

Modelling Default Risk of Borrowers: Evidence from Online Peer to Peer Lending Platforms in Australia

TAN Zhongming BAAH Alexander* DING Guoping OWUSU-ANSAH Patrick
AGYEMANG Kwabena
School of Finance and Economics, Jiangsu University,
Xuefu Road 301, Zhenjiang City, Jiangsu Province, China

Abstract

Peer to Peer lending has the capacity to transforming the mass banking industry worldwide but credit risk modelling remains the core challenge of the platform. The general objective of this study is to analyse the credit default risk of borrowers of Peer to Peer online lending platform based in Australia. Specific objectives include the following;

1. To identify the loan information applicants provide to request for a loan facility,
2. Using RateSetter.com published data on loans to predict the likelihood of credit risk of the platform.

In this article, we employed binary logistic regression model to assess the likelihood of loan default. Based on the mathematical approach and the nature of dependent variable, we grouped variables into categorical, numerical-continuous as well as binary. The dependent variable is dichotomous whilst real-life dataset was retrieved from a popular and competitive online lending platform based in Australia from 2014-2017. We identified that early repayment, no mortgage tenant; car, debt consolidation, investment, major events, professional services, 3-year loan duration, 4-year loan duration, interest rate and income have significant influence on borrowers' likelihood to default. Our empirical coefficients suggest that, there is 83.4% likelihood of borrowers default rate and hence recommended a critical examination of borrowers' information presented to the platform. This paper fulfills the need to examine the credit information provided by loan applicants. Similarly, it endeavors to predict the possibility of borrowers default risk and the reasons contributing to online lending credit default risk.

Keywords: Credit Risk, Peer To Peer Online Lending, Binary Logistic Regression

DOI: 10.7176/RJFA/10-2-01

1. Introduction

Over the past decade, social lending likewise acknowledged as peer to peer lending has become a striking field of interest among researchers and financial experts. Online P2P lending is a financial innovation that combines the internet and private lending marketplace to provide unsecured loans to borrowers. According to studies conducted by Bachmann et al. (2011) and Tao, Dong et al. (2017) P2P lending is the provision of unsecured peer-to-peer loans from lenders to borrowers through online marketplace and absolutely free from the involvement of financial intermediaries. Although newest and evolving, it is designed to supplement main stream financial institutions lending by closing borrowing deficit among Small Scale Medium Enterprises as well as individual borrowers.

Social lending has the potential to transform the mass banking model worldwide. Since the marketplace is basically organized on the internet, online P2P lending has the ability to connect every potential lender and borrower. Prodigious majority of borrowers with low net-worth and borrow comparatively modest amounts, currently have the chance to get finance via the P2P marketplace platform. It also encourages, intra-community, regional and global lending and borrowing. Invariably, the sustainability of online lending depends on the flexibility of service providers to devise a strategy of addressing risk management practices due to information asymmetry that arises among lenders and borrowers. The trustworthiness of both parties is unknown to each other.

The existence of market imperfect information between borrowers and lenders cause an adverse selection problem and subsequently misguide lenders to fund higher-risk sub-prime borrowers, Akerlof (1970), (Shen, Krumme et al. 2010). In an attempt of limiting such phenomenon, online platforms however have adopted "unconventional means" of credit rating system to judge creditworthiness of loan applicants. Borrowers' transaction and credit information available on Lending Club, Yue-Bao, Renrendai.com, Alibaba to mention few are mostly used as the yardstick to determine either the borrower is a "good" or "bad". In U.S and other Western countries there exist a standard credit rating system, where historical financial information of the borrower are directly provided by specialized and independent credit rating agencies like FICO (Tao, Dong et al. 2017), (Malekipirbazari and Aksakalli 2015, Lin, Li et al. 2017). P2P platforms prevent market losses due to NPL by maintaining certain funds. Zopa Safeguard Trust (ZST) for instance is founded by Zapo and provides a total sum of £9.6m (\$10.56m) to mitigate credit risk and NPL. Again, the opportunity granted by peer-to-peer marketplaces actually allow lenders and loan applicants to participate in pools of loans offered on the platform

and has further limited market risk that occurred in 2007-2008. As most internet finance platforms are still sprouting and improving their pricing and risk management models, addressing credit risk is quite often complicated. Credit facilities provided by these marketplace platforms are not securitized, scrutinized and unsecured making credit risk assessment the central concern for borrowers and lenders. Recent studies however focus on the auditing mechanism of peer-to-peer lending platform as well as the assessing borrowers' characteristics that affect credit default rate using information provided by borrowers.

Based on this context, we modeled default risk of the borrowers using dataset from a P2P lending platform in Australia. The study employed binary logistic regression model to evaluate the borrower's credit information (both soft and hard financial information) and other contributing factors in order to shed light on the standard credit risk assessment method for lending platforms and lenders. We therefore look critically into various variables and how they generate a repayment problem which is a major departure of works conducted by scholars in literatures reviewed. From the theoretical perspective, the study adds up to the existing literature concerning credit risk screening problems in P2P lending. The expedition has therefore broadened the knowledge of researchers about the ongoing credit risk analysis. From the practical perspective, the empirical findings serve as yardstick to lenders by making informed and optimal investment decision strategy. Hence, understanding the unique characteristics of P2P lending platform that boost the credit score of borrowers and lenders who take the advantage of lending booms for online loans. The rest of this paper is organized as follows. In Section 2, we present the related empirical literature and institutional framework of P2P lending. Section 3 describes the methodology and presents the empirical results for measuring the default risk of borrowers. Finally, Section 4 discusses the implication of the empirical result and concludes the whole article.

1.1. Literature Review

Fundamentally, credit may be referred to as any loan facility lend to a potential borrower or any monetary instrument (fixed coupon bond) that involves pre-determined fixed payments over a given time period. According to Anita (2008), credit risk is the possible loss of valuable assets caused by probable deterioration of the creditworthiness of counterparty or its inability to meet contractual obligations (Manab, Theng et al. 2015). Cantor (2001), identified credit risk as a dominant risk for financial institutions as the core mandate for banks are lending and deposit activities. Additionally, security firms may face credit risk as well due to their involvement in derivatives market, borrowing or lending securities and making fringe loans to customers. Thus, credit risk depends on the inability of borrowers to generate sufficient cash flows through operation, earnings, or asset sales to meeting future interest and principal payment of the outstanding debt.

The proliferation of countless number of growing P2P lending supported by the internet in recent years has gingered the interest of researchers. A number of studies therefore focus on the advantages and economic impact of – the evolving online P2P lending. Culkin, Murzacheva et al. (2016), identified the emergence of online P2P lending platforms as the provision of direct alternative source of funds to facilitate the growth and development of SSMEs. Fraser, Bhaumik et al. (2015), suggested that P2P lending bridges the financial gap created by the traditional banking system and information flow among players on the platform. They were on the other hand indistinct about the role of service providers in the context of the regulations. There are existing body of literature that focus on the determinants of funding success, the interest, and or credit risk particularly default rate of borrowers.

Mild, Waitz et al. (2015) and Guo, Zhou et al. (2016) have shown profound interest in the on-going debate of online lending and therefore proposed credit risk assessment models of the online borrowers. A study conducted by Klafft (2008) on Prosper online platform illustrates that borrowers' credit ratings and verified bank account information significantly influences the probability of successful funding. Empirically, they revealed credit rating and debt-to-income ratios as the fundamental factors determining interest rates charged on loans. They however, discovered that verified bank account information and home ownership insignificantly impact on interest rates. Notwithstanding the financial credit rating categories, Klafft (2008) suggested three guiding principles that mitigate risk of a default. This investment decision rule follows: (1) investors must only lend to borrowers without any delinquent accounts. (2) investors must satisfy Rule 1 and that they must have a debt-to-income (DTI) ratio less than 20% and (3) by satisfying Rule 2 such that the borrower has no credit inquiry reports during the last 6 months. Empirical studies conducted by, Lin, Prabhala et al. (2013), Duarte, Siegel et al. (2012), Emekter, Tu et al. (2015), Navarro-Galera, Rayo-Cantón et al. (2015), and Chen, Li et al. (2017) however argued strongly against the above findings of (Klafft 2008). They utilize new dataset from renowned lending platforms such as Renrendai.com in China, LendingClub in U.S.A, Zopa to mention few. They found that unpredictable socio-economic and financial factors have much repercussion on the likelihood of loan funding.

Besides, there is substantial number of articles that shed views on the ongoing credit risk analysis by scholars in the field of finance in recent time. Tao, Dong et al. (2017), analysed data from Renrendai.com and found that borrowers who receive higher income and own a car (vehicle) are more likely to be funded. Their

findings also suggest that such borrowers pay less interest rate and are likely to default.

Lin, Li et al. (2017), empirically investigated borrowers default risk of P2P in China using Logistic regression and found that gender, age, marital status, educational level, working years, company size, loan amount, monthly payment, debt-to-income ratio and delinquency history have significant contribution in modelling loan default risk among borrowers.

Ma and Wang (2016), however spend quality time to examine factors that determine credit risk by looking at the entire P2P platform, the borrower and the regulatory policies using Interpretative Structural Modelling (ISM). Empirical results of their study expose the existence of direct relationship between auditing mechanism of P2P platforms and credit risk. They further explored that borrowers' moral level, job stability and the policy environment also affect credit risk.

Although many studies have been conducted to analyse credit default risk of borrowers of P2P lending platforms, many of them fail to critically examine borrowers' information and how they generate repayment difficulties. This study will add to the existing literature by providing evidence from RateSetter.com, a typical P2P online platform based in Australia.

1.1.1 RateSetter P2P

RateSetter was established in United Kingdom in 2010 and launched Australia 2014, with the goal of providing the opportunity by dressing investment gap between cash and shares by making better returns accessible to everyone in a simple and modern way supported by information technology. Over £2,782,454,205 have been provided by investors and earned £111 million in interest. The platform can boast of a team of 200 people based in London, Leicester and across the whole of U.K. Its equity investors comprised major investment fund managers, Woodford Investment Management and Arteris. Lenders on the other hand can choose to invest a minimum of £10 up to £1,000,000s and requires the rolling market flexibility while the 1 year and 5-year market offers long term growth and income opportunities. RateSetter is the first P2P online platform to launch provisional fund that safeguard against NPL. The platform has therefore maintained a 100% track record since 2010. Investors on the other hand, receive the expected returns accordingly and no lender has lost a coinage.

1.1.2 RateSetter Borrower Default and Recovery Team

A number of debt recovery procedures are put in place to recover defaulted loans. With respect to late repayment, RateSetter may embark on any of the following processes:

Serving the borrower repayment notice,

Re-issuing direct debit instrument to the borrower's bank account

RateSetter debt collection team contact the borrower to discuss circumstances and to ascertain whether they are experiencing hardship;

Developing alternative repayment arrangements and matching repayment amounts to expected cash flows or by submitting default reminder to directors and third-party guarantors to repay amount due;

If appropriate reporting non-payment to credit bureaus and thereafter appointing an external collection agency to vigorously pursue payment.

Circumstances under which the borrower defaults;

RateSetter engages in legal processes where necessary court actions,

Conditions whereby the loan facility is secured, RateSetter appoints a specialized collection team to exercise the security interest, repossess and sell the relevant property.

2. Methodology

2.1 The mathematical model

Generally, in modelling loan default prediction, this paper adopted binary response denoted by Y such that,

$$Y = \begin{cases} 1 \dots \text{Bad_borrower} \\ 0 \dots \text{Good_borrower} \end{cases}$$

Mathematically,

$$Y = g(X) + e_i \tag{1}$$

Where $X = (X_1, X_2, \dots, X_k)$ the explanatory vector, k is the number of independent variable and e_i is the stochastic term in the model. Function $\{g(X)\}$ is assumed to exist, it is abstract in practice. In situation where $\{g(X)\}$ is known, it shall have predictive errors such that each $X = x$, there is a distribution of Y -values in general.

$\{g(X)\}$ is therefore defined as

$$g(X) = E(Y|X = x) \tag{2}$$

Where $E(Y|X = x)$ represents the expected values of Y for a particular realization X of the independent variable X. Function $\{g(X)\}$ specifies the probability of the loan default, which can be expressed as:

$$g(X) = E(Y|X = x) = P_r(Y = 1|x = x) \tag{3}$$

We then modeled a default estimate of $\{g(X)\}$ denoted by

$$E[(Y - \hat{g}(x))^2 | X = x] = [g(x) - \hat{g}(x)]^2 \tag{4}$$

The right-hand side of Eqn 4 represents the reducible error term that is inherent in Y. Hence, the learning problem is to find a good estimate of $\{\hat{g}^*(x)\}$ that minimizes the reducible error term in the model. Based on the mathematical model adopted, binary logistic regression model is considered for analyzing likelihood of loan default because the dependent variable of general regression model is not dichotomous. Our explanatory variable is categorized into non-default and default borrower which serve the solution to classification problem usually typical of logistic regression. In the binary logistic regression, the explanatory variables show the probability of default or otherwise. We therefore assumed z as unobservable continuous number and hence depict the likelihood of a default. All other things being equal, z value is positively related to default probability. In order to transform the continuous number into likelihood function, we adopted the binary logit model. The model takes the format of the function in Eqn 5;

$$p = \frac{1}{1 + e^{-z}} \tag{5}$$

In Eqn 6 below, k and z are said to be linearly correlated as described in the model;

$$z = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \dots + \beta_kx_k \tag{6}$$

With p representing the probability of default, x = explanatory variable, $\beta_k =$ the parameter estimates and k is the number of explanatory variable according to (Hosmer and Lemeshow 2000)

2.1.1 Data Description

This section of the article describes and summarizes the descriptive statistics of the data used in the study. Our dataset was retrieved from RateSetter.com website via www.ratesetter.com.au which is the most popular Peer-to-Peer platform in United Kingdom (U.K) specifically in Australia. The dataset starts from 2014:M10 to 2017:M12. Approximately, 15,776 borrowers had been assigned loans amounting \$208,078,046 (\$208.1m) within the 4 year-period. The dataset comprised loan transactions either issued or reached maturity and thus contains the financial information about the empirical creditworthiness of the applicants. Respective loan attributes were then allotted to each borrower depending on the credit information provided (Soft information, hard information and loan characteristics). Through data preprocessing and data screening, we excluded the loan status of attribute “borrower I.D” and “contract date” since they do not have effect on the results of this study. To identify empirically credit worthiness of the borrower, we considered ‘loan status’ of attributes ‘<30 days’, ‘>30 days’, ‘in-default’, ‘Hardship’, as “Bad or Default Borrower” whereas ‘fully paid’ and ‘On-schedule’, attributes was classified as “Non-Default or Good Borrower”. According to the available data, features used in our predictive model of each case comprised of borrowers age, employment history, annual income, purpose of the acquired loan facility, loan amount, interest rate, gender, early repayment by the borrower and borrower’s status (whether the applicant/borrower defaulted or non-defaulted otherwise known as good or bad borrower.) Among these variables, employment history, loan purpose, gender, early repayment, housing status, borrower’s residence was categorized. Specifically, gender, repayment status, early repayment was classified as binary variables; whereas borrowers’ residence was divided into Australian Capital Territory (ACT), New South Wales (NSW), Northern Territory (NT), Queensland (QLD), South Australia (SA), Tasmania (TAS), Victoria (VIC) and Western Australia (WA); housing status contains own a home(mortgage), Own a home (no mortgage), Tenant (mortgage) and Tenant (no mortgage). The categorization above increased the reliability and stability of the model thereby reducing the computational complexity.

3. Data Presentation and Analysis

3.1. Preliminary (Econometric Approach)

Multicollinearity is possible when the correlations between IVs are high. This makes it impossible to isolate individual effect on the DVs. High multicollinearity increases VIF of OLS estimator and reduces the significance of the estimate. To avoid data redundancy, skewing and the occurrence of TYPE II error, we assessed multicollinearity by examining VIF and T. The T and VIF as shown in Table 1 vary from 0.576 to 0.991 and

1.009 to 1.682 respectively. In OLS results, we found that the collinearity statistics-T and VIF were greater than 0.2 ($T > 0.200$) and less than 2.5 ($VIF < 2.500$) respectively, indicating low threshold for the multicollinearity problem and will not bias the results. In our dataset, the pairwise correlation between two variables ranges from -0.283 to 0.259. Our pairwise correlation is much lower than the 0.8 critical threshold for serious multicollinearity problem as suggested by (Gujarati 2009).

Table 1. Preliminary check and data reliability

Model	T	VIF
Constant		
Repayment Record	0.920	1.087
Gender	0.982	1.019
Housing Status	0.849	1.177
Employment History	0.930	1.075
Borrowers State	0.991	1.009
Loan Purpose	0.672	1.489
Age	0.846	1.182
Loan Duration	0.659	1.440
Interest Rate	0.576	1.682
Income	0.962	1.039
Loan Amount	0.761	1.315

Collinearity statistics-Tolerance (T) & Variance Inflation Factor (VIF)

We included eleven (11) independent variables in the model for precision. The standard deviation, skewness, standard errors, kurtosis, mean are presented in Table 2. Table 3 on the other hand, shows the correlation matrix. We noticed that the correlation between the variables was relatively low (less than 0.7).

Table 2;

Model/Variables	1	2	3	4	5	6	7	8	9	10	11
Loan Amount	1.000										
Borrower province	0.018	1.000									
Employment history	-0.034	-0.001	1.000								
Sex	-0.047	-0.029	-0.002	1.000							
Borrower income	0.005	0.017	0.067	0.003	1.000						
Repayment record	0.003	0.013	-0.015	0.006	-0.035	1.000					
Housing status	0.052	0.079	0.036	0.006	0.070	0.067	1.000				
Age	-0.037	0.021	-0.187	0.021	0.045	-0.009	0.259	1.000			
Terms of repayment	-0.067	0.011	0.024	-0.022	0.018	0.138	0.026	0.041	1.000		
Purpose of loan	0.211	-0.022	-0.047	0.082	0.071	-0.067	0.167	-0.127	-0.017	1.000	
Interest rate	-0.149	-0.019	0.012	0.021	-0.072	-0.022	-0.046	-0.048	-0.283	0.220	1.000

Pairwise correlation

Notwithstanding, the descriptive statistics of applicants' fund request data as captured in our study is depicted in table 3. Based on the sample of 15,776 loan applicants, 10,085 people are male accounting for 63.9% while the remaining 5,691 applicants comprised female representing 36.1%. About 0.2% of the sample size fall below 20 years, 32.5% is within 30-39 years followed by 40-49 years representing 27.3%. According to age structure of the applicants, most loan applicants are within age 20-59 years. About 55.1% borrowers live within the income bracket of \$50,000-\$100,000 annually, 19.4% earn from \$100,000 to \$150,000 whilst 9.1% and 16.4% of the entire sample earn.

Table 3

Variables	N	Min	Max	Mean	S.D	Skewness	S.E	Kurtosis	S.E
Repayment Record	15776	1	2	1.33	0.470	0.730	0.020	-1.467	0.039
Gender	15776	1	2	1.64	0.480	-0.580	0.020	-1.664	0.039
Housing Status	15776	1	4	2.37	1.407	0.169	0.020	-1.858	0.039
Employment History	15776	1	6	2.40	1.135	2.547	0.020	5.107	0.039
Borrowers State	15776	1	8	4.77	2.260	0.055	0.020	-1.475	0.039
Loan Purpose	15776	1	13	5.51	2.919	0.179	0.020	-1.499	0.039
Age	15776	1	8	3.71	1.181	0.443	0.020	-0.206	0.039
Loan Duration	15776	1	10	3.72	1.638	0.471	0.020	-0.575	0.039
Interest Rate	15776	2.87	11.95	6.7969	2.53786	0.132	0.020	-1.750	0.039
Income	15776	1	4	3.13	1.131	-0.939	0.020	-0.651	0.039
Loan Amount	15776	110	268982	13252.94	10693.053	3.140	0.020	34.942	0.039
Borrower Status	15776	1	2	1.98	0.141	-6.829	0.020	44.645	0.039

Descriptive Statistics: Note; N, Min and Max are Total dataset; Minimum and Maximum respectively. S.D and S.E depicts Standard deviation and Standard Errors

Below \$5,000 and above \$150,000 per annum respectively. Intuitively, majority of RateSettler P2P applicants are situated in QLD, NSW and VIC. Statistically, 4172 accounting for 26.4% of all borrowers live in QLD, 4147 representing 26.1% also live in NSW and 3999 borrowers accounting for 25.3% reside in VIC. The share of the remaining five major states in Australia comprised 3,458 loan applicants accounting for 22.2% borrowers which is less than the applicants in one of the states afore mentioned. About 82.7% of borrowers are full time employees, 6.1% are part-time workers and 7.4% are self-employed applicants. It was found that, 0.1% borrower's house-persons and 2.0% borrowers are either retired or engage in self-supporting economic venture other than as mentioned early on. In terms of borrower default history, 15,458 applicants have no default history and only 314 borrowers had default behaviors in the past. The minimum loan duration is 6 months with 7 years' maximum. Borrowers prefer 1 year, 2 years, 3 years, and 5 years' loan duration to 6 months, 9 months, 1.5 years, 4 years and 7 years.

3.1.1 Results

According to Table 4 where we reported the results of the regression, 7 out of 11 regressors contribute to loan default due to their statistically significant t-statistics. Observing Hosmer and Lemeshow's X^2 test statistics (11.913) the model is significant in explaining the phenomena. Additionally, the 0.027 and 0.153 are the Cox & Snell and Nagelkerke R^2 respectively. S.E (S. E<0.2) of the estimated coefficients as found in Table 3 further suggest that our regressors are free from multicollinearity problem. In addition, the estimated coefficients in our model are significant at the 1% level, 5% and 10% levels excluding variables which are insignificant. Of course, our results depict several implications. The empirical findings show that, gender, borrower's state (where the borrower lives), employment history, age of the borrower and loan amount insignificantly affect the likelihood of loan default. Although Housing status generally contributes significantly to borrower default, No-mortgage tenant specifically is significant in modelling default rate of the borrower. This implies that for every 1 unit increase of the predictor (No-mortgage Tenant) the odds are changing by 0.463units. That is, the intention of the 'No-mortgage Tenant' to default decreases the logit of loan default by -0.769 units than mortgage tenant.

The estimated coefficients illustrate that loan purpose is positively related to default risk. The default risk for acquiring loan facility for investment and to purchase a car/vehicle is relatively higher than the original. However, a percentage (1%) increase in any of the following IVs (1) debt consolidation, (2) major events, (3) Investment and (4) professional development/services, their respective log odds increases by 0.637 units, 1.588 units, 0.812 units and 0.794 units in that order.

The interest rate of a borrower is likewise related to the default rate significantly. As interest rate increase by 1%, the default rate will reduce to 32.7%. Nevertheless, repayment plan by the borrower is associated with default risk. Any attempt of the borrower to delay repayment by additional month, the borrower is 6.096 more likely to default than the borrowers who pay on time. It illustrates that, a borrower who had defaulted in the past is 0.998 more likely to default all things being equal. Finally, income is generally significant in predicting default risk of the borrower. A borrower whose monthly salary is from \$50 to \$100k is 2.262 more likely to default than the former whereas borrower whose salary is \$100-\$150k is 2.896 more likely to default payment. The results on the other hand show that, an increase in borrowers' income by 1% it is expected that the default rate will increase by 3.319 units' times the original.



Table 4: Presentation of results

Variables	β	SE	Wald	Sig	Exp(β)
1. Early repayment	1.808	0.233	60.173	0.000***	6.096
2. sex	-0.069	0.129	0.287	0.592	0.933
3. Housing Status			25.968	0.000***	
Own a home (No Mortgage)	-0.562	0.350	2.584	0.108	0.570
Tenant (mortgage)	-0.304	0.273	1.235	0.267	0.738
Tenant (No mortgage)	-0.769	0.153	25.231	0.000***	0.463
4. Employment History			6.933	0.226	
Employed full time	-0.15	0.516	0.01	0.977	0.985
Employed part time	0.342	0.585	0.341	0.560	1.407
House person	16.647	8041.932	0.000	0.998	16974045.33
Retired	-0.632	0.882	0.513	0.474	0.532
Self-employed	-0.447	0.548	0.667	0.474	0.639
5. Borrowers State			5.417	0.609	
NSW	-0.217	0.522	0.173	0.677	0.805
NT	-0.882	0.644	1.875	0.171	0.414
QLD	-0.384	0.523	0.538	0.463	0.681
SA	-0.591	0.558	1.121	0.290	0.554
TAS	-0.232	0.727	0.102	0.750	0.793
VIC	-0.409	0.521	0.617	0.432	0.664
WA	-0.257	0.537	0.228	0.633	0.774
6. Loan Purpose			24.757	0.016**	
Car/Vehicle	1.053	0.289	13.301	0.000***	2.867
Debt Consolidation	0.637	0.269	5.594	0.018**	1.891
Home Improvement	0.313	0.305	1.057	0.304	1.368
Investment	1.588	0.731	4.724	0.030**	4.893
Major Event	0.812	0.321	6.401	0.011***	2.252
Major Purchase	0.204	0.403	0.255	0.614	1.226
Other Consumer Loans	0.439	0.297	2.177	0.140	1.551
Professional Services	0.794	0.341	5.408	0.020**	2.213
Solar Energy Battery	15.205	14169.574	0.000	0.999	4011392.495
Solar Energy Equipment	17.392	7662.574	0.000	0.998	35739261.75
Solar Energy Panels	15.209	2660.587	0.000	0.995	4027703.632
Solar Water Heater	17.058	40192.969	0.000	1.000	25587921.44
7. Age			5.035	0.656	
20-29	-0.249	1.035	0.058	0.810	0.779
30-39	-0.210	1.036	0.041	0.839	0.811
40-49	-0.227	1.038	0.048	0.827	0.797
50-59	-0.071	1.044	0.005	0.946	0.931
60-69	0.356	1.093	0.106	0.745	1.428
70-79	0.070	1.328	0.003	0.958	1.073
80-89	-1.500	1.624	0.852	0.356	0.223
8. Loan Duration			76.343	0.000***	
1 year	0.214	0.454	0.233	0.637	1.239
2 years	0.249	0.414	0.363	0.547	1.283
3 years	4.313	0.628	47.171	0.000***	74.701
4 years	5.767	0.770	56.040	0.000***	319.732
5 years	5.116	0.744	47.296	0.000***	166.675
6 months	16.624	3790.376	0.000	0.997	16587012.38
6 years	18.240	12286.193	0.000	0.999	83507622.92
7 years	16.020	4381.294	0.000	0.997	9064541.458
9 months	16.501	6196.480	0.000	0.998	14659617.59
9. Interest Rate			87.442	0.000***	
Income	-1.117	0.119	36.199	0.000***	0.327
>\$150k	1.200	0.312	14.821	0.000***	3.319
\$100-\$150k	1.063	0.218	23.732	0.000***	2.896
\$50-\$100k	0.816	0.151	29.174	0.000***	2.262
10. Loan Amount			0.003	0.959	
Constant	0.000	0.000	35.232	0.000***	1.000
	8.654	1.458			5731.500
Chi-square (X^2)		X^2 :11.913	Sig:	0.155	

Note: ***, **, denote 1% and 5% significant levels correspondingly. K=1,000, > greater than; Australian Capital Territory (ACT), New South Wales (NSW), Northern Territory (NT), Queensland (QLD), South Australia (SA), Tasmania (TAS), Victoria (VIC) and Western Australia (WA); (\times^2) are Hosmer and Lemeshow's Test (2000)

Based on the estimated coefficients of the results in table 4 above, we can obtain the likelihood function in Eqn 6 as follows;

$$z = \beta_0 + \beta_1(\text{Early Repayment}) + \beta_2(\text{No Mortgage Tenant}) + \beta_3(\text{Car}) + \beta_4(\text{Debt Consolidation}) \\
 + \beta_5(\text{Investment}) + \beta_6(\text{major events}) + \beta_7(\text{Professional Services}) + \beta_8(\text{3 years}) \\
 + \beta_9(\text{4 years}) + \beta_{10}(\text{Interest Rate}) + \beta_{11}(> \$150k) + \beta_{12}(\$100 - \$150k) \\
 + \beta_{13}(\$50 - \$100k) + \varepsilon_i$$

$$(z = 8.654 - 0.769 + 1.053 + 0.637 + 0.313 + 1.588 + 0.812 + 0.794 - 1.500 + 4.313 \times 0.25 \\
 + 5.7670.333 + 5.116 \times 0.417 - 1.117 + 1.200 \times 0.150 + 1.063 \times 0.120 + 0.816 \times 0.150)$$

$$z = 1.617, \quad p = \frac{1}{1 + e^{-1.617}} = 0.834$$

NB: To simplify results, years were converted in months, income into units whereas the final z value was converted into units to simplify the analysis.

Ceteris paribus, there is 0.834 likelihood of borrower with the aforementioned characteristics to default. These characteristics have the high propensity on borrowers' default rate and hence must be dully checked by lenders and financial experts of the platform.

3.1.3 Discussion

Credit risk is an important concern for P2P lending market platforms. To broaden knowledge and understanding of default risk otherwise known as credit risk evaluation model, we employed binary logistic regression in modelling P2P funding success. The study computes the likelihood of choosing an alternative as a function of the attributes of all the alternatives available as discussed in section 2. First of all, contrary to our findings, related studies reveal that women in general exhibit common character traits when it comes to financial decision making them less likely to default (Powell and Ansic 1997, Chen, Li et al. 2017). Borrowers who are tenants and do not mortgage are less likely to default than the former as aforementioned in section 3. As the costs of rent generally hike, borrowers spend greater proportion of the borrowed loan on it and hence increase the likelihood of overdue repayment in future. The estimated coefficients however show that, mortgage influences the likelihood of successful funding and repayment. It can be mentioned that an increase in mortgage will decrease the probability of loan default. Intuitively, the higher the cost of rents, so does the borrower's default rate. Applicants of P2P platforms in Australia may have informed knowledge of rent control measures. Besides, borrowers who live in their own apartment as well as borrowers who live in rented apartment and mortgage do not contribute significantly to default and hence classified as good borrowers. The intention of the borrower to repay borrowed amount on schedule increases the logit of the estimate of repayment by 6.096 units.

Our research provides empirical evidence that some of the loan purpose variables contribute significantly to default risk of borrowers. It is obvious from the estimated coefficients in Table 4 such that, professional services, major events, vehicle acquisition, debt consolidation and major purchases are prime causes of credit risk among borrowers. Thus, we confirm that our results are consistent with (Herrero-Lopez 2009) who identified loan decision variables as the mediation between borrower characteristics and the likelihood of successful funding. Given the variations in application process, interest rate, loan amount, and term length across loan products, it is apparent that each option presents a unique set of pros and cons. Peer-to-peer loans offer the benefits of expedited application processing, smaller loan amounts, and shorter periods, but borrowers pay for these inconveniences in form of higher interest rates. In spite of the explained phenomenon, borrowers might want to fake their financial identity provided to the platform which may result in misclassification of good borrowers and vice versa. Although borrowers' residential state is statistically insignificant in explaining the model, the platform may find it difficult to monitor and the track the economic activities of these scattered clients. The exclusion of third-party's like the experienced traditional banks to further verify the borrowers' financial information is likely to have repercussions on the borrowers default risk. Mitigating against the above risk, RateSetter is compared to extensively use UKs premier credit reference agencies and business credit rating agencies like Graydon and Dunn & Bradstreet to assess borrowers' financial information. Additionally, the industry uses some kind of sophisticated identification mechanisms like CIFAS and other tools of combating financial fraud.

3.1.4 Summary and Conclusion

Economic and business opportunities as you might expect depend on affordable and reliable information technology (I.T). Sustainable development and growth of internet speed and access has enormously improved globally in recent years. It has therefore facilitated the creation of larger, faster and among all geographically diverse online lending marketplaces. The advent of online P2P marketplaces has successfully managed financial intermediaries' problem created by financial institutions. These platforms demand much time, effort and are riskier than mortal-and-bricks lending scenarios. Our study focuses on P2P lending of which individual lenders make unsecured loans to meet the financial request of potential borrowers. Getting true trustworthy and the degree of creditworthiness of a potential loan borrower is paramount for healthy functioning of social lending markets. In an attempt to determining the risk score of an individual borrower, financial features such as past

financial history, borrowers' age, employment history, annual income, purpose for the acquired loan facility, loan amount, interest rate, gender, and various other financial features are used. Borrowers who are recently unemployed, experiencing bankruptcy, and have had loan repayment difficulties are more likely to be offered unattractive interest rate as compared to borrowers with more positive credit report. The assumed imperfect information that exist in the fin-tech market compound information gathering problem. Again exclusively relying on online verification might pave way for debtors to undertake unproductive venture after acquiring the credit facility. We therefore suggest the use of double sword strategy. That is the adoption of online and offline information gathering to abjure information deficiencies and bottlenecks to modelling credit default risk.

Despite the encouraging empirical findings, the study suffers from a number of limitations. First, the study analyzed dataset from RateSetter.com website for a limited time range which denies the chance to exploit the structural dynamics among borrowers, lenders and lending platforms. Secondly, other micro and macroeconomic variables such as exchange rate, disaster (i.e. fire outbreak, death etc), government regulations, banking regulations and taxes, fraud etc may have aftershocks on our empirical estimations. We therefore propose the use of binary logistic regression model by importing data about borrower characteristics of other online lending platforms for further studies for a long-time range.

References

- Akerlof, G. A. (1970). "QUALITY UNCERTAINTY AND THE." *The quarterly journal of economics* **84**(3): 488-500.
- Anita, C. (2008). "Credit risk management: Prescription for Indian banks." *Credit risk models: new tools of credit risk management*: 3-15.
- Cantor, R. (2001). "Moody's investors service response to the consultative paper issued by the Basel Committee on Bank Supervision "A new capital adequacy framework"." *Journal of Banking & Finance* **25**(1): 171-185.
- Chen, D., X. Li and F. Lai (2017). "Gender discrimination in online peer-to-peer credit lending: evidence from a lending platform in China." *Electronic Commerce Research* **17**(4): 553-583.
- Culkin, N., E. Murzacheva and A. Davis (2016). "Critical innovations in the UK peer-to-peer (P2P) and equity alternative finance markets for small firm growth." *The International Journal of Entrepreneurship and Innovation* **17**(3): 194-202.
- Duarte, J., S. Siegel and L. Young (2012). "Trust and credit: The role of appearance in peer-to-peer lending." *The Review of Financial Studies* **25**(8): 2455-2484.
- Emekter, R., Y. Tu, B. Jirasakuldech and M. Lu (2015). "Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending." *Applied Economics* **47**(1): 54-70.
- Fraser, S., S. K. Bhaumik and M. Wright (2015). "What do we know about entrepreneurial finance and its relationship with growth?" *International Small Business Journal* **33**(1): 70-88.
- Gujarati, D. N. (2009). *Basic econometrics*, Tata McGraw-Hill Education.
- Guo, Y., W. Zhou, C. Luo, C. Liu and H. Xiong (2016). "Instance-based credit risk assessment for investment decisions in P2P lending." *European Journal of Operational Research* **249**(2): 417-426.
- Herrero-Lopez, S. (2009). *Social interactions in P2P lending*. Proceedings of the 3rd Workshop on Social Network Mining and Analysis, ACM.
- Hosmer, D. and S. Lemeshow (2000). *Applied Logistic Regression*, 2nd edn.(Wiley: New York.) NY, USA.
- Klaft, M. (2008). "Online peer-to-peer lending: a lenders' perspective."
- Lin, M., N. R. Prabhala and S. Viswanathan (2013). "Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending." *Management Science* **59**(1): 17-35.
- Lin, X., X. Li and Z. Zheng (2017). "Evaluating borrower's default risk in peer-to-peer lending: evidence from a lending platform in China." *Applied Economics* **49**(35): 3538-3545.
- Ma, H.-Z. and X.-R. Wang (2016). "Influencing factor analysis of credit risk in P2P lending based on interpretative structural modeling." *Journal of Discrete Mathematical Sciences and Cryptography* **19**(3): 777-786.
- Malekipirbazari, M. and V. Aksakalli (2015). "Risk assessment in social lending via random forests." *Expert Systems with Applications* **42**(10): 4621-4631.
- Manab, N. A., N. Y. Theng and R. Md-Rus (2015). "The determinants of credit risk in Malaysia." *Procedia-Social and Behavioral Sciences* **172**: 301-308.
- Mild, A., M. Waitz and J. Wöckl (2015). "How low can you go?—Overcoming the inability of lenders to set proper interest rates on unsecured peer-to-peer lending markets." *Journal of Business Research* **68**(6): 1291-1305.
- Navarro-Galera, A., S. Rayo-Cantón, J. Lara-Rubio and D. Buendía-Carrillo (2015). "Loan price modelling for local governments using risk premium analysis." *Applied Economics* **47**(58): 6257-6276.
- Powell, M. and D. Ansic (1997). "Gender differences in risk behaviour in financial decision-making: An experimental analysis." *Journal of economic psychology* **18**(6): 605-628.

-
- Shen, D., C. Krumme and A. Lippman (2010). Follow the profit or the herd? Exploring social effects in peer-to-peer lending. *Social Computing (SocialCom)*, 2010 IEEE Second International Conference on, IEEE.
- Tao, Q., Y. Dong and Z. Lin (2017). "Who can get money? Evidence from the Chinese peer-to-peer lending platform." *Information Systems Frontiers* **19**(3): 425-441.