Research Journal in Business and Economics
ISSN: 3078-381X
DOI: https://doi.org/10.61424/rjbe.v3i3.502
Journal Homepage: https://bluemarkpublishers.com/index.php/RJBE

RJBE
BLUEMARK PUBLISHERS

| RESEARCH ARTICLE

Evaluating Borrowers' Default Risk in Peer-To-Peer Lending: Evidence from A Lending Platform in Estonia (Europe)

Fosu Acheampong[1] ✉ and Albert Junior Nyarko[2]
[1]Department of Mathematics, College of Arts and Sciences, Illinois State University, Normal, IL 61761, USA; facheam@ilstu.edu
[2]Department of Communication, College of Communication and Information, University of Kentucky, Lexington, KY 40506, USA; any233@uky.edu
Corresponding Author: Fosu Acheampong, **E-mail**: facheam@ilstu.edu

| ABSTRACT

This study uses data from the Estonian Bondora platform to examine the factors that influence borrower default risk in peer-to-peer (P2P) lending. To investigate the impact of personal traits, loan qualities, and financial factors on default probability, 103,326 loan requests from March 2009 to June 2022 were examined. The results of binary logistic regression demonstrate that while debt-to-income ratio and length of employment did not significantly affect loan default, age, gender, marital status, education level, loan amount, monthly payment, work experience, and interest rate do. While single and divorced borrowers were at a higher risk, married and female borrowers had lower default rates. Reduced default chance was linked to both stable income levels and higher levels of education. These results provide P2P lending platforms with valuable insights to inform loan issuance and borrowers to make informed decisions that enhance both parties' assessment of credit risk.

| KEYWORDS

Peer-to-peer (P2P) lending, binary logistic regression, loan default, risk assessment

| ARTICLE INFORMATION

ACCEPTED: September 21, 2025 **PUBLISHED:** November 13, 2025 **DOI:** 10.61424/rjbe.v3.i3.502

1. Introduction

1.1 Financial Technology, Default Risk, Peer-to-Peer Lending, and Information Asymmetry

The field of financial technology (FinTech) integrates finance, technology management, and innovation management (Arner et al., 2016). Entire business models or even new businesses are frequently opened as a result of FinTech initiatives (Gimpel et al., 2021, as cited in OECD, 2017). FinTech is a potential industry for businesses, as evidenced by its rapid growth: global FinTech funding climbed to £88.2 billion in 2018, up from £40.08 billion in 2017 (Pollari & Ruddenklau, 2018). Two hundred and twenty-two FinTech businesses received an average investment of £15 million, or a total of £2.9 billion, in the years before 2017 (Gulamhuseinwala et al., 2015). As a result, online social media platforms like Facebook, LinkedIn, and MySpace have been used as the basis for new business models that have the potential to upend established trade in various industries. According to Zion Market Research (2018), the global market for mobile phone payment technology was worth £2,660.87 billion in 2024.

Financial lending is one field that has become a prime example of "social commerce" (the use of social networks for commercial activities), with peer-to-peer platforms transforming traditional lending models by directly connecting borrowers and investors online (Bachmann et al., 2011; Kwon, 2012; Chen & Han, 2012). One of the oldest financial processes is lending, but FinTech is changing how it operates, including who can obtain loans and how risks are evaluated. However, there are still a few fundamental ideas that underpin lending, and default risk is among the most important. Default risk is the risk that a lender assumes when a borrower fails to make the required payments on their debt obligation (Emekter et al., 2015). Almost every loan extension increases the risk of default for lenders and investors. Higher necessary returns translate into higher interest rates as default risk increases. When a lender extends credit to a borrower, there is a risk that the loan will not be fully repaid. The metric that takes this potential into account is the default risk. As FinTech is becoming increasingly crucial in lending decisions, the probability of borrowers defaulting on payments is also high due to the large number of participants on the platform seeking loans.

The global financial sector could undergo significant changes thanks to social lending. Since the market is largely set up online, online peer-to-peer (P2P) lending can connect any lender and borrower. P2P lending is an internet extension of the decision-making process based on personal credit. This innovative financial sector has experienced rapid growth in numerous other nations, including the United States, Denmark, Japan, and China, since the establishment of the first online peer-to-peer lending platform. In the online peer-to-peer lending market, borrowers and lenders can connect digitally and carry out loan transactions without the involvement of traditional financial institutions. Although it is new and evolving, the goal is to reduce the borrowing gap between small and medium-sized businesses and individual borrowers, thereby augmenting the loans provided by mainstream financial institutions. Peer-to-peer lending has been suggested as a new e-commerce trend in the financial industry, facilitated by online platforms. Economic efficiency could be greatly enhanced by it (Berger & Gleisner, 2009; Milne & Parboteeah, 2016). The vast majority of borrowers with low net worth and small loan amounts now have the opportunity to obtain funding through P2P marketplace platforms (Bachmann et al., 2011; Lin et al., 2017; Freedman & Jin, 2014). Moreover, P2P promotes lending and borrowing within neighborhoods, regions, and global markets.

The information asymmetry between borrowers and lenders, which prevents lenders and the online platform from understanding whether or not borrowers are trustworthy, is a significant problem in the online peer-to-peer lending sector. Individual lenders, unlike banks, lack the tools and experience to effectively exploit borrowers' credit information. Borrowers might provide a personal statement with other facts in addition to financial information in the hopes of increasing the likelihood of receiving credit. Moral hazard and adverse selection are still significant issues in digital lending platforms due to information asymmetry (Akerlof, 1978; Stiglitz & Weiss, 1981). Such difficulties can be addressed in different ways. Peer-to-peer lending, which eliminates the need for conventional intermediaries like banks, has revolutionized the way people borrow and lend money. P2P platforms have provided more intelligent, technologically advanced methods to safeguard all parties, in contrast to traditional lenders, who require a co-signer or mountains of paperwork. Imagine not requiring a formal guarantor by having your online community vouch for you or by using your Bitcoin as automatic collateral. By combining money to offset possible losses, some platforms even establish safety nets.

According to Guo et al. (2016), the default rate in P2P lending marketplaces may be significantly higher than that of traditional financial intermediaries because lenders do not undergo the screening process and have less experience managing credit risk. The study of P2P default risk is complicated by information asymmetry. Banks do not have blind spots like P2P lenders do. P2P platforms frequently rely on limited, self-reported data, whereas banks examine your complete credit history, income verification, and spending patterns. This leaves a dangerous gap, similar to attempting to assess the condition of a secondhand car based solely on its gleaming appearance. High gains or compelling stories may convince investors to ignore warning signs, which could result in more defaults than anticipated. This could eventually poison the well since only the riskiest borrowers remain as defaults increase, driving out legitimate users. P2P lending runs the risk of turning into a haven for bad loans in the absence of greater transparency, keeping regular investors in the dark.

Research on peer-to-peer lending reveals that, whereas international markets rely more on soft data, U.S. studies are dominated by FICO-based models (Emekter et al., 2015). Lin et al. (2017) discovered that in China, marital status and education are more important than financial measures. Debt-to-income ratios are universally important, according to cross-market assessments (Chen & Han, 2012; Chen et al., 2020), but in areas where credit bureaus are deficient, employment stability and social data become more important. Given this, the objective of this paper is to assess the default risk of P2P borrowers using loan information from an Estonian peer-to-peer lending platform. The study's primary goal is to examine borrower default risk using loan data from the Bondora peer-to-peer lending platform in Estonia. Specifically, the paper examines the relationship between borrowers' default risk and their personal characteristics, loan characteristics, and other factors.

The current study was guided by the following hypotheses: H_1: The default status and age of borrowers will have a statistically significant relationship. H_2: The default status and amount borrowed will have a statistically significant relationship. H_3: The default status and interest payment on money borrowed will have a statistically significant relationship. H_4: The default status and monthly pay of borrowers will have a statistically significant relationship. H_5: The default status and debt-to-income ratio of borrowers will have a statistically significant relationship. H_6: The default status and gender of borrowers will have a statistically significant relationship. H_7: The default status and educational level of borrowers will have a statistically significant relationship. H_8: The default status and income level of borrowers will have a statistically significant relationship. H_9: The default status and marital status of borrowers will have a statistically significant relationship. H_{10}: The default status and duration of work of borrowers will have a statistically significant relationship. H_{11}: The default status and work experience of borrowers will have a statistically significant relationship. The study seeks to answer the question: What are the personal characteristics, loan characteristics, and other contributing factors of borrowers that influence the default rate?

1.2 Relationship between Personal Characteristics and Default Risks

Using a random sample of 100 chicken farmers in Ijebu Ode Local Government Area of Ogun State, Nigeria, Oni et al. (2005) explored and examined the causes of loan default among these farmers. They discovered that a farmer's age, income, and educational attainment have a substantial impact on loan repayment default. Lin, Li, and Zheng (2017) examined the variables that affect default risk based on the demographic features of borrowers using data from a significant P2P lending platform in China. Their empirical findings showed that loan defaults were significantly influenced by factors such as gender, age, marital status, level of education, years of employment, firm size, monthly payment, loan amount, debt-to-income ratio, and delinquent history, according to Lin et al. (2017).

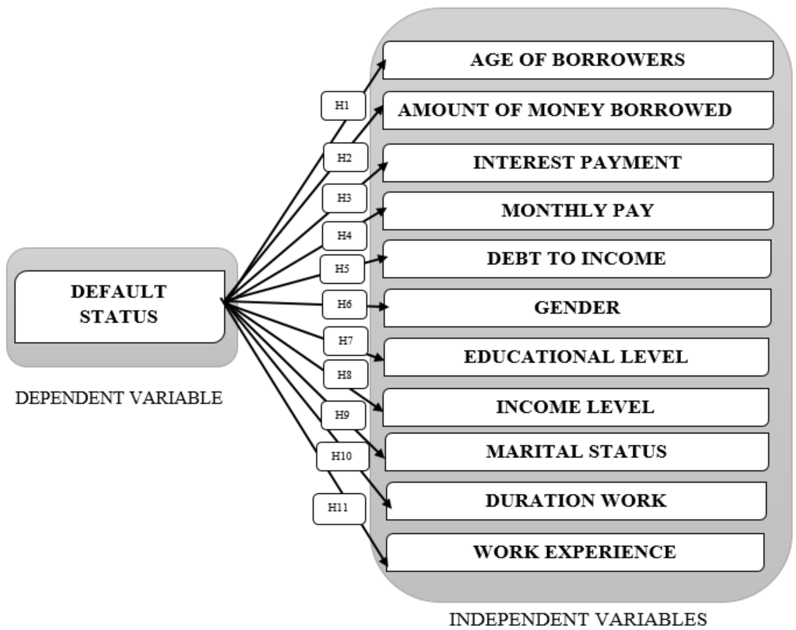
Duarte, Siegel, and Young (2012) found that borrowers who presented themselves as more reliable had higher credit scores, better odds of loan approval, and lower default rates. The importance of social networks is a characteristic shared by all peer-to-peer lending platforms: a loan should be less risky if a person who knows the borrower well can attest to the borrower's creditworthiness. According to a study by Navarro-Galera et al. (2015), population, societal, and financial characteristics affect the likelihood of default, and creating a loan pricing model to help local governments calculate the interest rate in line with their credit risk premium.

Younger borrowers typically have better success rates and tend to get funded more easily and repay reliably, but a racial disparity exists for younger borrowers between the investment success rate and the interest rate paid (Barasinska & Schäfer, 2014; Ravina, 2008). Although their analysis indicated that women received systematically lower interest rates, suggesting latent bias in pricing despite equal access to credit, Barasinska and Schäfer (2014) found no significant gender differences in funding success rates on a major German P2P platform. Ma and Wang (2016) found a direct connection between credit risk and the auditing system of a P2P lending platform. Eid et al. (2016) found that borrowers who choose to round their income declarations are more likely to fail and less likely to repay than borrowers who accurately disclose their income.

Using logistic regression, Wiginton (1980) and Steenackers and Goovaerts (1989) analyzed unsecured personal loans. According to Steenackers and Goovaerts (1989), significant predictors of default include borrower age, length of employment, occupation, monthly income, homeownership, and loan term. In addition to income and occupation, Tsai et al. (2009) found that several measures of money attitude, as measured by a survey administered to borrowers, are significant predictors of default. Borrower age, credit inquiries, credit use, and credit limit significantly predict default (Carling, Jacobson, & Roszbach, 1998). Loan size was found to be statistically insignificant in predicting default risk in Roszbach's (2004) examination of unsecured personal loans, defying accepted lending assumptions. Similar findings were obtained across several risk assessment models (Roszbach, 2004).

Gorton and Winton (1995) examined the empirical support for financial intermediaries and advocated a wide range of roles for banks. Two of these positions that are closely related to the problem of information asymmetry are banks acting as information providers and banks acting as delegated. If P2P lending is to succeed, these positions must be filled in some way. P2P borrowers frequently share non-financial personal details to increase the chances of getting a loan, unlike users of conventional financial intermediaries. According to Chen and Han's (2012) analysis of P2P lending practices in China and the USA, lenders in China rely more on "soft" credit information. However, "hard" and "soft" credit information both have significant effects on loan outcomes in these two countries. Considering available data, this study will consider these variables from the Bondora P2P lending platform of each loan case, which include age, gender, marital status, educational level, and employment duration with current employer, work experience, loan amount, income, monthly payment, debt-to-income ratio, interest, and loan default status.

Figure 1: Conceptual Framework



Note. The conceptual framework used for analyzing loan default status. The framework shows the relationship between the dependent and independent variables.

2. Method

2.1 Data Collection

The information utilized in this research is drawn from all loan postings published on the Estonian P2P lending platform Bondora between March 2009 and June 2022.

2.2 Measures

The variables for each loan case, according to the provided data, included the borrower's age, gender, marital status, level of education, length of employment with the current employer, work experience, loan amount, income, monthly payment, debt-to-income ratio, interest rate, and loan status. Among these variables, borrowers' gender, educational level, marital status, employment duration with current employer, work experience, income, and loan default status were categorical variables. The borrower's gender was specifically categorized as male or female. Educational level was divided into five categories: 1 (Primary education), 2 (Basic education), 3 (Vocational education), 4 (Secondary education), and 5 (Higher education). Marital status also had five levels: 1 (Married), 2 (Cohabitant), 3 (Single), 4 (Divorced), and 5 (Widowed). Employment duration with current employer was divided into nine levels: 1 (trial period), 2 (up to 1 year), 3 (up to 2 years), 4 (up to 3 years), 5 (up to 4 years), 6 (up to 5 years), 7 (more than 5 years), 8 (retirees), and 9 (other). Borrowers' working experience was categorized as follows: 1 (less than 2 years), 2 (2 to 5 years), 3 (5 to 10 years), 4 (10 to 15 years), 5 (15 to 25 years), and 6 (more than 25 years). Also, monthly income level of the borrower was divided into seven levels: 1 (less than 1000), 2 (1001–2000), 3 (2001–5000), 4 (5001–10,000), 5 (10,001–20,000), 6 (20,001–50,000), and 7 (more than 50,000), and lastly, loan default status had two levels, 1 (if the loan has been defaulted), and 0 (otherwise). However, the following variables were not grouped: age, loan amount, interest, monthly payment, and debt-to-income. Refer to Table 1 for the detailed description of the variables used in this paper.

Table 1*: Names of variables and their definitions*
Note. Key variables assessing borrower characteristics and loan performance.

Variable	Name	Definition
Age	Age	Age of a borrower (in years)
Gender	Male	1 if a borrower is male, and 0 otherwise (female)
Educational Level	Education	Level of education of a borrower is divided into five levels: 1 (Primary education); 2 (Basic education); 3 (Vocational education); 4 (Secondary education); 5 (Higher education)
Marital Status	Married	Married status of a borrower is divided into five levels: 1 (Married); 2 (Cohabitant); 3 (Single); 4 (Divorced); 5 (Widow)
Employment Duration with Current Employer	Duration	Years of employment of a borrower is divided into nine levels: 1 (Trial period); 2 (Up to 1 year); 3 (Up to 2 years); 4 (Up to 3 years); 5 (Up to 4 years); 6 (Up to 5 years); 7 (More than 5 years) and 8 (Retirees), and 9 (Other).
Work Experience	Worktime	Borrowers' working experience: 1 (less than 2 years); 2 (2 to 5 years); 3 (5 to 10 years); 4 (10 to 15 years); 5 (15 to 25 years), and 6 (More than 25 years).
Monthly Income Level	Income	Monthly income level of the borrower is divided into seven levels: 1 (less than 1000); 2 (1001–2000); 3 (2001–5000); 4 (5001–10,000); 5 (10,001–20,000); 6 (20,001–50,000), and 7 (more than 50,000).
Loan Amount	Amount	Loan amount requested by the borrower
Interest	Interest	The annual interest rate that a borrower pays on the loan
Monthly Payment	Payment	Monthly payment made by the borrower
Debt To Income	Debt	Debt to income ratio of the borrower
Loan Default Status	Default	1 if the loan has been defaulted, and 0 otherwise

2.3 Data Description
From the above table, we summarize and discuss the descriptive statistics of the paper's data, including the borrower's characteristics and the condition of the loan. The 103,326 loan requests used in this paper were obtained from https://www.bondora.com/en/public-reports, a significant Estonian peer-to-peer lending platform. We acquired 103,326 valid loan instances after excluding loan cases with missing value variables through data preprocessing and filtering. The combined total of all of these loans is around €270 million. Table 2 displays the loan status for each of the requested loans. The platform offered a total of 103,326 loans, of which 9,614 had late payments, resulting in a loan default status. It is conceivable for these defaulted loans to be converted into a loan default rate of 9.82 percent; however, this rate may be skewed lower because the default rate rises continuously as loans mature (Emekter et al., 2015).

2.4 Data Analysis and Presentation
The study proposes a credit risk assessment model to assess the default risk of borrowers based on their demographic characteristics and associated loan data from the default rate perspective. Specifically, the paper employs a quantitative approach, more precisely a binary logistic regression method, to examine the relationship

between borrowers' default risk and their personal characteristics, loan characteristics, and other contributing factors. This will establish a model that provides an effective method for assessing credit risk on lending platforms.

This straightforward binary classification problem is effectively solved by logistic regression. The findings of the regression model are crucial, as they will help determine the likelihood that a certain borrower will default or the probability that the dependent variable will take the value of 1. The possibility of the event occurring is represented by the dependent variable in the Binary Logistic Regression (BLR). It is equivalent to a default in this instance, supposing z is an unobserved continuous number. It indicates the likelihood of a default occurring. As a result, a greater value for z denotes a higher possibility of default. This continuous number must be changed into a number between 0 and 1 to be usable. We utilized the conversion as follows:

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

where p is the likelihood that a default will occur. Additionally, assuming that the k explanatory variables in the BLR have a linear correlation with z, the following description can be used for this model:

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_k x_k + \mu \tag{2}$$

The random error term is μ where k is the number of explanatory variables, x is the explanatory variable (Hosmer & Lemeshow, 2000). The empirical findings of the previously discussed BLR model are shown in Table 5.

3. Comparative Analysis and Results

The study specifically examined the relationship between borrower default risk and personal characteristics, loan characteristics, and other factors. We used the Statistical Package for Social Sciences (SPSS, Version 28.0.1.1) for data analysis. The data analysis method employed was logistic regression analysis.

Table 2*: Loan Distribution of Borrowers by Loan Status*

Loan Status	Number of Loans	Percent	Amount	Percent
Current	67873	65.69	182,624,032	67.49
Late	9614	9.30	26,572,929	9.82
Repaid	25839	25.01	61,401,825	22.69
Total	103326	100.00	270598786	100.00

Note. The distribution of loans by loan status, number of loans, percentage of loans, and amount of loans for borrowers in each category.

Table 2 illustrates that 9.82 percent of all loan requests, totaling approximately €27 million, may not be repaid because borrowers are not making their monthly payments or paying them back on time. Additionally, about €61 million, or 22.69 percent of these loans, have been fully returned. Additionally, there are already about €183 million worth of debts, or 67.49 percent outstanding. Lenders' top priority is whether borrowers will fail in the future; therefore, this is their biggest worry. If any distinctive characteristics of the borrower can be used to forecast their likelihood of failing, lenders will benefit.

3.1 Descriptive Statistics

Tables 3 and 4 display the descriptive statistics of the loan data used in this study, together with the general characteristics of borrowers and loans. According to Table 3, regarding the 103,326 loan requests, 51% are from males and 49% are from females. The average age of a borrower is about 39.67 years. Regarding borrowers' education level, 12.2% had primary education, comprising a total of 12,576 borrowers, 4.6% had basic education, comprising a total of 4,742 borrowers, and 29% had vocational education, comprising a total of 29,945 borrowers. 32.8% of borrowers had secondary education, totaling 33,859, and 21.5% had higher education, totaling 22,204. With marital status of the borrowers, 75.1% of them were married, making up a total number of 77,622 borrowers, 9.1% were cohabitants, making up a total number of 9,372 borrowers, 12% were also single, making up a total of 12,366 borrowers, 3.3% divorced, making up a total number of 3,366 and the remaining 0.6% of the borrowers were widowed, making up a total of 600.

Table 3: *Descriptive Statistics*

Variables	Frequency	Percentage
Gender:		
Male	52676	51.0
Female	50650	49.0
Education:		
Primary	12576	12.2
Basic	4742	4.6
Vocational	29945	29.0
Secondary	33859	32.8
Higher	22204	21.5
Marital status:		
Married	77622	75.1
Cohabitant	9372	9.1
Single	12366	12.0
Divorced	3366	3.3
Widow	600	0.6
Employment duration:		
Trial Period	525	0.5
Up to 1 year	18755	18.4
Up to 2 years	5066	2.7
Up to 3 years	4169	2.3
Up to 4 years	2897	1.5
Up to 5 years	22076	24.9
More than 5 years	37049	35.9
Retirees	4741	4.6
Other	5733	5.5
Work experience:		
Less than 2 years	2089	5.7
2 to 5 years	5123	14.0
5 to 10 years	7834	21.4
10 to 15 years	6900	18.9
15 to 25 years	7903	21.6
More than 25 years	6675	18.3
Income level of borrower:		
less than 1000	33644	32.6
1001–2000	42861	41.5
2001–5000	23730	23.0
5001–10,000	1605	1.6
10,001–20,000	1026	1.0
20,001–50,000	358	0.3
More than 50,000	100	0.1

Note. Descriptive statistics for various borrower attributes.

Concerning working duration with current employer, it was observed that the majority (60.8%) have worked with their current employers for 5 years or more, while the least (0.5%) were on trial periods. Regarding work experience, it was observed that the majority (about 62%) have been working for 5 to 25 years, while the smallest group was those who have worked for less than 2 years, representing 5.7%. For the income level of borrowers, the majority (41.5%) earned between €1,000 and €5,000, while only 0.1% earned more than €50,000 per month. Additionally, 9.82% of the borrowers had a history of defaulting, and 91.18% of them had never defaulted (refer to Table 2). On average, borrowers earned €2,118.02 per month (refer to Table 4). The average loan amount was approximately €2,618.88 (see Table 4), and the monthly payment was typically €105.67 (refer to Table 4). Debt-to-income was around 10.095 (refer to Table 4).

Table 4: *Maximum, minimum, skewness, and kurtosis statistics of the variables used in the study*

Variable	N	Min	Max	Mean	SD	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	S.E	Statistic	S.E
Age	103326	0	77	39.67	12.211	.399	.008	-.628	0.015
Amount	103326	6	10632	2618.88	2133.68	1.350	.008	2.059	.015
Interest	103326	2	264	27.20	17.416	5.591	.008	50.576	.015
Monthly payment	96686	0	2369	100.55	104.852	3.911	.008	32.176	.016
Income	103326	0	800900	2118.02	12099.7	52.582	.008	3056.096	.015
Debt to income	103276	0	198	10.09	17.883	1.769	.008	2.149	.015
Gender	103281	0	2	.52	.556	.440	.008	-.862	.015
Loan status	103326	0	1	.09	.290	2.802	.008	5.850	.015
Education	103326	1	5	3.47	1.225	-.644	.008	-.357	.015
Marital status	103276	1	5	1.45	.867	1.824	.008	2.344	.015
Current employment duration	101011	1	9	5.60	2.157	-.631	.008	-.844	.015
Work experience	36430	1	6	3.92	1.493	-.191	.013	-1.006	.026

Note. Statistical characteristics of different key borrower variables.

4. Results

We examined the factors in this case that affect the likelihood of loan default. To precisely determine the impact of each variable on the default loan rate, we used Binary Logistic Regression (BLR), which included all the parameters in the current work. Because the dependent variable in general regression cannot be classified as binary, we used the BLR model. The BLR is a solid solution to this simple binary classification issue. The findings of the regression model are crucial, as we anticipated finding the likelihood that a certain borrower would default, or the chance that the dependent variable would take the value 1. The dependent variable in the BLR represents the possibility of the event occurring. It is equivalent to a default in this instance. The empirical findings of the previously discussed BLR model are shown in Table 5.

Table 5: *Binary Logistic Regression Analysis*

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
Age	0.0270	0.002	173.942	1	0.000**	1.027
Amount	0.0001	0.000	67.842	1	0.000**	1.000
Interest	0.0250	0.001	1057.481	1	0.000**	1.025
Monthly payment	-0.0003	0.000	6.468	1	0.011*	1.000
Debt to income	-0.0010	0.001	2.128	1	0.145	0.999
Gender: Reference: Female						
Male	0.2110	0.028	55.115	1	0.000**	1.235
Education: Reference: Higher						
Primary	0.7970	0.125	40.719	1	0.000**	2.218
Basic	0.2800	0.047	35.216	1	0.000**	1.323
Vocational	0.6020	0.038	250.140	1	0.000**	1.826
Secondary	-0.0310	0.036	0.766	1	0.381	0.969
Marital Status: Reference: Widowed						
Married	0.2290	0.114	4.069	1	0.044*	1.257
Cohabitant	0.0110	0.117	0.010	1	0.922	1.012
Single	0.4570	0.116	15.574	1	0.000**	1.579
Divorced	0.4470	0.117	14.482	1	0.000**	1.564
Duration of employment: Reference: More than 5 years						
Trial period	-0.0220	0.116	0.037	1	0.848	0.978
Up to 1 year	-0.0540	0.042	1.619	1	0.203	0.948
Up to 2 years	0.0460	0.045	1.077	1	0.299	1.048
Up to 3 years	-0.0830	0.048	2.963	1	0.085	0.920

Up to 4 years	0.1000	0.053	3.569	1	0.059	1.105
Up to 5 years	0.0590	0.052	1.294	1	0.255	1.061
Work experience: Reference: More than 25 years						
Less than 2 years	-0.0450	0.093	0.236	1	0.627	0.956
2 to 5 years	0.1800	0.072	6.233	1	0.013 *	1.197
5 to 10 years	0.1650	0.062	7.051	1	0.008 **	1.179
10 to 15 years	0.1930	0.055	12.365	1	0.000 **	1.213
15 to 25 years	0.1460	0.046	9.917	1	0.002 **	1.157
Level of income: Reference: 50001 to 10000						
Less than 1000	0.7490	0.034	496.202	1	0.000 **	2.115
1001 to 2000	0.2750	0.036	57.258	1	0.000 **	1.316
2001 to 5000	-0.3180	0.201	2.489	1	0.115	0.728
10001 to 20000	-0.1220	0.442	0.076	1	0.782	0.885
20001 to 50000	0.3960	1.444	0.075	1	0.784	1.485
More than 50000	0.0350	1.047	0.001	1	0.973	1.036
Constant	-4.0600	0.178	517.647	1	0.000 **	0.017

Note. Results of binary logistic regression analysis examining factors influencing loan default status. *p <.05; **p <.01.

Age; Amount; Interest; Monthly Payment; Debt To Income; Education is categorized as Primary, Basic, Vocational, Secondary and Higher for 1,2,3,4 and 5 respectively, with 5 being the reference; Marital Status is categorized as Married, Cohabitant, Single, Divorced, and Widowed for 1,2,3,4 and 5 respectively, with 5 being the reference; Current Employment Duration is grouped as Trial Period, Up To 1 Year, Up To 2 Years, Up To 3 Years, Up To 4 Years, Up To 5 Years, and More than 5 years for 1,2,3,4,5 6 and 7 respectively, with 7 being the reference; Work Experience is grouped as Less Than 2 Years, 2 To 5 Years, 5 To 10 Years, 10 To 15 Years, 15 To 25 Years, and More than 25 years for 1,2,3,4,5 and 6 respectively, with 6 being the reference; Level of Income is categorized as Less than 1000, 1001 to 2000, 2001 to 5000, 10001 to 20000, 20001 to 50000, More than 50000 for 1,2,3,4,5 6 and 7 respectively, with 7 being the reference.

Nine variables (age, gender, marital status, educational level, loan amount, work experience, monthly payment, income, and interest) out of the 11 factors in the BLR model statistically substantially affected loan default. All of the estimated coefficients' results were 1% significant. Debt-to-income and employment duration, however, were statistically insignificant. With a chi-square value of 241.973 and a significance of 0.000, Hosmer and Lemeshow's test also revealed that the model was adequate for explaining loan default (Hosmer & Lemeshow, 2000). Additionally, because all standard errors (SEs) of the estimated coefficients were substantially smaller than two, the estimated BLR model did not have a multicollinearity problem. Similarly, the Cox and Snell R^2 and Nagelkerke R^2 for this BLR model were 14.4% and 20.2%, respectively. As can be seen from Table 5, we arrived at numerous conclusions:

i. With this data sample, neither debt-to-income ratio nor duration of employment had a significant effect on the loan default rate.
ii. Female borrowers had a lower default rate than male borrowers.

iii. Borrowers' default rates were significantly impacted by their single or divorced status.

iv. The probability of default is highly correlated with education level. Specifically, the secondary education level did not have an impact on the default rate, but the others did.

v. Borrowers with up to 4 years of current employment duration were at risk of default. The default rate dropped by 8.7% of the initial rate for each year that a borrower with up to three years of employment increased their present employment. For up to 4 years of employment duration, the default rate increased by 9.2% in that respect.

vi. The probability of a borrower defaulting on a loan increased with age. For instance, a borrower's risk of default rose by 2.8% for each year of age.

vii. High loan amounts were associated with a high default rate.

viii. The risk of default increased as a borrower's monthly repayment obligation increased.

ix. A borrower's low income level was also strongly correlated with the rate of loan default. The default rate rose as income levels dropped and vice versa.

Using the following model and the estimated coefficients shown in Table 5, we can determine the likelihood that a specific loan will default based on the regression results. When the estimates from Table 5 are substituted into (1), we obtain (2) as stated below.

$$\begin{aligned} Z = &-4.0600 + 0.02700(AG) + 0.00010(\mathrm{LA}) + 0.0250(IN) \\ &-0.0003(\mathrm{MP}) - 0.0010(\mathrm{DI}) + 0.211(Ma) + 0.797(PR) \\ &+0.280(BS) + 0.6020(VC) - 0.03100(SR) + 0.2290(MA) \\ &+0.110(CO) + 0.4570(SG) + 0.4470(DV) - 0.0220(TP) \\ &-0.0540(1Y) + 0.0460(2Y) - 0.08300(3Y) + 0.1000(4Y) \\ &+0.0590(5Y) - 0.0450(M1) + 0.1800(M2) + 0.193(M3) \\ &+0.193(M4) + 0.1930(M5) + 0.7490(X1) + 0.2750(X2) \\ &-0.3180(X3) - 0.1220(X4) + 0.3960(X5) + 0.035(X6) \end{aligned}$$

That is, anytime we want to estimate the probability of a borrower defaulting, provided the borrower's characteristics for the variables in the model are known, that can be done using the estimated model. However, the variables in the equation and their symbols are presented in Table 6.

Table 6: Variables in the Equation and their Symbols

Variables	Abbreviation
Age	AG
Loan amount	LA
Interest	IN
Monthly payment	MP
Debt to income	DI
Gender:	
Male	Ma
Education:	
Primary	PR
Basic	BS
Vocational	VC
Secondary	SR
Marital status:	
Married	MA
Cohabitant	CO
Single	SG

Divorced	DV
Employment duration:	
Trial Period	TP
Up to 1 year	1Y
Up to 2 years	2Y
Up to 3 years	3Y
Up to 4 years	4Y
Up to 5 years	5Y
Work experience:	
Less than 2 years	M1
2 to 5 years	M2
5 to 10 years	M3
10 to 15 years	M4
15 to 25 years	M5
Income level of borrower:	
Less than 1000	X1
1001–2000	X2
2001–5000	X3
5001–10,000	X4
10,001–20,000	X5
20,001–50,000	X6

Note. Variables used in the binary logistic regression model and their respective abbreviations.

5. Discussion

Credit risk is a major concern in P2P lending systems. In this paper, we examined several factors that influence default risk and proposed a model to predict the likelihood of default for different peer-to-peer (P2P) loans. According to the paper, borrowers' default risk characteristics included age, gender, marital status, level of education, length of time employed in the current job, loan amount, income, monthly payment, and interest rate. The findings imply that, in addition to the loan applicants' friendships and social connections (Freedman and Jin, 2014; Lin, Prabhala, and Viswanathan, 2013), nine of the eleven indicators listed above are crucial in determining the chance that loans may fail. Additionally, the results indicate that women have a considerably lower default rate than males, which is consistent with Chen, Li, and Lai (2017). Thus, women tend to exhibit less risk-seeking behavior than males when making financial decisions (Powell & Ansic, 1997).

Married borrowers have reduced default probabilities, according to P2P research conducted in both Eastern (Lin et al., 2017) and Western markets (Navarro-Galera et al., 2015), because divorced debtors must care for their families by themselves, which may put them in a precarious financial situation and increase the likelihood that they would default on their loans in the future. Because the default rate for highly educated borrowers is significantly lower, it is suggested that higher educational levels will reduce the likelihood of loan default and vice versa. Highly educated individuals are more aware of their position and respect, and are therefore more inclined to pay their debts on time. On the other hand, a person who understands the value of borrowing may be more likely to hold a higher degree (Lin et al., 2017; Oni et al., 2005).

Work history and age have a U-shaped association with default risk; it is higher for both early entrants and near-retirees, while it is lowest for established mid-career borrowers (35–50 years old with 5–15 years of experience) (Emekter et al., 2015). Surprisingly, P2P borrowers with higher debt-to-income ratios often default to a lesser extent than expected. Emekter et al. (2015) found this happens because platforms offset the risk with stricter terms (similar to higher rates), creating a counterintuitive safety net. This unusual phenomenon is most likely explained by the fact that debtors with low debt-to-income ratios may not be concerned about the financial consequences of default, and that the cost of late payment is negligible. A borrower's risk of defaulting on future loans increases significantly if he has a history of doing so. This paper suggests that lenders should not offer credit to borrowers with a history of delinquency; it also highlights the value of credit and the high costs associated with loan default for borrowers.

5.1 Limitations and Further Directions

Using loan data from an Estonian P2P lending network, this study evaluates borrowers' default risk by focusing on soft information: personal characteristics, loan details, and other borrower-provided factors that help assess repayment likelihood. While these insights are valuable, we recognize their limitations. Estonia's lending landscape, though insightful, cannot fully represent the diverse financial behaviors across Europe or beyond. That careful borrower in Tallinn might manage debt very differently from an entrepreneur in Lisbon or a gig worker in Berlin. To strengthen these findings, future research should expand to include more countries, to enable generalizing results and uncover cultural or economic nuances in borrowing behavior. This broader approach would provide richer data for policymakers and lenders alike. Additionally, while our dataset offers meaningful patterns, it is incomplete, thus neglecting rejected loan applicants, for instance, which remain a hidden piece of the risk puzzle. A larger, more diverse sample, spanning multiple markets, would make findings more robust and widely applicable. Beyond geographic expansion, future studies could delve deeper into practical metrics such as expected profit, loss, repayment ratios, as well as tools that would help lenders make more informed decisions. Even with these limitations, however, our study provides a solid foundation, using real P2P lending data to shed light on what drives borrower defaults. The goal is not just better risk models, but fairer, more transparent lending systems that work for borrowers and investors alike.

6. Conclusion

Table 5 shows several possible outcomes. Debt-to-income, employment duration, and income level did not significantly affect the loan default rate in this dataset. Female borrowers had a lower default rate than male borrowers. Being a single or divorced person had a greater impact on the likelihood of a borrower defaulting. The risk of default was highly correlated with the level of education. Borrowers who had been employed for up to 4 years were at risk of default. A borrower with up to three years of current employment saw an 8.7% reduction in the default rate for every additional year. For up to 4 years of employment, the default rate increased by 9.2% in that respect. The probability of a borrower defaulting on a loan increased with age. For instance, a borrower's risk of default rose by 2.8% for each year of age. Loans with large loan amounts had a higher default probability. Moreover, the risk of default increased as a borrower's monthly repayment obligation increased. For each monthly payment of €100, a borrower's chance of default decreased by 0.003%. The borrower's debt-to-income ratio had a minimal impact on the default rate. The default rate reduced to 0.1% of the initial amount as the debt-to-income ratio rose by 100%.

Fosu Acheampong https://orcid.org/0009-0006-2149-2552
Albert Junior Nyarko https://orcid.org/0000-0001-7422-1170
Conflict of interest: We have no known conflict of interest to disclose.

References

[1] Akerlof, G. A. (1978). The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235–251). Academic Press.
[2] Arner, D. W., Barberis, J., & Buckley, R. P. (2016). 150 years of Fintech: An evolutionary analysis. *Jassa, 3*, 22–29.
[3] Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, M., ... & Funk, B. (2011). Online peer-to-peer lending: A literature review. *Journal of Internet Banking and Commerce, 16*(2), 1.
[4] Barasinska, N., & Schäfer, D. (2014). Is crowdfunding different? Evidence on the relation between gender and funding success from a German peer-to-peer lending platform. *German Economic Review, 15*(4), 436–452.
[5] Berger, S. C., & Gleisner, F. (2009). Emergence of financial intermediaries in electronic markets: The case of online P2P lending. *BuR Business Research, 2*(1), 39–65.
[6] Carling, K., Jacobson, T., & Roszbach, K. (1998). Duration of consumer loans and bank lending policy: Dormancy versus default risk (No. 70). *Sveriges Riksbank Working Paper Series.*

[7] Chen, D., & Han, C. (2012). A comparative study of online P2P lending in the USA and China. *Journal of Internet Banking and Commerce, 17*(2), 1.

[8] Chen, D., Li, X., & Lai, F. (2017). Gender discrimination in online peer-to-peer credit lending: Evidence from a lending platform in China. *Electronic Commerce Research, 17*(4), 553–583.

[9] Chen, S., Gu, Y., Liu, Q., & Tse, Y. (2020). How do lenders evaluate borrowers in peer-to-peer lending in China? *International Review of Economics & Finance,* 69, 651–662.

[10] Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies, 25*(8), 2455–2484.

[11] Eid, N., Maltby, J., & Talavera, O. (2016). Income rounding and loan performance in the peer-to-peer market (MPRA Paper No. 72852). University Library of Munich, Germany. https://ideas.repec.org/p/pra/mprapa/72852.html

[12] Emekter, R., Tu, Y., Jirasakuldech, B., & Lu, M. (2015). Evaluating credit risk and loan performance in online peer-to-peer (P2P) lending. *Applied Economics, 47*(1), 54–70.

[13] Freedman, S., & Jin, G. Z. (2014). The signaling value of online social networks: Lessons from peer-to-peer lending. *National Bureau of Economic Research*. https://doi.org/10.3386/w19820

[14] Gimpel, H., Hosseini, S., Huber, R., Probst, L., & Röglinger, M. (2021). Entrepreneurial, institutional and financial strategies for FinTech profitability. *Financial Innovation, 8*(15). https://doi.org/10.1186/s40854-021-00325-2

[15] Gorton, G. B., & Winton, A. (1995). Bank capital regulation in general equilibrium. *National Bureau of Economic Research Working Paper*. https://doi.org/10.3386/w5244

[16] Guo, R., Zhang, H., & Yu, L. (2016). Instance-based credit risk assessment for investment decisions in P2P lending. *Electronic Commerce Research and Applications, 18*, 50–61. https://doi.org/10.1016/j.elerap.2016.07.001

[17] Gulamhuseinwala, I., Bull, T., & Lewis, S. (2015). FinTech is gaining traction and young, high-income users are the early adopters. *Journal of Financial Perspectives, 3*(3).

[18] Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). Wiley-Interscience.

[19] Kwon, O. (2012). A new ensemble method for gold mining problems: Predicting technology transfer. *Electronic Commerce Research and Applications, 11*(2), 117–128.

[20] Lin, M., Prabhala, N. R., & Viswanathan, S. (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.

[21] Lin, X., Li, X., & Zheng, Z. (2017). Evaluating borrower's default risk in peer-to-peer lending: Evidence from a lending platform in China. *Applied Economics, 49*(35), 3538–3545.

[22] Ma, H.-Z., & Wang, X.-R. (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. https://doi.org/10.1080/09720529.2016.1178935

[23] Milne, A., & Parboteeah, P. (2016). The business models and economics of peer-to-peer lending. *ECRI Research Report No. 17*.

[24] Navarro-Galera, A., Rayo-Cantón, S., Lara-Rubio, J., & Buendía-Carrillo, D. (2015). Loan price modelling for local governments using risk premium analysis. *Applied Economics, 47*(58), 6257–6276. https://doi.org/10.1080/00036846.2015.1068924

[25] Oni, O. A., Oladele, O. I., & Oyewole, I. K. (2005). Analysis of factors influencing loan default among poultry farmers in Ogun State, Nigeria. *Journal of Central European Agriculture, 6*(4), 619–624.

[26] Pollari, I., & Ruddenklau, A. (2018). The pulse of fintech 2018: Biannual global analysis of investment in fintech. *KPMG International*. https://home.kpmg/xx/en/home/insights/2019/02/pulse-of-fintech-h2-2018-global.html

[27] Powell, M., & Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology, 18*(6), 605–628.

[28] Ravina, E. (2008). Love & loans: The effect of beauty and personal characteristics in credit markets. *The Review of Financial Studies, 21*(3), 1211–1250. https://doi.org/10.1093/rfs/hhm060

[29] Roszbach, K. (2004). Bank lending policy, credit scoring, and the survival of loans. *Review of Economics and Statistics, 86*(4), 946–958.

[30] Steenackers, A., & Goovaerts, M. J. (1989). A credit scoring model for personal loans. *Insurance: Mathematics and Economics, 8*(1), 31–34.

[31] Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review, 71*(3), 393–410.

[32] Tsai, M. C., Lin, S. P., Cheng, C. C., & Lin, Y. P. (2009). The consumer loan default predicting model – An application of DEA–DA and neural network. *Expert Systems with Applications, 36*(9), 11682–11690.

[33] Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis, 15*(3), 757–770.

[34] Zion Market Research. (2018, November 30). Global mobile payment technology market will reach USD 3,371.6 billion by 2024. *GlobeNewswire*.