I, but it changes over time (monthly) based on Gamma distribution under Basel II.

The primary explanations behind using the progressive censoring are: the time of study is constant, the borrowers can join the study any time during the fixed period of investigation, and the borrowers could default or not before the end of the study. Practically, the objective of this paper is to estimate the Loss Given Default (LGD) using parametric models based on survival time. The used parametric models are fitted to the right-censored data obtained from a bank in Jordan for the period from January 2010 until December 2014. The portfolio capacity is 4393, and the overall number of defaults during the 5-year period is 495. The sample size is same as number of default. A borrower is declared as default if he is unable to pay his cash installment in a period of 3 months. The estimated LGD is then used for predicting the performance of credit risk of a corporate portfolio under Basel II. For estimating LGD default used this proxy:

$$LGD \text{ for each borrower} = \frac{\text{outstanding amount}}{\text{Amount of borrowing}}$$

But with the effect of survival time, the used hybrid LGD monthly approach is:

$$\text{LGD for each month}=\frac{\sum_{i=1}^{n} \text{LGD for each borroweri}}{n} \tag{1.1}$$

where i represents the number of borrowers in the same month. For more details to calculate LGD, see the following example a company borrows 1 million dollars for 5 years with interest rate about 15 % yearly and payments monthly see Table 1.

Table 1. Example for calculate LGD

Months	**1**	**2**	**3**	—	**58**	**59**	**60**
Beginning Amount	1000000	988710	977279	—	69616	46696	23490
Interest	12500	12359	12216		870	584	294
Cash paid	(23790)	(23790)	(23790)	—	(23790)	(23790)	(23790)
Outstand	988710	977279		—	46696	23490	0
LGD	0.9887 (=988710/ 1000000)	0.9773 (=977279/ 1000000)	0.9657 (=965705/ 1000000)	—	0.0467 (=46696/ 1000000)	0.0235 (=23490/ 1000000)	0.0000 (=0000/ 1000000)

In the previous example, the LGD is higher in the beginning months then degreased gradually to become lower in the last months same as (Chen 2017). In this study, we consider LGD at time of default (T=3) equals 0.9657, LGD at time of default (T=58) equals 0.0467, and so on. Then the average of LGD takes in each time.

Theoretically, LGD is defined as the key component for expected loss for credit portfolio, which is central to credit risk management. The expected loss (EL) or risk weight asset (Bauer & Agarwa 2014) for a credit portfolio is defined as:

$$\text{The expected loss } (EL)=\sum_{i=1}^{n} \text{Probability of default* Loss given default * Exposure value at the time of default}$$

LGD is used to find the expected loss that is key to determine minimum capital required based on IRB approach in the Basel committee. Moreover, it may be used to calculate the premium in credit insurance for corporations.

The rest of this paper proceeds as follows. In the next section, the framework of estimating LGD will be discussed. Section 3 presents parametric distributions for fitting the data, and Section 4 describes regression models for determining which financial variables significantly affect the LGD for our sample data. We present the sample data and the results in Section 5, and the final section will be the conclusion.

2. Proposed framework for estimating LGD

In figure 1.1, LGD data related with time extracted from credit portfolio from Jordanian bank. The right censored data approach is used to extracted LGD data for default borrowers during the 5-year period. The sample size of default which is declared as unable to pay financial obligations in a period of 3 months is 495 from 4393 borrowers. Then LGD data is fitted probability density function (pdf) with parametric models associated with time such as exponential, Weibull, Gamma, Gompertz models. Gamma distribution is better than other parametric models.

The LGD is assumed to be survival function S(t) as a result of heavy tail. We will consider Gamma distribution that has two parameters namely; shape 2 and rate. To find the financial variables that affect on LGD as covariate, Gamma regression model is used to estimate equation of rate parameter with constant shape 1. In this case, we consider time as dependent variables and financial variables as independent variable to estimate mean in Gamma regression.

Then we used Gamma distribution with two parameters; the rate parameter from Gamma regression model and the shape 2 parameter from a simulation. The gamma distribution has the cumulative distribution, survival distribution, and other characteristics more than Gamma regression model. Finally, Gamma distribution with the best link function for rate from Gamma regression selected based on statistical criteria. The methodology of this paper can be summarized in the following flow chart.

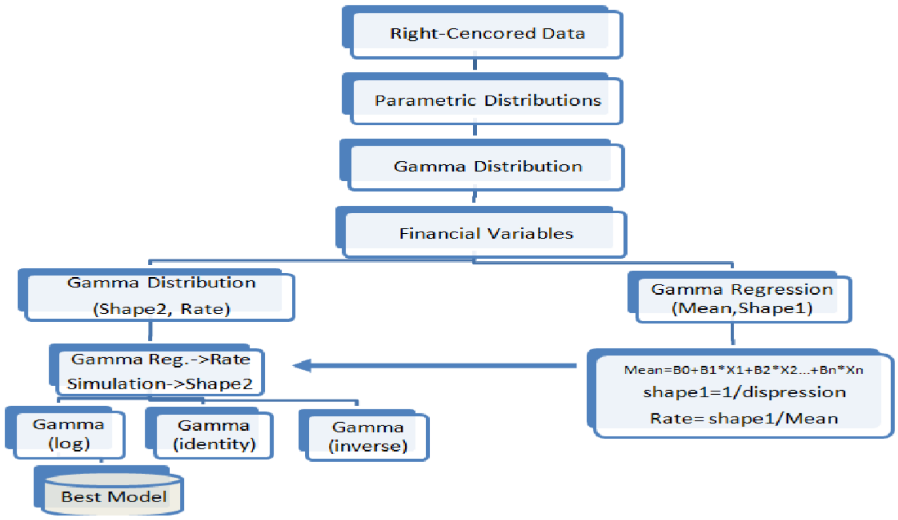

Figure 1: The flowchart of the WT forecasting.

3. Modelling LGD with parametric distributions

Common parametric distributions for survival analysis are considered. Table 1 provides the density and survival function for the distributions considered in this study, which are Exponential, Gamma, Weibull, Gompertz. Generalize Gompertz, and Gompertz-Makeham see Table 1 in next section.

We use five types of accuracy criteria to select the best model; mean square error (MSE), root mean square deviation (RMSD), mean absolute error (MAE), Nash–Sutcliffe model efficiency coefficient (NSE), and mean absolute percentage error loss (MAPE). The mean square error

$$\text{(Ramsey 1999)} = \frac{\sum_{i=1}^{n} (actual - predicted)^2}{n},$$

where n is the sample size. The root mean square deviation (RMSD) $=\sqrt{MSE}$. The RMSE values can be used to distinguish model performance in a calibration period with that of a validation period as well as to compare the individual model performance to that of other predictive models. The mean absolute error

$$(MAE) = \frac{\sum_{i=1}^{n} |actual{-}predicted|}{n},$$

where$n$ is the sample size. The MAE is the average vertical and horizontal distance between each actual and predicted points. The other criteria is Nash–Sutcliffe model efficiency coefficient

$$(NSE) = 1 - \left[\frac{\sum_{i=1}^{n} (actual_i{-}predicted_i)^2}{\sum_{i=1}^{n} (actual_i{-}E(actual))^2}\right]$$

where n is the sample size. The NSE is a normalized statistic that determines the relative magnitude of the residual variance "noise" compared to the measured data variance "information" (Nash & Sutcliffe 1970). The Nash–Sutcliffe efficiency can range between $-\infty$ and 1.0 (1 inclusive), with NSE $= 1$ being the optimal value. Values between 0.0 and 1.0 are generally viewed as acceptable levels of performance, whereas values $\leq$0.0 indicates that the mean observed value is a better predictor than the simulated value, which indicates unacceptable performance. (Servat & Dezetter 1991) also approved NSE to be the best objective function for reflecting the overall fit of a hydrograph. (Legates & Mccabe Jr 1999) suggested a modified NSE that is less sensitive to high extreme values due to the squared differences, but that modified version was not selected because of its limited use and resulting relative lack of reported values. Mean absolute percentage error loss

$$(MAPE) = \frac{100\%}{n} \sum_{t=1}^{n} \left|\frac{actual_t{-}predicted_t}{actual_t}\right|,$$

the MAPE can be indicated on how much error in predicting compared with the actual value.

4. Modeling LGD with Gamma Regression

In our study we use Gamma regression for fitting the LGD data with covariates. The reason for fitting Gamma regression is that the distribution is well-known with time. Gamma distribution is the best model for fitting the sample data of our case study. Therefore, we consider several Gamma regressions, by using different link functions, for fitting the LGD data with covariates.

Table 2: parametric models for LGD

Models	Parameters	Probability density function $f(t)$	Survival function $S(T)$
Exponential distribution	Rate(λ)	$\lambda e^{-\lambda t}$	$e^{-\lambda t}$
Gamma distribution	Shape (α) Rate (β)	$\frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t}$	$1-\frac{\gamma(\alpha,\beta t)}{\Gamma(\alpha)}$ $\gamma(\alpha,\beta t) = \int_0^{\beta t} z^{\alpha-1} e^{-z}.dz$
Weibull (proportional hazards)	Shape (α) Rate (β)	$\beta t^{\alpha-1} e^{-\beta t^\alpha}$	$e^{-\beta t^\alpha}$
Gompertz distribution	Shape (α) Rate(λ)	$\lambda e^{\alpha t} e^{\frac{-\lambda}{\alpha}(e^{\alpha t}-1)}$	$e^{\frac{-\lambda}{\alpha}(e^{\alpha t}-1)}$
Generalize Gompertz distribution (GGD)	Shape ($\boldsymbol{\alpha}$) Shape ($\boldsymbol{c}$) Rate ($\boldsymbol{\lambda}$)	$\theta \lambda e^{ct} e^{\frac{-\lambda}{c}(e^{ct}-1)} \left[1-e^{\frac{-\lambda}{c}(e^{ct}-1)}\right]^{\theta-1}$ $\lambda, \theta > 0, c \geq 0; t \geq 0$	$1-\left[1-e^{\frac{-\lambda}{c}(e^{ct}-1)}\right]^{\theta}$
Gompertz–Makeham	Shape (α) Shape (β) Rate(λ)	$\left(\alpha e^{\beta t}+\lambda\right) e^{-(\lambda t+\frac{\alpha}{\beta}(e^{\beta t}-1))}$, $\alpha, \beta, \lambda > 0\ ; t \in \mathcal{R}$	$e^{-(\lambda t+\frac{\alpha}{\beta}(e^{\beta t}-1))}$

Let random variable Y follows a Gamma distribution $Gamma(\alpha,\beta)$,where the parameters $\alpha,\beta > 0$.The mean and variance of Y are, respectively $E(y) = \alpha/\beta$ and $Var(y) = \frac{\alpha}{\beta^2}$. Joint modeling of the mean and the shape parameters in gamma regressions were proposed by (Cepeda-Cuervo 2001). With the re-parameterization of the gamma distribution as a function of the mean, $\mu = E(Y)$, and the shape parameter, α, as proposed in (Cepeda-Cuervo 2001) and (Cepeda & Gamerman 2005), setting $\beta = \frac{\alpha}{\mu}$, the gamma density function can be written as

$$f(y) = \frac{1}{y\Gamma(\alpha)}\left(\frac{\alpha y}{\mu}\right)^{\alpha} e^{-\frac{\alpha}{\mu}y}$$

Under this re-parameterization, we use $Y \sim Gamma(\mu,\alpha)$ to denote that Y follows a gamma distribution with $E(y) = \mu$, $Var(y) = \frac{\mu^2}{\alpha}$ and α as a shape parameter. From this re-parameterization of the gamma distribution, the joint mean and shape gamma regression was proposed in (Cepeda-Cuervo 2001), under classic and Bayesian methodologies.

Let $y = (y_1,\ldots,y_n)^T$be a random sample, where$y_i \sim Gamma(\mu_i,\alpha)$, $i = 1,\ldots,n$. In the gamma regression model with a constant shape parameter, the mean regression structure is defined by

$$g(\mu_i) = \eta_i = f\left(x_i^T;\beta\right) \tag{4.2}$$

where g is the link function, $\beta = (\beta_1,\ldots,\beta_n)^T$a vector of covariate unknown regression parameters which are assumed to be functionally independent, $\beta \epsilon\, \mathbb{R}^n$, η_i is a predictor. Some usual link functions in the gamma regression are:

log$g(\mu) = \log(\mu)$; identity $g(\mu) = \mu$; and inverse $g(\mu) = 1/\mu$.A rich discussion of link function can be found in (Mccullagh 1989).

The parameter of gamma regression can get from maximum likelihood estimator (MLE). The log-likelihood function for this class of gamma regression models has to form

$$\ell(\beta, \alpha) = \sum_{i=1}^{n} \ell_i(\mu_i, \alpha) \tag{4.3}$$

where $\ell_i(\mu_i, \alpha) = (\alpha - 1)\log y - \log\Gamma(\alpha) - \alpha\log\mu + \alpha\log\alpha - \left(\frac{\alpha}{\mu}\right) y$;

$\mu_i = g^{-1}(\eta_i)$, as defined in equations (3), are functions of β and α, respectively. The estimated parameters can be obtained by maximizing the likelihood in (4). The derivatives $U_k(\beta, \alpha) = \partial\ell(\beta, \alpha)/\partial\beta_k$obtained parameters beta regressionsβ's . For more detail about a prove gamma regression refer to (Cepeda-Cuervo 2001).

5. Results

The sample data of credit portfolio obtained for this study are collected from a bank in Jordan. The monthly data of the credit portfolio were collected from January 2010 until December 2014. The size of the portfolio is 4393, while the total number of defaults throughout the 5-year period is 495. The sample size is same as number of default. In this section, LGD data related with time. For the sample data, a borrower is declared default when his/her cash installment is not paid within 3 months. In this study, LGD is estimated by the default sample only.

Table 3 shows the risk exposures (number of loans at risk) and the number of defaults in each year. The highest number of defaults occurred in the second year, and the highest number of defaults per exposure also occurred in the same year (168 defaults from 1125 exposures). Table 4 provides summary statistics for the monthly data. The average monthly exposure is 73, while the average number of defaults per month is 8.

Table 3: Number of exposures and defaults in each year

Year	exposure	# of defaults	% (# of defaults per exposure)
2010	1265	137	10.83
2011	1125	168	14.93
2012	783	67	8.56
2013	652	41	6.29
2014	568	82	14.44
Total	4393	495	-

Table 4: Summary statistics for credit data (monthly)

	Exposure per month	# of defaults per month	# of censored per month
Min	29	0	23
Max	272	33	254
Mean	73.22	8.25	64.97
Std. dev.	42.88	6.35	38.93
Total (N)	4393	495	3898

Table 5 provides the estimated parameters, together with the MSE, RMSD, MAE, NSE, and MAPE. The best model is chosen based on the smallest MSE, RMSD, MAE, MAPE and the largest NSE. The results in Table 8 show that the Gamma distribution is the best parametric model since it has the smallest MSE, RMSD, MAE, MAPE and the largest NSE.

Table 5: Estimated parameters and goodness-of-fit criteria for parametric models

Models	Parameters est.	No. of iteration	MSE	RMSD	MAE	NSE	MAPE
Exponential	rate (λ)=0.00901	1000	0.00766	0.08755	0.06751	0.73329	0.11290
Gamma	shape (α)=1.59 rate (β)=0.02	50	0.00662	0.08138	0.05804	0.76956	0.09999
Weibull (proportional hazards)	Shape (α)=1.4 rate (β)=0.0021	50	0.00669	0.08181	0.05857	0.76708	0.10059
Gompertz	shape (α)=0.00680 rate (λ)=0.01380	50	0.00702	0.08376	0.06318	0.75584	0.10615
Generalize Gompertz distribution (GGD)	shape (α)=0.0180 shape (c)=0.99 rate (λ)=0.006	50	0.00702	0.08381	0.06194	0.75558	0.10497
Gompertz–Makeham	shape (α)=0.0301 shape ((β))=0.00510 shape (λ)=0.00200	50	0.00712	0.08436	0.06367	0.75236	0.10689

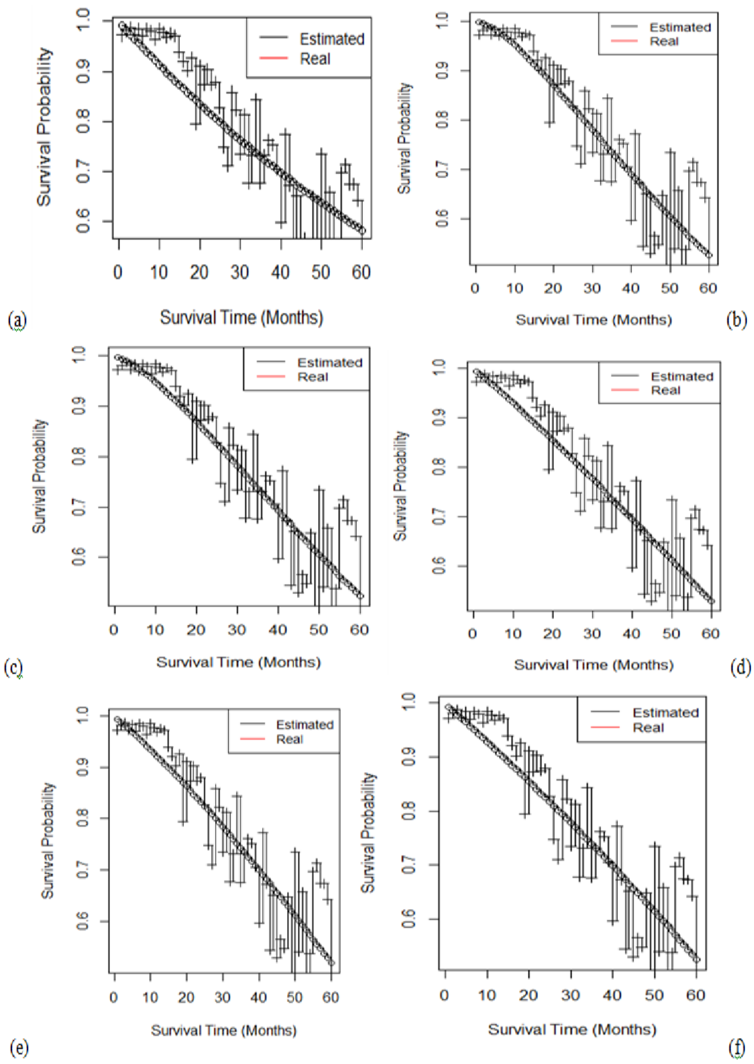

Figure 2: Fitted survival function for parametric models. (a) Exponential distribution, (b) Gamma distribution, (c) Weibull distribution, (d) Gompertz distribution, (e) Generalize Gompertz distribution, (f) Gompertz -Makeham.

For further comparison, Figure 2 shows the curve of survival function for all of the fitted parametric models. The graphs illustrate that the survival curve of Gamma model is closest compared to other models.

Table 6. Summary statistics for explanatory variables as covariate

Explanatory variables	Mean	Std. dev.	Minimum	Maximum
The service pricing policy(O.E/ T.L)	2.3063	.6324	.9578	3.1188
Operating efficiency (T.O.E/T.O.R)	.9208	.1318	.6048	1.2214
Liquidity ratio (Current ratio)	1.0168	.1408	.7783	1.5508
Profitability ratio (ROE)	.5839	.5838	-1.1598	1.7039

Table 7: Gamma regression distributions with different link function $Gamma(\mu, \alpha_1)$,

Models	parameters	Equations	Beta	Std. Error	t value
Gamma regression (log)	μ	(Intercept)	3.2932	0.0611	53.8850***
		(O.E/T.L)	-0.2989	0.0726	-4.1150***
		(T.O.E/T.O.R)	-0.2481	0.0627	-3.9550***
		Current ratio	-0.1495	0.0760	-1.9670.
		ROE	0.3796	0.0750	5.0630***
	α_1	=1/dispression	=1/0.2241019		
		Chi.test	12.3256	**AIC**	490.02
		p-value	0.0155	**BIC**	502.5905
				Iteration	**10**
Gamma regression (Identity)	μ	(Intercept)	36.1	3.4	10.8***
		(O.E/T.L)	0.7	0.7	1
		(T.O.E/T.O.R)	-15.5	2.6	-6***
		Current ratio	8.8	3.5	2.5.
		ROE	17.8	2.4	7.5***
	α_1	=1/dispression	1/0.416465		
		Chi.test	215.8659	**AIC**	538.7
		p-value	0.001397218	**BIC**	551.2635
				Iteration	25
Gamma regression (Inverse)	μ	(Intercept)	-0.0558	0.0225	-2.5.
		(O.E/T.L)	0.0121	0.0042	2.8*
		(T.O.E/T.O.R)	0.0368	0.0126	2.9*
		Current ratio	0.0458	0.0157	2.9*
		ROE*	-0.0222	0.005	-4.5***
	α_1	=1/dispression	1/0.2527023		
		Chi.test	13.89824	**AIC**	499.6
		p-value	0.009505567	**BIC**	512.1716
				Iteration	5

Table 6 explains the descriptive statistics for explanatory variables. The mean of service pricing policy (operating expenses (O.E) to total liability (T.L)) is 2.31 and standard deviation is 0.63. In addition, the mean of operating efficiency (Total operating expenses (T.O.E)/Total operating revenue (T.O.R)) is 0.92 with standard deviation is 0.13. Furthermore, the mean of liquidity ratio (Current ratio=current asset/current liability) is 1.02 with standard deviation is 0.14. Finally, the mean of profitability ratio (return on equity ROE=Net Income/Owner Equity) is 0.58 with standard deviation is 0.58.

Estimating LGD based on $Gamma(\alpha_2, \beta)$ in this section. The parameters rate β estimated based on $Gamma(\mu, \alpha_1)$, where mean μ estimated from financial variables namely; standardized profitability ratio, standardized liquidity ratio, standardized operating efficiency ratio, and standardized service pricing policy. The different link function used to estimate this parameterμ such as log, identity, and inverse with gamma regression explain in table in next table.

Conclusion

In the context of credit portfolio losses, loss given default (LGD) is the proportion of the exposure that will be lost if a default occurs. Uncertainty, regarding the actual LGD is an important source of credit portfolio risk in addition to default risk. In this study, we have used several parametric models to estimate LGD with time variables in the aim to evaluate the performance of a credit risk portfolio samples. The financial variables models are fitted to a sample data of credit portfolio obtained from a bank in Jordan for the period from 2010 - 2014. The used models were selected using several goodnesses of fitting criteria to compare the performance of varied distributions. The used models were selected using several goodnesses of fitting criteria to compare the performance of the varied distributions. The results show that the Gamma distribution and Gamma regression are the best parametric models for estimating LGD with time of default based on the following tests: MSE, RMSD, MAE, NSE, and MAPE.

Acknowledgments

The authors would like to thank the editor and the reviewer for their valuable comments and constructive reviews to further improve this paper. The authors are also grateful and would like to acknowledge the financial support granted by the Ministry of Higher Education (MOHE) and Universiti Kebangsaan Malaysia (UKM) in the form of research grants (FRGS/1/2019/STG06 /UKM/01/5 and GUP-2019-031).

References

[1] E. I. Altman, *Measuring corporate bond mortality and performance*, The Journal of Finance, 44 (1989), 909-922.

[2] J. Bauer, V. Agarwal, *Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test*, Journal of Banking and Finance, 40 (2014), 432-442.

[3] T. Bellotti, J. Crook, *Modelling and estimating Loss Given Default for credit cards*, University of Edinburgh Business School, 2008.

[4] T. Bellotti, J. Crook, *Loss given default models incorporating macroeconomic variables for credit cards*, International Journal of Forecasting, 28 (2012), 171-182.

[5] J. Beran, A.-Y. K. Djaïdja, *Credit risk modeling based on survival analysis with immunes*, Statistical Methodology, 4 (2007), 251-276.

[6] M. Bruche, C. González-Aguado, *Recovery rates, default probabilities, and the credit cycle*, Journal of Banking and Finance, 34 (2010), 754-764.

[7] R. Calabrese, *Estimating bank loans loss given default by generalized additive models*, UCD Geary Institute Discussion Paper Series, WP2012/24, 2012.

[8] R. Cao, J. M. Vilar, A. Devia, *Modelling consumer credit risk via survival analysis*, SORT: statistics and operations research transactions, 33 (2009), 0003-30.

[9] K. Carling, T. Jacobson, K. Roszbach, *Duration of consumer loans and bank lending policy: dormancy versus default risk*, Sveriges Riksbank Working Paper Series,1998.

[10] E. Cepeda-Cuervo, *Modelagem da variabilidade em modelos lineares generalizados*, Unpublished Ph. D. Tesis, Instituto de Matemáticas. Universidade Federal do Río do Janeiro, 2001.

[11] E. Cepeda, D. Gamerman, *Bayesian methodology for modeling parameters in the two parameter exponential family*, Revista Estadstica, 57 (2005), 93-105.

[12] H. Z. Chen, *A new model for bank loan loss-given-default by leveraging time to recovery*, Journal of Credit Risk, 14 (2018), 1-29.

[13] A. C. Cohen, *Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples*, Technometrics, 7 (1965), 579-588.

[14] J. Dermine, C. N. De Carvalho, *Bank loan losses-given-default: a case study*, Journal of Banking and Finance, 30 (2006), 1219-1243.

[15] G. Giese, *The impact of PD/LGD correlations on credit risk capital*, Risk-London-Risk Magazine Limited, 18 (2005), 79-80.

[16] D. Glennon, P. Nigro, *Measuring the default risk of small business loans: a survival analysis approach*, Journal of Money, Credit and Banking, 15 (2005), 923-947.

[17] X. Huang, C. W. Oosterlee, *Generalized beta regression models for random loss given default*, The Journal of Credit Risk, 7 (2011), 45-50.

[18] S. A. Klugman, H. H. Panjer, G. E. Willmot, *Loss models: from data to decisions*, vol. 715, John Wiley and Sons, 2012.

[19] B. Kollar, T. Kliestik, *Simulation approach in credit risk models*, in 4th International Conference on Applied Social Science (ICASS 2014), Information Engineering Research Institute, Advances in Education Research, (2014), 150-155.

[20] S. Krüger, D. Rösch, *Downturn LGD modeling using quantile regression*, Journal of Banking and Finance, 79 (2017), 42-56.

[21] D. R. Legates, G. J. McCabe Jr, *Evaluating the use of "goodness of fit" measures in hydrologic and hydroclimatic model validation*, Water resources research, 35 (1999), 233-241.

[22] G. Loterman, I. Brown, D. Martens, C. Mues, B. Baesens, *Benchmarking regression algorithms for loss given default modeling*, International Journal of Forecasting, 28 (2012), 161-170.

[23] M. Malik, L. C. Thomas, *Modelling credit risk of portfolio of consumer loans*, Journal of the Operational Research Society, 61 (2010), 411-420.

[24] P. McCullagh, Nelder, J., *Generalized linear models*, Seconded, Chapman and Hall, London, 1989.

[25] J. E. Nash, J. V. Sutcliffe, *River flow forecasting through conceptual models part IA discussion of principles*, Journal of hydrology, 10 (1970), 282-290.

[26] M. Pykhtin, *Recovery rates: unexpected recovery risk*, Risk-London-Risk Magazine Limited, 16 (2003), 74-79.

[27] M. Qi, X. Zhao, *Comparison of modeling methods for loss given default*, Journal of Banking and Finance, 35 (2011), 2842-2855.

[28] J. B. Ramsey, *The contribution of wavelets to the analysis of economic and financial data*, Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences, 357 (1999), 2593-2606.

[29] E. Servat, A. Dezetter, *Selection of calibration objective fonctions in the context of rainfall-ronoff modelling in a Sudanese savannah area*, Hydrological Sciences Journal, 36 (1991), 307-330.

[30] M. Stepanova, L. Thomas, *Survival analysis methods for personal loan data*, Operations Research, 50 (2002), 277-289.

[31] Y. Tanoue, A. Kawada, S. Yamashita, *Forecasting loss given default of bank loans with multi-stage model*, International Journal of Forecasting, 33 (2017), 513-522.

[32] L. Thompson, T. Brandenburger, *Survival analysis methods to predict loss rates in credit card portfolios*, 2019.

Accepted: 29.02.2020