

The interest rate variability within the P2P lending market

Author: Mostafa Ahmed
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,

Peer-to-peer (P2P) lending is an innovative alternative to traditional banking, it connects borrowers directly with individual lenders through online platforms. This research investigates the determinants of interest rate variability in the P2P lending market, by focusing on borrower background, loan term, and loan purpose. By using a regression analysis on data from the Dutch P2P platform Geldvoorelkaar, the research finds that loan term significantly affects interest rates, with longer-term loans associated with lower interest rates. On the other hand, borrower background and loan purpose do not show statistically significant impacts on interest rates, which is quite surprising. These findings provide insights for both borrowers and lenders to optimize their financial decisions in the P2P lending market.

Graduation Committee members:

Huang, Xiaohong. Shen, Lingbo

Keywords:

Interest rates, Peer-to-peer lending/P2P, loan term, loan purpose, borrower background, credit scores, Regression analysis.

1. INTRODUCTION

1.1 Background:

Peer-to-peer lending, also known as social lending or marketplace lending, is a form of online platform that connects borrowers directly with individual lenders, it is a widely known concept recognized as P2P. Peer-to-peer lending has emerged rapidly as an alternative to traditional banking systems, allowing borrowers to access finance effortlessly apart from the traditional methods used and providing investors with genuine opportunities to earn returns by funding loans. (Havrylchuk et al., 2016)

The concept of P2P lending was originated back in the 2000s with platforms like, Zopa (founded mid 2000s in the UK), Faircent (founded in 2013 as India's largest P2P) or Funding Circle (founded in 2010 in the UK), these peer-to-peer organizations aim to connect lenders with credit worthy borrowers. P2P platforms implement various risk management strategies which include credit risk assessment, diversification, risk-based pricing which all play a role in interest rates setting, while minimizing risk for lenders, (Irwanto, 2023)

This model facilitates direct lending between borrowers and lenders without the need for any financial intermediaries such as banks, venture capital, etc.... While peer-to-peer lending seems to act in the same ways, it is crucial to understand the differentiation between both business models. P2P aims to directly connect borrower and lenders by making lenders evaluate loan listings to decide which loans to fund. Lending platforms aim to disintermediate the lending process by eliminating the need for traditional financial institutions, instead borrowers can access funding directly from individual investors/lenders, who get the opportunity to earn returns on said funding.

The significance of this research is important, as it clarifies the various factors involved in determining interest rates. This helps individuals profit from both sides, as it allows lenders to make a return, while providing opportunities for borrowers. The understanding behind the variability helps enhance individuals critical thinking, to smartly/profitably borrow or lend.

1.2 Process of setting interest rates

The process of setting interest rates for P2P lending platforms involves a combination of data analysis, risk assessment, market insights and some regulatory considerations. The main goal of P2P platforms is to strike a balance between attracting borrowers with competitive rates (credit risk) and providing lenders with high returns while effectively managing risks within its ecosystem to mitigate the risks associated with lending money, (Deitrich, 2020). Platforms could potentially use algorithmic pricing models to determine the optimal rates at any given moment, by using advanced mathematical models and data analysis. This method considers many variables such as supply and demand, borrower's characteristics, loan term and loan purpose which is relevant to the determination of interest rates. These variables play an important role in this context as they are the main influential factors for the determination of interest rates for borrowers, and based on the lenders requirements the platforms connect the lender to the right borrower.

1.3 Motivation

Within this topic it is crucial to understand how interest rates are set and how they can be influenced by several different factors.

P2P lending platforms all differ depending on the loan purpose; thus, it is critical to understand how they set their rates considering the multiple factors that depend on the loan characteristics. The understanding of these characteristics is critical, as it contains the main drivers behind interest rates making it important to apprehend the individual effect it can have upon the interest rates. Prior studies have analyzed the P2P determinants but have not mentioned the difference in the importance of the variables, this analysis should allow me to fill a gap, to enable borrowers to understand the importance of each different variable, (Deitrich, 2020).

I would like to study this further to have a better understanding of the factors that affect interest rates, mainly to comprehend the main drivers behind interest rates, how the different variables correlate and affect the rates differently.

These effects are of great importance as they allow trusted borrowers to obtain lower interest rates based on their personal financial history, while allowing less trusted borrowers to secure loans at higher interest rates, this is critical since it helps borrowers understand why they were given certain rates. understanding them allows borrowers and lenders to maximize the benefits when lending or requesting a loan. I will analyze the different impact each variable has on determining interest rates, this will help individuals focus on factors that are more valuable to achieve better rates if in need of a new loan.

Understanding how interest rates work in P2P is crucial for several reasons, it allows borrowers to make better financial decision making, it is helpful for lenders as they can make strategic decisions to maximize their returns while managing risk. Risk management is a part of the process, as understanding risk management measures is essential for effective financial planning and risk management. Finally, it is important for the borrower to understand the reason behind his individual interest rate, to allow him to ameliorate for future loans. This depends on borrower's background thus allowing individuals to enhance their financial history before requesting a loan, to minimize rates while maximizing the potential opportunities that come from loans.

Despite the existing literature, there are several research gaps remaining in understanding the determinants of interest rates. Firstly, while previous studies have examined borrower and loan characteristics, (Sun et al., 2023), there is limited research on the influence of loan purpose on interest rates, which makes me want to validate the gap, so that borrowers can have a better understanding of how loan purpose has a significant effect on interest rates. As it can significantly impact borrower risk and loan terms, yet its effect on interest rates remains underexplored in the literature. The existing research mostly focuses on specific P2P lending platforms which limits the generalizability of findings. This research aims to address these gaps, by using one platform while seeking to provide insights into the factors driving interest rates variability and contribute to the existing literature on this topic.

Research question:

What are the key determinants impacting interest rate variability across individual loans in the P2P consumer lending market?

2. LITERATURE

In this paper I will focus on factors affecting interest rates in P2P lending, as they depend on multiple aspects ranging from borrower background, loan term and the purpose of loan, (Yuliyani & Ulrich, 2023). I will analyze these variables to comprehend their individual effects on interest rates, since based on relevant readings they might be the main drivers that influence the fluctuation of interest rates for borrowers.

The maturity of loan is important, as peer to peer lending would usually use fixed maturity, which I will be focusing on, since variable maturity may have terms that allow for loan term adjustment & interest rates which will be irrelevant for this research. The focus on fixed maturity loans is more logical as it is based off predetermined and constant repayment periods which should make the analysis less demanding, as if I include variable maturity the results will be complex. This makes the length of the loan another independent variable that will be accounted for in the analysis, (Deitrich, 2020).

2.1 factors influencing interest rates:

The borrower background involves many factors, the most significant one is the credit score which allows lenders to assess the worthiness of a borrower, higher credit score usually result in lower interest rates. Then there's income and employment, as it is a way to determine borrowers' stability and a way to avoid defaults, (Avgeri & Psillaki, 2023). Finally, we have the debt, which will be analyzed through the credit scores, then we have the individual's income, which is also going to be represented through credit scores, as there we're many difficulties in obtaining such data. The credit scores are crucial as they present a rough estimate of one's income and debt which are needed for this analysis.

In P2P lending information asymmetry could occur when borrowers have more information about their creditworthiness than lenders. This could lead to difficulties in setting interest rates, as high-risk borrowers are more likely to seek loans at lower interest rates, and then engage in an unworthy behavior post-loan. On the other hand, lenders might demand higher interest rates to compensate the perceived risk, which would lead to increased interest rate variability and potentially higher defaults, (Izang et al., 2024). These issues could reduce the efficiency of P2P lending and can discourage new lenders to participate.

Loan term is the length of the loan, this factor is usually a less significant factor in determining interest rates, (Deitrich, 2020). Although that matter, I will still be analyzing this variable because it still has a relevant effect which is usually measured based on the maturity of the loan, (Bhardwaj & Sengupta, 2011). It is important to understand how the length of the loan might affect the interest rates, since short term loans do not have the same risks and returns as long-term loans. Usually, shorter loans come with less risk because of the shorter repayment period, making them have lower interest rates, and since longer-term loans have an extended repayment period, they usually have higher interest rates because of the risk involved. Considering all information above, rates might still differ based on the lenders preferences, which in this case makes this variable less significant, nevertheless it is still useful to acknowledge. It's effects, to further understand the impact of long maturity.

Loan purpose is one of the main factors, because depending on the purpose some lenders might reject a loan before checking if the borrower would be worthy of such loan or not. Lenders need to be convinced that they will be paid back to further proceed with a borrower, (Chen, 2014). This factor is very relevant as some individuals request loans for ridiculous purchases instead of using it as an opportunity to build more wealth or pay off certain debts. The purpose of a loan could potentially be anything, thus it is important that I analyze the most requested loans, like business loans, property loans, car loans, etc...

Figure 1: factors influencing interest rates.

Variables	Attributes	Data origin (country)	Article	Relevancy of factors (for analysis)
Borrower background	Credit score Income Debt	- United states	(Avgeri & Psillaki, 2023)	High
Loan purpose	Property loans Business expansion loans acquisition Vehicle loans	- China	(Chen, 2014)	High
Loan term	Long term Short term	- Switzerland	(Deitrich, 2020)	Medium
Loan amount	Large Small	- Estonia	(Kula, 2021)	Low
Market condition	Inflation Demand Economic uncertainty	- United states	(Klein et al., 2021)	Low
Platform fees	High fees vs low fees. Different platforms = different fees	- United states	(Avgeri & Psillaki, 2023)	Medium

The figure above presents the most impactful factors, since the determination of interest rates relies on multiple factors as mentioned before, for this analysis we will be focusing mainly on the variables mentioned in table 1 (shown in methodology), as there are multiple other factors that I will be disregarding (check figure 1) as they do not seem relevant enough. There are factors like the loan amount which I haven't discussed because this analysis will rely on a fixed loan amount range to understand the different variables affecting the rates because loan amount is too general of a variable. To study the actual significance of loan amount, a much larger sample is needed, which will be unattainable due to the difficulty in data availability. Platform fees and market condition will be disregarded as well, since I will be using one platform, making the fee irrelevant as they will differ from platform to another which could affect the data. As of market condition, it is important to consider it when trying to figure out the best time to request a loan, but for this research it seems irrelevant, since I will be analyzing the rates based on factors that are influenced by the borrower mainly, (Avgeri & Psillaki, 2023). The other main drivers behind interest rates should be within the scope of this analysis. The importance of the factors has been assessed through many academic literatures, which made me form a valid assessment on the relevance of the factors. (Amvella Motaze, 2022) The debt and income will be assessed under the variable credit score as mentioned before, since the debt and income will be based on the different credit score ranges which will be mentioned in the data section.

2.2 Hypothesis:

Based on prior literature I already know that the borrower's background can tend to influence interest rates depending on the selected sample. To make sure that the borrower's background in fact has a significant effect on interest rates in P2P, I will study its effect through the sample on hand. To further investigate I will compare the results between the different variables to see which variable has the most significant effect.

H1a: Borrower's background has a significant effect on interest rates.

The loan purpose should play a role in determining interest rates, as it is a factor that helps lenders perceive the risk present, it allows lenders to assess the borrower and understand their goals. The purpose of loans plays a role in the acceptance of a loan, since the effect behind the different purposes is crucial, as loans taken out for business expansion are set to have lower interest rates since they are often seen as income generating activities, this ensure to lenders that the risk is very minimal. On the other hand, loans taken out for personal expenses, such as wedding expenses, vacation funding or debt consolidation will generally have higher interest rates, as they increase the perceived risk for lenders. Purpose affects interest rates heavily, since it is a variable that mitigates risk for lenders, it can significantly impact the interest rates offered to borrowers depending on the perceived risk of the loan purpose.

H1b: Loan purpose has a significant effect on interest rates.

Loan term is supposed to be a very critical factor when considering interest rates, as it can affect interest rates in different ways. Usually depending on the length of a loan, interest rates are affected, since the risk involved differs when it comes to the terms of a loan. Longer term loans should have a higher interest rate because of the risk involved in having a longer payment plan to repay the loan, then we have the opposite effect where short-term loans have less risk since it is easier to repay. But on the other hand, longer term loans have lower monthly payments which can make it easier for individuals, thus this contradiction will draw the next hypothesis, to determine the actual effect loan term produces on interest rates. It is discussed in prior studies that longer term loans sometimes carry lower interest rates, since there is a longer repayment period which allows borrowers to comfortably repay their loan. (Deirich, 2020). But on the other hand, individuals with low credit tend to get rejected for longer term loans as it raises suspicion of default.

I will be investigating how longer-term loans have a lower interest rate, to assess whether there is a significant interest rate variability because of different loan terms.

H1c: long term loans have a lower interest rate than short term loans.

3. METHODOLOGY

I will be using a linear regression equation, since it is a common choice when exploring the relationship between variables, the goal is to model the relationship between a dependent variable and the independent variables and to estimate the different effects each variable has. It will be considered as a multiple regression analysis because I will be analyzing the relationship between interest rates and the multiple independent variables below.

Regression equation: (Depren et al., 2021).

$$\text{Interest rates} = \beta_0 + \beta_1 \times \text{credit score} + \beta_2 \times \text{property loans} + \beta_3 \times \text{business expansion loans} + \beta_4 \times \text{loan term} + \beta_5 \times \text{loan amount} + \beta_6 \times \text{start-up company} + \epsilon$$

The coefficients needed to be estimated are $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$, while ϵ is the error term.

3.1 Dependent variable:

Interest in P2P can also be influenced by market dynamics, which include changes in economic conditions, inflation, and instabilities in market. These factors are not directed by peer-to-peer platforms, which can create fluctuations into interest rates determinants. I will not be studying platforms as dummy variables since I will only be focusing on one platform which is Geldvoorelkaar, this is because the data needed for this analysis is based on this specific platform, so that I can accurately measure their effects.

The interest rate, as the dependent variable is typically measured as a percentage of the loan amount, that the borrower must pay to the lender. This rate is usually set over a specific period, it can be presented in various ways depending on the lenders preferences.

3.2 Independent variables:

The loan term will be expressed in years, it is a measure that presents how long it would take for the borrower to pay off his loan. Usually, loan term is discussed beforehand depending on the loan criteria and lenders preference, the payment increments are monthly, (Yuliyana & Ulrich, 2023).

The borrower's background will be represented by the credit scores, since these scores are usually an accurate measure of an individual's worthiness. This score is usually created based off the income, debt, employment status, car payment, and multiple other factors, to perfectly assess one's financial worthiness and his capabilities to repay his loan on time. The credit scores are given from the banking institutions, but P2P lending platforms re-asses these credit scores to make sure that no defaults occur. A low credit score is a representative of how affective past financial decision can be, while a high credit score represents financially responsible individuals.

As for the loan purpose I will be analyzing 4 different types of loans, to further investigate how business loans, known as money making loans differ from mortgage loans which usually involves property down payments, renovation, and other factors. I will not include vehicle loans, as it is not a common type of loan in the P2P lending market, which will make any analysis unreliable due to the lack of data available.

To measure the effect, dummy variables will be employed to categorize loans based on their purpose. Specifically dummy variables will be created for distinct loan purposes such as property loans, business expansion loans and loans for startup companies, these are only 3 variables out of the 4 since in regression analysis and other modelling techniques we use a (k-1) to encode this information. Each loan purpose of the three mentioned above will be represented by a binary dummy variable, with a value of 1 indicating the presence of the specific purpose and 0 otherwise. For the last purpose, which is Business acquisition loans, we can determine the value based on the other dummy variables already present. Because if all three variables are equal to zero, we will know that the loan on hand is a business acquisition loan. By using the k-1 dummy variables, I avoid redundancy and ensure that the regression model can be estimated without issues related to multicollinearity.

Loan amounts is a control variable which is measured in monetary form (Euro). The loans will have absolute values to further investigate the significance of the variable. Higher loan

amounts would usually mean that the risk is higher, which thus influences the interest rate based on the different amounts.

Table 1: Relevant variables

Variable	description	Unit
<i>Dependent:</i> Interest	Average interest rate on loan	%
<i>Independent:</i>		
Borrower background		N/A
- Credit score	worthiness of borrower	(0-100)
Loan purpose		
- Business loans	Dummy variable	(0 or 1)
- Property loans	Dummy variable	(0 or 1)
- Business expansion	Dummy variable	(0 or 1)
- Acquisition	Dummy variable	(0 or 1)
Loan term	Duration of loan, Long/short term	years

There are potential issues that may arise given the variables used for this analysis, firstly I must ensure that the sample size will be large enough since there are multiple variables within the analysis. Since there are multiple variables, it is crucial to consider multicollinearity because if the independent variables are highly correlated, they will make it difficult to interpret the individual effects of each of the independent variables on the dependent variable. After analyzing and understanding the different effects each variable has on interest rates, I will examine the coefficients estimated from the regression model to assess the impact of each independent variable on interest rate variability.

The regression model will allow me to estimate the effect of each variable, by including the relevant data from each variable, dummy coding will be used for loan purpose. The regression model enables lenders to have a full report on their potential borrower, to further aid with mitigating risks within the investments, (Bao et al., 2023), as higher risk borrowers typically receive higher interest rates to compensate lenders for the increased chance of default. The regression model will help predict the interest rates for future loans, based on different purposes of loans, loan term and loan amount as well. The model will help explain how the different variables affect interest rates differently. This will allow platforms to perfectly match borrowers and lenders in the future, so that everyone is satisfied.

3.3 Data collection:

I am collecting secondary data through the online platform Geldvoorelkaar, this platform is one of the most popular P2P platforms in the Netherlands, which was founded in 2011 making it one of the first peer-to-peer lending platform. I chose it as it had the most suitable data for this research due to the limited data available. The data will be collected through the platform directly, if there are issues, I will contact the platforms and analyze their previous loans that are within the requirements of this analysis. I will be using data extraction/ data scrapping methods to gather the data needed for the analysis. This data should allow me to further proceed with this analysis to understand the impact of each variable, and to have all the relevant data available to complete this research.

The quantitative data needed to complete this research are income information, debt information which will be studied through the credit scores, as there was limited data available. Based of previous literature, the credit score system that will be used according to the platform is a simplified system ranging from 0 to 100. (Kumar et al., 2019). I will extract the data from the platform “Geldvoorelkaar”, as this platform offer previous loan data which include factors like the interest rate, loan term, loan purpose, and risk classification, which is the related to the

credit score. This platform will enable me to extract all the data needed for this research. After investigating multiple platforms, I have realized that the use of one platform will be better because it will allow me to extract a large enough sample that is within the same country so that I can analyze the effects of the independent variables on interest rates. This will help me properly perform the regression analysis, as it will help me reflect on the hypothesis.

Platforms like the lending club, Zopa, and the funding circle have all parted ways with peer-to-peer lending because of the effects of Covid-19 crisis, because regulators have considered P2P a “high risk investment” since it caused a disruption for the industry and led to some players quitting the market, (Jones, 2021).

4. DATA

Based on the data I collected, I will be manually creating a data set, it will consist of all eligible loans needed for this analysis. Because of the low data availability, I was able to gather 51 different observations only. These observations won't all be part of the analysis, because to analyze the data properly, I have set some criteria's, which led to filtering out all the data not meeting the requirements for this analysis. The sample selection criterion in question is as follows, loans equivalent or under 200 thousand euros, as this will enable me to achieve a more precise analysis. As for the duration of the loan, I found it complex to study a small duration interval, as most loans are usually longer term. Thus, I included loans ranging from 2 years (24 months) to 10 years (120 months), to expand the research and ensure a good understanding on how longer-term loans might differ to shorter term loans. For the loan purpose I introduced 4 types of loans, business expansion loans, acquisition loans, property loans and start up loans. I decided to study these purposes as each purpose carries different risk levels, which is important when it comes to setting interest rates, (Sun et al., 2023).

I have used such criteria to enhance the precision of this analysis and to understand properly the different effects the independent variables have on the dependent variable. Based on the filters regarding loan amount, loan duration and loan purpose I have excluded 8 of the 51 samples, which leaves the complete data set for this analysis at 43 samples that meet all the requirements to further aid in this research. The sample period for this analysis ranges from 2022-2024, which helps limit the research and ensures the relevance of data because of how recent the sample period is.

4.1 Data description

I have used excel to analyze the data set, by creating a descriptive statistic table that summarizes all the data used in this analysis. I could already see that borrowers have a credit score of 53,6 on average, which shows that borrowers need to have an almost average score for such loans. The credit scores analyzed in this paper have a range of 70, with the lowest at 30, and the highest at 100. The lower the credit score the harder it is to attract lenders, including the potential for higher interest rates. On the other hand, the higher the credit score, the easier it is to gain lenders trust as the score represents a stable financial background, which indicates lower risk. the mean of credit score (53.6), shows that the average borrower in this analysis has a fair credit score, this shows us that borrowers within this analysis are mostly encountered with high debt. The mean of loan term lies at 4.9 years, which presents us the average loan duration for this analysis. Finally loan amount, with a mean of 129330 euros, shows us that most loans studied are relatively large when

depending on the range of loan amounts studied. Prior literature usually analyzes broader datasets, (Deitrich, 2020), thus the results might differ because of the lack of data available.

To analyze the borrower background, credit scores will be analyzed in this paper using a range from 1 to 100 which is a simplified system, it can vary significantly by country or even by a lender within a country depending on how specific it is needed. Geldvoorelkaar uses this model to assess the creditworthiness of potential borrowers. The score range are as follows; 1 to 39 is considered as poor credit, which means that there is a high risk of default based on the borrower's financial history, potentially significant debt issues. The scores 40 to 59 is considered as fair credit, since there is moderate risk with relatively high debt. The scores from 67 to 79 are considered good credit because there is low risk and generally all payments are made on time, with manageable debt levels that ensure no defaults. Finally, we have 80 to 100 which is excellent credit which means that there is very low risk including a strong history of on time payments with a very low debt level

The Correlation of independent variables with the dependent variable will be studied through scatter plots and they will also be studied through a correlation table which are both present in the appendix. Figure 2 Shows us the correlation between loan amounts and interest rates, we can see that the larger loan amounts do not have low interest rates and usually would have higher rates than for lower loan amounts. There are still some outliers which are present because of other effects affecting the interest rate. As for figure 3 we could see the loan terms effect on interest rates, I could visualize a pattern where longer-term loans have lower interest rates, this may be the case because there is a longer repayment period thus the loan installments are lower and easier to manage. Finally, we have Figure 4 which presents the effects of credit score on interest rates. Visually I can already tell that Credit scores do in fact matter a lot as I can see that the lower the credit score the more likely the interest rate is higher, while the higher credit score usually would get a lower interest rate. The analysis of these scatterplots already provides me with a clear view on how interest rates are correlated with the different independent variables.

Table 2: Descriptive statistics

	Credit score	loan term (years)	loan amount (€)	interest rates
Mean	53,628	4,907	129330	0,084
Standard Error	2,312	0,205	7818,795	0,002
Standard Deviation	15,163	1,342	51271,270	0,013
Sample Variance	229,906	1,801	2628743112	0,00
Kurtosis	2,470	9,143	-1,310	0,041
Skewness	1,056	2,223	-0,195	-0,893
Count	43	43	43	43

The descriptive statistic table above (table 2) represents an analysis done through excel to better understand the values of the analysis.

For the credit score analysis, I could see that there is a slight positive skew, because a skewness of 1.05 (shown in table 2) usually indicates that the data distribution is positively skewed. This means that right tail, which includes the higher values is larger than the left tail. This concludes that there are several credit scores that skew the distribution to the right.

A kurtosis of 2.47 is slightly below 3, which would mean that the distribution has a flatter peak and that it consists of lighter tails compared to a normal distribution. This shows us that there is fewer extreme value than in a normal distribution, which show consistency throughout the data. The credit scores within the analysis are evenly spread out without significant outliers. This explains that very low credit scores will usually be declined, as I have not found any loans that meet the criteria for this research, under the minimum credit score of 30.

As of the standard error, credit scores hold the value of 2.30 (shown in table 2), which means that the sample mean is expected to vary by 2.30 points from the true population mean. This reflects the accuracy of the analysis, because the standard error is considered small in the context of the data I gathered, which indicates that the sample mean is a precise estimate of the population mean regarding credit scores. In conclusion there is a moderate positive skew, and there is a near normal kurtosis for credit scores, with a relatively precise estimate. It is important to have a precise estimate, as credit scores are a significant factor when it comes to the borrower background.

For the loan term, a skewness of 2.20 (table 2) indicates a highly right skewed distribution, as I can see the average loan duration is about 4.9 years (around 59 months) which is relatively high regarding the data collected. It might seem nearly normal since it is under the median of 5 years (60 months) (shown in table 2), but it is positively skewed with heavy tails since the range is large, to precisely analyze the different effects it has on interest rates.

The kurtosis of 9.14 (shown in table 2) means that the distribution has more peaks and heavier tails compared to a normal distribution. This simply explains that there are more extreme values compared to a normal distribution, which makes sense since the range of the data analyzed is 96 months. The kurtosis suggests that there are much longer or much shorter loans than the average which indicates potential outliers. To deal with these outliers I have changed the loan term measure from month to years, to investigate if there are potential problems but it seems to have insignificant effect which is why I ended up measuring loan term in years.

The standard error of 0.20 indicates the average amount of years it might vary, but it also indicates that the sample mean is a precise estimate of the population mean. This is important as it considers the variability in loan terms regarding interest rates and will help me further investigate how long- and short-term loans have different effects on interest rates.

For the loan amount the skewness of -0.19 (table 2) indicates that the distribution of loan amounts is skewed to the left, this means that most loans carry lower amounts. This means that the left tail which are the lower loan amounts are slightly heavier than the right side which includes the higher loan amounts. The distribution is still considered close to symmetrical.

Loan amounts have a kurtosis value of -1.30 (table 2) which indicates that the distribution of the data is flatter which means that there are fewer extreme values compared to a normal distribution. This means that they are relatively spread out since they do not contain significant outliers.

I have dummy coded the different types into zeros and ones, to further analyze their effect. Table 2 presents the descriptive statistics of the dummy coded variables as well. The 4th variable known as Business acquisition loans is not on the table as it is

only measured when all other variables are equal to zero, because when u have categorical variables, you use k -1 to encode dummy variables. I will not go into detail analyzing the dummy coded variables as they do not present significant data, but as they are part of the regression, I made sure the data is available.

5. RESULTS

The objective of this research is to determine the interest rate variability within the peer-to-peer lending market. A regression analysis has been conducted which includes (Credit score, Loan term, Loan amount) and the dummy-coded variables (Business expansion, Business acquisition, Property loans, Start-up company loans) that represent the different loan purposes. The regression analysis includes multiple variables, this will allow me to identify each of their effects on interest rates.

Based on the regression analysis results, I will evaluate the hypothesis, and share the findings of these analysis, to identify factors that affect interest rate variability within peer-to-peer lending market.

5.1 Regression analysis

Table 3: Regression statistics

Multiple R	0,659
R Square	0,434
Adjusted R Square	0,340
Standard Error	0,010
Observations	43

The regression model shows an R-squared value of 0.434 (Table 3), this indicates that approximately 43.4% of the variance in interest rates can be explained by the independent variables included in this model, (Credit score, Loan term, Loan amount). The F-statistic is equal to 4.60 with a f-statistic of 0.0014 (shown in table 5), these values indicate that the overall regression model is statistically significant. Which also demonstrates that the model provides a good fit to the data and that the predictors are collectively influencing the interest rates. To further investigate we can see that the adjusted R square is equal to .34 which is a more accurate measure of the variability of the dependent variable, especially since there are multiple predictors involved. While 33.4% is not extremely high, it still indicates that there is some significance to the model.

Table 4: Regression analysis results

	Coefficients	Standard Error	P-value
Intercept	0,109	0,011	3,46E-12
Credit score	0,000	0,000	3,23E-01
loan term (years)	-0,005	0,001	3,70E-04
loan amount (€)	0,000	0,000	4,24E-01
property	-0,003	0,005	5,62E-01
start up company	0,009	0,005	6,56E-02
business expansion	0,001	0,005	8,89E-01

The intercept represents the expected interest rate when all variables are equal to zero and since loan term cannot be zero, the intercept will serve as a baseline value for the interest rates.

The p-value of the intercept (shown in table 4) is statistically significant which indicates that the baseline is different than zero.

The coefficient for the credit score variable (shown in table 4) presents the value for each one-point increase in the credit score, which means that for each one-point increase in the credit score, the interest rate decreases by 0.012%, (shown in table 4). This negative relation is not statistically significant with a p-value of 0.32 (Shown in table 4), which clearly indicates that higher credit scores are associated with lower interest rates, but the effect is not statistically significant. The confidence interval (shown in table 6) for the credit scores does in fact cross zero, this shows the insignificance of the results. The borrower's background is represented by the credit score in the regression analysis. The coefficient -0.00012 and p-value of 0.32 (shown in table 4) indicate no statistical significance, thus there is enough evidence to reject hypothesis 1a since the borrower's background has no significant effect on interest rates.

The coefficient for loan term (shown in table 4), suggests that for each additional year in the loan term, the interest rate decreases by 0.47%. I can see that this variable also has a negative relationship which is statistically significant with a p-value of 0.0004, indicating that the longer loan terms relate to lower interest rates, and shorter-term loans associate with higher interest rates. The confidence interval for loan term (table 6) supports the significance of this variable. Since the interval does not cross zero. I could mention that longer term loans will always benefit from better interest rates, compared to shorter term loans. The scatter plot (Figure 3) indicates that there are multiple outliers, and that the data is not consistent. This helps reinforce the fact that longer term loans do not have the same interest rates as shorter-term loans. Loan term is part of the regression analysis with a coefficient of - 0.0047 and a p-value equal to 0.0003, (shown in table 4) this suggests that there is a significant negative effect. This result does in fact support the hypothesis that suggests that long term loans have a lower interest rate than short term loans. We have enough evidence to not reject hypothesis 1c, as longer-term loans do in fact have lower interest rates.

The coefficient for loan amount (shown in table 4), shows insignificant effect on the interest rate. The p-value for loan amounts of 0.42 suggests that the variable is not statistically significant, this means that the loan amount does not significantly influence the interest rates within the P2P lending market. The confidence interval for loan term (shown in table 6) includes the zero which also proves the insignificance of this predictor. The scatter plot (Figure 2) shows us that the data is almost perfectly spread, this shows consistency and proves that the loan amount in fact is not a significant effect as discussed before.

The coefficient for property loans is -0.003 (shown in table 4), which suggests a slight negative effect on interest rates compared to the business acquisition loans (the intercept). But however, the p-value equals to 0.56 which indicates that the effect is not statistically significant. The confidence interval for this dummy variable includes the zero which further supports the insignificance of this predictor.

The coefficient for start-up company loans is 0.0086 (shown in table 4), this value indicates a potential increase in the interest rate compared to business acquisition loans. The p-value of 0.065 suggests that this potential effect is hardly significant, since it does not meet the threshold to be considered significant. The confidence interval includes the zero indicating that the effect can also sometimes be neutral at 0.

The coefficient for business expansion loans (shown in table 4), suggests that there is an insignificant effect on the interest rates compared to business acquisition loans. The p-value is equal to 0.88 (shown in table 4), which indicates that this effect is not statistically significant. Finally, the confidence interval for this variable includes zero, confirming the insignificance of this predictor.

Loan purpose is analyzed through the dummy coded variables (property, start-up company, business expansion and business acquisition as the intercept). Based on the findings of the regression analysis which included the 3 dummy coded variables, there was no significant effect, except for start-up company loans which had a marginally significant effect. Even though there might be some significance to loan purpose as a whole, because of the intercept, it is not enough to consider the effects of loan purpose as significant. Thus, there is enough evidence to reject hypothesis 1b since loan purpose does not have statistically significant effect on interest rates.

Based on these results, I can see that there is a significant effect on interest rates by the loan term, but not by the loan purpose, loan amount or borrower background. For the alternative hypothesis, I failed to reject hypothesis 1c but had enough evidence to reject 1b and 1a. I rejected hypothesis 1a because there was no statistical significance from credit scores on interest rates. I have also rejected hypothesis 1b as the loan purpose was not significant enough, this means that while loan purpose does have some impact, the specific purposes used in this analysis do not significantly differ from business acquisition loans and their effect on interest rates. In conclusion the only hypothesis I had no evidence to reject was 1c, which states that longer term loans have lower interest rate.

6. DISCUSSION

This research aimed to identify the interest rate variability in the peer-to-peer lending market by recognizing the main determinants in play. By analyzing the results of the regression which represented (loan term, credit score and loan amount) and the dummy-coded variables representing different loan purposes (property, start-up company, business expansion and business acquisition as the reference category). I will be able to discuss the findings, understand their relevance and effects.

The regression analysis performed showed more statistical significance in the variables compared to the dummy coded-variable as they had a more significant effect. This already had revealed how much more significant loan term are compared to loan purpose, loan amount and borrower background. The loan purpose intercept had a moderate statistical significance, but the dummy variables had marginal to zero significance. This indicates that the specific loan categories used in this research seem to have similar risk which would explain why there was no statistical significance between each of the dummy variables. As for the loan amount, the coefficient and p-value (shown in table 4) allows me to determine that there is no statistical significance. This means that loan amount does not influence interest rates variability in the peer-to-peer lending. These findings do not seem strange especially since the loan amount range analyzed in this research was very small. But in the case of borrower's background, it seems strange to identify no significant effect, this might be due to simplified credit score system, (Avgeri & Psillaki, 2023).

The loan term coefficient and p-value (shown in table 4) showed statistical significance. This means that the loan term in fact has

a more significant effect than loan purpose, loan amount and borrower background. This implies that longer term loans do in fact have lower interest rates than shorter term loans, which supports the hypothesis. This leads me to believe that prior studies have not studied well the loan term as some mention that longer term loans usually have higher rates since there's a longer repayment period, (Deitrich, 2020).

The hypothesis 1a and 1b, have enough evidence to be rejected, but as of hypothesis 1c there is not enough evidence to reject the hypothesis. Thus, making the loan term the main interest rate determinant in this research, as it presents the maturity of the loan. This shows that loan term has a more significant effect compared to the borrower background, loan purposes and loan amount. Regarding the dummy-coded variables there was not a significant enough of an effect, thus the hypothesis relating to loan purpose was rejected (1b).

This research presented very different results than what I had expected, the independent variables that have significant effect seem to be very important when applying for a loan, like the loan term, which has a significant effect. However, variables like loan amount, loan purpose and the borrower's background have no relevant statistical significance, making them negligible. This is crucial to know as these findings can help individuals obtain higher interest rates loan by focusing on the longer-term loans.

6.1 LIMITATIONS

Within this study, I might encounter many limitations when analyzing the variables, since P2P lending platforms often rely on limited data of borrower compared to traditional banks, this makes it harder to accurately assess borrowers background to determine appropriate rates. Unlike banks peer to peer lending platforms do not have historical data and a reliable risk assessment model, which may lead platforms to face challenges in assessing borrower risk. This can result in either under estimating risk, or on the other hand over estimating risk which may lead to difficulties when setting rates.

P2P lending platforms often have their own regulations and policies to setting interest rates, these policies may not always align with market conditions or lenders preferences which leads to interest rates not being priced correctly, (Klein et al., 2021). The unexplained variance in both the regression analysis could be associated with platform regulations, lender preferences or other third-party factors. These factors have not been addressed in prior studies, thus, to further investigate these effects future research could help develop a base on how the third-party factors play a part in interest rate variability within the peer-to-peer lending market.

The loan amounts analyzed in this research we're two hundred thousand euros and under, this can be ameliorated by analyzing a larger data set which includes higher loan amounts. To ensure that loan amounts do not in fact have any statistical significance on interest rates in peer-to-peer lending. The debt was not mentioned a lot in this paper, as the credit score factor represent it, the debt estimated might not be as accurate as needed which can raise suspicion. Future studies could have more specific findings by reinforcing their analysis by included variables such as the debt itself instead of representing it by credit scores. I used the same approach for the income, as the credit score is usually estimated based on many factors which includes income to debt ratio.

The borrower background was represented by the credit score, and since I was using a simplified credit score system which ranges from 0-100, used by the platform, I might not have the

most accurate presentation of the effects of borrower background. To further analyze the effects a more detailed analysis will help provide with a more accurate and specified effect of borrower background.

6.2 CONCLUSION

This research aimed to identify the key determinants impacting interest rate variability in the P2P lending market, by focusing on borrower background, loan term, and loan purpose. The findings of this paper reveal that loan term is the most significant factor, with longer-term loans benefiting from lower interest rates. In contrast, borrower background, which is represented by credit scores, loan purpose and loan amounts showed no statistically significant effects on interest rates. These results challenge some of the traditional assumptions about interest rate determinants and highlight the unique dynamics of the P2P lending market. The research suggests that borrowers seeking lower interest rates should consider applying for longer-term loans, while lenders and P2P platforms should enhance their risk assessment models to better capture the effects of borrower background and loan purposes. Future research should explore broader datasets and incorporate more detailed borrower information to validate these findings and enhance the understanding of interest rate variability in P2P lending.

7. APENDIX

Figure 2: Loan amounts on interest rates

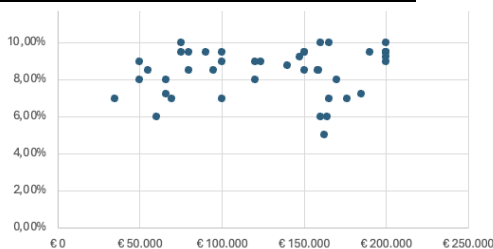


Figure 3: Loan term on interest rates

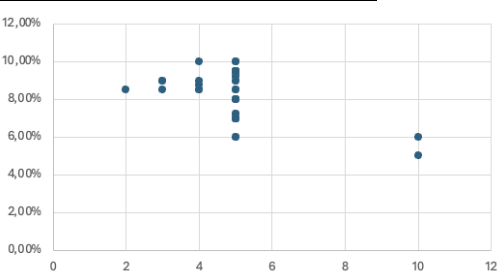


Figure 4: Credit score on interest rates

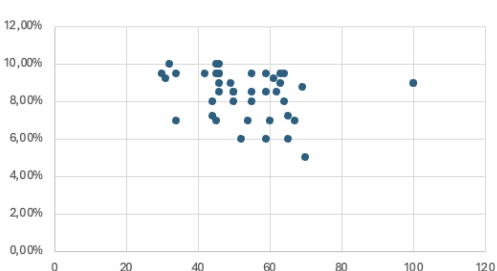


Table 5: Anova Table

ANOVA	df	SS	MS	F	Significance F
Regression	6	0,003	0,000	4,605	0,001
Residual	36	0,004	0,000		
Total	42	0,007			

Table 6: regression results including confidence intervals.

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0,109	0,011	10,220	3,46E-12	0,087	0,130
Credit score	0,000	0,000	-1,003	3,23E-01	0,000	0,000
loan term (years)	-0,005	0,001	-3,930	3,70E-04	-0,007	-0,002
loan amount (€)	0,000	0,000	0,809	4,24E-01	0,000	0,000
property	-0,003	0,005	-0,585	5,62E-01	-0,014	0,007
start up company	0,009	0,005	1,899	6,56E-02	-0,001	0,018
business expansion	0,001	0,005	0,140	8,89E-01	-0,009	0,010

Table 7: Correlation

	Credit score	loan term (years)	loan amount (€)	property	start up company	business expansion	interest rates
Credit score	1						
loan term (years)	-0,104	1					
loan amount (€)	0,000	-0,056	1				
property	0,462	0,039	0,204	1			
start up company	-0,310	0,044	0,140	-0,342	1		
business expansion	-0,017	-0,039	-0,187	-0,323	-0,365	1	
interest rates	-0,237	-0,482	0,153	-0,279	0,373	-0,055	1

8. REFERENCES

- Chen, Dongyu. (2014, December). (PDF) A trust model for online peer-to-peer lending: A lender's perspective. *Information Technology and Management* 15(4):239-254 [10.1007/s10799-014-0187-z](https://www.researchgate.net/publication/272018846)
<https://www.researchgate.net/publication/272018846> A trust model for online peer-to-peer lending a lender's perspective
- Setyaningsih, T., Murti, N. W., & Nugrahaningsih, P. (2019). Fintech based peer to peer lending: An opportunity or a threat? *Riset Akuntansi Dan Keuangan Indonesia*, 4(3), 122–133.
<https://doi.org/10.23917/reaksi.v4i3.8577>
- Irwanto, M. A., & Azis, A. (2023). Risktaking Behavior in Fintech users peer to peer lending. *Journal of Social and Industrial Psychology*, 12(1), 25–31.
<https://doi.org/10.15294/sip.v12i1.70053>
- Deitrich, Andreas. (2020, March). *Determinants of Interest Rates in the P2P Consumer Lending Market: How Rational are Investors?*
<https://www.researchgate.net/publication/>

[340172917 Determinants of Interest Rates in the P2P Consumer Lending Market How Rational are Investors](#)

- Bao, T., Ding, Y., Gopal, R., & Mohlmann, M. (2023, July 18). *Throwing good money after bad: Risk mitigation strategies in the P2P lending platforms - information systems frontiers*. SpringerLink. <https://link.springer.com/article/10.1007/s10796-023-10423-4>
- Foo, J., Lim, L.-H., & Wong, K. S.-W. (2017, October 31). *Macroeconomics and Fintech: Uncovering latent macroeconomic effects on peer-to-peer lending*. arXiv.org. <https://arxiv.org/abs/1710.11283>
- Nguyen, L., Ahsan, M., & Haider, J. (2024, March 5). *Reimagining peer-to-peer lending sustainability: Unveiling predictive insights with innovative machine learning approaches for loan default anticipation*. MDPI. <https://www.mdpi.com/2674-1032/3/1/12>
- Avgeri, E., & Psillaki, M. (2023, September 5). *Factors determining default in P2P lending*. Journal of Economic Studies. <https://www.emerald.com/insight/content/doi/10.1108/JES-07-2023-0376/full/html>
- Jones, Rupert. (2021, December 11). *Zopa exits peer-to-peer lending to focus on banking*. The Guardian. <https://www.theguardian.com/money/2021/dec/11/zopa-peer-to-peer-lending-p2p-money#:~:text=However%2C%20the%20coronavirus%20crisis%20and,some%20players%20quitting%20the%20market.>
- Yuliyani, M., Ulrich, S. (2023, May 18) *On the relationship between borrower and bank risk*. https://www.researchgate.net/publication/370848566_On_the_Relationship_between_Borrower_and_Bank_risk
- Business loan calculator*. Funding Circle. (n.d.). <https://www.fundingcircle.com/uk/small-business-loans-calculator/>
- Zopa annual statement*. Annual statement. (2020, August). <https://www.datocms-assets.com/23873/1658822618-zopa-group-limited-annual-report-and-financial-statements-2020.pdf>
- Vinod Kumar L, Natarajan S, Keerthana S, Chinmayi K M, & Lakshmi N. (2016). *Credit risk analysis in peer-to-peer lending system*. *2016 IEEE International Conference on Knowledge Engineering and Applications (ICKEA)*. <https://doi.org/10.1109/ickea.2016.7803017>
- Klein, G., Shtudiner, Z., & Zwilling, M. (2021, June 17). *Why do peer-to-peer (P2P) lending platforms fail? the gap between P2P lenders' preferences and the platforms' intentions*. Electronic Commerce Research. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8210520/>
- Sun, Q., Wang, J., Zhang, H., & Wen, T. (2023, September 7). *How loan descriptions affect the likelihood that borrowers obtain loans in P2P networks? -an empirical analysis based on the "Renrendai" platform*. PloS one. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10484447/>
- Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). *Determinants of default in P2P lending*. *PLOS ONE*, 10(10). <https://doi.org/10.1371/journal.pone.0139427>
- Bhardwaj, G., & Sengupta, R. (2011). *Credit scoring and loan default*. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2038599>

- Depren, O., Kartal, M. T., & Depren, S. K. (2021). Macroeconomic determinants of interest rates in BRICST, mint, and fragile five countries: Evidence from quantile regression analysis. *Studies in Business and Economics*, 16(1), 51–67. <https://doi.org/10.2478/sbe-2021-0005>
- Dr. S., Kumar, Dr. A., Ramesh, Dr. S., & S., Mr. R. (2019). A technology on Credit Score System - assessing public perception in Bengaluru City. *International Journal of Innovative Technology and Exploring Engineering*, 8(12), 1929–1934. <https://doi.org/10.35940/ijitee.I2890.1081219>
- Amvella Motaze, S. P. (2022). The determinants of the lending interest rate in a cost-based approach: Theoretical Model and empirical analysis. *The Quarterly Review of Economics and Finance*, 83, 36–51. <https://doi.org/10.1016/j.qref.2021.10.003>
- Izang, A. A., Ajayi, O. F., Junaid, O., Nwigwe, B., & Onyeka, P. O. (2024). Design and evaluation of a peer-to-peer student lending platform to mitigate information asymmetry and credit risk. *Ingénierie Des Systèmes d'Information*, 29(3), 885–894. <https://doi.org/10.18280/isi.290309>