TÍMEA ÖLVEDI

THE NEW FORM OF FINANCIAL INTERMEDIATION: KEY ISSUES OF PEER-TO-PEER LENDING

Institute of Finance

Supervisor: Barbara Dömötör, PhD, Associate professor

CORVINUS UNIVERSITY OF BUDAPEST

DOCTORAL SCHOOL OF BUSINESS MANAGEMENT

THE NEW FORM OF FINANCIAL INTERMEDIATION: KEY ISSUES OF PEER-TO-PEER LENDING

PHD DISSERTATION

TÍMEA ÖLVEDI

Budapest, 2023

TABLE OF CONTENTS

LIST OF CHARTS

LIST OF TABLES

1. Introduction

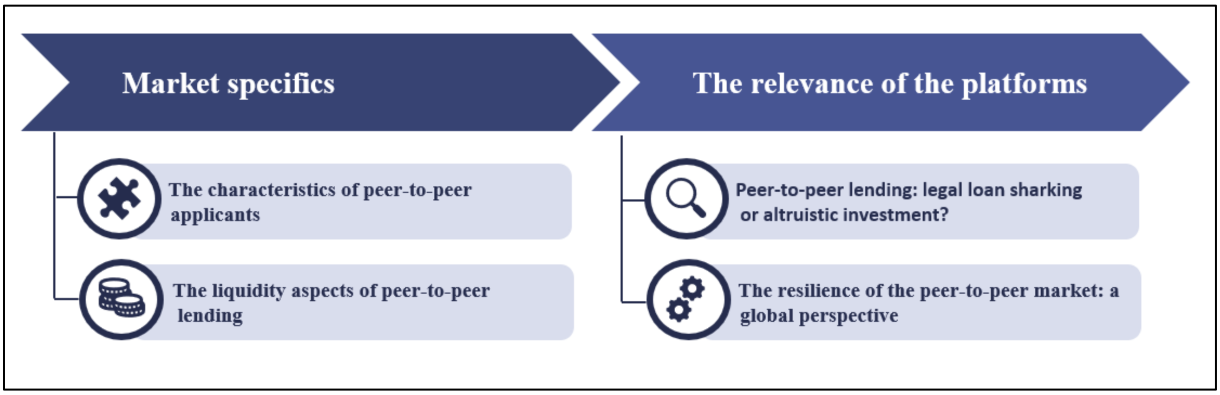
In the previous years, fintech transformation and innovative technological solutions had a significant impact on the economy. This trend reached the financial market as well. Conventional financial institutions introduced a wide range of online products and services. Besides that, a disintermediation tendency has started, and alternative funding models appeared on the market, performing capital allocation. One of this new form is peer-to-peer lending (P2P) or marketplace lending, where the traditional intermediary role is left out of the process. The main idea of the business model is that being more cost efficient, the online platforms offer more beneficial conditions for borrowers compared to a bank loan. From investor perspective the expected return is promised to be higher than a bank deposit yield. However, the risk associated with this investment is also significantly higher. After the first platform was launched in 2005, the segment showed a robust expansion and several new players appeared on the market in many different countries. The strong market growth raises several questions regarding the future of financial intermediation, the role of the platforms on the financial market and their interaction with commercial banks.

The purpose of this dissertation is to gain a comprehensive understanding regarding the relevance of the peer-to-peer lending platforms and to examine the market specific features. The four papers presented in the subsequent sections aim to answer the following research questions:

- Which macroeconomic factors explain the demand for P2P lending in the US market? Which borrower groups can be differentiated based on similar patterns?
- How liquid is the secondary market based on the three liquidity dimensions of selling rate, average selling time, and discount? Which variables have a significant impact on the selling outcome, selling time, and discount rate? How does the secondary market react to external shocks (indicated by a pandemic situation)?
- Do P2P lending platforms have an advantage in information processing compared to traditional banks due to the incorporation of alternative information? What is the performance profile of P2P investments for the lenders?
- Are pre-COVID economic variables applicable to classify countries according to their market reaction to the pandemic outbreak? How does the P2P market respond to an external shock in the different types of economies?

The papers can be divided into two groups. The first two studies are exploring the market specific features, covering the characteristics of the P2P applicants and the examination of the secondary market. The remaining two papers investigate the relevance of the platforms and their potential role on the financial market. Specifically, the information processing advantage of the marketplaces is analyzed and the market reaction in case of an external shock in the different types of economies. Figure 1 presents the above-mentioned concept.

1. Figure: The key studies presented in the dissertation



Source: by Author

The structure of the dissertation is the following:

I. The overview of peer-to-peer lending[1]
   1. What is peer-to-peer lending?
   2. Market statistics
   3. Comparison with conventional bank lending
   4. Main research directions
   5. Regulatory framework

II. Four research papers
   1. Research: The characteristics of peer-to-peer applicants
   2. Research: The liquidity aspects of peer-to-peer lending
   3. Research: Peer-to-peer lending: legal loan sharking or altruistic investment?
   4. Research: The resilience of the peer-to-peer market: a global perspective

[1] The main thoughts of the theoretical part were published in selected journals. For more information, please see Section 7.

In the first part of the dissertation, peer-to-peer lending is defined and embedded into the literature of financial intermediation. Then current market figures are described. Later the main features of the platforms are presented in comparison with the conventional banking sector, covering the relevant market imperfections. After that, the main research directions in peer-to-peer lending literature are summarized. Finally, the theoretical part ends with an overview of the segment's regulation.

In the second part the four research papers are presented. The first study explores the platforms from borrower perspective. To obtain a comprehensive understanding of the fast-growing market of peer-to-peer lending, the demand coming from potential debtors has to be investigated, which contributes and maintains the segment's continuous spread. The sample used for the analysis is a unique, manually collected database, including more than 135,000 loan applications from Prosper, which is one of the market leader platforms in the United States. The applications were published on the webpage of the platform in textual format between 2014 and 2020. After manual data collection, the information was transformed into table format using data manipulation techniques. LASSO regression analysis was applied to find relation between a wide range of macroeconomic indicators and the demand for peer-to-peer lending by state. Then cluster analysis was performed to identify borrower groups with similar patterns.

According to the results, the portion of delinquent mortgages by state has the highest impact on the demand for P2P loans. Furthermore, the most frequently declared loan purpose is debt consolidation, which suggests that applicants intend to utilize the P2P funding to refinance their overdue claims. Peer-to-peer customers can be classified into four different groups based on their patterns and a significant portion of them are eligible for bank funding and owns a good credit history. Therefore, it can be assumed that the debtor group of the platforms overlap with bank clients and the marketplaces complement bank lending only in a small customer segment in case of the US.

The second paper examines the market from liquidity point of view. The basis of the analysis is a large secondary market dataset, covering more than 5 million listings from a noted Estonian platform, Bondora. Three liquidity dimensions were defined, specifically the selling rate, average selling time and the level of discount. These dimensions were examined with heatmap and different regressions to identify which variables have significant impact on them. The previous years were characterized with economic prosperity and robust market expansion for the marketplaces; therefore, it is also relevant to investigate how the secondary market reacts in case of an external shock, caused by a pandemic situation. The second part of this study is also an introduction to the fourth paper, discussing the early analysis of the COVID-19 crisis. The data was split into two sub-samples based on the outbreak of the pandemic in March 2020 to examine the effect of these uncertain conditions on the secondary market.

According to the results, the average selling time is around 1.5 days and there is a robust demand for performing loans which are not overdue, though the discount rate is relatively high, which assumes that the cost of liquidity is incorporated in the price which must be paid by the seller. The key variables which have impact on the successful resale and selling time are the discount rate, the principal amount, the country of the borrower and the month passed after the loan was originated. Regarding the COVID outbreak, the results suggest that there was a structural break in the trend of the secondary market. Investors attempted to liquidate their money with high discount, thus it can be inferred that the market is sensitive to external shocks.

The third study investigates the potential advantage of the marketplaces regarding information processing. The platforms assess applicants based on their own scoring methodology using standard variables and they also include alternative information which is not applied in case of traditional bank lending. Investors often rely on these internal ratings during their funding decision; therefore, it is expected that the ratings properly reflect the credit risk of the potential borrower. To investigate the performance of the platform's model, a benchmark model was prepared using only standard variables which are frequently applied by banks. The modelling sample covers more than 107 000 observations of an Estonian platform, Bondora between the period of 2012 and 2019. Furthermore, the historical return of investors is also estimated to gain better view regarding the expected loss of these investments.

According to the results the platform's credit assessment model can be considered as adequate. Alternative information could improve the models, but our analysis could not confirm that the platforms utilize such data. The average internal rate of return (estimated on the closed transactions of the sample) is -4.17% and more than 42% of the loans ends with a negative IRR. The analysis concludes that in the European market, P2P lending serves to supply high-risk borrowers. However, investors are not compensated for the credit risk even by the extremely high, loan-sharking level interest rates.

The last research examines the segment in the period of economic downturn represented by the COVID-19 crisis. The pandemic outbreak had a significant impact on the economy and the financial sector, and the long-term consequences of the crisis are still unpredictable. The previous literature mostly focused on the time of economic prosperity, while there is limited research on the downturn period. The purpose of this paper is to contribute to the understanding of the segment in the case of economic distress. The dataset covers 61 countries with diverse economic background. Unsupervised machine learning approach was applied to classify countries based on their pre-COVID economic profile. Then the clustering performance was tested in terms of the P2P market reaction of each group after the pandemic outbreak, using one-way analysis of variance (ANOVA) test.

According to the results, three types of economies can be differentiated. The outcome suggests that the model using pre-COVID economic variables is applicable to classify countries regarding their P2P market behavior after the outbreak. In addition to that, the results imply that the growth rate of marketplace lending is in line with the economic instability of the country. In the case of developing countries with low economic performance and weak banking system, the P2P lending volume showed a robust growth. These economies are already struggling with high default rates in their banking system, and due to interrelations on the financial market, the possible non-payments in the P2P segment might further deepen the lending and liquidity issues.

Overall, the market of peer-to-peer lending is a relatively new form of financial intermediation. The segment experienced a huge expansion in the last decade and raised the attention of the supervisory authorities and relevant market participants. Based on the above presented papers, it can be suggested that marketplace lending brings significantly higher risk, compared to conventional lending and investors are not necessarily compensated. The segment is still vulnerable in case of an external shock and the trust needs to be build towards this alternative funding model. From the other side, the platforms serve mostly underbanked customers. Therefore, there is a potential in them to have an important contribution from social perspective as they provide funding for less credible borrowers. However, in order to maintain long term successful operations, it is essential to introduce a comprehensive regulatory framework which enhances the credit risk management, investor and borrower protection and transparent operation of the platforms.

2. The overview of peer-to-peer lending

The purpose of this part is to present the existing literature and to provide a comprehensive overview of peer-to-peer lending. The structure is the following: section 2.1. defines peer-to-peer lending and embed it into the literature of financial intermediation. In section 2.2. the most significant market players and their loan volumes are presented. In 2.3 the main features of the platforms are introduced and compared to traditional bank lending and the market imperfections of bank theory are analyzed. In section 2.4. the leading directions of the literature on platform lending and their key findings are summarized. Finally, in section 2.5. the main regulations of the segment are introduced.

2.1. What is peer-to-peer lending?

In the previous years, fintech transformation had a significant role in the economy and this trend impacted the financial market as well. Conventional financial institutions introduced a wide range of online products and services. Besides that, a disintermediation tendency has started, and alternative funding models appeared on the market. One of this new form is peer-to-peer lending or marketplace lending, where the traditional intermediary role is left out of the process. The initial idea of the business model is that the online platforms offer more beneficial conditions for borrowers compared to a bank loan and for investors the expected return is higher than a bank deposit yield. From the other side, the risk associated with this investment is also significantly higher compared to conventional funding. Besides that, the segment is barely regulated, and the default rates are excessive. After the first platform was launched in 2005, the segment showed a robust expansion and several other players appeared on the market in many different countries. The strong market growth raises several questions regarding the future of financial intermediation, the role of the platforms on the financial market and their interaction with commercial banks.

The emerge and evolve of fintech in the financial sector has an extensive literature. As Frame et al. (2018) highlighted, technological developments led to financial innovations, which enhance cost reduction, risk mitigation, and improve social prosperity. According to the estimations of Philippon (2014) the cost of financial intermediation in the US has been stagnating at 2% in the previous 130 years, leaving a space for further cost-efficient solutions. Varga (2017) highlights that the key value drivers for fintech companies are the enhanced user experience, disruptive business models and cheap access. Goldstein et al. (2019) emphasizes that the current fintech revolution raised the attention of market participants as now the introduction of innovative solutions accelerated and the main changes are coming from technology firms outside of the financial sector, disrupting the incumbents and resulting a sharper competition. Thakor (2020) described four main areas which are covered in fintech: credit and capital-raising services, payment services, investment services and insurance. However, as noted by Arner et al. (2015), the regulation of the segment, often referred as 'RegTech' is still not developed and the operations of these non-conventional institutions are probably not compliant with financial regulations.

Online lending is one of the main financial innovations of fintech revolution, impacting credit and capital-raising services. In order to understand the phenomenon of peer-to-peer lending, first the expression of crowdfunding has to be explained. According to Belleflamme et al. (2014), “Crowdfunding involves an open call, mostly through the Internet, for the provision of financial resources either in the form of donation or in exchange for the future product or some form of reward to support initiatives for specific purposes.” In line with the literature, different types of crowdfunding can be specified. Based on the classification of Walter (2019), four main groups can be classified. The first is the donation-based version, where investors do not receive any financial profit or return. The second is reward-based crowdfunding, where some material or immaterial return is offered. The third is equity-based crowdfunding, in which investors receive part of the firm's equity. The last form of crowdfunding is peer-to-peer lending. The market of alternative finance is further investigated by the Cambridge Centre for Alternative Finance (2021) whose aim is to provide a comprehensive report on the digital finance activities which are not covered by the traditional financial market. Due to the continuous development of the segment, different model types can be determined. As a high-level classification, CCAF differentiates debt, equity and non-investment categories. Debt models cover P2P lending and other platforms which are not taking deposit and the loan they facilitate can be secured or unsecured. In case of equity models, investors buy shares of a business, usually a start-up. Non-investment models are reward-based and donation-based crowdfunding, where the investors do not receive any financial return. Table 1 summarizes the main alternative finance models and their market share and volume for the year of 2020 (categories under 1% of market share are excluded).

1.Table: Market share and volume of alternative financial models in 2020

Class	Business model	Model description	Market share (%)	Volume ('000 USD)
P2P/Marketplace lending	P2P Consumer Lending	Individuals or institutional funders provide a loan to a consumer	31	34 740 386
	P2P Business Lending	Individuals or institutional funders provide a loan to a business borrower	14	15 374 366
	P2P Property Lending	Individuals or institutional funders provide a loan, secured against a property, to a consumer or business borrower	3	3 073 502
Balance sheet lending	Balance Sheet Business Lending	The platform entity provides a loan directly to the business borrower	25	28 018 497
	Balance Sheet Consumer Lending	The platform entity provides a loan directly to a consumer borrower	11	13 025 246
	Balance Sheet Property Lending	The platform entity provides a loan, secured against a property, directly to a consumer or business borrower	2	1 808 250
Crowdfunding	Donation-based Crowdfunding	Donors provide funding to individuals, projects or companies based on philanthropic or civic motivations with no expectation of monetary or material	6	7 002 990

	Real Estate Crowdfunding	Individuals or institutional funders provide equity or subordinated debt financing for real estate.	2	2 777 136
	Equity-based Crowdfunding	Individuals or institutional funders purchase equity issued by a company	1	1 520 444
	Reward-based Crowdfunding	Backers provide funding to individuals, projects or companies in exchange for non-monetary rewards or products.	1	1 250 683
Invoice trading	Invoice trading	Individuals or institutional funders purchase invoices or receivables from a business at a discount	3	3 882 363

Source: CCAF, 2021, page 31 and 41, Table 1.4.

CCAF defines three forms of P2P lending: consumer lending offering loan for retail borrowers, business lending for SMEs and property lending where the loan is secured with a property. As it can be seen P2P consumer lending has the highest market share among the alternative finance models with 31% and $34.74 billion volume. However, it has to be mentioned, that the segment showed a decline compared to 2019, where it owned 59% of the market. The drop can be explained with the decrease of the Chinese P2P consumer lending. China played a significant role on the market, although from 2018 due to regulatory changes and market developments a remarkable decline can be observed (CCAF, 2021). As a comparison, the sum of the consumer credit volume in the US, Europe and China reached approximately 6 815 billion USD as of 2020, meaning that the portion of alternative finance is less than 1% (FED, Eurofinas, Statista 2020). After gaining on insight on the different alternative financial models, the form of peer-to-peer lending needs to be described. There is no unified definition for marketplace lending, however these intermediaries are often referred as loan based-based crowdfunding in the literature. The Financial Conduct Authority of the UK applies the following explanation in its Policy Statement (2019): "People and institutions use these types of platforms to lend money directly to consumers or businesses, to make a financial return from interest payments and the repayment of capital over time."

The first peer-to-peer lending platform was launched in the United Kingdom in 2005, called Zopa. In the previous 16 years, several players appeared on the market worldwide and the segment showed a robust growth. Different business models evolved: some of them provide loans only to retail borrowers while others are specialized on business lending for SMEs. The investor side is also diverse, originally retail investors lent, however later institutional investors stepped into the market, first in the United States, which led to a more robust expansion. The geographical scope of the marketplaces is also varied, large part of them allows cross-border lending, others permit only domestic participants. The products offered by the platforms are limited to different loan types such as personal loan, student loan, SME loan or property loan. The loan terms range between 3 and 5 years (Prosper, 2020; Bondora, 2020).

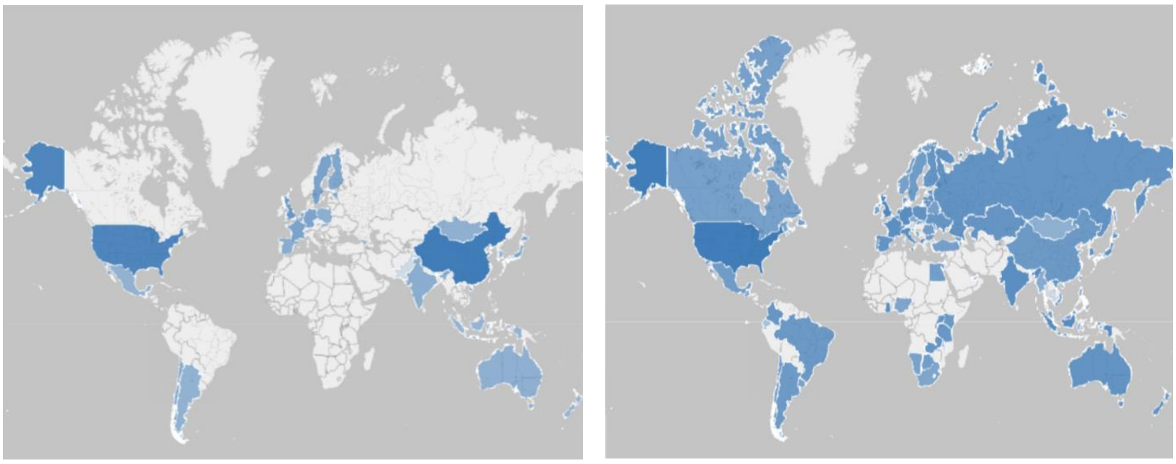
The platforms might have specialties in their operating model; however, the general process is the following. The applicant sets the requested loan amount and the interest rate that he/she is is willing to pay on the webpage of the platform. Then the potential borrower has to provide personal and financial information, which is usually cross-checked and completed with Credit Bureau data. The platform performs credit risk assessment of the applicant based on the mentioned data and assigns a credit rating to each borrower in line with its internal credit risk methodology. As a standard practice, they incorporate alternative information, besides the conventional variables (Jagtiani & Lemieux, 2019). As a next step, the request is listed on the webpage of the platform, including financial and sociographic information and the rating of the platform. Potential investors review the listings and select based on their risk appetite. The marketplaces usually publish their loan dataset to investors in case they would like to analyze historical returns. It is possible for lenders to finance only small portion of the loan and make a diversified portfolio. Generally, investors bid and set the minimum return what they would like to earn on their investment. In case one request receives more offer in the online auction, the lowest interest rates are selected. As an alternative for the previously described manual selection process, most platforms offer portfolio manager service for their investors, meaning that the program assorts the loan listings based on preliminary determined criteria (Bondora, 2021). The platform usually charges an origination fee which is approximately 1-5% of the total loan amount (Morse, 2015). From loan origination perspective, two models are in practice, depending on the legal environment. In the UK, the marketplace originates the loan and the cash flows through the customer accounts. In the US and some parts of Europe, solely a licensed bank is authorized to originate the loan. Consequently, in this model a bank also participates in the process as an additional intermediary. It means that the bank provides the loan and sells it to the platform. The two approaches are similar in a way that the marketplace only conducts a brokerage activity and does not bear the risk (Lenz, 2016).

The related literature discusses several benefits and drawback of this alternative funding model. According to Milne and Parboteeah (2016), one of the main advantages is that the platforms could offer better conditions both for borrowers and investors compared to a bank product. They facilitate debtors who are not eligible for bank funding to obtain credit. Besides that, they use innovative technology, provide fast and convenient user experience. From the other side, the risk is greater, as the loans are generally unsecured. The segment is barely regulated, in case of non-payment, the consequences are limited, and the risk management framework is rudimentary. There is no official investor and borrower protection and there is no cap for the interest rate which they offer for the potential debtors. Furthermore, the transparency in case of loan recovery and also the operation process in general has to be enhanced. The platforms introduced risk mitigating measures in the previous years, e.g. some marketplaces offer buyback guarantee to investors, meaning that they propone to pay back the loan, sometimes the expected return as well in case of a default. Loans with buyback guarantee offer lower yield reflecting reduced level of risk. In case of non-payment, first the platform attempts to collect the loan, then hand over the claim to a third-party debt collector agency. Besides, preliminary criteria are set for applicants e.g. minimum threshold of FICO score to filter out debtors with poor credit quality. Most of the platforms launched secondary market in order to enhance liquidity, where investors can pass on their claims. While these measures are promising, the overall operational model of marketplace lending still brings significant risk compared to traditional intermediaries.

2.2 Market overview

As mentioned, the first platform was introduced in the United Kingdom in 2005. After that, the segment showed a robust expansion, and nowadays marketplace lending is present in all continents. Different statistics are published regarding the size of the market and it is hard to find accurate data. According to the report of CCAF on the global alternative finance market (2021), the consumer segment of peer-to-peer lending is around \$34.7 billion as of 2020, and the business segment specialized on business purpose loans, reached \$15.3 billion dollars. The total volume of consumer loans in the US is around $4 142 billion dollars in Q3 as of 2020 (FED, 2021). As we can see, the portion of peer-to-peer lending is around 1,2%, which is still not significant compared to the whole US credit volume, however the market had a steep growth rate in the previous years. Figure 2 presents the expansion of the segment from 2014 to 2020.

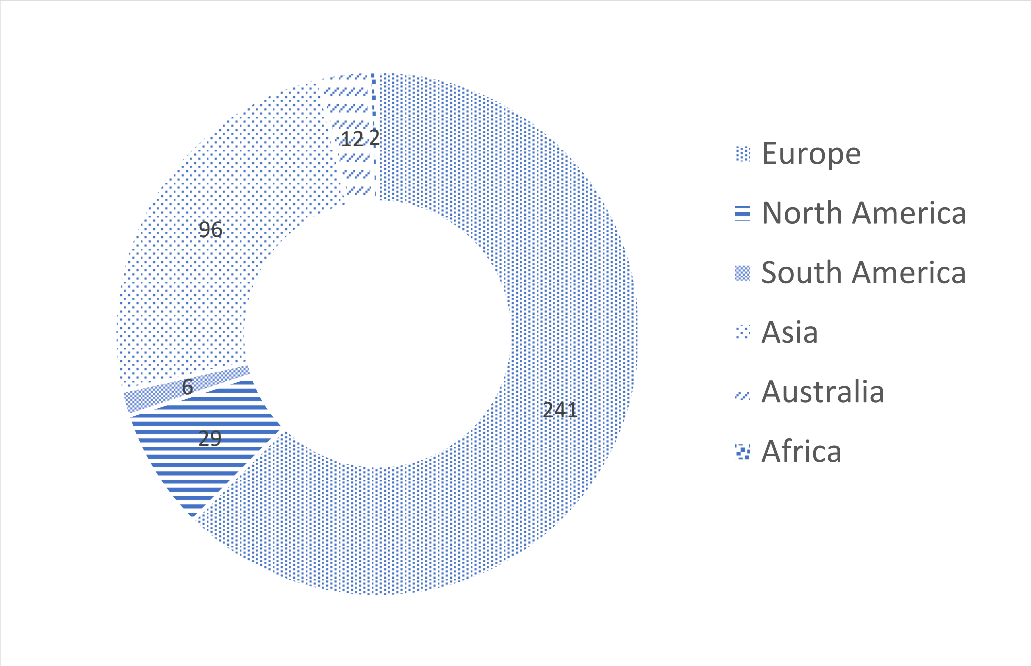
2. Figure: The expansion of the market between the period of 2014 and 2020



Source: Author's editing, based on CCAF, 2021
Note: the left figure presents the market volumes in 2014 and the right one in 2020

As a next step, the number of platforms is examined. Figure 3 presents the distribution of the marketplaces by region. Europe is leading with 241 platforms due to the large concentration in the United Kingdom and the prosperity of social lending in the Baltic region, both having developed market with several participants. It is followed by Western Europe, where France, Germany, Switzerland, and the Netherlands are operating the highest number of marketplaces. In the South part, Italy and Spain owing twenty-six platforms together. Easter Europe run the lowest number of social lending platforms, where Czech Republic has the most extended market. Although, the number of platforms in North America is roughly one eighth of Europe, it has to be mentioned, that the loan volume is significantly higher. It is mostly due to presence of institutional investors, boosting the overall loan originations. The United States is the most significant region, operating eighteen platforms with diverse business models. It is followed by Mexico with seven marketplaces. In terms of Asia, India and Indonesia are leading the market, then comes South Korea. South America, Australia and Africa have modest number of players and the segment is still not extended there (P2PMarketData, 2021).

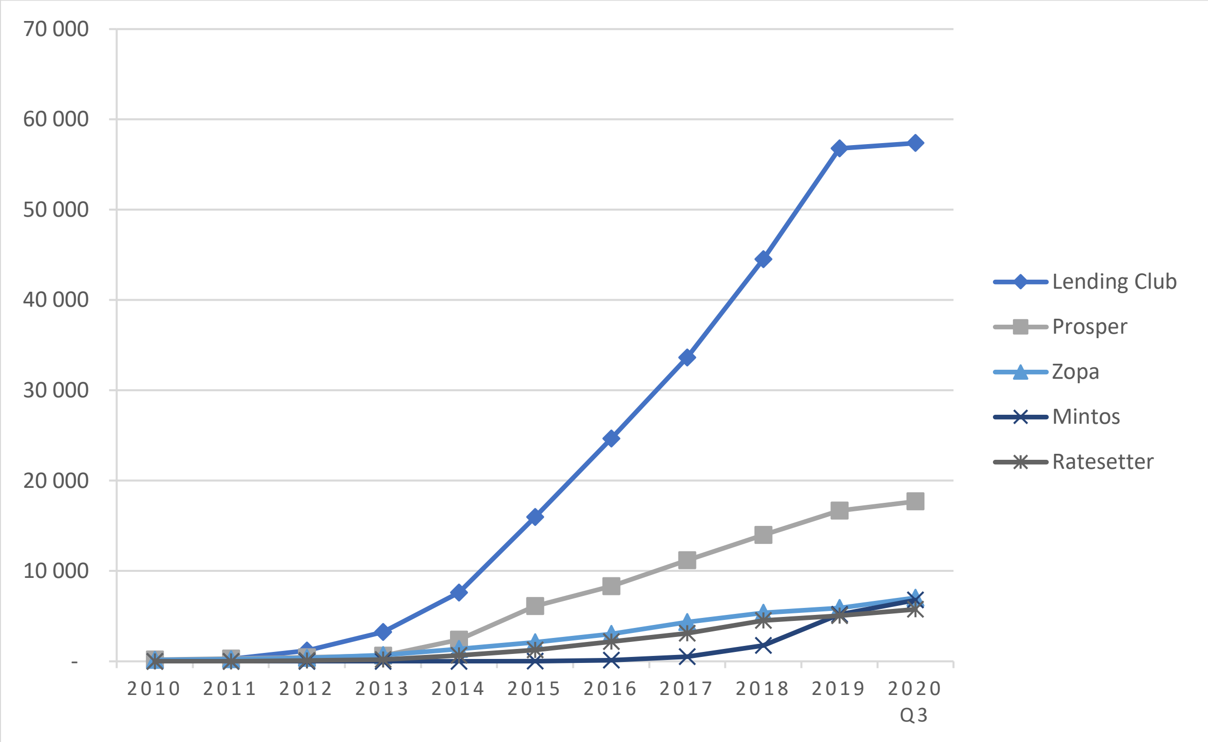
3. Figure: The number of peer-to-peer platforms by region as of 2022



Source: Author's editing, based on p2pmarketdata.com, 2022

In terms of lending volume, Figure 4 presents the aggregated loan amounts from origination between 2010 and 2020, in case of the top five market players. Based on the figure, LendingClub is the market leader with a cumulated loan amount of $57.3 bn dollars as since its foundation. LendingClub was launched in 2006 as one of the first platforms of the US. The firm's robust growth and reputation are supported by its introduction to the New York Stock Exchange as of 2014 and its large base of institutional investors (LendingClub, 2021). Its main competitor in the US market is Prosper, facilitated roughly $17.7 bn dollars. Prosper is followed by Zopa, which was the first peer-to-peer platform, implemented in the United Kingdom in 2005, owing $7 billion dollars of aggregated loan amount. The subsequent site is Mintos with $6.7 bn dollars. Mintos is a relatively new platform, launched in 2015, however the Latvian site has become the most dominant marketplace of the Baltic region in the recent years. The fifth is Ratesetter with $5.7 bn dollars volume, founded in 2009 in the United Kingdom. The mentioned platforms demonstrated a robust expansion in the previous ten years; however, it seems that their growth rate slowed down in 2020.

4. Figure: The aggregated volume of the top five peer-to-peer lending platforms between 2010-2020 (million dollars)[2]



Source: Dömötör&Ölvedi, 2021, page 784, Figure 1

As of 2020 an interesting trend could be observed on the market. Some of the above-mentioned market players announced to change their business model and obtain banking license. Initially, Zopa the first peer-to-peer platform received full banking license from the UK authority and currently operates as Zopa Group, which covers the P2P platform itself and a new digital bank, offering saving accounts and credit card, besides the loan products (Zopa, 2020). Ratesetter announced to sell its loan portfolio to Metro Bank and the platform closed (Ratesetter, 2021). LendingClub, the market leader of the previous years decided to retire its peer-to-peer platform and acquire Radius Bankcorp, which is the holding company of Radius Bank. According to the official announcement, the aim is to become a "fintech marketplace bank and the first public U.S. neobank" (LendingClub, 2021). It is still early to predict the future of the marketplaces, however it seems that 2020 was a turning point for the segment, as the market's dynamic growth slowed down and key participants decided to transform their business model towards traditional

[2] The Chinese market is not presented in this figure. The reason is that previous studies (e.g. Milne & Parboteeah, 2016; Morse, 2015) usually consider the statistics without China, as the market has different characteristics which would lead to distortion in the results.

banking. Currently, it is not obvious if this trend is indicated by the pandemic, or it can be considered as the natural evolving of this type of business model.

2.3 Comparison with conventional bank lending

To gain a better view on the business model of the marketplaces, their main features should be analyzed in the light of the traditional lending. According to Freixas and Rochet (2008), the main function of the conventional commercial banks is to provide loans and receive deposits from the public. These financial institutions mostly utilize the deposit to finance loans. They are private entities; however, they ensure a public good, specifically accessible, and effective payment system. Based on the banking theory, four main bank functions can be differentiated:

- Providing liquidity and payment services
- Transforming assets
- Managing risk
- Processing information and monitoring borrowers

Additionally, the resource allocation role of the banks, supporting economic growth is widely discussed in the literature (Gerschenkron, 1962; Greenwood & Jovanovic; 1990; Hellwig, 1991). Besides that, banks also have significant contribution in mitigation of macroeconomic shocks and risk sharing (Allen & Gale, 1995). Additionally, owing to their procyclical nature, banks can potentially intensify the impact of economic cycles (Banai, 2014).

The operation of the peer-to-peer segment is less complex, and it is barely embedded in the economic cycle, therefore there are significant differences between the two financial intermediators. Some of the core banking functions are not applicable for the platforms, specifically, they do not provide payment services, neither collect deposit or perform asset transformation in the conventional way. However, risk management and information processing appear in their operations. Platforms do not hold capital in their balance sheet, their functioning is more like a brokerage activity. Due to the lack of capital, they are less exposed to default risk and the capital regulation framework applies to banks is not relevant for them.

Dömötör and Ölvedi (2021) made a comparison of traditional bank funding and peer-to-peer lending features, based on the relevant aspects of financial intermediation. Their analysis is presented in Table 2.

2. Table: The comparison of bank and P2P features

	Bank	P2P Platform
Regulation	Since banks finance loans from collected deposits, they are exposed to several risk factors. In order to manage risks appropriately, international standards were introduced for banking regulation (the Basel Accords).	There is no comprehensive regulatory framework in the P2P segment. However, some initial regulations have appeared in Europe, specifically in the United Kingdom, Lithuania, and Switzerland. The main issues that these laws cover are transparency and risk management. Platforms are not collecting deposit, instead they grant the technical framework for intermediation. Therefore, the related regulations are not that extensive, compared to banks, meaning that the regulatory compliance cost is lower as well.
Investors	Depositors are protected. Different authorities are responsible for deposit insurance in each country.	In general, investors bear the risk as the funding decision is on their side. There is no regulation on investor protection. Some platforms have introduced investor protection measures, such as buyback guarantees, and payment guarantees. The guarantee means that in case of non-payment, platforms take over the claim and reimburse the remaining portion to the investors. Defaulted loans are often passed to a third-party debt collector, where the recovery rates range between 7-12% (Prosper, 2023).

Capital allocation and economic role	Banks play a crucial role in the efficient allocation of capital in the economy. A few decades ago, financial institutions were the only entities to offer these services in the market. Banks also provide a buffer in the economy in case of an external macroeconomic shock. However, the network effects and potential contagion are also essential to consider (Csoka & Jean-Jacques, 2018)	These platforms apply innovative technologies, such as machine learning and artificial intelligence, in the credit assessment process in order to improve efficiency. The funding process is based on an online auction, in which supply, and demand determine the final interest rate. Due to the moderate level of regulation, the platforms have broader resources for IT developments. Such platforms also serve the underbanked-customer segment with lower credit ratings.
Liquidity and payment services	Banks offer money-changing services and manage the savings of depositors. Further, through their clearing activity, they facilitate the money transfer process.	Platforms do not provide money-changing and payment services. Some platforms operate in a secondary market to ensure liquidity for the lenders.
Asset transformation	Banks perform several asset transformations. They can transform the size of the product, e.g., by providing loans in small amounts from a large deposit investment. Quality transformation refers to the fact that the risk-return characteristics of the deposits and loans can differ.	Platforms do not perform asset transformation in the classic interpretation of the term. However, they enable investors to compile diversified portfolios. In addition, large loan requests can be funded with small investments. There is no possibility for maturity transformation, as the platforms do not own capital and funding is given directly, without

	Maturity transformation is also essential, as it turns short-term deposits into long-term loans. While banks have interest gain on maturity transformation, there is a cap for the maximal level of interest rate they can offer.	intermediary transformation. However, there is no cap for the maximal level of interest rate they can set.
Risk management	Banks have a comprehensive risk management framework that covers credit risk, liquidity risk, market risk as interest rate risk, and operational risk. Regulators carefully monitor risk management practices.	Platforms apply different risk management measures, e.g., credit assessment, secondary market, buyback guarantee, and rejection for those under a certain score, etc. Currently, market risk is not relevant to them, while credit and liquidity risks are passed to the investors. However, operational risk, including risks regarding cybersecurity, is high for these platforms.
Information processing and monitoring	Due to information asymmetry between investors and borrowers, banks play a key role in monitoring potential applicants and their performances. In terms of client data, banks are not allowed to request sensitive information from the applicants.	Platforms as well as investors perform monitoring. Platforms perform credit risk assessment based on the information provided by the borrower and on data from credit bureaus. The decision to fund is made by the investor, thus some platforms set their portfolio data to be publicly available to investors, so that they might investigate historical performance. Further, more alternative data is used for information processing, such as the picture of the applicant. Platforms

	cannot request sensitive information either, however they can incite applicants to provide such data on voluntary basis.

Source: Based on Table 2 in Dömötör&Ölvedi, 2021, page 395

2.3.1 Market imperfections

The reason behind the existence of financial institutions is often explained with different market imperfections which is considered as a widely discussed area within bank theory. In practice, financial markets are not frictionless, therefore perfect diversification is not possible. Due to the lack of complete markets, banks play a crucial role in the economy in order to overcome these imperfections (Freixas&Rochet, 2008). Financial institutions mitigate information asymmetry, reduce risk and transaction cost and provide liquidity, which leads to more efficient resource allocation. In order to examine if peer-to-peer lending sites have any added value compared to conventional financial institutions, the market frictions should be investigated from the platforms' perspective.

Different transaction costs - both physical and technological – increased the need for the appearance of financial institutions. Due to the indivisibility of financial products and the economies of scale, financial institutions act as intermediaries between lenders and borrowers which results more beneficial outcome for both parties compared to direct finance. Banks reduce fixed cost and maintain profitable operations. The cost is partially coming from searching and monitoring activity which banks can perform in a more efficient way in contrast to individual market participants (Pringle, 1975). In order to quantify the unit cost of financial intermediation, Philippon (2014) examined the US market and found that it has been stable around 2% since more than 100 years, in spite of the technological developments e.g. the digitalization of the banking segment. Bazot (2018) had the same conclusion regarding the European market. Financial intermediation costs cover the sum of all fees and spreads paid to financial institutions for their services.

In terms of social lending, as discussed previously, one of their main advantage is the low operational cost, which is coming from their less complex, mostly brokerage nature activity, the lack of regulations in the segment, the absence of branch network and the use of advanced techniques in their processes. From revenue perspective, the marketplaces charge origination fee and servicing fee on lenders and borrowers. Based on these charges, Wolfe et al. (2017) estimated the intermediation cost of the platforms. According to their calculations, the cost is between 2% and 3%, which is in line with the previously mentioned bank expense.

Asymmetric information between lenders and borrowers is another commonly discussed feature of the traditional financial markets. It arises from the fact that the debtor has clear information about his or her own credit quality, in contrast, lenders have limited knowledge. The existence of information asymmetry might lead to adverse selection and moral hazard (Akerlof, 1978). In the first case, the lender is not able to differentiate between borrowers with varied credibility due to the lack of information. Therefore, conditions are set to match with weak credit quality borrowers which keeps away the high-quality ones. In case of moral hazard there is no incentive for one party to fend off risk, as the other party will bear the cost (Stiglitz&Weiss, 1981). The mentioned frictions might lead to credit rationing on the loan market, meaning that the borrower is not able to obtain a loan even if he or she is willing to pay higher interest rate than requested. In spite of the overdemand for loans, lenders are not increasing interest rate, as it would diminish the stake of the borrower in the project and the probability of repayment (Tirole, 2010). Banks have a significant role in the mitigation of information asymmetry through their monitoring activity. As a financial intermediator, it collects and processes information in a cost-efficient way, therefore the monitoring is delegated to the bank (Diamond, 1984). It has to be mentioned that state subsidy could also mitigate this impact and reduce the moral hazard (Berlinger et al; 2017).

As a first sight, the information gap between lenders and borrowers seems to be a more crucial issue in case of platform lending, where the investors bear the credit risk instead of the platform, who evaluates borrowers. From one side, the marketplaces can obtain less information regarding the potential debtor and the credit assessment methodology of the platforms is not publicly available, therefore it is a “blackbox” from investor perspective. From the other side, they often complete the risk assessment with alternative information, besides the traditional variables. Alternative information might cover optional narratives provided by the applicant, pictures, or social capital. Some marketplaces allow participants to join to networks within the community of the platform, which can be considered as a signal regarding the credit quality of the borrower. This relationship related information is utilized by investors when making their finding decisions, furthermore borrowers with social network are less likely to default (Lin et al; 2013). Besides that, the use of social collateral and soft information enables the platforms to fund applicants with low risk and low assets, who would have been rejected by financial institutions (Liu et al; 2020). However, it has to be noted, that the platforms are not lending from their own capital, thus the signaling is less convincing in their case, compared to a bank. Overall, the platforms are mostly operating in the retail segment, which is heavily impacted by information asymmetry, therefore it suggests that they might possess some advantage in information processing.

Financial institutions also handle the problem of liquidity issues. Fundamentally, investors prefer short-term investments, due to the possible threat of a liquidity shock, therefore long-term investments would not obtain funding without an intermediary. To overcome this deficiency, banks provide liquidity services and investor protection offers. Large part of the marketplaces operates secondary market in order to enhance liquidity and offer buyback guarantee conditions, which means that the platform purchases the claim from the investor in case of a default. In spite of these measures, it has to be mentioned that in the beginning of the pandemic outbreak, several investors withdraw their investments from the platforms, which suggests that the trust towards social lending is still not solidified.

2.4 Main research directions

The segment of peer-to-peer lending has a relatively short history, as the first platform was launched in 2005. Even so, numerous papers were published in order to investigate the different aspects of marketplace lending. The vigorous interest towards the segment was further supported by the publicly available loan books of the platforms, serving as a basis for empirical analysis. The intention behind the data availability was to enhance transparency and the trust of investors who had the opportunity to perform their own estimates on historical data. Nowadays, only a few sites provide open access to their loan book, which might be due to intensifying market competition.

The papers investigating social lending can be classified into three main groups:

- Business model and the role of the platforms
- Portfolio performance and default risk
- Determinants of funding

The articles related to the first topic, are examining the potential role of the platforms on the financial market and the reason behind their fast-growing trend. They also analyze the interaction with commercial banks. The second stream focuses on the information processing method and the scoring model of the marketplaces, covering the determinants of default prediction. The last field investigates the main factors which impact investor decision when provide funding to the potential borrowers. The following sections summarize the key findings of the literature related to each research area.

2.4.1 Business model and the role of the platforms

This field focuses on the most quintessential questions, specifically the reason behind the emerge and the robust expansion of the marketplaces and their possible role on the financial market. This stream also covers the existing business models, regulatory aspects and the potential future of the platforms. One of the key questions discussed by the papers is the segment's relation with conventional banks. Different scenarios are forecasted, e.g. the two intermediators might compete as there is overlap in their customer segment. It is also questionable if the platforms have complementary or more like substitute function to bank lending. From the other hand, it is also plausible that they build strategic partnership which brings mutual benefit for both parties.

De Roure et al. (2021) investigated the competition between banks and platforms through their theoretical model which they tested on German market data. They simulated an external shock, indicated by regulatory cost increase, which impacts only the banking sector. According to their findings, prior to the shock, platforms had complementary function in half of the cases (where the level of risk of P2P loans was higher than the risk of bank loans). After the shock, the unaffected banks raise their lending, but only in case they are well-capitalized. In contrast, if the unimpacted banks are not well-capitalized, marketplaces have advantage to replace the missing credit supply, which results an aggregated decrease in market share of banks. Similarly, Tang (2019) simulated the impact of an external shock demonstrated by a regulatory change on the bank lending supply which indicated tighten funding criteria. The results support the substitute role of the platforms as there is overlap in their customer base. However due to the low fixed cost of the platforms, they have advantage in providing small amount loans, therefore they can act as complements in this specific segment. Cornaggia et al. (2017) also identified dual function of the platforms. In case of high-risk borrowers, alternative funding substitute conventional lending, however for low-risk ones, it might have a complementary role and supports credit expansion to credit constrained applicants.

On the contrary, Cole at al. (2019) examined the link between bank failures and alternative finance using a large US dataset. Although their analysis covers all crowdfunding platforms not only peer-to-peer marketplaces, their results are relevant to consider. They found that a bank failure lead to lower number of crowdfunded projects in a particular country. Therefore, the two intermediators are complements. Milne and Parboteeah (2016) presume that banks and platforms will enter into a cooperative relationship instead of rivalry. According to their view, banks own competitive advantage due to their access to money market funds, furthermore they are more flexible in providing liquidity services. Consequently, the marketplaces' role is more complementary. It is probable that banks will collaborate with them e.g., let platforms offer funding to existing bank customers or to borrowers who face constraints to obtain bank funding. The complementary approach is further supported by Liu et al. (2020) who highlighted the importance of social collateral and soft information (e.g. social capital or behavior) utilized by the platforms during credit risk assessment. As a result, platforms overcome some deficiencies of the traditional credit market and serve a customer segment who is underbanked. These are low risk borrowers who own little asset, and their credibility is justified through soft information which is not applied by banks. Therefore, P2P segment has a value creation ability through extending the credit access to unserved borrowers. Molnar (2018) share the complementary view and highlights that unlike banks, marketplaces are not performing maturity transformation. Therefore, maturity mismatches exist and there is a potential for high systemic risk in case the platforms offer more complex products in the future.

Jagtiani and Lemieux (2018) strengthen the concept that the platforms' core customer base is the underbanked clients. According to their empirical analysis based on US data from LendingClub, the marketplaces have stronger presence in regions where the branch network is weaker. Besides that, the extension of alternative loans is more significant in areas where economic performance is modest. This reasoning is further supported by Havrylchyk et al. (2017) who ascertained that the extension of social lending is related to modest density of branch network and weaker bank concentration based on their analysis on Prosper and LendingClub data. In line with their view, branch density is related to the advertising strategy of banks and brand loyalty. Therefore, the outcome suggests that brand loyalty towards banks is weaker in the mentioned areas.

The paper of Oh and Rosenkranz (2020) is in line with the previously presented studies. They investigated the determinants behind the segments robust expansion based on a dataset of 62 countries. They explored that marketplace lending is more widespread in economies where the access to conventional financial institutions is constrained e.g. due to the lack of physical infrastructure. Furthermore, they found that financial literacy also has a positive impact on the spread of the segment. The importance of financial knowledge in the expansion of P2P lending is further supported by Han et al. (2019) in case of the Chinese market. They also highlight that the robust growth of the segment can be increased with financial literacy education.

Hemer (2011) identified a somewhat unique perspective regarding the role of crowdfunding. Namely, crowdfunding has become a widespread alternative form of finance, especially in creative and innovative segments. Therefore, it can motivate private investors to mobilize their capital and probably support the funding of early stage start-ups in the future. Overall, it completes traditional finance as it supports borrowers at the initial stage of their operations when they are not yet bank eligible.

Another exciting field of this research stream is covering the risks related to the business model of peer-to-peer lending. Morse (2015) raise the attention, that a potential financial distress would impact the loan portfolio of the platforms. Furthermore, in this model investors are pooled – as more lenders finance one loan – therefore they are exposed in case of an unexpected liquidation. Besides that, the lack of transparency in case of credit scoring method and general operations of the platforms need to be handled in the future. Currently, the lending practice and operating model of the platforms is quite diverse, therefore the introduction of a comprehensive segment regulation is relevant. Davis (2016) also emphasizes the importance of the legislative framework of the segment – with special focus on the Australian market - as the platforms perform different financial and economic activities (e.g. credit risk assessment, manage the delivery of the loan etc.), which needs to be reflected in the regulation. The current laws cover these activities separately, thus a broaden approach would be desirable, which specifically address the risk factors arise from the operation of the platforms. Namely, this type of investment is relatively new for investors who are not informed properly, the platforms credit assessment quality is questionable, agency risk between the platform and investors.

The nature of the segment is further discussed by Käfer (2018) who concluded that platform lending is riskier than conventional bank lending. This ascertainment is confirmed with the following arguments: soft information used by the sites might be advantageous in credit assessment, however this kind of data is usually unverified. Over and above, the existence of herding behavior, the potential risk of platform default and the growing presence of institutional investors - which enhances the possibility of systematic risk - all support this finding. Besides that, the business model and operation of the platforms meet the definition of shadow banking. Specifically, marketplaces are outside of the regular banking system, external backstops are needed in their operations, they don't have access to central bank reserves etc. Based on the conclusion of the paper, the segment can be considered as part of the shadow banking system. In line with the previously mentioned concerns, Kirby and Worner (2014) highlights that due to the robust growth rate, there is a potential in the segment to become remarkable investment possibility in the future. However, due to its emerging embeddedness in the economy, it might increase the threat of systematic risk. It is also emphasized that possible insolvency issues in case of cross-border lending is another question which should be addressed with a unified regulation.

Finally, it has to be mentioned that peer-to-peer lending is primarily related to the financing, however, the segment is also relevant to other fields due to its unique features which distinct the marketplaces from banks. Caldieraro et al. (2018) highlights the existing information asymmetry from marketing perspective, while Wang et al. (2015), Au et al. (2020) and Chen et al. (2014) examines the technology management aspects through the different development stages and strategies of the platforms.

2.4.2 Portfolio performance and default risk

This stream forms another significant part of the marketplace lending literature, based on loan book data of the platforms. The papers investigate the determinants of default prediction, the performance of the scoring models applied by the platforms and the significant factors, which have impact on the pricing. As mentioned, the regulatory framework of the segment is still in its early stage, therefore the risk management framework is not regulated, which results higher default ratios compared to the defaults observed in case of traditional bank portfolios. Moreover, the consequences are limited in case of non-payment which might has a negative impact on the willingness to pay. From the other side, the platforms often apply advanced methods during credit risk evaluation and incorporate the previously mentioned soft information or alternative data, which overall has the potential to enhance their risk management practice.

Notable part of the papers investigated the relation between the probability of default and different credibility related financial variables, e.g. pervious loan history. Emekter et al. (2015)

examined the default prediction using the US LendingClub portfolio. Based on their empirical analysis, the probability of default is significantly impacted by the low FICO score, low credit grade, high revolving credit utilization and high debt-to-income ratio. Besides that, it seems that the higher interest rate set for the high-risk borrowers is not sufficient in the light of their non-payment probability. Comparing P2P borrowers to an average US debtor, debt-to-income ratio is higher, and the average income is lower for marketplace participants. Lin et al. (2017) examined the same question on the Chinese peer-to-peer market, where there is a lack of unified credit score like FICO in the US. Their results suggest that low debt-to-income ratio, good previous payment history, long employment history and high level of education are all contributing the low probability of default.

Apart from the conventional variables which are frequently applied by commercial banks and mostly related to the financial background of the potential debtor, marketplaces often utilize alternative information. Jagtiani and Lemieux (2018) explored that the correlation between P2P ratings and FICO score declined from 80% to 35% in the period of 2007-2015, suggesting that the contribution of nonconventional data in the marketplace assessment increased significantly. Interestingly, the platform ratings proved to be a good indicator in the loan performance prediction based on two years' time horizon from loan origination. Overall, the utilization of alternative data enables a particular group of borrowers, who would have been classified as subprime according to bank criteria, to obtain funding. The importance of alternative data is also supported by Croux et al. (2020) who explored significant role of this information in determining loan default. They also identified that certain loan purposes, specifically medical expenses and SME loans proved to have higher probability of default than other categories. Ge et al. (2017) examined self-declared social media information, covering the activities and the network of the debtors on the social site for borrowers who provided their account during the P2P application process. Their analysis proves that adding the social media account has positive relation with low default probability. Besides that, the extended friendship network and the active behavior on the social site also has a positive impact on the likelihood of non-payment. The paper suggests that the underlying reason behind the role of the social media presence is the reputation risk of the borrower.

Besides the borrower related information, it seems that macroeconomic variables also have notable impact on the probability of default and pricing. Wang and Ni (2020) examined the US P2P loan book on aggregated level instead of the previously used transaction level approach.

They modelled the trend of the default rate and using long short term memory (LSTM) model, their results suggest, that including unemployment rate significantly improved the performance of the prediction model. The role of macroeconomic data is further evidenced by Foo et al. (2017), who investigated the relation between the marketplace segment and the general economy. The outcome suggests that the disparity in P2P spreads is correlated with the non-default rate, market uncertainty and the fundamental value of the equity market. The significance of macroeconomic factors on the variety of lending rates is also supported by Dietrich and Wernli (2016) who performed regression analysis on Swiss marketplace data and explored that unemployment rate and three-year government bond yields proved to have a positive correlation with increasing platform interest rates.

Another interesting investigation is related to the methodology of precise borrower classification, specifically the distinction of good and bad debtors. Several techniques are demonstrated which can be utilized by investors and by the platforms as well in order to improve their credit risk assessment. Bhuvaneswari and Segalini (2020) raised that the disproportion between good and bad debtors might lead to distortion in the accuracy of the prediction model, which is the case for P2P portfolios. The authors suggest advanced methods to handle this issue. Using class rebalancing techniques and mixing different probability prediction approaches would improve the accuracy of the estimation. Byanjankar et al. (2015) suggest the utilization of data mining methods, namely the application of artificial neural networks in the credit scoring methodology to classify performing and defaulted debtors. Based on the estimations conducted on the dataset of a leading Estonian platform, the outcome implies that the proposed approach has a good classification and prediction performance. In line with this finding, Ahelegbey et al. (2019) advise the application of network-based method and highlight the importance of a systematic risk aspect of the segment, therefore the key is to explore the interrelation among borrowers. According to their reasoning, the investigation of network structures allows the identification of deeper relations and latent factors within the community. Although their analysis was performed on a P2P dataset of SME loans, the conclusion is in line with previous papers, as the model led to better prediction performance.

Finally, the expected credit loss and profitability are essential questions from investor perspective and serves as a basis when the investment strategy of P2P lenders is determined. Klafft (2008) identified a few investment rules which would enhance the profitability of these alternative investments. Specifically, it is advisable to avoid delinquent borrowers with credit inquiries in the last six months and a debt-to income ratio lower than 20%. The paper highlights that the average return is quite attractive in case of the three strongest rating categories, exceeding the highly rated US treasuries. Contrary, lower ratings showed negative performance.

2.4.3 Determinants of funding

The most extensive stream of the peer-to-peer lending literature is related to the determinants of investor decision. As mentioned, information asymmetry is a commonly discussed issue of the lending markets. Though, in the segment of platform lending, information gap is even more crucial. On the one hand investors own limited knowledge regarding the credibility of the applicant. On the other hand, the platforms' credit risk assessment methodology is not transparent, and therefore its reliability might be questionable. Lender’s decision generally relies on two types of information: hard and soft data. The first is mandatory, e.g. the income of the applicant, while the latter is optional and mostly unverified. This field investigates the determinants of successful funding, covering the previously mentioned data types with special attention on the possible contribution of alternative information, which might have signaling effect regarding the credibility of the protentional debtor. The results can vary based on the application data of each platform, which might be impacted by the economic background, social features, or the culture of the country where the marketplace operates.

As a general remark, Chen et al. (2014) emphasizes that the willingness of financing is highly dependent on the trust of investors in borrowers and in intermediaries. They conducted a survey in cooperation with one of the leading Chinese platforms to examine lenders' perspective. The results of the questionary suggest that the first factor, namely the trust in borrowers is more crucial in order to enhance the inclination of funding. Besides that, the increased trust in applicants has a positive impact on the lenders’ opinion regarding intermediaries. Consequently, applicants should bestow high-quality information, which should be coupled with high quality services and appropriate security background ensured by the platforms.

As it might be expected, the commonly examined variables related to the financial stability of the applicant are investigated in the literature. Herzenstein et al. (2008) evidenced the role of the credit score and the loan characteristics in investor decision using application data from the US Prosper. According to the results, the likelihood of a successful funding decision is higher in case of applicants with stronger credit score and lower debt-to-income ratio. The pricing of the loan is in line with these observations. Over and above, the outcome implies that lenders prefer lower loan amount which carries less risk. These results are in line with Gavurova et al. (2018) who examined the same question in case of Bondora, a leading platform in the Baltic region, founded in Estonia. Besides credit rating and debt-to-income ratio, the level of education proved to be a relevant factor in the European market.

Apart from the conventional factors, the role of soft information is a widely investigated area from investor decision perspective. One of the platform specific soft information is coming from narratives. Applicants have the opportunity to provide a short description as part of their listing, where they expound their loan purpose and add any relevant information which might enhance their credibility. Herzenstein (2011) et al. denoted the importance of narratives using Prosper application data. They defined different identities based on the context of the narrative, e.g. hardworking or religious. The outcome suggests that the number of identities positively correlates with the funding success, however interestingly these borrowers paid back the loan with a lower probability. Content wise, the use of trustworthy and successful expressions led to higher funding success. The relevance of narratives is further evidenced by Larrimore et al. (2011) who found that lenders prefer detailed description with concrete statements. In contrast, personal and emotional details decrease the trust of investors.

Generally, first impression and appearance are remarkable factors in case of labor market decisions. The literature investigates this aspect from marketplace lending perspective. Platforms enable potential debtors to attach picture to their application, which is another nonconventional source of soft information. Duarte et al. (2012) examined a wide range of pictures provided by Prosper applicants to analyses the role of appearance in P2P transactions. They involved independent third party to adjudicate on the trustworthiness of each applicant based on their photograph, focusing on the willingness to pay, not the ability. Using this independent judgement, they concluded that applicants rated as trustworthy have higher probability to obtain funding and they receive lower interest rate compared to ones who are not considered credible at first sight. Interestingly these borrowers have higher credit rating and lower change of non-payment. Ravina (2019) examined the pictures from attractiveness perspective and proved that good-looking borrowers obtain loan with 11.7% higher probability then average looking ones, having the same financial background. Though, the event of non-payment is more common among attractive borrowers. Applicant photos are further analyzed by Gonzalez and Loureiro (2014) who focused on personal characteristics. Their research implies that investors are indifferent regarding the attractiveness of the applicant in case the perceived age is around middle age which suggests trustiness. However, when borrower and lender have the same gender, the attractiveness of the applicant might be detrimental on the outcome.

It seems that personal judgement of investors has a significant impact on the funding decision along with financial metrics. While the signaling effect of soft information can be beneficial to enhance the credibility of the borrower, it also brings the possibility of discrimination. Pope and Sydnor (2011) examined this question on US data focusing on the race of the applicant. According to their conclusion, in case of black applicants there is a 25-35 percentage lower change to obtain funding compared to white debtors having the same characteristics. Besides that, the average interest rate set for black borrowers is also higher compared to white applicants. Barasinska and Schäfer (2014) analyzed the presence of discrimination from gender perspective using data from one of the leading German platforms. According to their finding, male and female applicants with similar credit background have equal chance to obtain funding. In contrast, Chen et al. (2017) evidenced the incidence of gender discrimination in case of a remarkable Chinese platform. The results denote that investors prefer female applicants; however, they receive higher interest rate. In line with this observation, female borrowers have a significantly lower non-payment rate compared to male. Interestingly, in case of the US market, using Prosper data, Kuwabara and Thébaud (2017) explored that females requesting loan for business purpose are less likely to receive funding. This observation suggests stereotype that running a business is more like a male aspiration. Barasinska (2011) investigated the impact of gender differences on the risk appetite among investors using transaction data from a German marketplace. Based on the standard deviation of the projects' expected return, there is no evidence for any difference in the risk-taking attitude between male and female lenders. Although the paper points out some disparity regarding loan term and loan purpose. Female investors prefer short-term loans and customer loans instead of business purpose.

Finally, the impact of social collateral is another interesting aspect of investor decision. Some platforms enable participants to form groups and creates social networks within the platform, e.g. borrowers who belong to the same employer or university alumni can indicate their credibility this way. The membership is optional, and the group leader provides access based on verification. Hildebrand et al. (2010) investigated this market specific soft information and found that group membership has a positive impact on the funding decision. Furthermore, debtors obtain lower interest rate and the probability of non-payment is also lower. Lin et al.

(2013) supported the relevance of friendship networks and proved that investors consider membership as a sign of credibility, therefore it led to higher likelihood of successful finance and more beneficial interest rate. Freedman and Jin (2017) denoted that groups are supposed to screen the members and incites them to repay the loan. However, the existence of these social relations outside of the platform is unverified, therefore it does not serve as exact evidence for investors regarding the credibility of the debtor. Using US Prosper data, they proved that lenders prefer applicants with more social connections, which is in line with the previous findings of the literature. Although, they note that better repayment performance was observable only in case the friend who bestow endorsement also provided funding to the request.

2.5 Regulation

However, the market of peer-to-peer lending demonstrated a robust expansion in the previous decade, its long-term maintenance and potential to become a permanent participant of the financial market is highly dependent on the implementation of a comprehensive regulatory framework. Currently, the segment is scarcely regulated. Although recently the segment raised the attention of supervisory authorities and market participants. Consequently, more and more marketplace specific regulations were developed especially in the regions where social lending is the most extended. Most of these regulations emphasize the risk factors and challenges in terms of the business model of P2P lending and provide recommendations instead of exact guidance. The purpose of this section is to summarize the current regulations in force focusing on the prime markets of platform lending.

In terms of the US, Prosper and LendingClub (as mentioned in section 2.4., LendingClub announced to close its retail P2P segment and transform into a hybrid digital bank at the end of 2020, however its operations in the last fourteen years heavily influenced the current regulatory framework of the US) owned the vast majority of the market and the two intermediaries operated with similar business model. In practice, it means that investors do not fund the loan directly, instead there is a third-party bank who originates the loan. After that, the platform purchases the claim and issues a security – which is called as note – to the investor (Prosper, 2021). Therefore, lenders invest into securities instead of a substantive loan and the payment of the note depends on the repayment of the underlying loan. As the originated notes are considered as securities, they fall within the scope of federal securities regulation, specifically Securities Act of 1933 (Douglas & Bates, 1933), regulated by the Securities and Exchange Commission (SEC) of the US. According to the Act, all securities traded in the US must be registered on SEC and a preparation of a prospectus is also required in order to duly inform investors. 2008 is counted as a decisive year for the US P2P market, as the SEC ascertained that Prosper did not meet the mentioned registration requirements, therefore issued a cease-and-desist order against the platform (SEC, 2008). Platforms were struggling to comply with the regulatory requirements, which resulted a momentary shut down of the marketplaces after the order was issued. The registration imposed a heavy administrative burden on the platforms which also implied significant cost increase for them. On the other hand, the regulation improved the portfolio performance of the marketplaces, where the non-payment reached around twenty-four percent previous to the issuance of the SEC order. Besides that, the registration requirements deterred other platforms from Europe and Asia to expand their operations and enter into the social lending market of the US (Magee, 2011). As of 2010, the Dodd–Frank Wall Street Reform and Consumer Protection Act requested a regulatory recommendation for the segment. Prosper argued to nominate the Consumer Financial Protection Bureau (CFPB) as the responsible authority for the segment and to be exempted from the SEC regulation. However, this suggestion was refused as it would have been perilous to rely solely on the expertise of one authority, furthermore it would enhance the existing concerns regarding the P2P market (Chaffee & Rapp, 2012).

In case of the UK the market is remarkably developed with several participants, therefore it is relevant to discuss it separately from Europe from regulatory perspective. The Financial Conduct Authority (FCA) was appointed as the responsible entity to regulate marketplace lending. The body has continuously monitored the peer-to-peer industry in the previous decade. Apart from the official supervisory entities, P2PFA was established as of 2011, which is a self-regulatory body urged by the marketplaces of the UK. Its aim is to determine targets for capital reserves in order to protect investors. P2PFA cooperates with the regulatory entities and engaged to meet anti-fraud requirements. The first official regulation was launched by the FCA as of 2014. Their approach was to incorporate the segment's regulation into the existing laws. The rules are mostly in line with the P2PFA endeavor, focusing on lender protection. The regulation required the introduction of capital reserves; however, the platforms found the defined amount of capital overly strict, therefore FCA lowered the requirement (Rogers & Clarke, 2016). The most recent policy statement was published in 2019, specifically tailored to peer-to-peer lending and crowdfunding platforms. The regulation’s focus is mostly on the deepens of investor protection. The rule enhances the transparency as more detailed information need to be posted for investors to support their funding decision. Besides that, the lenders' knowledge and experience in social lending has to be assessed. Furthermore, the guidance advice on the risk management and credit risk assessment practice of the sites and the fair valuation method they apply. For investors who are new in the P2P business and having less experience, an investment cap was introduced, meaning that maximum 10% of their investable asset can be placed in peer-to-peer lending (FCA, 2019).

In terms of Europe, there was no comprehensive regulatory framework until the European Commission published the Regulation on European Crowdfunding Service Providers (ECSP) for business as of 2020 November. After a 12-month transition period, the law is applicable for all EU members. According to the Commission, the EU market is underdeveloped compared to other regions and the lack of common rules was the main obstacle for the European players to operate efficiently. The new rules introduce the clear definition of crowdfunding, specify the authorization process, define the risk management and governance practices, and enhance investor protection (European Commission, 2020). Besides the unified EU framework, some states introduced regulations on country level. Specifically, they ordain for the platforms to obtain license and operate in compliance with the requirements of the 2nd Markets in Financial Instruments Directive (MiFID II) (Jorgensen, 2018). Latvia is a prominent example in this respect, as their regulator, namely the Financial and Capital Market Commission (FCMC) agreed with leading market players, e.g. Mintos, Twino and VIAINVEST to start the Investment Brokerage License application. In practice it means that the supervisory authority will regularly review the operation of the marketplaces who have reporting obligation. Besides that, the measures enhance investor protection, including a state-guaranteed program up to 20 000 EUR per investor in case of platform default. It provides guidance for the platforms how to assess the financial knowledge of lenders to ensure that they understand the risk of this type of investment. Apart from this, some capital requirements are also applicable for the marketplaces (Mintos, 2020; Twino, 2021). Overall, the regulation of the European lending platforms shows a promising trend, although the implementation of the EU wide rules and its harmonization with country level laws might bring some challenge for the member states in the upcoming period.

3. The characteristics of peer-to-peer applicants[3]

In recent years, different forms of social lending have become a widely researched area. One of the most extensive business models is peer-to-peer lending (P2P), an online platform connecting lenders and borrowers. The segment's rapid growth has attracted the attention of market participants, and demand has arisen for a deeper understanding of this new form of financial intermediation. The purpose of this paper is to contribute to the existing literature by examining the borrower side of P2P lending. The analysis is based on a unique, manually collected dataset from a market leader platform of the United States. LASSO regression is used to examine the relationship between applications and a wide range of local microeconomic and socioeconomic indicators. Then, k-means cluster analysis is applied to identify borrower groups with similar characteristics. The results indicate there is a strong positive correlation between the portion of mortgage delinquency and demand for P2P funding. Furthermore, the platform's customer base significantly overlaps with bank clients.

Keywords: Peer-to-peer lending, P2P, Financial intermediation, Borrower groups, Microfinance

JEL Codes: G21, G29

[3] Ölvedi, T. (2022). The Characteristics of Peer-To-Peer Applicants. The Journal of Alternative Investments, 25(2), 66-86.

3.1 Introduction

In recent years, social lending (peer-to-peer, or P2P lending) has appeared in the market as a new form of financial intermediary. It is an online, platform-based solution, linking potential borrowers and lenders to facilitate funding. According to the original concept, the purpose of this business model is to provide more favorable conditions for both parties, compared to commercial bank offers. The first platform was implemented in the United Kingdom in 2005; therefore, the segment is relatively new in the financial markets. Nonetheless, in subsequent years it has garnered the attention of supervisory authorities and conventional financial institutions, due to its rapid global expansion. The total volume of P2P lending in the retail consumer lending segment was 34 billion dollars in 2020 (CCAF, 2021).

Due to the segment's robust expansion and growing notoriety, several studies have investigated this new form of financial intermediary. Since the risk is significantly higher, compared to bank funding, most of the papers focus on portfolio performance and the key factors impacting successful funding. The purpose of this paper is to contribute to the existing literature by examining the characteristics of peer-to-peer applicants and exploring the motivation of debtors for turning to this alternative funding source. The segment's ongoing spread is highly dependent on the demand from potential debtors; therefore, it is relevant to investigate their specifics and intention. The following research questions are examined in this study:

1. Which local economic indicators explain the demand for P2P lending in the US market?
2. Which borrower groups can be differentiated based on similar patterns?

The relationship between different economic indicators and P2P expansion is an essential question in order to understand the drivers behind the segment's growth. Previous studies also discussed this connection focusing on various aspects: Havrylchyk et al. (2020) explored reverse relation with P2P demand and bank concentration, Oh and Rosenkranz (2020) identified that financial literacy positively correlates with P2P lending, using a sample of 62 economies. Jagtiani and Lemieux (2018) found correlation between P2P debtors and credit gap measures. Using a theoretical framework, Polyzos et al. (2021) explored relation between peer-to-peer expansion and economic instability under different scenarios.

The paper contributes to the existing literature of social lending from several aspects. The unique sample covers more than 135,000 manually collected applications that were published on the webpage of Prosper.com in textual format. (Prosper was the first platform in the United States and is currently the market leader.) After manual data collection, the information was realigned in table format using data transformation techniques. Because the sample includes records from 2014 until the end of 2020, the period covers the economic downturn due to the appearance of COVID. In contrast, the time horizon of previous studies included times of economic prosperity only. Besides the unique dataset, the impact of local microeconomic and socioeconomic variables on P2P applications is examined. A wide range of economic indicators is reviewed, to identify the social groups of applicants. Previously, macroeconomic variables were analyzed in association with loan default probability, and thus included only a few indicators. Cluster analysis was also previously utilized for default prediction; however, in this paper it is applied to understand the motivations of potential borrowers.

Analysis proceeded as follows. First, the number of applications by state is examined, using different microeconomic and socioeconomic variables in regression analysis to find associations between economic indicators and demand for social lending. Then, k-means cluster analysis is performed to identify borrower groups with similar patterns. The results indicate that the portion of delinquent mortgages has the highest impact on the demand for P2P funding, and the most commonly declared loan purpose is debt consolidation, which suggests that borrowers probably need alternative funding to refinance their overdue mortgage debts. According to the literature, loan delinquency and refinance using alternative funding is related to the lowering regulatory burden resulted by fintech expansion (Cornaggia et al. 2018, Buchak et al. 2018). The results are in line with Ramcharan and Crowe (2013) who explored that falling house prices negatively correlate with P2P conditions, speeding up delinquency. Overall, the contribution of this paper is that it investigates the demand for marketplace lending instead of the loan conditions. Besides that, previous studies focusing on the spread of the segment, examined the relation with macroeconomic variables (Jagtiani & Lemieux, 2018; Havrylchyk et al. 2017), while this paper includes an extensive dataset, covering a wide range of microeconomic and socioeconomic indicators. Furthermore, the analysis was performed on a unique sample, including data from the period of economic turbulence due to the pandemic.

In the second part of the analysis, four types of P2P customers can be differentiated, and a large fraction of them are eligible for bank funding. The results suggest that the customer group of P2P platforms overlaps with bank customers and P2P platforms supplement bank lending only in a small segment and for most of the cases it substitutes bank funding, especially in the lower end of the score distribution. Previous literature explored that platform lending is more accessible for bank eligible clients (Tang, 2019), however according to the author's best knowledge, this is the first paper which applied cluster analysis to investigate the possible complementary or substitutional role of the platforms. Besides that, more granular customer groups are differentiated, compared to the previously identified high-risk and low-risk borrower classes (Roure et al. 2016).

The structure of the paper is as follows. Section 3.2 provides an overview of the existing literature on P2P lending and highlights its key findings. Section 3.3.1 presents a high-level overview of Prosper application data, while section 3.3.2 analyzes the statistics from a bank eligibility perspective. Then, section 3.3.3 details the LASSO regression analysis of economic indicators on a state-by-state basis and, finally, section 3.3.4 provides a k-means cluster analysis. Section 3.4 concludes and offers possible practical implications.

3.2 Literature review

Even though the first marketplace was implemented only 16 years ago, P2P lending has an extensive literature. The relevance of these social platforms and their influence on the financial markets is still not fully revealed. Besides retail investors, the segment has also gained the attention of supervisory authorities and conventional financial institutions in several countries, and demand has risen for researchers to provide a deeper understanding about this new form of financial intermediary (Lenz, 2016; Macchiavello, 2014).

This paper contributes to this research, by examining the main reasons behind the emergence of the platforms, focusing on the borrower side. Rubanov et al. (2019) investigated the extension of alternative financing (covering different types of crowdfunding) by region in the 31 countries with the highest lending volume. According to their results, the volume of P2P lending has the highest separating force and impacts the global alternative finance market the most. Regarding the role of the marketplaces, Milne and Parboteeah (2016) presume that the platforms will serve as complementary funders to banks, since conventional financial institutions have comparative advantages in the market, e.g., in terms of liquidity. De Roure et al. (2016) further elaborate the role of the platforms based on borrower groups. According to their results, high-risk debtors who apply for P2P loans are likely to be unserved by traditional banks. Jagtiani and Lemieux (2018) examined U.S. data and found that P2P lending mostly impacts underbanked areas where economic performance is lower. Havrylchyk et al. (2017)

found that P2P lending is more extended in regions where bank concentration and the density of branch networks is weaker.

This paper also relates to the substantial part of the P2P literature that examines default predictions and credit risk assessments of debtors. The scope of those studies is the proper borrower classification. Several advanced techniques are applied to achieve more accurate predictions, which can be utilized by investors. Bhuvaneswari and Segalini (2020) investigated the credit risk assessment of borrowers through secondary aspects. Machine learning techniques and clustering methods were applied to improve the identification of good and bad debtors. Ahelegbey et al. (2019) examined the outcome of the scoring model for SME loans using a factor-based classification method, splitting the sample into network communities. According to the results of Croux et al. (2020), alternative data has a significant role in the determination of potential default, and loans used for medical expenses and small businesses hold higher risk. Emekter et al. (2015) proved that traditional data, specifically FICO score and debt-to-income (DTI) ratio, impact the default probability. Wang and Ni (2020) studied the probability of default on an aggregated level of data to explain default rate trends and found that incorporating the unemployment rate can enhance the model's performance. Foo et al. (2017) also proved that macroeconomic factors correlate with the P2P market, specifically with the credit spread. The role of macroeconomic variables on risk-based pricing is further supported by Dietrich and Wernli (2016), based on evidence from the Swiss social lending market.

The purpose of this study is to contribute to the existing literature, by examining social lending platforms from the borrower's perspective. To gain more comprehensive view on demand side of the platforms, it is essential to identify the different borrower groups and their motivations for using this alternative form of finance. Furthermore, based on the literature, it seems that economic indicators impact the performance of P2P portfolios. Therefore, it is relevant to analyze which local economic factors have a significant influence on the volume of applications, which supports the identification of different social groups who serve as pools of social lending customers.

3.3. Analysis of Prosper data

As can be seen from the foregoing discussion, the relevance of P2P lending is a commonly researched question. The intension behind the increased demand for social finance is still not fully understood. In order to answer this question, it is essential to compare the conditions offered by the platform with conventional banking opportunities. Furthermore, the social groups of P2P applicants have to be identified.

The first part of this section provides an overview of the data; the second part looks at candidates from a bank eligibility perspective; the third outlines the social groups of P2P borrowers from an economic point of view; and the last part attempts to identify the groups with cluster analysis.

3.3.1. Data

As a representative sample, the dataset of Prosper was chosen. Prosper is currently the market leader lending platform, operating in the United States. In previous years, Lending Club dominated the market, with a strong institutional investor background and a constantly growing lending volume. However, in November of 2020 Lending Club announced that it would close its platform to retail investors at the year end and transform into a fintech marketplace bank (Lending Club, 2020). As a result, Prosper has the strongest presence in the P2P market, with a total lending volume of $18 billion dollars from origination. Prosper was founded in 2005, as the first entrant into the U.S. market, and it facilitates small-sized retail loans between $2,000 and $40,000 (Prosper, 2021). Applicants must provide information regarding their financial and socioeconomic status. Then, the data are supplemented with information from a credit bureau regarding the borrower’s historical payment performance. Based on this information, the platform performs a credit risk assessment, and assigns a rating and the corresponding interest rate the applicant will pay.

The full portfolio of Prosper is not publicly available, however its listings are posted daily to its webpage in a report format. The listings are also available on the U.S. Securities and Exchange Commission (SEC) website. The reason behind is that notes are considered as securities and fall under the scope of the federal securities regulation. Therefore, each note must be registered on the SEC in order to properly inform investors. In 2008, Prosper obtained a cease-and-desist order from the SEC as the platform was not fulfilling the regulatory requirements. After a temporarily shut down, the marketplace started to report all loan applications to the supervisory, however the order imposed a heavy administrative burden on the segment (Magee, 2011).

The data of each listed note was collected manually from the webpage of Prosper (Prosper, 2020), under latest SEC filings. All available listings were downloaded from the site of the marketplace between 2014 and 2020. The information is incorporated into the report in a text format. All listed notes have a unique identifier. Besides that, the reports cover data regarding the loan characteristics, e.g., the requested amount, the term, the monthly payment and the investor yield and servicing fee. In terms of borrower characteristics, the FICO score, previous and current payment performance, employment related information, income, occupation, location, and loan purpose are presented. For a listing example, please see a snapshot in Appendix I. Sales reports are also included in the disclosure documents which accompany the notes which are sold to investors. After manual collection, the data was restructured in a table format using automatic data transformation techniques. After that, several data cleaning steps were performed, to eliminate records with missing or invalid fields. In addition, duplicate fields with different name formats were unified (e.g., categorical variables such as the state of the borrower), and when applicable categories with a few records were merged to facilitate interpretation (e.g., when the "loan purpose" covered several different purposes referring to the same collective group).

The final dataset covers the period between January 1, 2014 and December 31, 2020, and after data cleaning the sample contains 135,592 observations. Approved and rejected applications are both included in the dataset (Prosper does not publish information about cancelled records). Table 3 summarizes the main statistics of the data by the income stated by the borrower.

3. Table: Data overview of Prosper applicants by stated income between the period of 2014 and 2020[4]

Stated Income ($)	Portion of total (%)	Mean interest rate (%)	Mean loan amount ($)	Mean investor yield (%)	Mean revolving credit balance ($)	Mean bankcard utilization (%)	Mean monthly payment ($)	Mean FICO score	Mean debt-to-income (DTI) ratio (%)	% of financed applicants
1–24,999	3	22.02	3,421	19.30	6,502	51	125	692	31	69
25,000–49,999	29	19.57	8,953	16.96	9,201	51	294	699	30	70
50,000–74,999	30	18.46	12,161	15.91	13,584	52	386	703	27	73
75,000–99,999	18	17.54	14,361	15.02	18,649	53	450	705	25	74
100,000+	20	17.16	16,721	14.65	33,438	55	523	708	21	73
Total	**100**	**18.46**	**12,287**	**15.90**	**17,022**	**52**	**391**	**703**	**27**	**72**

Source: Author's estimation based on Prosper listing and sales data, 2014-2020.

[4] Prosper verifies the borrower's stated income for a portion of the applicant pool, based on an internal algorithm. For this verification, official documents, e.g. bank statements, are required. (Prosper, 2020).

Note 1: Income was provided in the listing as a range in the report.
Note 2: Interest rates were provided as a range in the report. The average interest rate was estimated as the average of the top and the bottom of the range.
Note 3: FICO score was provided as a range in the report. Average FICO score was estimated as the average of the top and the bottom of the range.

As can be seen from the summary table, the applicants are mostly evenly distributed between the different income brackets except for the lowest one, which is barely represented. The interest rate charged decreases in a monotonous way, meaning that applicants with stable income receive lower rates and investors realize lower returns on these investments as the risk is lower. The average FICO score increases in line with income, which is reasonable. On a portfolio level, most of the applicants (72%) received funding. The average applicant has a revolving credit balance of $17,022, a 52% bankcard utilization, and a debt-to-income (DTI) ratio of 27%. Additionally, 91% of the applicants are employed, including part-time and self-employed candidates.

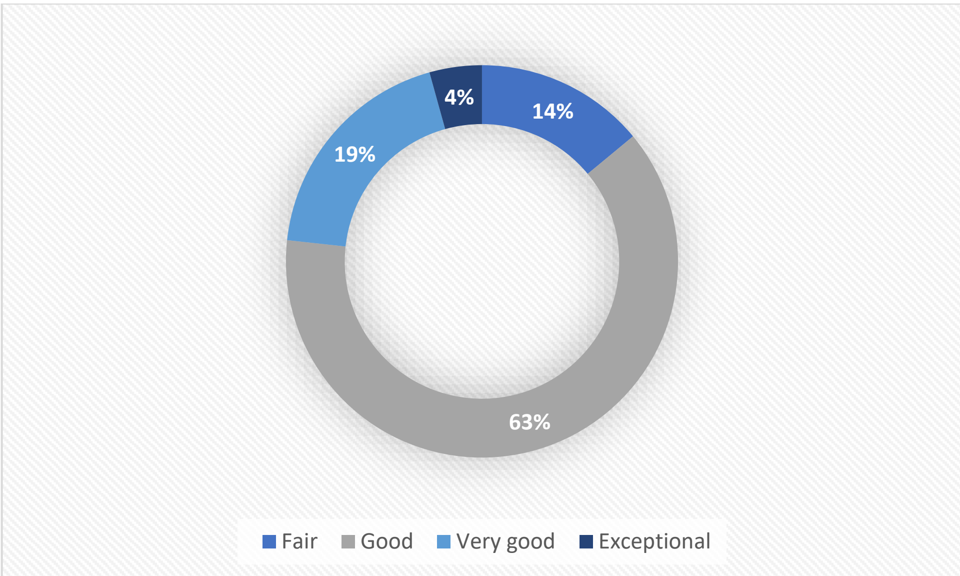
3.3.2. Statistics by bank eligibility

In order to acquire a better view regarding the credit background and bank eligibility of P2P applicants, FICO scores and interest rates are examined in this section. A FICO score is one of the most commonly applied metrics used by commercial banks for credit risk measurement. A FICO score includes a previous payment history, based on national credit bureau data calculated from the credit reports of each debtor. The score was implemented by Fair Isaac Corporation in 1989, and different versions have been released over the years. It has a range between 300 and 850 points (Experian, 2020). A FICO score has five key components, with the following weights (FICO, 2020):

- Payment history – 35%
- Amount of debt – 30%
- Length of credit history – 15%
- New credit – 10%
- Credit mix – 10%

Prosper listings include the credit score of each applicant. Based on the literature, the score is considered to be one of the main factors impacting a funding decision (Herzenstein, et al., 2008; Gavurova et al., 2018). Overall, it is relevant to examine the score distribution of social lending applicants, compared to the national average. Prosper sets a minimum threshold for an application, which is FICO score of 640 (Prosper, 2020); thus, borrowers with poor credit scores are not eligible for a loan. Prosper pulled FICO scores from one of the main credit Bureaus, Experian, however from 2017.13.31. switched to TransUnion data (Prosper, 2020). The score interval of the two credit Bureaus is the same, however there might be differences in their method. As the time horizon of this sample covers data from two different credit Bureaus, the credit quality should be interpreted together. According to Experian (2020), borrowers with scores ranges of 300–579 are considered very poor, while 580–669 is fair, 670–739 is good, 740–799 is very good, and 800–850 is exceptional. In case of TransUnion (2021) the ranges are the following: 300-600 is very poor, 601-657 is fair, 658-719 is good, 720-780 is very good and 781-850 is exceptional. To examine the score distribution of Prosper applicants, the portion of each range is calculated for Experian and TransUnion scales separately and then the weighted average of the two is estimated. Figure 5 presents the distribution of Prosper applicants by credit score, classified according to the previously mentioned method.

5. Figure: Distribution of Prosper applicants by credit score (based on the mixture of Experian and TransUnion credit scale)



Source: Author's estimation based on Prosper listing data 2014-2020.

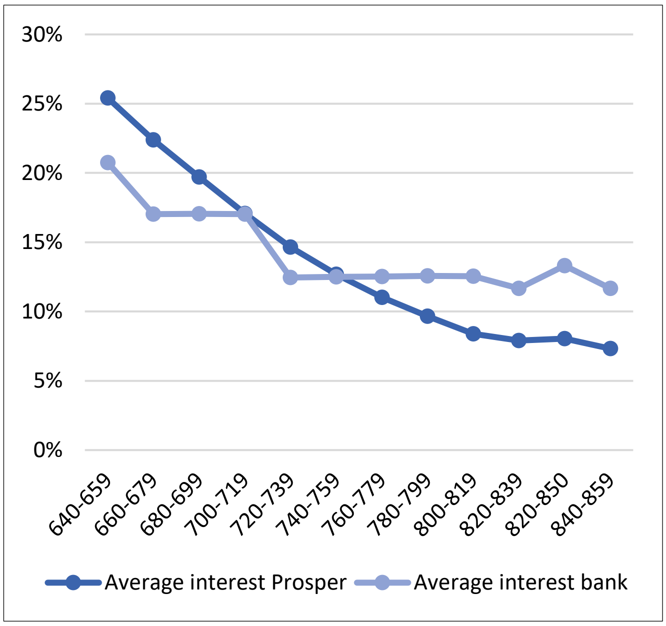
As can be seen, the majority (63%) of Prosper's applicants are classified as good, 19% as very good, and 4% as exceptional, which suggests that a large number of those applicants would be eligible for bank loans, as well. This favorable credit distribution is partially explained by the fact that Prosper rejects applicants with poor scores as part of the platform's credit risk management practice.

As a next step, interest rates and loan conditions are compared to bank loans. Prosper's price is risk-based and determined by the platform's internal rating. However, as can be seen from Table 4, the average Prosper interest rates are in line with the FICO range. For bank interest rates, credit card rates were chosen, as they can be a good benchmark due to their unsecured nature. Furthermore, P2P loans are generally small consumer loans, and those expenditures are usually financed (if not through a P2P platform) by credit card debt. Using credit cards as a benchmark to fintech loans is further supported by the study of Jagtiani and Lemieux (2018). A credit card interest rate was assigned to each record based on the year of application and FICO score of the borrower, and an average rate was estimated. Table 4 and Figure 6 presents the average Prosper rates and the average bank rates by FICO range.

4. Table: Interest rates by FICO

FICO range	Average interest: Prosper	Average interest: Bank
640–659	25	21
660–679	22	17
680–699	20	17
700–719	17	17
720–739	15	13
740–759	13	13
760–779	11	13
780–799	10	13
800–819	8	13
820–839	8	12
820–850	8	13
840–859	7	12

6. Figure: Interest rates by FICO



Source: Author's estimation based on Bureau of Consumer Financial Protection 2019 and Prosper listing data 2014-2020.
Note: Credit card rates were used for bank values as a benchmark from the corresponding years

It seems that borrowers with high FICO scores (above 740) can receive P2P loans at a lower interest rate than offered by banks. Therefore, their motivation to seek P2P funding might be the lower price. For borrowers below 740, P2P rates are less favorable. However, it might be that a portion of these applicants are not eligible for bank funding at all, and thus had to find alternative form of finance. According to Experian (2020), applicants having a FICO score around the top of the "Fair" category (roughly 670) might face difficulties with a bank loan request. Apart from the credit score, other triggers can cause not-eligible status for bank funding (e.g., fraud indicators).

To further investigate candidates' motivations, the loan purpose, which was provided by them during the application process, is analyzed. In Table 5, applicants are split by bank eligibility status, with a FICO score of 670 set as a threshold. Applicants below this limit are considered as not eligible for bank funding. As can be seen, debt consolidation was marked as the main reason for the loan request (72%). Home improvement was listed as the second most common purpose (10%), while the rest of the categories are insignificant. This is true for both credit groups and in total as well. Regardless of their financial background, people are utilizing social funding to repay outstanding debt.

5. Table: Applications by loan purpose (%)

Loan purpose	Not bank eligible	Bank eligible	Total
Auto / Motorcycle / RV / Boat	1	1	1
Business	1	1	1
Debt Consolidation	74	71	72
Home Improvement	9	10	10
Household Expenses	1	1	1
Large Purchase	2	2	2
Medical / Dental	3	4	3
Other	8	8	8
Taxes	1	1	1
Vacation	1	1	1

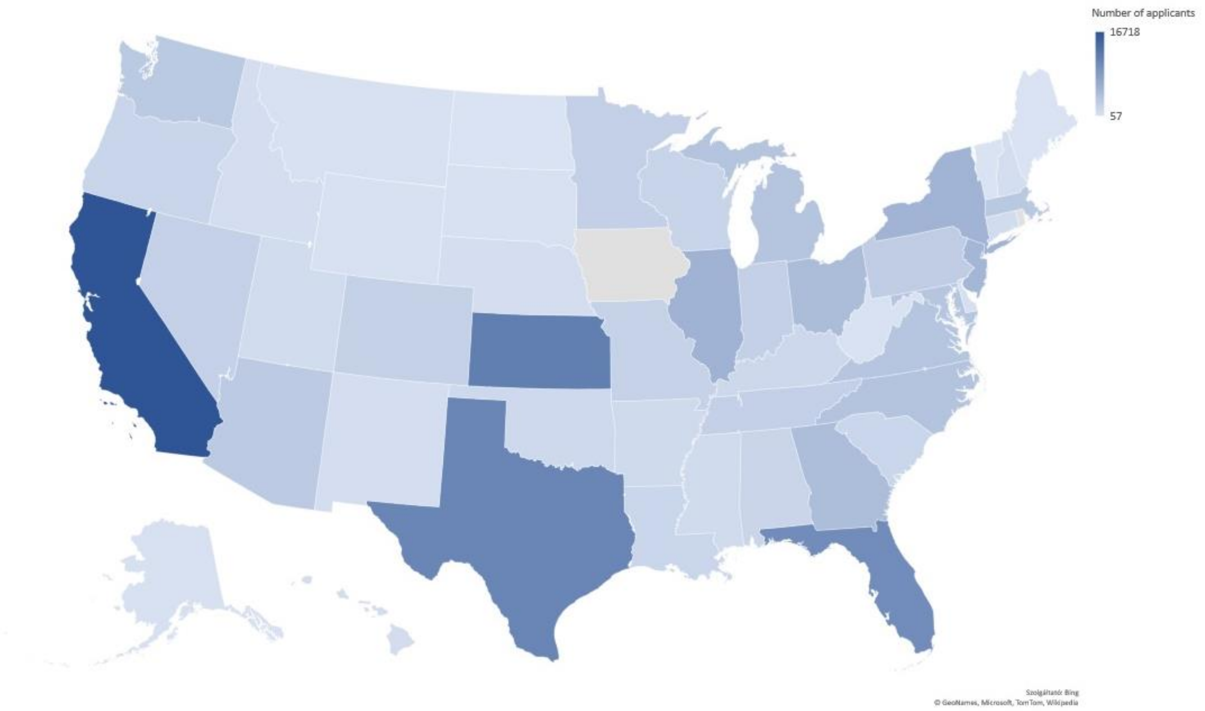
Source: Author's estimation based on Prosper listing data 2014-2020.
Note 1: Applicants with FICO scores above 670 are considered as bank eligible and below this threshold are classified as not eligible
Note 2: According to the website of Prosper (2021), 21 loan purposes can be selected. However, a few categories were combined in Table 3, as their portion was insignificant separately. Engagement Ring Financing, Wedding Loans, Baby & Adoption, Green loans and Special Occasion were classified as "Other". Motorcycle, RV and Boat were grouped under "Auto / Motorcycle / RV / Boat". There was no example in the dataset for the following categories: Cosmetic Procedures, Personal loan, Student use.

3.3.3 Applicants by state

In this section the number of applicants by state is analyzed using LASSO regression, in order to gain a deeper understanding of the potential customers of P2P platforms and their motivation.

Based on the literature, there are some preliminary expectations we can assume in terms of the results. Emekter et al. (2015) found that on average P2P borrowers have lower income compared to bank debtors. Jagtiani and Lemieux (2018) ascertained that the marketplaces principally serve underbanked customers. Havrylchyk et al. (2017) highlighted that the expansion of social lending is related to the lower density of branch network In addition, Oh and Rosenkranz (2020) and Han et al. (2019) explored that financial literacy, and the level of education demonstrates a positive relation with the spread of the segment. Overall, the theory implies that from one side the limited access to bank finance and from other side financial knowledge and education might have a significant role on the motivation of the applicants.
The most detailed geographic information that we can obtain from Prosper data is the state of residence of the applicants. The number of applications by state are presented in Figure 7. Based on the heatmap, most of the applicants are coming from California, Kansas, and Texas, followed by states along the East Coast.

7. Figure: Distribution of Prosper applicants by state between 2014-2020



Source: Author's estimation and editing based on Prosper listing data 2014-2020.
Note: Prosper lending is not available in Iowa (Prosper, 2021); therefore, there is no data for Iowa.

To perform a regression analysis, several state level national statistics were collected for the period of 2014-2020, covering, e.g., average credit profile, indebtedness, and financial and social features of the applicants' households. The mentioned six years are not split into further sub-samples as this time horizon is a prospering period for the emerge of marketplace lending, which could serve as a basis for robust outcome. A comprehensive microeconomic and socioeconomic database was produced and linked to the Prosper dataset aggregated by state and by year. The main question of the analysis is which economic indicators have a significant impact on the number of applications. The purpose of this examination is to outline which social groups serve as a pool for social lending.

The variable selection for the analysis is supported by previous studies focusing on the fintech market. Buchak et al. (2018) used demographic data covering education, ethnic information, income, poverty etc. to identify relationship with shadow banks on the residential lending market. Several papers discuss the role of socioeconomic data on the P2P market as a driver for investor decision (Ravina, 2019; Duarte et al. 2012; Pope&Sydnor, 2011; Barasinska&Schäfer, 2014). Other studies proved the importance of credit score, financial background, indebtedness, and loan history (Herzenstein et al. 2008, Gavurova et al. 2018). As the mentioned variables proved to be significant indicators from investor perspective, it is relevant to examine them from the borrower side. Hidajat (2021) explored the financial literacy of borrowers, based on their regular monthly spending and basic knowledge of finance. Agarwal et al. (2020) also examined the expenses among other financial stability variables when investigating default prediction with alternative scoring approaches. Bassani et al. (2019) found correlation between health expenditures and crowdfunding platforms specialized on healthcare. Chen et al. (2019) analyzed platform lending from house price perspective, focusing on the home value of borrowers. Ramcharan and Crowe (2013) explored correlation between house prices and loan conditions on the US P2P market. Havrylchyk et al. (2017) examined the role of the branch networks on platform lending.

Table 6 lists the selected variables and their description. The descriptive statistics are presented in Appendix II.

6. Table: The description of local microeconomic and socioeconomic variables by state used for the analysis

Variable	Description
FICO score	Average FICO score. FICO scores are not publicly available at the state level before 2018, therefore the 2018 values were used as proxies for earlier years.
Student loan	Student loan debt level per capita.
Student loan delinquency	Percent of student loan debt more than 90 days delinquent.
Population	Total population.

Man	Proportion of men in the population.
Employed	Proportion of population over 16 years who are in the labor force.
Total debt	Total household debt level per capita.
Mortgage debt	Mortgage loan debt level per capita.
Mortgage delinquency	Percent of mortgage loan debt more than 90 days delinquent.
Auto debt	Auto loan debt level per capita.
Credit card debt	Credit card debt level per capita.
Credit card delinquency	Percent of credit card debt which is more than 90 days delinquent.
Race_White	Proportion of White people in the population.
Race_Black	Proportion of Black people in the population.
Race_Latino	Proportion of Latino/Hispanic people in the population.
Median age	Median age in the population.
Total expenditures	Total personal consumption expenditures in million dollars.
Total expenditures per cap	Total personal consumption per capita.
Household expenditure	Total expenditures for household services.
Health expenditure	Total health related expenditures.
Financial services expenditure	Total expenditures for financial services and insurance.
Education_high school	Proportion of the population that has at least a high school degree. Data for 2020 are not yet available; therefore 2019 values were used for 2020.
Education_bsc	Portion of the population that has at least a four-year college degree. Data for 2020 are not available yet, therefore 2019 values were used for 2020.
DTI	Household debt-to-income ratio.
Poverty rate	Proportion of people living in poverty. The poverty income threshold for a family with two adults and one child is $20,578 per year. This is the most commonly applied threshold and is an official measurement used by the U.S. Government. Data for 2020 are not yet available; therefore 2019 values were used for 2020.
Median sales price	Median sales price of houses. Data for 2020 are not yet available; therefore 2019 values were used for 2020.
Mean household income	Mean household income by state. Data for 2020 are not yet available; therefore 2019 values were used for 2020.
Number of branches	Number of branches of FDIC-insured institutions, based on FDIC Deposit Market Share survey (2020).

Source: Author's collection from Economic Inclusion (2020); Experian (2020); Federal Reserve Bank of New York (2020); Federal Reserve (2020); United States Census Bureau (2020); Kaiser Family Foundation (2020); Bureau of Economic Analysis (2020); and Federal Deposit Insurance Corporation (2020).

Note: The variables are collected for each year between the period of 2014 and 2020.

LASSO (Least Absolute Shrinkage and Selection Operator) regression, which has become a widely used method in recent years, is applied to the dataset. LASSO is considered a useful technique when the number of observations is small and there are several potentially explanatory variables. It can improve the accuracy of the estimation, while producing easily interpretable results. The algorithm applies shrinkage, meaning that coefficients are shrunk towards zero and variables with zero coefficients are excluded from the regression. The concept is based on linear regression; however, it applies a penalty term to the sum of the absolute value of the coefficients. LASSO performs the following minimalization:

$$\sum_{i=1}^{N}\left(y_i-\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)\right)^2+\lambda\sum_{j=1}^{p}\left|\beta_j\right| \tag{1}$$

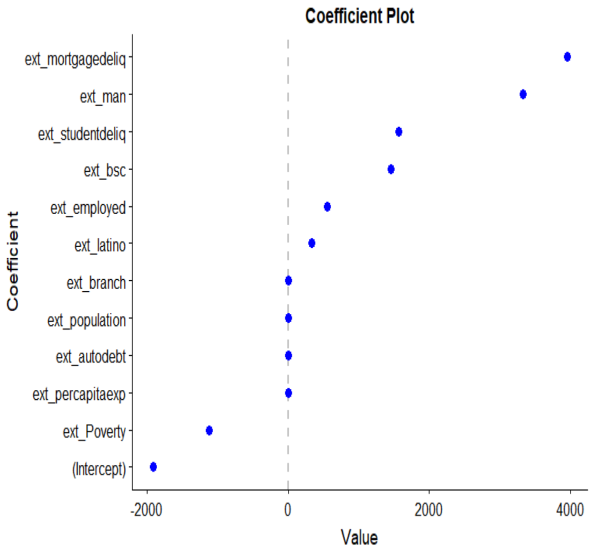
where y = dependent variable; x = explanatory variables; β = coefficients; and λ = tuning parameter. The tuning parameter controls the weight of the penalty term, and a higher λ indicates a stricter variable selection coupled with smaller coefficients. The optimal value of λ is set using k-fold cross-validation. The algorithm estimates the coefficients and the mean square error (MSE) using several λ values separately and finds the optimal balance for bias and variance (Békés & Kézdi, 2021).

The dependent variable in the analysis is the number of applications in each state and the potentially explanatory ones are the above presented 28 variables. First, the data are split into two sub-samples where 80% is the training sample and 20% is the test sample. Then, the optimal value of λ is estimated with k-fold cross-validation and the regression is performed with the optimal tuning parameter. The variables kept in the analysis, where the coefficients are not shrunk to zero, are listed below. All variables were standardized to avoid distortion in the results.

7. Table: Result of LASSO regression

Variable	Coefficient
(Intercept)	-1 913.602
ext_autodebt	-0.003
ext_mortgagedeliq	3 950.343
ext_population	0.000
ext_bsc	1 458.809
ext_poverty	-1 118.645
ext_studentdeliq	1 560.321
ext_man	3 330.643
ext_latino	334.881
ext_employed	557.372
ext_percapitaexp	-0.015
ext_branch	0.106
R^2	77.21%
No. of observations	336

8. Figure: Coefficients of LASSO regression



Source: Author's estimation based on Prosper listing data 2014-2020
Note: The description of the above presented variables is the following:
ext_autodebt: auto debt, ext_mortgagedeliq: portion of mortgage delinquency, ext_population: population, ext_bsc: portion of bsc degree, ext_poverty: poverty rate, ext_studentdeliq: portion of student delinquency, ext_man: portion of man, ext_latino: portion of latino population, ext_employed: portion of employed residents, ext_percapitaexp: personal expenditures per capita, ext_branch: number of branches

Overall, eleven variables were kept in the final model. Mortgage delinquency has the highest positive coefficient, which implies that the number of P2P applicants is higher in states where the mortgage delinquency rate is high. Therefore, it suggests that applicants probably utilize the P2P loan to refinance their overdue mortgage debts from a commercial bank. This assumption is supported by the previously presented fact that debt consolidation is the most frequent loan purpose. This result is in line with the research of Cornaggia et al. (2018) who found that loan delinquency and charge-off activity performed by small commercial banks is in line with peer-to-peer expansion. According to their explanation banks lowering credit requirements due to increased competition with the platforms and provide entry to less credible debtors. The finding is also supported by Maggio et al. (2017) who examined the relation between credit market and local economic factors related to mortgages and credit card balances. According to their conclusion, economic uncertainty due to high-risk borrowers is positively correlated with housing market illiquidity. Furthermore, the result is consistent with Buchak et al. (2018) who found that shadow banks, including fintech lenders gained a significant market share in refinancing in the residential lending market, mostly due to their lower regulatory burden.

The portion of men in a state is the second variable with a high coefficient, followed by the portion of student loan delinquencies. Student delinquencies is in line with the finding that applicants use the loans to refinance their current debt. The next variable showing positive relation is the proportion of people holding a college degree. A higher level of education implies a higher financial consciousness. It suggests that this subclass of applicants found this alternative form of finance beneficial, even though as it was mentioned previously, 86% of them are eligible for conventional funding. A higher level of employment also correlates positively with the number of applications, which is reasonable as 91% of the candidates in the sample were employed. The portion of Latino/Hispanic applicants shows a positive correlation as well. This result suggests that part of the applicants might belong to this ethnicity. The relation between the race of an applicant and funding success was researched in several studies (Pope and Sydnor, 2011; Harkness, 2016; Herzenstein, et al., 2008), therefore it might be a relevant borrower attribute in this segment. On the other side, poverty correlates negatively with the number of applicants, which can be explained by Prosper's borrower requirements—specifically, a minimum FICO score of 640, which excludes people with weak financial backgrounds. Furthermore, it is also possible that regions with higher poverty rates have lower infrastructure for online applications. Overall, the high geographical concentration of the applicants could be explained with the significance of the mentioned microeconomic and socioeconomic variables.

This model was applied on the test sample (which represents 20% of the original sample) for prediction. The goodness of fit of the model was measured with R^2 at 77.21%, implying a relatively good performance.

3.3.4 Clustering

In order to gain a deeper understanding of P2P applicants, a cluster analysis was performed. The purpose of this investigation is to identify patterns in the population and form groups based on those patters. The point of the analysis is both to find similarities within the groups and differences between them. The K-means clustering method was chosen, which is one of the most commonly used approaches.

K-means clustering aims to find the number of clusters, represented by *K,* based on the mean (or centroid) of the groups. *K* can be determined in advance, showing the requested number of clusters in our analysis. Each point of the population is then assigned to the nearest centroid and each group, based on these points, can be considered as a cluster. To determine distance, Euclidean distance is usually applied. The performance of the clustering is measured by the sum of squared error (SSE). In case of two classification options, we choose the one with a lower SSE, which is estimated with the following formula:

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} dist(c_i, x)^2 \qquad (2)$$

where dist = Euclidean distance; x = observation point; C_i = clusters; c_i = centroid of each cluster; and K = number of clusters. This iterative process is repeated until the final clusters are classified and the SSE is minimized (Steinbach et al., 2005).

Only continuous variables can be included in the cluster analysis; categorical variables must be eliminated. Table 8 presents the variables used for the estimation.

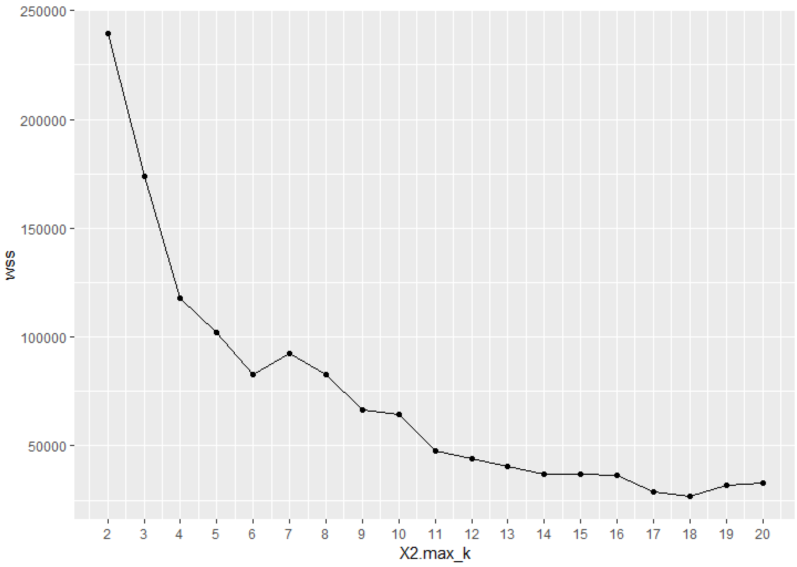
8. Table: The description of variables included in cluster analysis

Variable	Description
Revolving credit balance	Revolving credit balance of the borrower
Current delinquencies	Number of currently delinquent claims of the borrower.
FICO average	Average FICO score of the borrower, estimated as the average of the minimum and the maximum FICO range.
Amount	Loan amount requested by the borrower.
Amount delinquent	Amount of currently delinquent loans by the borrower.
Delinquencies last 7 yrs	Number of borrower delinquencies in the previous seven years.
Total credit lines	Number of total credit lines of the borrower.
Bankcard utilization	Bankcard utilization of the borrower.
Monthly payment	Estimated monthly payment for the requested loan.
DTI	Debt-to-income ratio of the borrower, calculated as the current aggregate monthly payments, divided by monthly income.
Interest average	Average interest rate the borrower will pay for the loan, calculated as the average of the minimum and the maximum of the interest rate range.

Source: Author's editing based on Prosper listing data 2014-2020.

For technical reasons only 100,000 observations can be included in the analysis, which were randomly selected out of the 135,592 records available. After variable selection, correlation between the variables was checked using a rank correlation table. The summary of the descriptive statistics and the correlation matrix are presented under Appendix III and Appendix IV. In order to avoid distortion in the results, in cases of significant correlation one of the variables was eliminated. A relative threshold of 0.3 was chosen for the correlation cut-off, in line with previous literature (Gavurova et al. 2018, Li et al. 2020). If variables with a high correlation are included in the analysis, the same information will have a higher weight in the outcome. After filtering, we have the following three variables: *revolving credit balance*, *current delinquencies* and *FICO average*. These variables must be standardized, meaning that we rescale their values, as they all are measured in different units. Scaling is essential as we need to unify the variables to be able to properly determine the distance. After scaling, the optimal number of clusters must be ascertained, using the elbow method, which shows the level of within-cluster sum of squares (WSS) for each cluster number. This represents the sum of distances between the observations and the centroids for each group. Figure 9 shows the results.

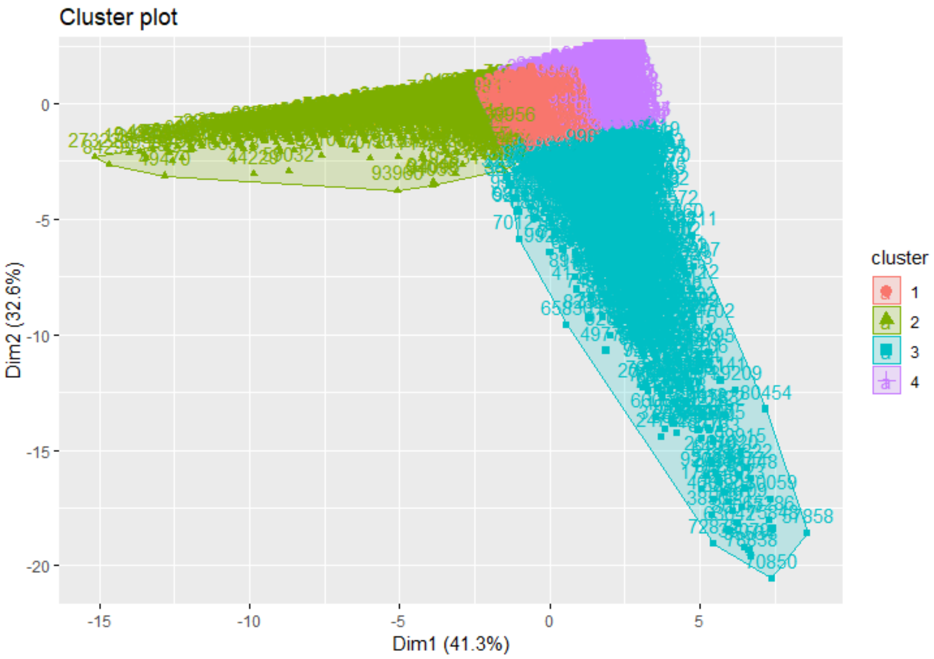
9. Figure: The number of clusters based on WSS



Source: Author's estimation based on Prosper listing data 2014-2020.
Note: The number of clusters is presented on the x-axis and the within-cluster sum of squares is on the y-axis.

Based on Figure 9, four clusters seem to be the optimal choice as there is a steep drop in the WSS between two and four clusters; after that, the line becomes less steep, meaning that adding another cluster does not cause a significant decrease in WSS. Therefore, the value of *K* was set to four and k-means clustering iteration was performed. To assess the performance of the clustering, the BSS/TSS ratio was examined, which reflects the between-cluster sum of squares divided by the total sum of squares. It examines the internal cohesion and the well-separation of the different clusters. The ratio is 60.6%, meaning that 60.6% of the total variance in the data is explained by the clustering. The final clusters are presented in Figure 10.

10. Figure: The distribution of Prosper applicants based on k-means clustering



Source: Author's estimation based on Prosper listing data 2014-2020.
Note: The clustering is based on the following variables: revolving credit balance, current delinquencies, and FICO average.

As can be seen the groups are separated from each other. To further interpret the results, the mean values of the variables for each cluster are presented in Table 9.

9. Table: Borrower statistics for each cluster

Group	Mean FICO	Mean Revolving credit balance ($)	Mean Current deliquencies (%)	Mean Amount ($)	Mean Total credit lines	Mean Delinquencies last 7 yrs	Mean Bankcard utilization (%)	Mean Interest rate (%)	Mean Amount deliquent ($)	Mean DTI	Portion of each group (%)
1	757	13,971	0.03	14,124	25	0.42	0.33	12.30	12	0.27	29
2	682	14,202	0.13	11,597	24	3.66	0.60	20.10	58	0.27	64
3	712	129,807	0.06	17,594	35	1.02	0.75	17.37	31	0.30	3
4	669	9,122	3.25	10,488	28	10.29	0.51	21.31	1476	0.23	5

Source: Author's estimation based on Prosper listing data 2014-2020.

Based on these statistics, it seems that four types of applicants can be differentiated, having the following characteristics:

- Group 1 applicants have with strong credit backgrounds and good credit history, with low levels of current claims. They are bank eligible; however, they might choose social funding due to more favorable conditions.
- Group 3 applicants seem to be bank eligible based on their FICO score; however, their revolving credit balances and bankcard utilization are quite high, and they request the highest loan amounts. The DTI ratio is also the highest, although their income covers the revolving debt. It can be suggested that they have immediate liquidity needs which might be funded faster by the platform compared to a bank process.
- Group 2 applicants probably face constraints upon bank funding, as their average FICO is relatively lower and they have delinquencies in their previous loan history; thus, obtaining a social loan could be easier than conventional funding.
- Group 4 applicants are likely not bank eligible as their average FICO is the lowest, and they have significant delinquencies in their loan history; therefore, they must necessarily explore alternative forms of finance.

Overall, the results suggest that P2P lending has dual function on the credit market: it supplements bank lending for a small segment of customers and for most of the cases it substitutes bank funding. Major part of the applicants is bank eligible based on their FICO score; however, large part of them (64%) are in the lower end of the score distribution with an average score of 682. The bank eligible clients prefer platform lending either due to more favorable conditions or easier access to funding. These conditions could be offered due to their lower operation costs enabled by the less complex, brokerage nature activity and the moderate regulation in the segment (Milne and Parboteeah, 2016; Lenz, 2016). The lower interest rate is attractive for the high-quality borrowers while the more accessible funding is engaging for the lower end of the bank borrowers.

The mixed role of the platforms on the consumer credit market is further supported by Tang (2019) who identified the importance of the loan size in this matter. According to his study, platforms are complements for banks in case of small loans and substitutes contrarily. This is in line with the results of the cluster analysis, as Group 2 and 4 have the lowest FICO score, coupled with smaller requested loan amount. The complementary approach is supported by Cole et al. (2019) who concluded that bank failures are associated with decrease in the volume of crowdfunding. However, their study focused on project finance instead of retail lending, where the complementary function of alternative finance is probably more dominant.

3.4 Conclusion

The P2P lending market has experienced a rapid expansion in recent years. As previous studies mostly focused on the credit risk perspective, there is space for further research to investigate the platforms from the borrower's side. The purpose of this paper was to identify the P2P lending customer groups and to understand their motivation for using this alternative form of funding. A unique, manually collected dataset was analyzed from the largest marketplace in the U.S. to examine the relationship between the number of applications and a wide range of local microeconomic and socioeconomic variables, then a k-means cluster analysis was performed. The results indicate that delinquent mortgage loans had the highest impact on the demand for marketplace lending, and that a majority of applicants declared debt consolidation as the loan purpose. This suggests that applicants probably need a P2P loan to refinance overdue mortgage debts owed commercial banks. Furthermore, four types of P2P customers can be differentiated, based on the results of the cluster analysis. A large portion of them is eligible for bank funding. Overall, it seems that the core customer group of P2P lending are not poor credit debtors, but potential bank clients. Therefore, the platforms supplement bank lending only for a small segment of customers who have poor credit quality, and in a majority of cases they compete with banks for the same group of customers. Based on the results, borrowers diversify their debt structure and utilize the P2P loan to refinance their obligation from a commercial bank. The development of the P2P segment might have an indirect effect on the banking sector, especially in case of an unexpected shock. Therefore, as a practical implication, following-up the trends of marketplace lending might be relevant for the supervisory authorities. It also proves the need for a comprehensive regulation of the segment in order to enhance transparency and to improve the risk management framework of the platforms, with special attention on the investor and borrower protection.

4. The liquidity aspects of peer-to-peer lending[5]

Purpose – The purpose of this paper is to investigate investor behavior in the market of peer-to-peer lending, which is an alternative form of finance, from the liquidity perspective.

Design/methodology/approach – Liquidity metrics and regressions are used to identify the trend of the market and the variables that significantly impact the successful resale, selling time and discount rate in the secondary market. Structural break analysis is used to examine the impact of COVID on the market.

Findings – There is a high demand for performing loans that are sold quickly; however, the discount rate is also high, which reflects the price of liquidity. Based on the results obtained from regressions, the main factors that impact the investors' decisions are discount rate, borrower's country, principal and the month passed after loan origination. Furthermore, it can be concluded that the pandemic has led to a structural break in March 2020, and investors have started to liquidate their claims.

Practical implications – This paper's purpose is to add to the research that examines the secondary market in social lending. It also contributes to the understanding of an investor's decision and behavior, which are key parts of the segment's long-term sustainability from the demand perspective. The comprehensive understanding of a lender's behavior is also essential for supervisory authorities and other participants of the financial market.

Originality/value – Previous studies mostly focused on credit risk aspects, whereas this paper contributes to the modest research of liquidity features. The added value of this paper is further supported by the use of a large European secondary market data set, including more than 5 million transactions, covering an 18-month horizon. Moreover, the market's sensitivity is analyzed in the case of an external shock. In the beginning of 2020, the COVID outbreak caused an economic shutdown in many European countries. The paper examines how these uncertain economic conditions impact the secondary market.

Keywords: Alternative finance, Investor behavior, Peer-to-peer lending, P2P, Secondary market, Social lending

JEL Codes: G21, G29

[5] Ölvedi, T. (2022). The liquidity aspects of peer-to-peer lending. Studies in Economics and Finance, 39(1), 45-62.

4.1 Introduction

Peer-to-peer (P2P) lending originated from the crowdfunding model as a form of sharing economy. The platforms match individual borrowers and lenders for the provision of loans. It is a relatively new field: the first platform was introduced in 2005 in the UK. Since then, a wide range of platforms have been developed worldwide, and the market has expanded with different players, showing robust growth. This rapid market expansion was also supported by the growing number of underbanked clients who had no access to bank funding after the financial crisis in 2008. However, due to the short history of P2P platforms, it is too early to observe their impact on the traditional lending and financial market. On the other hand, P2P platforms have several risk factors owing to their insecure nature. Most of the studies have focused on the credit risk aspect; however, one should be aware of the liquidity view of this alternative investment from the investors' perspective. This paper contributes to the current literature by analyzing the liquidity factor of social lending by using a large secondary market dataset from a significant Estonian platform named Bondora. In recent years, several platforms shut down their secondary markets due to a low number of transactions; however, after the COVID-19 outbreak, the demand for liquidity increased in the beginning of 2020. Therefore, this study was conducted to investigate the following three research questions:

RQ1. How liquid is the secondary market based on the three liquidity dimensions of selling rate, average selling time and discount?

RQ2. Which variables have a significant impact on the selling outcome, selling time and discount rate?

RQ3. How does the market react to external shocks (indicated by a pandemic situation)?

The investigation involved a large sample of more than 5 million observations to obtain a robust outcome. According to the results, the market is relatively liquid, with the average selling time being around 1.5 days. The highest demand is for performing loans with 0 days past due (DPD); however, the discount rate on the market is quite high, which suggests that the price of liquidity has to be paid by the seller. The liquidity dimensions are not dependent on the rating, investors consider the DPD instead. The role of days past due on the secondary market is reasonable as it can be considered as the most up to date information, reflecting the actual performance, while rating is related to historical experience.

Based on the regressions, the main factors impacting the successful resale and the selling time are discount rate, country of the debtor, principal and the months passed from loan origination. Discount rate is impacted by the number of days in debt, number of times the loan is listed on the secondary market, country of the debtor and principal. Though, model fit was low for selling time and discount rate. However, recent years have showed economic prosperity and robust growth in social lending volume, which is why it is relevant to examine the impact of an external shock on the market. Based on regressions performed in two sub-samples before and after the appearance of COVID in March 2020, the economic shutdown due to the pandemic was introduced in many countries, resulting in a structural break in the trend of the secondary market. However, it is still too early to predict the long-term impact of these extraordinary circumstances based on a five-month observation period. Additionally, the market seems to be sensitive to external shock, as investors started to liquidate their claims when the pandemic started in Europe.

Liquidity risk is a commonly discussed issue in the field of banking theory. Banks are considered as intermediary financial institutions, and one of their main activities is to transform illiquid assets into liquid deposits (Diamond and Dybvig, 1983). Furthermore, banks monitor the particular borrower on the behalf of the depositor to manage imperfect information (Freixas and Rochet, 2008). Liquidity issues also relate to the fragility of the financial system through the contagion effect and spreading of risk (Yellen, 2013). Regulators specify different requirements with respect to liquidity to enhance the resistance of the banking system in the case of an external shock. In contrast, peer-to-peer platforms do not own capital but have a mediatory function. Platforms perform credit risk assessment for a particular borrower; however, the task of monitoring and the associated risks are that of the investor. Furthermore, the peer-to-peer lending market is barely regulated by any supervisory authorities, and, thus, they are less exposed to systemic risk because several individual investors provide small funding amounts. Overall, the nature of peer-to-peer lending is quite different from conventional financial institutions with respect to risk management and regulations, and it is essential to pay attention to peer-to-peer lending platforms' operations to understand their possible roles in the financial market.

The paper is divided into two sections. The first section provides an overview of the peer-to-peer lending market, including the main features of the market, and an introduction of growing literature. The second part investigates the secondary market using transactional data from one of the main European platforms to examine the liquidity dimensions, the factors considered in the case of these dimensions, and the impact of an external shock on the market.

4.2. Peer-to-peer lending

Peer-to-peer lending is a platform-based solution that links borrowers and lenders to facilitate funding without any intermediary institution. The essence of this model is to benefit both the parties; potential debtors have access to loans at lower interest rates than those offered by banks, and the lenders can expect a high return on their investment in comparison to a conventional deposit. Another benefit of P2P lending is the low operational cost and lack of additional administrative expenses due to the absence of an intermediary institution in the process. This business model is further supported by the customer experience and innovative technology, e.g. the use of artificial intelligence in the borrower assessment process.

Despite its various advantages, the whole operating model, in general, has its risks. In a way, the loans are usually unsecured as there is no collateral security, and the consequences are rather modest in case of non-payment. Additionally, platforms are barely regulated by any supervisory authorities, and their operations are less transparent compared to monitored credit institutions. Thus, to attract investors and maintain their operational model over a long term, the P2P platforms have introduced different risk management measures. They assess the credibility of an applicant based on the provided data and assign a credit rating to them, which serves as a basis for the pricing. They also include data from the credit bureau registry that covers the applicant's loan history and payment behavior. In addition, in case of a default, the platform attempts to collect the obligation from the borrower, but only after a specific period, when the case is transferred to a debt collector agency. From the liquidity perspective, some platforms offer a buyback guarantee and operate in the secondary market.

Examining the alternative finance market, P2P platforms that specialize in consumer lending have the highest market share of 36% (CCAF, 2020). In case of the large marketplaces, commonly in the USA, institutional players appear on the investors' side, resulting in a higher total lending volume. According to the recent statistics, Europe has the

highest number of P2P platforms, i.e. 144 sites (58% of total), followed by North America with 35 (14%). In terms of lending volume, the total funding reached 19.3bn EUR in the European Union and 68.8bn EUR in the USA in July 2020 (P2PMarketData, 2020). The portion of social lending is still insignificant, approximately 2% of the total consumer credit, but the growth rate of the segment is quite robust.

4.2.1. Literature review

Despite the fact that the first platform appeared only 15 years ago, the literature on peer-to- peer lending has been rapidly growing. This paper relates to the main research area, which investigates how particular borrower characteristics influence the funding success through investors' decision on the primary market, while the scope of this study is the same in the secondary market. In the primary market, Herzenstein et al. (2008) and Gavurova et al. (2018) highlighted the impact of the financial stability of a debtor on successful funding, e.g. the importance of debt-to-income ratio and credit rating. Herzenstein et al. (2011) and Larrimore et al. (2011) found relevance in narratives provided by borrowers, while Duarte et al. (2012) and Ravina (2019) examined the appearance and trustworthiness of the debtors. Pope and Sydnor (2011) and Barasinska and Schäfer (2014) focused on the borrower's social features and discrimination. Lin et al. (2013) and Freedman and Jin (2008) investigated the role of borrowers' social relations.

Most of the studies on social lending are credit risk related, but this paper contributes to the research focusing on the secondary market and liquidity management perspective of P2P platforms. Caglayan et al. (2019) proved the existence of mispricing in the peer-to-peer secondary market using the Bondora data set. They observed cases when investors failed to sell a high-quality loan and, inversely, when a low-quality loan was sold successfully. According to their reasoning, mispricing is caused by the loan's different valuation from the buyers and sellers' perspective. Mispricing is further evidenced by Harvey (2018), who examined Lending Club data and highlighted the possibility of arbitrariness in the secondary market. However, it was concluded that the trading volume is too small to meet the arbitrary conditions. Byanjankar et al. (2020) proved, with empirical analysis, that the discount rate and number of days in default significantly impact an investor's decision. Reher (2014) highlighted the relation between the implementation of a secondary market and the level of interest rates in the primary market. According to an empirical study performed by Byanjankar et al. (2020), this kind of expansion lowers interest rates and can have a risk impact on the main market.

This paper's purpose is to add to the research that examines the secondary market in social lending. It also contributes to the understanding of an investor's decision and behavior, which are key parts of the segment's long-term sustainability from the demand perspective. The comprehensive understanding of a lender's behavior is also essential for supervisory authorities and other participants of the financial market.

4.2.2 Secondary market

The general opinion regarding peer-to-peer lending, as an alternative investment, is quite divisive. According to the recent studies and historical experience, investors are concerned about the high risk, possibility of non-payment and lack of sufficient transparency (Milne and Parboteeah, 2016). It seems that there is room to further improve the credit risk perspective of these platforms; however, it is also necessary to consider the liquidity aspect.

As a result of the growing loan volumes originated by these platforms, large-sized marketplaces introduced secondary markets that serve as a protection measure for investors, enhancing the liquidity of this type of investments. It presents an opportunity for the investors to sell their loans, but they must be aware that it is risky to buy a second-hand loan. To ensure diversity, an investor can sell only a portion of a particular loan or even a whole portfolio. Automatic functions are usually implemented to facilitate matching the preferences of the buyers and sellers. Investors determine pricing, which may not reflect the fair market value. Based on their decisions, loans can be offered with a premium on the original price or with a discount. Loans with the same characteristics and performances might be priced differently in the secondary market, and, sometimes, an investor’s behavior happens to be irrational.

It needs to be mentioned that the role and existence of the secondary market in the peer- to-peer industry is still questionable. In recent years, more leading marketplaces ceased their secondary market operations due to a low number of transactions (P2P-Banking, 2016). In contrast, due to COVID-19, many investors are attempting to sell their investments, causing liquidity issues. Consequently, some platforms have shut down their secondary market or introduced an additional fee for exit (MoneyWeek, 2020). It is still too early to predict the impact of stressed economic conditions on social lending; however, it seems that the demand for secondary market is growing, which might enhance the need for wider liquidity options. The next section presents the liquidity metrics in the secondary market. It investigates the main variables that impact these metrics and examines whether the appearance of COVID caused a structural break in the number of resales.

4.3. Secondary market analysis

The transaction data of an Estonian platform, Bondora, were used for this analysis. The platform was launched in 2009 and has been expanding rapidly ever since; currently, Bondora is operational in Finland, Spain, and Slovakia. As one of the main peer-to-peer platforms in

Europe, Bondora reported a total loan amount of 387.46m euros from its origination in the primary market as of October 2020. The marketplace facilitated investment for more than 140,000 investors, offering a wide range of products with different conditions. The secondary market was introduced in 2013, providing an opportunity for investors to liquidate their money. There is no extra fee for buyers and sellers in the secondary market. However, the platform gains the attention of lenders, and trading in the secondary market involves a significantly higher risk compared to the primary one. Claims can be sold manually or through the Portfolio Manager, which provides automatic trading function based on previously set preferences (Bondora, 2020). This section provides an overview of the secondary market portfolio through descriptive statistics to offer a better understanding of the market dynamics. Then, three different analyses are presented. The first one covers liquidity dimensions of the market; the second shows the tested variables that have a significant impact on the successful resale, selling time and discount rate of a second-hand loan to completely understand an investor's behavior; and the third section investigates the impact of COVID-19 on the market.

4.3.1. Data

The marketplace provides a publicly available data set, which is updated daily. The data cover the actual portfolio table and the historical secondary market transactions in two separate tables. The portfolio table contains information regarding the loan characteristics issued on the primary market, the social features and the financial background of the borrower. The secondary market table covers the loan characteristics of the claim on the secondary market and its performance since listing. The sources can be merged together with a unique loan ID to include more variables in the analysis, 13 altogether. Most of the variables have been drawn from the secondary market database and from the portfolio table, while four additional variables were created. The description of each variable is presented in Table 10. The time interval of the sample comprises 18 months of historical data, from 01.02.2019 to 01.08.2020.

10. Table: Variables Used in the Analysis

Variable	Description
Interest	The interest rate of the loan cannot be changed during secondary market listing; thus, this is the original interest rate of the claim.
Discount rate	Discount rate is determined by the seller, expressed in percentage (%) term. A positive value denotes a premium, and a negative value

	reflects price reduction. A discount is estimated with the following formula: resale value / original loan value – 1.
Debt days at start	Days in delay from the time when the loan was listed on the secondary market
Debt days at end	Days in delay at the time when the loan was removed from the secondary market
Principal at start	Outstanding principal when the loan was listed on the secondary market
Principal at end	Outstanding principal when the loan was removed from the secondary market
Country	Country of the borrower; categorical variable: EE (Estonia), ES (Spain), FI (Finland), and SK (Slovakia)
Rating	Rating of the borrower assigned by the platform; categorical variable: AA (1), A (2), B (3), C (4), D (5), E (6), F (7), and HR (8)
Until maturity	Created variable, the term of the loan until maturity in months
Selling time	Created variable, the difference between the start date and end date of the listing
From origination	Created variable, the months passed from the loan origination
Default flag	Created variable, dummy: if days past due are higher than 90 (1), otherwise (0). Days past due means the number of days in payment delayed from the time when the loan was listed on the secondary market.
Application number	Created variable, the number of times when the same loan was offered on the secondary market; it is possible that investors are unable to sell their loans the first time and, thus, attempt to sell them again, e.g., with a higher discount rate.

Source: Bondora Public reports, August 2020

Data cleaning was performed to eliminate outliers and invalid records. All the missing and invalid values were filtered out. Furthermore, the 3% cut-off was applied at the tails for continuous variables. The cut-off is based on expert judgement: due to the large sample 1% is considered low; however, 5% would result excessive sample reduction; therefore, 3% was chosen. A cap and a floor were set for the interest and discount rates to include only a valid range of observations. For interest rate it means between 0% and 100%, for discount rate between 100% and 100%. Borrowers from Slovakia were extracted from the sample, as their portion is less than 1% in the whole data set. After data cleaning and variable transformation, the final database included 5,112,566 observations, with 2,956,611 successful resales and 2,155,955 failed resales.

As the next step, a high-level portfolio overview was obtained, as presented in Table 11. The summary shows the number and portion of total loans, their average interest rates and principal amounts, and the time until the maturity date split by the days past due (the actual days past due at the time when the loan was listed on the secondary market). The number, portion and the average discount rate are also presented for the successfully sold portfolios.

11. Table: Secondary Market Portfolio Overview by days past due (DPD) Buckets

	All loans					Successfully Sold		
DPD	**Number of Loans**	**%**	**Average Interest Rate**	**Average Discount Rate**	**Average Time Until Maturity (in month)**	**Number of Loans**	**% of all loans**	**Average Discount Rate**
0	2 961 275	58%	29%	-0.41%	40	2 138 721	72%	-4.3%
1-30	351 385	7%	32%	-5.32%	37	170 471	49%	-12.2%
31-60	124 065	2%	35%	-9.44%	37	56 410	45%	-18.8%
61-90	85 064	2%	37%	-13.28%	37	36 057	42%	-23.1%
91-120	68 537	1%	38%	-14.65%	37	24 609	36%	-23.0%
120+	1 522 240	30%	33%	-18.14%	21	530 343	35%	-25.8%
Total	5 112 566					2 956 611		

Source: Author's estimation based on the Bondora secondary market database as of August 2020

According to the aforementioned data, 58% of the loans registered on the secondary market were performing well with 0 DPD, while another significant portion, i.e. 30%, were not performing at all with a DPD of more than 120 days. The rest of the buckets are insignificant. In terms of successful selling, the highest demand is shown for the best performance with 0 DPD, where 72% loans were sold. The portion of successfully sold loan volume decreases when the days past due increase.

Interest rate is mostly in line with performance, which refers to appropriate pricing at loan origination in the primary market. As mentioned, the interest rate cannot be changed for a secondary market listing; thus, it is based on the initial credit assessment and rating of the borrower. As expected, the discount rate increases when the days past due increase. Discount rate is set by the seller, and even the best-quality loans are offered with an average discount of 0.41%. Based on the discount rate of successfully sold loans, a significant price reduction is necessary to resell the claims. Overall, it seems that investors are willing to sell their loans at lower prices to minimize loss or to liquidate their investment. On average, the time until maturity of the loan is roughly 3 years. The relatively short expiry for the 120 bucket can be explained with the fact that the maturity date has passed for some of these loans; thus, their negative value decreases the category's mean.

Table 12 presents the portfolio overview by the country of the borrower. As it can be seen, most of the debtors are from Estonia (69%), which is followed by Finland (24%) and Spain (7%). The average interest rate and discount rate also differ in each country. The average discount rate is the highest in Spain (9.86%) with the portion of relatively high successfully sold loans (57%). Claims with Estonian borrowers have slightly better selling rate and the loans are offered with a small discount (5.66%). The high interest rate in Spain (50%) suggests lower portfolio performance compared to Estonia, where the interest rate is half (25%).

12. Table: Secondary Market Portfolio Overview by the Country of the borrower

Country	All loans					Successfully Sold		
	Number of Loans	%	Average Interest Rate	Average Discount Rate	Average Time Until Maturity (in month)	Number of Loans	% of all loans	Average Discount Rate
Estonia	3 530 251	69%	25%	-5.66%	35	2 091 205	59%	-8.57
Spain	375 733	7%	50%	-9.86%	32	213 547	57%	-9.92
Finland	1 206 582	24%	41%	-8.56%	33	651 859	54%	-11.22
Total	5 112 566					2 956 611		

Source: Author's estimation based on the Bondora secondary market database as of August 2020

4.3.2. Liquidity dimensions

The liquidity of a certain market can be measured in several ways. According to the literature, the three main aspects which should be considered regarding liquidity are the traded volume, time horizon of the transaction and loss indicated by pricing (Szűcs and Váradi, 2014). The purpose of this section is to present different figures from the secondary market, in line with liquidity dimensions. The selected variables are the successful selling rate, average selling time and discount rate. Each variable was examined in a matrix based on the days past due and rating. The rating is based on the platform's internal credit risk assessment process, using data provided by the applicant and also information about historical performance from the Credit

Bureau system. Average annualized expected loss range is estimated for each application based on standard credit risk parameters and a rating is assigned to each range (Bondora, 2020).

Table 13 presents the heatmap of the portion of the successfully sold loans (AA reflects the best rating classification and HR is not performing). As can be seen, there is a demand for the best performing claims, which are not overdue, and the proportion of selling decreases in line with worsening performance. Furthermore, it is also interesting that the selling rate does not depend on the rating; instead, investors use the information regarding days past due. Rating has significant role on the primary market, and it mainly serves as a basis for investor funding decisions (Herzenstein et al. 2008; Gavurova et al. 2018). However, the rating methodology is a "blackbox", its accuracy was examined in several studies (Jagtiani and Lemieux 2019; Bhuvaneswari and Segalini, 2020; Byanjankar et al. 2015). The role of days past due on the secondary market is reasonable as it can be considered as the most up to date information, reflecting the actual performance, while rating is related to historical experience. For more information on the number of loans in each category, please see Appendix V.

13. Table: Portion of successfully sold loans split by rating and days past due (DPD), which means the number of days in payment delay from the time when the loan was listed on the secondary market

DPD	AA	A	B	C	D	E	F	HR
0	67%	58%	67%	72%	75%	75%	77%	75%
1-30	41%	41%	50%	51%	52%	48%	44%	49%
31-60	43%	40%	50%	49%	46%	41%	48%	48%
61-90	38%	35%	43%	46%	46%	42%	40%	40%
91-120	38%	31%	43%	40%	35%	39%	32%	33%
120+	33%	27%	35%	34%	32%	35%	39%	34%

Source: Author's estimation based on the Bondora secondary market database as of 2020 August

As the next step, the mean selling time was analyzed to see the average time it takes to convert P2P investments into cash. The sample consisted of only the successful sales, and the selling time was calculated as the difference between the date when the loan was sold and the date it was posted on the secondary market. Based on the results, the average selling time of the whole sample was 1.5 days. The detailed period of purchase for each rating category is summarized in Table 14 for the rating and DPD dimensions.

Overall, the average selling time is seen to be independent of the rating, and there is no significant difference between the average selling time of each category. However, the best DPD bucket and the worst ones above 90 DPD can be liquidated in the shortest period. A relatively short resale period is further supported by the platform's practice. Investors have the opportunity to conduct their own investigation regarding the secondary market supply and manually choose the loan or portfolio they would like to buy. Additionally, they can use the service of a portfolio manager, who buys and sells loans automatically on the secondary market, based on previously set parameters for speeding up the process (Bondora, 2020).

14. Table: Average selling time split by rating and days past due (DPD), which means the number of days in payment delay from the time when the loan was listed on the secondary market

DPD	AA	A	B	C	D	E	F	HR	Standard deviation
0	0.83	1.65	1.13	1.45	1.52	1.50	2.06	1.94	6.49
1-30	5.05	6.87	4.36	4.54	4.35	3.67	3.90	3.63	11.08
31-60	4.77	5.51	3.35	3.55	3.61	3.15	3.22	2.51	10.25
61-90	2.03	3.07	2.63	2.09	1.90	2.10	2.08	1.84	7.99
91-120	0.72	0.87	2.05	1.79	1.34	1.67	1.23	0.86	7.11
120+	1.12	0.91	0.84	0.75	0.94	0.86	0.70	0.67	5.72

Source: Author's estimation based on the Bondora secondary market database as of August 2020

Table 15 presents the level of the discount only for the sold offers. The amount of fare reduction increases in a monotonous way, while the impact of the rating is less dominant, which is in line with the observations above. Discount rates are relatively high and performing loans are sold with a discount, which suggests that the price of liquidity has to be paid by the seller. As a comparison, bank sale discounts reached ~28% in Europe for NPL portfolios. It has to be emphasized that is hard to compare the P2P discount rates with bank sale discounts, as numbers vary based on the selected time period and portfolio.

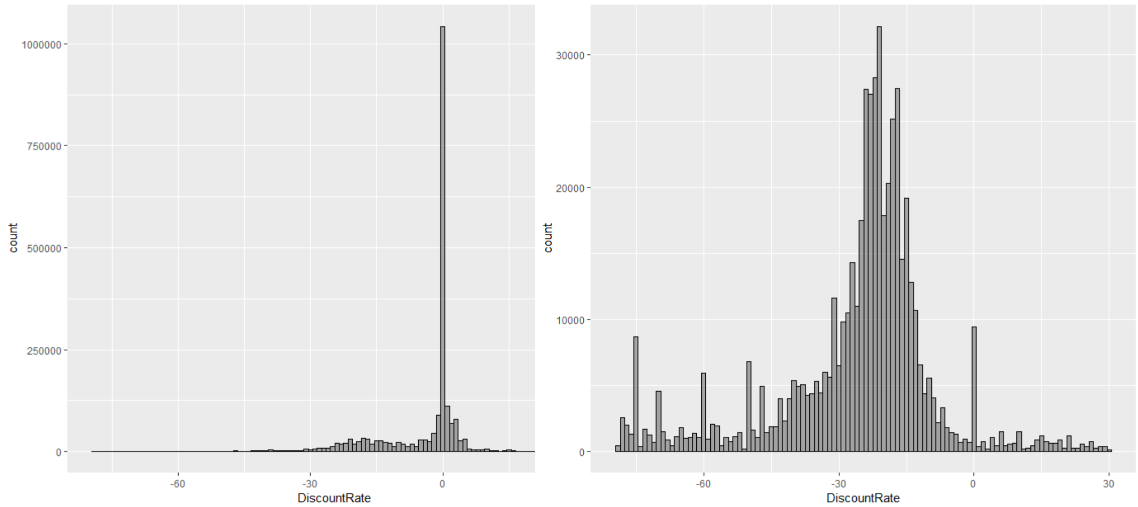
15. Table: Average discount rate split by rating and days past due (DPD), which means the number of days in payment delay from the time when the loan was listed on the secondary market

DPD	AA	A	B	C	D	E	F	HR
0	-5.78%	-6.13%	-6.77%	-4.96%	-3.45%	-3.17%	-2.94%	-4.17%
1-30	-9.15%	-7.75%	-12.00%	-12.03%	-13.09%	-14.03%	-9.90%	-12.54%
31-60	-12.99%	-12.99%	-20.31%	-20.71%	-21.14%	-20.17%	-14.08%	-21.23%
61-90	-16.53%	-19.24%	-22.77%	-23.52%	-26.74%	-26.20%	-17.70%	-21.76%
91-120	-19.97%	-21.08%	-24.28%	-21.57%	-24.59%	-28.71%	-17.55%	-22.40%
120+	-23.50%	-24.46%	-25.92%	-26.48%	-27.44%	-26.61%	-22.46%	-27.65%

Source: Author's estimation based on the Bondora secondary market database as of August 2020

For a better look at the distribution of discounts, Figure 11 presents the histogram for the 0 DPD and the 120 DPD claims, as these two have the highest weight in the whole secondary market portfolio. For performing loans, the original price is the most common selling price, while, for 120 DPD, the scale is wider, and there is a high concentration between 20 and 30% discount.

11. Figure: The distribution of claims for the 0 DPD (left) and the 120+ DPD (right) buckets



Source: Author's estimation based on the Bondora secondary market database as of August 2020
Note: the width of each bin is 1

Based on the liquidity figures, it can be concluded that investors consider days past due, as the differences between the rating categories were not significant in either case. This behavior is reasonable, as borrowers were rated when the loan was originated in the primary market, while days past due reflect the actual performance. Therefore, investors are able to avoid the impact of potential mispricing and incorrect assessment on the primary market. Another inference is that there is a high demand for performing loans with 0 days past due, and their selling time is quite short. However, in general, these performing loans can be sold with discount, even if they ensure a stable return. The price of liquidity has to be paid by the seller in this way.

4.3.3 Analysis and results–regression analysis

After presenting an in-depth overview of the secondary market from the previously presented descriptive statistics, this section focuses on the determinants of a successful resale, selling time and discount rate to understand investor behavior better. Different regressions were performed for the analysis, according to the nature of the data and the research problem. The question here is whether the variables significantly impact investor behavior during a successful transaction using the liquidity dimensions from the previous section.

As an initial step, the normality was examined using the Kolmogorov–Smirnov test. According to the results, the variables in the sample were not distributed normally. Thus, the Spearman rank correlation was chosen as a nonparametric test to select the range of variables for the regression. The results showed that a strong association between the variables for a few pairs. Eliminating one of the pairs with a correlation coefficient above 0.35, the final sample was reduced to four variables. The correlation matrix is presented in Table 16.

16. Table: The results of the Spearman rank correlation

	I	DR	DDAS	DDAE	PAS	PAE	C	R	UM	FO	D	AN
I	1.00											
DR	-0.03	1.00										
DDAS	0.12	-0.46***	1.00									
DDAE	0.12	-0.44***	0.96***	1.00								
PAS	0.06	-0.03	0.01	0.00	1.00							
PAE	0.07	-0.04	0.01	0.00	0.99***	1.00						
C	0.53***	-0.08	0.21*	0.21*	-0.06	-0.06	1.00					
R	0.94***	-0.07	0.21*	0.21*	0.06	0.06	0.55***	1.00				
UM	0.13	0.23*	-0.46***	-0.46***	0.19*	0.19*	-0.06	0.08	1.00			
FO	-0.15*	-0.32**	0.61***	0.61***	-0.17*	-0.17*	0.05	-0.02	-0.74***	1.00		
D	0.10	-0.43***	0.89***	0.90***	0.01	0.01	0.20*	0.18*	-0.43***	0.57***	1.00	
AN	-0.12	-0.08	0.22*	0.21*	-0.03	-0.03	-0.18*	-0.09	-0.30*	0.41***	0.20*	1.00

Source: Author's estimation based on the Bondora secondary market database as of August 2020
*Note: significance levels at 10% *, 5% **, 1% ****
I: interest, DR: discount rate, DDAS: debt days at start, DDAE: debt days at end, PAS: principal at start, PAE: principal at end, C: country, R: rating, UM: until maturity, FO: from origination, D: default flag, AN: application number

To understand an investor's perspective, it is essential to know the factors that are considered in case of a successful loan resale. Logistic regression was applied for the first analysis based on the nature of the historical database. The dependent variable was the result of selling, which can have two outcomes: success or failed.

To investigate the impact of each variable on the likelihood of resale, a logistic regression function was used; y shows the probability of an event to befall – in this case, the resale – between zero and one, and it is described with the formula below:

$$f(y) = \frac{1}{1+e^{-y}} \tag{3}$$

The equation for the probability of the successful resale with multiple independent variables is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + .. + \beta_n x_n \qquad (4)$$

where y represents the dependent variable, b 0 the constant, b i the regression coefficients, xi the explanatory variables, and n is the number of variables included in the model (Wooldridge, 2012).

After the rank correlation analysis, the following four variables were included in the regression model: Discount rate, Principal at start, Country and From origination. Variable From origination and Principal at start were log transformed due to their distribution, and, for the latter, the quadratic term was taken. All these variables proved to be statistically significant; the results are presented in Table 17.

17. Table: The results of logistic regression for the probability of successful resale as a dependent variable

Variable	**Coefficient**	**Marginal Effects**	**Std. Error**	**z-value**	**Pr(>|z|)**
(Intercept)	3.2140		0.0045	705.17	0.0000 ***
DiscountRate	-0.0413	-0.0098	0.0000	-611.81	0.0000 ***
log_PrincipalAtStart_2	-0.0104	-0.0024	0.0004	-25.61	0.0000 ***
CountryES	-0.3305	-0.0808	0.0038	-86.35	0.0000 ***
CountryFI	-0.2284	-0.0552	0.0023	-98.33	0.0000 ***
log_from_origination	-0.5567	-0.1331	0.0007	-747.36	0.0000 ***
No. of observations	5 112 566				
AUC	0.75				

Source: Author's estimation based on the Bondora secondary market database as of August 2020
Note1: In case of Country, which is a categorical variable, class "EE" was taken as the reference, which means Estonia
Note2: Discount rate is in percentage (%) format.
*Note3: *** indicates the significance at the 99% level, respectively*

According to the regression coefficients, the month passed from initial loan origination negatively correlates with the dependent variable, which means that investors do not prefer loans with long maturity. Those loans tend to be overdue, having been presented on the market for a long time. The country coefficients were interpreted in comparison to category EE, which was chosen as a reference, and it means that the borrower is from Estonia. There is a lower chance for successful resale in case the borrower is outside of Estonia, specifically from Spain and Finland, with 0.33 and 0.23 β values, respectively. As the platform is Estonian and the majority of the participants are domestic, cross-border lending might be less preferred. The discount rate proved to impact the selling with a 0.041 β negatively, which is reasonable–the lower the price, the higher the selling success. The standard error coefficients obtained are quite low, showing that the multi-collinearity in the model is irrelevant. Based on the results from the logistic regression, the impact on the successful resale can be estimated with the following equation:

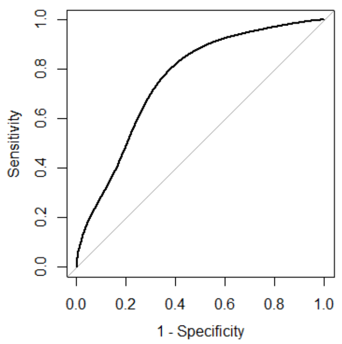
$$y = \beta_0 + \beta_1 DiscountRate + \beta_2(\log(1 + PrincipalAtStart))^2 + \beta_3 \log(1 + FromOrigination) + \beta_4 Country \quad (5)$$

As an example, we consider a claim with a 5% discount rate, a 100 euro principal with 36 months from loan origination and a borrower from Spain. After logistic transformation, the probability of a successful resale in this case is 89.80%. The equation is as follows:

$$Probability = \frac{1}{1+e^{-(3.2140-0.0413*(-5)-0.0104*\log(1+100)^2-0.5567*\log(1+36)-0.3305)}} = 0.8980 \quad (6)$$

To assess the classification power of this model, the receiver operating characteristic (ROC) curve was estimated. The curve can capture the true positive rates (sensitivity) and false positive rates (1-specificity), based on the model's ability to classify observations between the two categories correctly. Figure 12 presents the ROC curve of the model – the y-axis shows the true positive rate, and the x-axis shows the false positive rate.

12. Figure: Receiver operating characteristic curve for the logistic regression examining successful resale



Source: Author's estimation based on the Bondora secondary market database as of August 2020

The most common parameter in interpreting the curve is the area under it (AUC). The AUC of the estimated model is 0.7566, meaning that there is a 75.66% probability that the model can correctly classify the two classes.

As an additional test, besides the selling outcome, the selling time and the discount rate were also examined using linear regression. In the case of selling time, only the successful transactions and previously presented variables were included in the analysis (after variable elimination based on the correlation matrix in Table 16), and variable selling time was log transformed. The results were in line with the determinants of the selling success, and all variables showed a significant impact on the selling time. Specifically, a higher premium added to the price results in a longer selling time. Higher loan amount also lengthens the selling time because investors prefer lower amounts, as shown by the previous analysis of selling success. However, the overall model fit is low. The details of the linear regression are presented in Table 18.

18. Table: The results of linear regression for selling time as a dependent variable

Variable	**Coefficient**	**Std. Error**	**t-value**	**Pr(>|t|)**
(Intercept)	0.5057	0.0014	350.13	0.0000 ***
DiscountRate	0.0036	0.0000	113.02	0.0000 ***
log_PrincipalAtStart_2	0.0111	0.0001	61.63	0.0000 ***
log_from_origination	-0.0457	0.0002	-178.73	0.0000 ***
CountryES	0.0227	0.0017	13.21	0.0000 ***
CountryFI	0.0030	0.0010	2.86	0.0042 **
No. of observations	2 956 611			
R^2	0.03			

Source: Author's estimation based on the Bondora secondary market database as of August 2020
Note1: In the case of Country, which is a categorical variable, class "EE" was taken as the reference, which means Estonia.
Note2: Discount rate is in percentage (%) format.
*Note3: *** indicates the significance at the 99% level*

In terms of discount rate analysis, it is worth to examine the pricing trends on the primary market to have some expectations on the relevant factors. Herzenstein et al. (2008) explored that loan characteristics have significant impact on the level of interest rate. Gleisner and Berger (2009) concluded that the rating has positive impact the final interest rate, while Weiss et al. (2010) found that the elevated number of bids lead to higher prices. Hildebrand et al. (2010) found the role of soft information - specifically social networks – as a significant indicator. Overall, on the primary market financial information, credit quality and soft data could have an essential part.

Discount rate analysis was performed on the whole sample, and the variables were reduced based on the correlation matrix presented in Table 16 The following variables proved to be significant: Debt days at start, Application number, Principal at start and Country. As previously mentioned, the discount rate is positive when premium is offered and negative when the price is reduced. The results are reasonable, as claims where the number of days in debt is lower can be offered with a premium. Furthermore, if a loan is listed multiple times in the secondary market, a higher discount is offered. Claims outside of Estonia are provided with a higher discount, which is in line with the previous results from selling success. Similarly, the overall model fit is low (Table 19).

19. Table: The results of linear regression for discount rate as a dependent variable

Variable	**Coefficient**	**Std. Error**	**t-value**	**Pr(>|t|)**
(Intercept)	-1.6420	0.0132	-123.72	0.0000 ***
DebtDaysAtStart	-0.0212	0.0000	-981.98	0.0000 ***
Application_number	-0.0024	0.0000	-77.79	0.0000 ***
log_PrincipalAtStart_2	-0.2031	0.0029	-70.02	0.0000 ***
CountryES	-0.8069	0.0275	-29.32	0.0000 ***
CountryFI	-0.5739	0.0170	-33.70	0.0000 ***
No. of observations	5 112 566			
R^2	0.17			

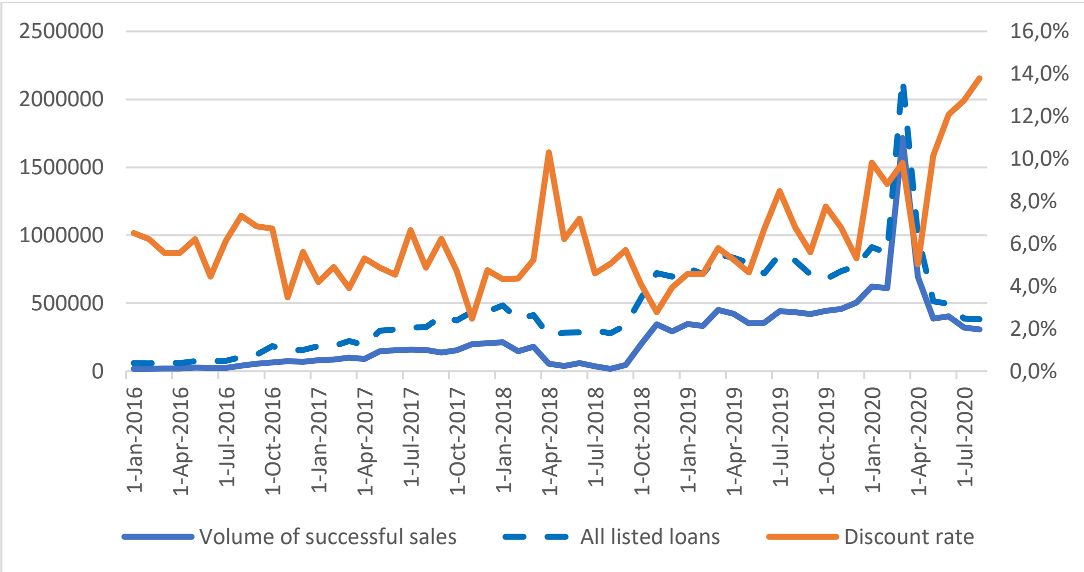
Source: Author's estimation based on the Bondora secondary market database as of August 2020
Note1: In case of Country, which is a categorical variable, class "EE" was taken as the reference, which means Estonia.
Note2: Discount rate is in percentage (%) format.
*Note3: *** indicates the significance at the 99% level, respectively*

4.3.4 Analysis and results: COVID-19 impact

The years preceding COVID-19 were marked by economic prosperity, and the social finance market showed rapid expansion. However, the pandemic resulted in an external shock to the economy from the beginning of 2020, causing forced shutdowns in many countries. It is still too early to predict the long-term impact of these extraordinary circumstances, but there is currently a five-month observation period from the time when the pandemic appeared in Europe. This section analyzes how COVID-19 has impacted investor behavior in the secondary market for peer-to-peer lending.

To obtain a high-level overview of the previous year's trend, Figure 13 presents the aggregated volume of all listings, aggregated successful resales and monthly average discount rate, from January 2016. In terms of resale volume, there was a significant increase in March 2020 when the pandemic started to grow rapidly across Europe, and the volume dropped in the next period. A strong co-movement can be observed between the volume of successful sales and all listings. Discount rate has a volatile trend; however, from April 2020, the level of fare reduction started to increase.

13. Figure: The volume of all listings, successful resales, and the level of discount rate between 01.01.2016 and 01.08.2020



Source: Author's estimation based on the Bondora secondary market database as of August 2020
Note: the right axis belongs to the discount rate and the left to the volumes

To examine whether an external shock causes any significant change in the secondary market, the data was split into two subsamples; the breakpoint was set as March 2020. Regressions presented in the previous section were performed again for the two separate periods (using the same variables which proved to be significant) to check how the relationship changed between the liquidity dimensions and the explanatory variables before and after the breakpoint. The coefficients of each regression are presented in Tables 20 and 21.

20. Table: The coefficients from the regressions before and after March 2020 with successful selling and selling time as dependent variables

Variable	Successful selling		Selling time	
	Before 2020.03.01.	After 2020.03.01.	Before 2020.03.01.	After 2020.03.01.
(Intercept)	2.9400	4.4805	0.5646	0.2373
CountryES	-0.3698	-0.2007	0.0162	0.0398

CountryFI	-0.2732	-0.0207	-0.0033	0.0150
log_PrincipalAtStart_2	-0.0017	-0.0141	0.0140	0.0036
DiscountRate	-0.0385	-0.0504	0.0024	0.0060
log_from_origination	-0.5608	-0.6590	-0.0584	0.0062
No. of observations	5 112 566	5 112 566	2 956 611	2 956 611
AUC / R^2	0.77	0.75	0.04	0.02

Source: Author's estimation based on the Bondora secondary market database as of August 2020

21. Table: The coefficients from the regressions before and after March 2020 with discount rate as a dependent variable

Variable	Discount rate Before 2020.03.01.	Discount rate After 2020.03.01.
(Intercept)	-0.4758	-4.4050
CountryES	-0.8011	-0.4182
CountryFI	-0.3513	-1.0680
log_PrincipalAtStart_2	-0.1997	-0.2900
DebtDaysAtStart	-0.0223	-0.0193
Application_number	-0.0009	-0.0021
No. of observations	5 112 566	5 112 566
R^2	0.19	0.14

Source: Author's estimation based on the Bondora secondary market database as of August 2020

Based on the results, the slope and the betas differ in the sub-samples for the three liquidity dimensions. However, all the independent variables remained significant at the 1% significance level for all the regressions. The difference suggests that there is a structural break in the data series in relation to the growth in pandemic. In case of discount rate, investors offered higher discount after the appearance of the pandemic, which has a greater impact on the selling success after 2020 March. The risk aversion of investors is further supported by the loan amount, as higher amount indicates lower selling success. These findings are in line with the results of the other two liquidity dimensions, selling time and discount rate. Overall, after the breakpoint, all the variables support a higher discount and longer selling time, which suggests that investors started to liquidate their money in a low demand environment.

Generally, it seems that the segment of peer-to-peer lending is sensitive to external shocks. As of March 2020, when the pandemic caused an economic shutdown in Europe, a peak could be observed in the volume of resales driven by an increased supply, followed by a drop and, subsequently, a rising trend in the discount rate with a one-month lag. This indicates that investors started to liquidate their money and were willing to provide higher discounts in the dropped demand environment. Overall, this recent external shock resulted in a structural change in the secondary market. However, it is still too early to predict if this trend is temporary or if it will continue in the future.

4.4 Conclusion

Peer-to-peer lending is a relatively new field. However, the market underwent rapid expansion, and several studies have been published in recent years. The purpose of this paper was to add to the modest research regarding the liquidity perspective of peer-to-peer lending, as previous research has mostly focused on the credit risk aspects. Using a data set from a large Estonian platform, which included more than 5 million observations, a secondary market analysis was performed. According to the liquidity metrics (selling rate, average selling time and discount), the average selling time is around 1.5 days, and performing loans are mostly sold 0 days past due. However, the discount rate is high, even for performing claims, which suggests that the liquidity of this type of investment has to be paid by the seller. Based on the results of the regressions, successful selling and selling time are impacted by the discount rate, borrower's country, principal, and number of months passed after loan origination. Furthermore, separate regressions before and after March 2020 suggest that there was a structural break in the trend of the secondary market, resulting in the market being sensitive to external shocks. Investors started to liquidate their money and offer higher discounts in line with the pandemic's expansion. Due to the changing economic circumstances, the current risk management practice of the platforms might change in the future, and the liquidity aspect needs to receive more attention.

5. Peer-to-Peer Lending: Legal Loan Sharking or Altruistic Investment? Analyzing Platform Investments from a Credit Risk Perspective [6]

This paper analyzes the performance of peer-to-peer investments, the potential benefits of their information processing and the investor returns, based on the entire portfolio of the Estonian platform Bondora. We found that the platform's scoring model relies on different default probabilities across countries and is weak at predicting default within countries. Alternative information could improve the models, but our analysis could not confirm this benefit of the platform. The average internal rate of return on closed transactions was -4.17%, and 42% of the loans end with a negative IRR. We concluded that P2P borrowers in the European market are mainly high-risk, bank ineligible clients, accepting even loan-sharking level interest rates, which excludes altruistic motives of investors. Even so, investors are not compensated for the credit risk.

Keywords: alternative finance, peer-to-peer lending, information asymmetry, credit rating, scoring model

JEL Codes: G21, G23, G32

[6] This paper is a joint work with Barbara Dömötör and Ferenc Illés.
Dömötör, B., Illés, F., & Ölvedi, T. (2023). Peer-to-peer lending: Legal loan sharking or altruistic investment? Analyzing platform investments from a credit risk perspective. Journal of International Financial Markets, Institutions and Money, 86, 101801.

5.1 Introduction

Rapid technological changes in recent decades have contributed to financial disintermediation trends. Alternative financing providers have emerged or started to increase their market share by promising cheaper services and better financial inclusion through the exclusion of intermediary institutions (Polasik et al., 2020).

There are still many questions about the relevance, potential benefits and future of direct lending through an online platform. While the removal of the intermediary layer—and, hence, the regulatory obligations on banks—may be cost-effective and help absorb underbanked customers, it is not evident whether investors are compensated for the high credit and liquidity risk or whether their altruistic motives are needed to sustain the business. On the borrower side, it is also unclear whether better access to finance for subprime customers is in their interest or an access contributes to higher default rates and greater difficulties for the segment (Gosztonyi &Havran, 2021).

This paper is linked to two strands of research in the literature. The first is the role of alternative information in the risk assessment of platforms, which has so far been studied mainly in the US market. The second is whether online platforms substitute or complement banks and serve to improve financial inclusion.

We analysed the loan level data and all cash-flows of Bondora (https://www.bondora.com/en), the sixth largest peer-to-peer platform in Europe (p2pmarketdata.com). Bondora allows to their clients investing in personal debt starting at €1 and enables borrowers to receive funding directly from investors. The investments are denominated in EUR and are available for retail investors mainly from the EU. At the time of our analysis, loans were available to borrowers in 4 countries: Estonia, Finland Slovakia and Spain. The platform itself provides the online marketplace and some services, such as credit rating of the applicants, without taking any credit risk.

In this study we address the following research questions:

1) Do P2P lending platforms have an advantage in information processing compared to traditional banks due to the incorporation of alternative information?
2) What is the performance profile of P2P investments for the lenders?

Our contribution to the literature is twofold. First, while the majority of studies on P2P lending focus on the US market, we analyze the European market in Bondora's very detailed database, which includes loan-level portfolio tables and all cash flows since the platform was launched. Second, in order to assess the performance of the loans, we conduct an ex-post cash-flow analysis, which to our knowledge has never been done before. By analyzing the cash-flow data of the platform, we can investigate not only the default defined by the platforms, which ignores all further possible payments, but also the losses realized ex-post.

We found that the platform's scoring model could not outperform our model, although we used the standard, publicly available data of the platform. Both our model and the platform's model rely on differences in default by country of origin, but they are poor at predicting default within countries.

Although alternative information could improve the predicting power of the models, there is no sign of benefits from using them. We could not confirm that the platform was able to reduce information asymmetry better than traditional financial intermediaries. This contradiction between our findings and the results of the literature could be due to the difference between the US and the continental European markets.

The ex-post analysis of the cash-flows shows that the average internal rate of return (IRR) is negative, and more than 40% of all transactions end with a negative IRR; thus, there is a net loss of investment. From a regulatory perspective, our results suggest that platform investors bear an uncompensated, maybe unforeseeable credit risk. The high default rate of non-Estonian borrowers reflects a decrease in willingness to pay with distance, that can be due to lower cross border collection efficiency. This inefficiency combined with loan-sharking level interest rates (77.5% on average in the worst rating segment) may lead to adverse selection of the borrowers.

While harmonisation of regulation and improvement of investor protection is underway (Regulation (EU) 2020/1503 on European crowdfunding service providers for business is applied from November 2021), further requirements and specifications would be needed to improve the efficiency of cross-border collection and the transparency of platform investment performance. In addition to ex-ante data published by the platform, ex-post performance disclosure both in terms of returns and the accuracy of models could reduce information asymmetry.

The structure of the paper is as follows. The next section presents a summary of the literature, and then in Section 5.3, we present the analysis of the portfolio and cash-flow tables of Bondora. Finally, in Section 5.4, the conclusions are derived.

5.2 Literature review

The literature on platform lending is relatively new but quite extensive. The emergence of platform lending is frequently explained by the rapid technological changes of the last decades, which contributed to the disruption of many services in the economy (Goldstein et al, 2019). However, there is no evidence that the potential benefits of P2P lending platforms disqualify the relevance of banks.

The existence and the rationale of banks are consequences of market imperfections, such as transaction costs, liquidity shocks, and information asymmetry, which make perfect diversification impossible (Freixas & Rochet, 2008). Although P2P platforms are free of the considerable fixed costs of branch networks or employers, their cost efficiency is unsupportable if we compare their average cost of 3–4% of the intermediated amount (Morse, 2015) with the long-term stable intermediation cost of 2% that is present for banks (Philippon, 2016; Bazot, 2018). Platforms primarily only match investors and borrowers; thus, they also do not protect against liquidity shocks.

While P2P lending underperforms, compared with banks, with regard to transaction costs and liquidity insurance, these platforms may be more advantageous in that they reduce information asymmetry to a greater extent, as they use also soft and sensitive data given voluntarily and apply big data and artificial intelligence more flexibly than banks (Liu et al. 2020). Another advantage of these platforms is that they are still free of regulatory restrictions; thus, they can offer high-risk investment possibilities, without risking their own capital (Davis, 2016).

An often-emphasized argument for fintech companies such as peer-to-peer (P2P) platforms is their flexibility to apply the latest and most advanced data analysis methods, which ensures them a competitive advantage over traditional financial institutions (Duarte et al., 2012; Lin et al., 2013; Jagtiani &Lemieux, 2019; Feyen et al., 2021), however, this advantage can disappear, as banks are incentivized to improve their digital services to compete with the new challengers.

As our study relates to the importance and role of P2P lending, we focus on the strands of research that investigate the substitutive or complementary nature of this kind of alternative financing and the impact of alternative information in the lending process.

5.2.1 The role of platforms: substitutes or complements

Thakor (2020) summarizes the literature on fintech around four main questions, one of which is the role of marketplace lending in financial intermediation. He concludes that as banks are unique in providing deposit insurance, P2P lending can complement banks, mainly if banks are more capital constrained, serving clientele unable to pose collateral.

The empirical literature on the role of platform lending is quite widespread, investigating the evolution of the market and the characteristics of P2P loans and comparing them to bank loans. However, only a few theoretical models aiming to situate P2P lending in financial intermediation are present in the literature. In Merton and Thakor's (2019) model of financial intermediation, financial institutions are financed by two types of partners: investors and customers. Investors are willing to take on risk in exchange for an appropriate risk-adjusted return, while customers demand financial services free of credit risk. The optimal contractual design is determined by the cost of insulating customers from the credit risk of the intermediary. According to this concept, the distinction between banks and the market (direct lending) disappears once the above costs, called customer contract fulfillment costs, become sufficiently large. In the case of financing frictions, direct lending can be an attractive alternative for at least a section of the customers. Liu et al. (2019) incorporate both social collateral and soft information into their model and show that both can reduce information asymmetry, making financing available even to small borrowers with limited assets. In contrast to the traditional lending market, low-risk borrowers can crowd out high-risk borrowers, and P2P platforms complement traditional banks by serving those who are not targeted for bank lending.

The majority of empirical evidence shows that platform borrowers are mainly underbanked customers with limited access to bank finance (Das, 2019; Maskara et. al, 2021); the unavailability of other financing options encourages borrowers to turn to the platform. De Roure et al. (2016), examining the German market, also conclude that P2P lending platforms serve an underbanked segment of low-credit customers, which is out of the scope for conventional banks. This concept of collaboration is supported by Milne and Parboteeah (2016), who propose that P2P platforms have a complementary function in lending activity. Specifically, they supplement traditional banks because banks possess a few comparative advantages, which precludes platforms from competing with them. The two financing forms may cooperate in the future.

In contrast to the above results, Tang (2019) finds, when investigating the unsecured consumer loan market in the US, that a negative shock in the bank's credit supply lowers the quality of P2P platforms' credits. He concludes that the results confirm the role of platforms as a substitute for bank lending in serving infra-marginal bank borrowers while also complementing bank lending for small loans. These results also suggest that the credit expansion of P2P lending was based on borrowers who already have access to bank credit. Cornaggia et al. (2018) also highlight the substitutive role of P2P platforms, stressing that smaller banks suffer losses due to the decline in loan volumes, while large commercial banks are not affected. Following the restrictions of the COVID-19 pandemic, Najaf et al. (2022) argue, fintech P2P lending has become the most viable alternative credit option available to borrowers. Moreover, online services have the potential to augment or replace lending provided by traditional or conventional banking institutions.

From the investor's perspective, however, platform funding has higher risk and less transparency, and risk management is underdeveloped (Milne & Parboteeah, 2016). Although platform investing provides higher interest rates, risk-adjusted interest rates are comparable (De Roure et al., 2016), and it is questionable why unsecured P2P lending, which is not even covered by deposit insurance, is beneficial for lenders. The existence of the market under unfavorable lending conditions can be explained by investor preferences, as in the model of Berentsen and Markheim (2020), where altruistic investors are willing to finance even projects generating negative expected cash-flow.

5.2.2 Platforms' risk assessment: role of alternative information

Information asymmetry in financial intermediation relates to the problem of the lender having constrained knowledge of the borrower's creditworthiness. The consequences of information asymmetry can be considered to be specific forms of transactional costs, which can be reduced by monitoring and, hence, improving the efficiency of lending. Financial intermediaries create value by economizing monitoring costs (Diamond, 1984) and by having better access to borrowers' credit and account history or other public sources, such as bad debtor registry or legal processes. Banks, using their own capital to finance borrowers, also provide signals about the quality of the debtor. P2P platforms have little to no access to the previous financial history of the borrowers, and the verification of this information is also costly and, sometimes, even impossible. As P2P lending platforms do not offer credit from their own sources, the signaling effect is also less significant than in the case of banks. Additionally, these platforms have the advantage of applying big data analysis techniques and obtaining “soft information,” which banks are not allowed to gather (Havrylchyk & Berdier, 2018). Earlier, in the case of crowdfunding campaigns, borrowers and lenders were aware of one another, and the social relationship facilitated the screening of borrowers. In the present day, this kind of proximity is atypical among borrowers and lenders; however, the narratives submitted voluntarily by borrowers can contain sensitive information and may have a high impact on investors’ decisions.

The fact that platform lending is the most intense in consumer lending, where informational asymmetry is the highest, also confirms the importance of the information that banks are prohibited from collecting (Havrylchyk & Verdier, 2018). Information asymmetry and capital requirements make it unfeasible for banks to finance high-risk customers, even if they are willing to pay higher costs. By reducing information asymmetry, P2P platforms are seemingly able to reduce credit rationing. Although financial institutions have also started to use digital techniques to handle low-quality big data, the fact that machine learning tools work as a black box constrains their applicability for regulatory purposes (van Liebergen, 2017).

The information P2P platforms collect is mainly based on hard data on the borrower and the credit itself, as well as other local economic information, such as the location’s criminal statistics or employment rates (Jagtiani & Lemieux, 2018). Borrowers may provide other data sources, such as an account of the utilities availed, public reports, and alternative lending payment history. The narratives applicants provide regarding their goals and credit purposes are also a source of soft information; however, their usefulness is not confirmed (Herzenstein et al, 2008). Emekter et al. (2015) examine the credit risk and performance models of platforms based on the data of the biggest P2P lending platform Lending Club. They find that besides the platform’s credit grade, a few other variables, such as debt-to-income ratio, FICO score (credit score created by Fair Isaac Corporation), and revolving line utilization, also have significant explanatory power on loan default. The credit grading reflects the riskiness of the loan, but the higher interest rate charged is not enough to compensate the investors in the worst clientele.

Jagtiani and Lemieux (2019), in their analysis of the loans of Lending Club, find that the correlation between the assigned grade of the platform and the borrowers’ FICO score declined from 80% to 35% from 2007 to 2014, and the platform’s grades proved to perform better while predicting loan default. Their results confirm the benefits of alternative data used by fintech lenders. Having examined an extended sample (2007–2018) of the Lending Club consumer platform, Croux et al. (2020) also confirm the importance of alternative data on loan default. Das (2019) highlights the importance of alternative data in developing better credit models, which allows lenders to select creditworthy borrowers from the lower FICO score bucket, who are otherwise excluded from traditional financing. Cumming et al. (2020) also highlight the role of soft information, proving that the higher the risk the borrower faces (large amount needed or all or nothing financing form), the higher the length and readability of soft information provided. Hughes et al. (2022) compared the lending efficiency of LendingClub's with that of traditional financial institutions and they find that the platfom's credit evaluation to be more accurate, which they explain by the use of alternative data and complex modeling capability of fintech credit providers.

5.3 Performance analysis of peer-to-peer investments

Based on the theory and the empirical evidence on P2P platforms presented in the previous sections, we first examined the information processing of the P2P platform Bondora, an Estonian marketplace launched in 2009, to find evidence on the role of alternative information. Then, to answer our second research question, we analyzed the performance of the platform's investments.

Bondora provides different datasets, which are updated daily. We used two types of datasets: first, the raw data of the loan book containing all loans with different applicant-related and other variables (Dataset 1 and 2, detailed description in the next subsection) and, second, the historical payment table that includes all cash-flow series of each loan (Dataset 3). Table 22 contains the description of them.

22. Table: Description of the Datasets of Bondora used for the Analysis

	Dataset 1	Dataset 2	Dataset 3
Downloaded	October 5, 2020	March 15, 2022	May 24, 2022
Data type	Loan level data (112 variables)	Loan level data (112 variables)	Cash-flow data (4,343,194 payments)
Number of raw data	151,866 loans	222,978 loans	243,453 loans
Data used for	Investigation and in-sample testing of the scoring model	Out-of-sample testing of the scoring model	Performance from investor perspective (based on IRR)

Data cleaning	Inconsistent, missing and invalid observation were excluded. Origination date between December 10, 2012 and October 5, 2019	Inconsistent, missing and invalid observation were excluded. Origination date between October 6, 2019 and March 15, 2021	Missing values, loans with current or NA status were excluded.
Number of loans used for the analysis	107,588	50,251	62,537 + 42,598

Source: own elaboration based on Bondora datasets

The loan book data were downloaded on October 5, 2020, and on March 15, 2022. The data of the first period were used to build our benchmark model, while the data of the second period were used for the out-of-sample testing. Historical payments covering the principal and the interest amount paid in each month, were downloaded on May 24, 2022. These data were used to examine the performance of platform investments.

5.3.1 Risk assessment of the platforms

First, we analyzed the loan book containing all loans between 2012.12.10 and 2020.10.05. In the framework of data cleaning loans with prepaid status were filtered out, and inconsistent records were also eliminated (e.g., when the default date is missing, but the loan is more than 90 days past due). Additionally, missing and invalid observations were extracted from the dataset. Loans issued within a 12-month time period were also eliminated, which is a necessary requirement for our default definition. After data cleaning, 73,865 observations remained in the First, we analyzed the loan book containing all loans taken between December 10, 2012, and October 5, 2020. The original dataset covered 151,866 transactions and 112 variables, including different types of information: data regarding the characteristics, financial background and payment history of the applicant (e.g. the number and the amount of the pervious loans the borrower had before this application), standard information about the loan request (e.g. interest rate, loan amount, tenor) and technical data (e.g. loan ID, the form of bidding). The full list of the variables included in the original dataset is presented in the Appendix VI.

During data cleaning framework, inconsistent records were eliminated (when the default date is missing but the loan is more than 90 days past due). It impacted 3,029 rows. Additionally, missing and invalid observations were excluded from the dataset (we considered a record invalid in case its value was not defined in the Bondore data description). Loans issued within a 12-month period (from October 6, 2019) were also eliminated, as to our definition of default 12 months after issuance is relevant. After data cleaning, 107,588 observations (loan transactions) remained in the sample.

To support the decision of investors, the platform performs a credit risk assessment for each applicant and assigns a rating based on its internal evaluation. The methodology is not publicly available; only the final rating is shared with investors. According to the literature, the rating of the platform is one of the main factors that impact investors' decisions. Better rating generally results in greater success in funding (Herzenstein et al. 2008, Emekter et al. 2015, Gavurova et al. 2018). Therefore, it is crucial that the rating appropriately reflects the risk of the potential borrower. Furthermore, according to the platforms, one of the main advantages of P2P lending is their credit risk assessment process, as the platform applies alternative information for their assessment, besides the standard variables used by banks (among others Popescu, 2016; Croux et al., 2020; Hughes et al. 2022).

According to the platform's webpage, the current credit risk rating methodology was introduced in December 2014 to improve the previous rating practice and support risk-based pricing. As stated, the method is in line with the industry's best practices, commonly applied by the banking sector. Overall, eight rating bucket grades are determined on a scale from AA (lowest risk) to HR (highest risk). The rating classification is based on the expected loss intervals, calculated by the platform with the following formula:

$$EL\ (\%) = PD * LGD * EAD \tag{7},$$

where *PD* is the probability of default, *LGD* is the loss given default, and *EAD* is the exposure at default. The data is derived from three main sources: information provided by the applicant, an external credit bureau database, and behavioral information collected through the application process on the webpage (Bondora, 2021). It must be noted that this practice differs from banks' methodology, where the basis of rating assignment is the probability of default and not the expected loss.

The portfolio table provides information on the performance of the transaction by disclosing the number of days past due (DPD) and the date of default if the loan is in default. However, these default indicators are not fit our purpose, as industry practice considers a default to be a delay in payment of more than 90 days in the first 12 months after the loan is issued. Therefore, we created a default flag accordingly and considered a borrower to be in default if they were more than 90 days in arrears in the 12 months following the origination of the loan. In the following, we use this default definition for modeling and also when presenting ex-post default frequency. Since the portfolio table contains no information on the delayed amount, the materiality threshold is not included in the default definition. It is important to note that the default defined above does not represent the ultimate loss considered in subsection 3.4, as loans that are 90 days past due in the first year after issuance may recover and meet their obligations later. Similarly, loans that perform well in the first year may default later, leading to losses.

Table 23 below presents the distribution of the loans examined for the first period (2012–2019) based on the rating category provided by the platform at origination and the days past due (DPD).

23. Table: The Distribution of the Portfolio by Rating and DPD

DPD	AA	A	B	C	D	E	F	HR
0	71%	70%	65%	59%	50%	44%	33%	27%
1-7	3%	3%	3%	3%	3%	3%	2%	1%
8-15	3%	3%	4%	4%	4%	3%	1%	1%
16-30	5%	7%	7%	7%	8%	7%	6%	3%
31-60	3%	3%	3%	4%	5%	5%	6%	3%
61-90	2%	2%	2%	2%	3%	3%	3%	2%
91-120	0%	1%	1%	1%	1%	1%	1%	1%
121-150	1%	1%	1%	1%	2%	2%	1%	1%
151-180	1%	1%	1%	1%	1%	2%	2%	1%
180+	10%	10%	13%	18%	24%	30%	43%	61%

Source: Bondora webpage as of 2020 October

The platform's credit risk assessment seems to be mostly reasonable. Claims with worse ratings usually showed worse performance. As we had no external data on the borrowers, such as the FICO score in the US, we could not compare the platform's ratings with banking models, as Jagtiani and Lemieux (2019) do, but our data support their findings.

For a high-level portfolio overview, the loans' main characteristics were examined. Table 24 provides a summary of the main statistics by the platform's internal rating.

24. Table: Descriptive Statistics of Bondora's Portfolio by Rating as of 2020 October

Rating	Number of loans	Average loan amount (EUR)	Standard deviation of loan amount	Average interest	Standard deviation of interest rate	Average expected return	Number of defaulted loans	Average default rate
AA	2,686	1,390	1,452	11.50%	4.50%	9.58%	255	9%
A	5,381	1,575	1,661	13.56%	4.68%	10.46%	566	11%

B	12,986	2,003	1,969	16.15%	3.91%	10.82%	1,740	13%
C	17,332	2,481	2,304	21.81%	3.94%	12.27%	3,073	18%
D	18,079	2,761	2,305	28.51%	3.98%	13.50%	4,623	26%
E	17,624	2,879	2,298	35.14%	4.19%	14.37%	5,727	32%
F	19,701	3,367	2,259	53.17%	11.30%	17.87%	9,710	49%
HR	13,799	1,750	1,516	77.24%	50.90%	15.52%	8,113	59%
Total	107,588	2,531	2,201	36.62%	27.72%	13.93%	33,807	31%

Source: Bondora webpage as of 2020 October

The loan amounts are low, with an average of €2,531 and a maximum of €10,632. Loans are concentrated in the lower rating categories, with more than 80% of the loans falling into the C-rated or even lower one. Risk-based pricing is reflected in the average interest rate. Even for the AA grade, the interest rate is higher (11.50%) than the average bank interest rate in the Euro area, which is around 6.9% for the examined period in the retail consumer loan segment (ECB, 2020). Investors' expected returns are growing monotonously with the lowering of the rating, except for the HR rating, where the higher risk is not compensated with a higher expected yield. In comparison, the average bank deposit in the Euro area for the same period was around 0.9% for maturity over 2 years (Euro Area Statistics, 2020). Regarding the default rate (based on the above-described default definition), the portion of defaulted loans also increases with the worsening of the rating. The average default rate on the whole portfolio is 31%, which is extremely high, compared to commercial banks' retail portfolios.

5.3.2. Analysis of the scoring model

To investigate the performance of the platform's scoring model, we built a benchmark model using publicly available standard variables, usually included in the credit risk assessment process of a commercial bank. Then, we compared our results with the platform's estimation to find evidence of the role of potential alternative data used by the platform.

For our benchmark model, we selected 12 standard variables related to the financial position of the borrower, their previous loan history, and a few social features. The variables were selected taking into account the relevance of the data for scoring and their availability. We performed sanity checks to confirm that the dataset of each variable was complete and valid. The examined variables are listed in Table 25. For descriptive statistics, see Appendix VII. We examined the value set and the distribution of each variable and performed a few transformations where we found them reasonable.

25. Table: Description of the Variables of the Scoring Model

Variable	**Description**	**Transformation**
Age	Age of the borrower	No transformation was applied.
Country	Country of the borrower. A category type variable, it can take the following values: EE (Estonia), ES (Spain), FI (Finland), and SK (Slovakia).	Weight of evidence (WOE) transformation was performed. The WOE was calculated using the following equation: $WOE = ln\left(\frac{Portion\ of\ goods}{Portion\ of\ bads}\right)$ where good refers to the portion of borrowers who paid the claims, and bad is the portion of debtors who defaulted according to our definition of a default. The estimated WOE by buckets are assigned to each observation in the portfolio table, and they are used to perform binary logistic regression instead of the original values.
IncomeTotal	The sum of the debtor's total income	The natural logarithm of the total income was taken.
ExistingLiabilities	The number of current liabilities of the debtor	Above the first five categories, the other categories were merged and considered as one category (existing liabilities above 4).
LiabilitiesTotal	The total monthly liabilities of the debtor	We created a dummy variable. We calculated the deciles and merged the first two categories and assigned a 0 for them, while the rest got 1.
DebtToIncome	The debtor's monthly loan installments divided by the	No transformation was applied.

	monthly gross income. Expressed in percentage (%)	
NoOfPreviousLoansBeforeLoan	The number of previous loans taken before this loan was issued	Above the first five categories, the other categories were merged and considered as one category (the number of previous loans above 4).
AmountOfPreviousLoansBeforeLoan	The amount of previous loans taken before this loan was issued	We created a dummy variable. In case the borrower has zero amount of previous loans we assigned 0, otherwise 1.
PreviousRepaymentsBeforeLoan	The amount of previous loans repaid by the borrower	We created a dummy variable. In case the borrower has zero amount of previous loans we assigned 0, otherwise 1.
Employmentduration	Time spent with the current employer	No transformation was applied.
Education	The education level of the debtor. A dummy variable, it can take the following values: 1 (primary education), 2 (basic education), 3 (vocational education), 4 (secondary education), and 5 (higher education).	No transformation was applied.
Homeownershiptype	The type of the debtor's home ownership. A dummy variable, it can take the following values: 0 (homeless) 1, (owner), 2 (living with parents), 3 (tenant, pre-furnished property), 4 (tenant, unfurnished property), 5	We created a dummy variable. In case the borrower is an owner we assigned 0, otherwise 1.

	(council house), 6 (joint tenant), 7 (joint ownership), 8 (mortgage), 9 (owner with encumbrance), and 10 (other).	

Source: Bondora webpage as of 2020 October

Even if in a predictive model the variable X has a high explanatory power in the sense that the conditional distributions P(Y|X) (where Y is the dependent variable) vary with the values of X, it may not be suitable for logit regression, for two main reasons. The values of categorical variables are usually arbitrarily chosen integers that have no specific meaning, so their weighted sum would be meaningless in the model. Because the logit function has low values for low PD and high values for high PD, or vice versa, if PD is not a monotonic function of a variable, it performs poorly in the model. In such cases, some transformation must be applied to the variable to make it fit the model. The values of the variable are grouped into 'similar' categories (or split if the distribution is continuous), which means that the PD is more or less constant within each category, but varies widely between categories. If an intuitive high and low risk category (as in our case all liabilities) can be identified, a dummy variable is usually created. If there seem to be more than two homogeneous categories (in our case, the borrower's country), a WOE transformation is usually applied. The average PD of each category is calculated and plugged into the inverse logit function $WOE_i = ln((1 - p_i)/p_i)$, so that a linear relationship between the values of the variable and the log-odds is established, which is ideal for the logit model. Finally, if a variable has a "wild" distribution, a smooth transformation (e.g. logarithm) is usually used to handle extreme values and outliers. This is the case for the total income variable.

As a next step, we estimated the GINI coefficients that reflects the explanatory power of the variables, results are presented in Table 26.

26. Table: The GINI Value of the Variables used for the Scoring Model

Variable	GINI
Age	0.035
country	0.387
IncomeTotal	0.153
ExistingLiabilities	0.182

LiabilitiesTotal	0.076
DebtToIncome	0.114
NoOfPreviousLoansBeforeLoan	0.214
AmountOfPreviousLoansBeforeLoan	0.186
PreviousRepaymentsBeforeLoan	0.002
Employmentduration	0.005
Education	0.075
Homeownershiptype	0.149

Source: Bondora webpage as of 2020 October

By examining the default explanatory power of each single variable country proved to be the strongest, but income and indebtness characteristics are also impacting the default frequency. Although theoretically possible, it is not realistic to build strong models from weak variables, so we only included variables with sufficiently high explanatory power. There is no exact rule for when a variable is strong enough, but 10% is a good rule of thumb. There are two other variables close to this threshold, total liabilities and education, which are considered very important for retail lendings. We included these in the model and dropped the weakest three. Based on the correlation of the variables, we eliminated one pair of variables (with lower GINI) whose correlation coefficient was above 0.5. To build a scoring model, we ran different logistic regressions. We had a group of variables, measuring the indebtedness of the borrower (AmountOfPreviousLoansBeforeLoan, NoOfPreviousLoansBeforeLoan, PreviousRepaymentsBeforeLoan) and we tried each of them together and also separated along with other variables in the regression, however the GINI of the model was the same in the different versions. Therefore, we selected the final model based on intuition and economic interpretation of the betas. The final variables to be used for our scoring model are the country of the borrower; the type of home ownership; total income; the number of loans taken before the loan; and the total liabilities the borrower has. The results of the regression are presented in Table 27 below.

27. Table: Results of Binary Logistic Regression

Variable	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-0.312080	0.095584	-3.265	0.00109 **
country	-0.934619	0.010481	-89.171	< 2e-16 ***
IncomeTotal	-0.085385	0.013042	-6.547	5.86e-11 ***

LiabilitiesTotal	0.166665	0.017715	9.408	< 2e-16 ***
NoOfPreviousLoansBeforeLoan	-0.086781	0.005162	-16.811	< 2e-16 ***
HomeOwnershipType	0.397527	0.014481	27.451	< 2e-16 ***

Source: Bondora webpage as of 2020 October
**p< .1, **p< .05, ***p< .01.*

Based on the regression results, all variables proved to be significant, all of them at the 99% significance level. The impact of income and total liabilities is in line with our previous expectations, with higher income and lower liabilities reducing PD. Interestingly, if borrowers had more credit before the current one, their creditworthiness improves. Home ownership has a value of zero if the borrower owns (wholly, jointly owned, encumbered or unencumbered) the house and a value of one if the house is not owned by the borrower (e.g. living with parents or renting). Being an owner reduces PD as expected.

Using the beta coefficients of the final model, the score for each observation was estimated. This was transformed into the probability of default with the following formula:

$$f(y) = \frac{1}{1+e^{-y}} \quad (8),$$

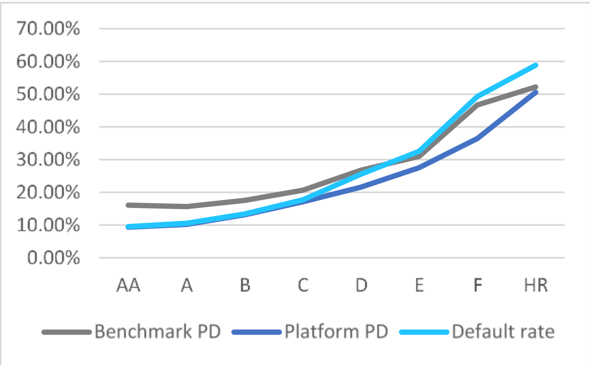
where

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + .. + \beta_n x_n \quad (9).$$

The average probability of default for each rating category based on our model (Benchmark PD), the PDs estimated by the platform at origination (ex-ante), and the real default rates using our default definition are presented in Figure 14.

14. Figure: Comparison of the ex-ante Model PDs and the Observed Default Rate

Rating	Benchmark PD (ex-ante)	Platform PD (ex-ante)	Default rate (ex-post)
AA	16.02%	9.32%	9.49%
A	15.66%	10.22%	10.52%
B	17.54%	13.15%	13.40%
C	20.58%	17.10%	17.73%
D	26.75%	21.56%	25.57%
E	30.97%	27.52%	32.49%
F	46.69%	36.37%	49.29%
HR	52.16%	50.51%	58.80%

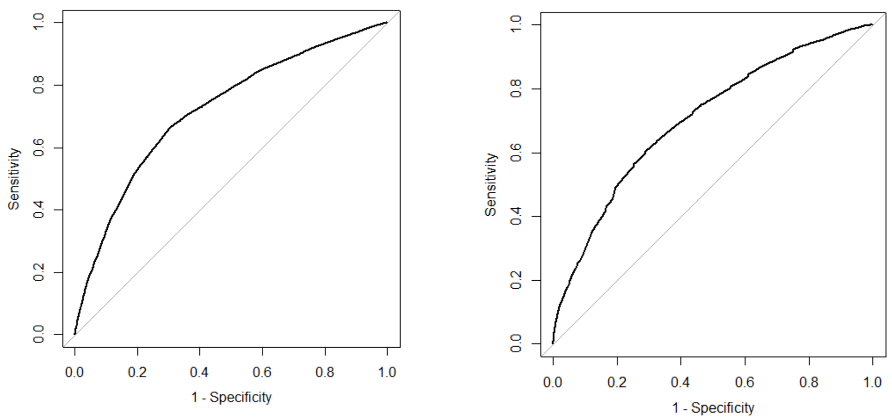


Source: Bondora webpage as of 2020 October

Based on our data, the platform underestimates the probability of default for all rating grades. Our model's in-sample estimation is a little closer to the observed default rates for the worse rating categories than the platform's results, however, we overestimated the probability of default for the best rating category.

We estimated the ROC curve to check the classification power of the models (presented in Figure 15). Our benchmark model's in-sample performance resulted in a GINI of 44.10%, while the platform's classification achieved a GINI of 41.08%. The goodness of these GINI values is hard to judge. In the case of a commercial bank, a retail scoring model is expected to achieve a GINI higher than 80%, but for special, high-risk portfolios, significantly lower GINIs may also be acceptable. On the other hand, Jagtiani and Lemieux's (2018) model obtained a GINI of 38% (Area under the ROC curve 69%), even for the best-performing variable set.

15. Figure: Comparison of the ROC (Receiver operating characteristic) curves in-Sample



Source: Bondora webpage as of 2020 October

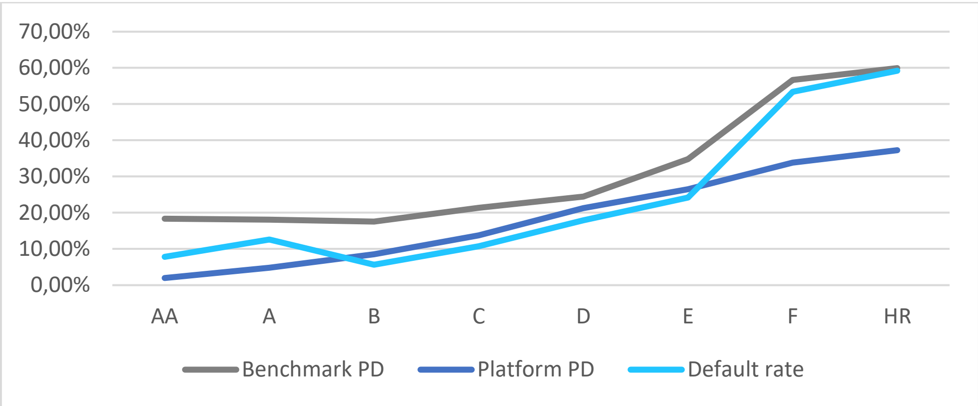
Note: The figure on the left shows our model's curve, and the one on the right is the curve of the platform's model.

The in-sample GINI we could achieve was only slightly higher than the GINI of the platform's model for the same period. Therefore, we can confirm that the platform's model performs appropriately. However, the performance of our benchmark model suggests that a similar result can be achieved based on a classic "banking-like" information dataset. Hence, in our investigation, we could not detect any sign of the benefits of using alternative data sources or information processing of fintech lenders.

5.3.3 Out of sample results

We also performed out-of-sample testing to check the default prediction performance on the independent data horizon. We performed the same data cleaning steps for the portfolio table as of March 15, 2022, as mentioned for the benchmark model. As the time horizon of the benchmark model ends on October 5, 2019 (we filtered out loans issued within 12 months), the validation sample consists of 50,251 loans starting from October 6, 2019. We used the parameters (WOE binning for the country variable and coefficients) of the original model and re-estimated the PDs for the new database. The results are presented in Figure 16, along with the platform estimation and the observed default rates.

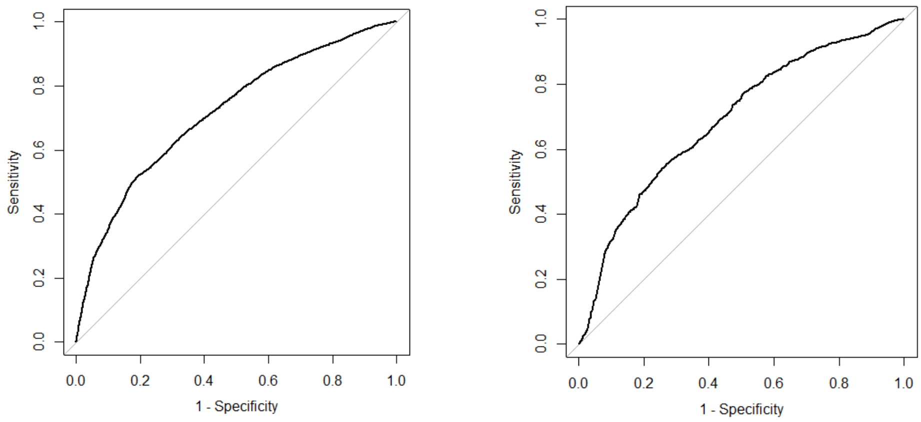
16. Figure: Comparison of the Out-of-Sample PDs and the Real Default Rate



Source: Bondora webpage as of 2022 March

The platform PDs underestimated the real default rates in the best and the worst rating categories, while our model overestimated them, in a manner similar to the previous period. The GINI of our model is 43.28%, only slightly below the in-sample value (44.10%). The analysis of the platform's PD resulted in a GINI of 37.92%, compared to the previous period's 41.08%. The ROC curves are presented in Figure 17. Interestingly, the drop in GINI was higher for the platform's estimation, even though their rating is given at the origination. Thus, this change is not due to the out-of-sample testing, as in the case of our benchmark model, but the changing market conditions.

17. Figure: Comparison of the ROC Curves out-of Sample

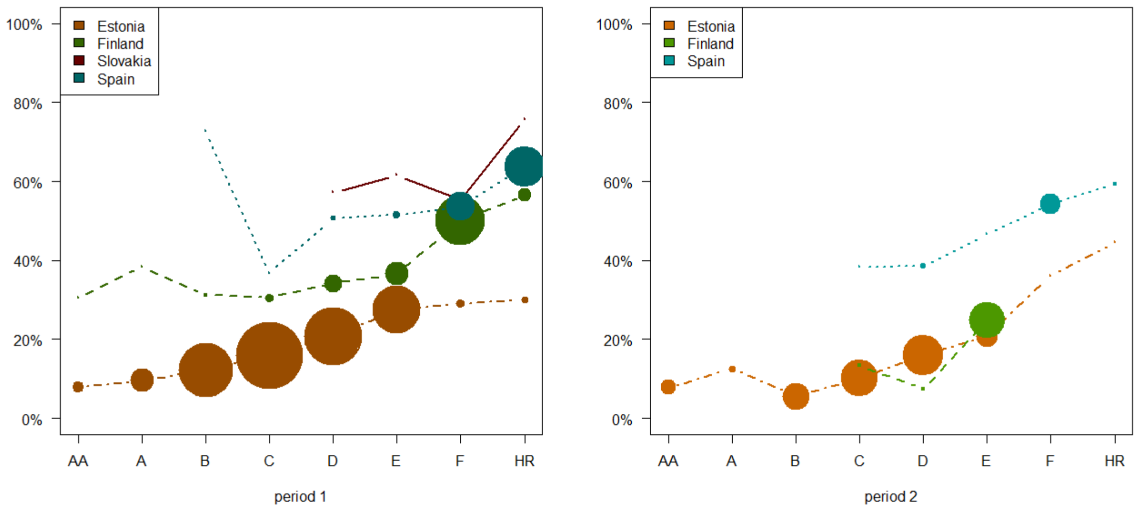


Source: Bondora webpage as of 2022 March

Note: The figure on the left shows our model's curve, and the one on the right is the curve of the platform's model

As the highest explanatory value—GINI—was classified by the country, it is worth examining the loans on a country level. Figure 18 presents the country-level default rates of the rating categories, separately for the first and the second periods. Ratings and risk assessment is adequate for Estonia, but for the other countries the default rate is not a monotonic function of rating. Loans outside Estonia (foreign origin) are much riskier, reflected in high default rates. There are no loans provided for Slovakian borrowers in the second period.

18. Figure: Default Rates by Country for the Two Periods



Source: Bondora webpage as of October 2020 (period 1) and March 2022 (period 2).

Note: The size of the bubbles reflects the number of loans. Period 1 contains loans issued between December 2012 and October 2019, Period 2 lasts from October 2019 to March 2021.

Loans out of Estonia not only have lower ratings, but also higher default rates in all rating categories (especially in the first period). This suggests that the payment discipline of borrowers of foreign origin is much lower, which can be due to the weaker cross border credit collection.

We tested the performance of our model for each country separately by calibrating a scoring model to country-level data. The predictive power of the risk models is essentially reduced at the country level; a summary of the GINIs is shown in Table 28. Our country-level models, which rely on the variables used for the benchmark model above except for the country (total liabilities, the type of home ownership; total income; the number of loans taken before the loan) performed poorly with GINIs of around 20%. The platform model performed slightly better in 5 out of 7 cases, but its discriminatory power is also very low.

28. Table: Default Rates and Performance (GINI) of the Scoring Models by Country

	Default rate		Number of loans		Platform model GINI		Benchmark model GINI	
	2012–2019	2019–2022	2012–2019	2019–2022	2012–2019	2019–2022	2012–2019 (In-sample)	2019–2022 (Out-of-sample)
Estonia	18.12%	12.54%	62,392	33,407	28.80%	25.13%	15.12%	22.35%
Finland	43.66%	22.54%	26,317	9,661	21.98%	16.87%	16.12%	17.13%
Spain	58.16%	51.34%	18,585	7,183	19.51%	14.70%	22.62%	11.99%
Slovakia	70.49%	NA	288	0	29.89%	NA	15.88%	NA

Source: Bondora webpage as of October 2020 and March 2022.

The decline in the performance of the scoring model at the country level is due to the fact that the rating of borrowers differs significantly across countries. Consequently, the score of each loan is determined by the country of the borrower and the other variables have much less explanatory power. It seems that in the high-risk segment, where platform lending is active, individual defaults are much less predictable. Since the platform model could not significantly outperform our naive models, and both performed rather poorly, it seems that additional - alternative - information could help to build a better model. However, based on the data analysed, there is no evidence that platforms use and exploit alternative information. It is also important to note that the high default frequency suggests that P2P loans differ significantly from the loans acceptable for traditional financial institutions. Thus, using traditional techniques for their risk assessment is not appropriate.

To better understand the characteristics of P2P loans and the motives of P2P lending from the lenders' perspective, in the following section, we analyze the ex-post performance of platform investments.

5.3.4 Return of peer-to-peer investments

P2P lending represents high-risk, bank-ineligible loans. Due to the strict regulatory requirements and reputational risks, banks are unable to provide financing for this segment, but individual or institutional investors may benefit if the risk is compensated by high interest rates.

To understand the effective performance of P2P loans, we examined the dataset of historical payments, with the principal and the interest amount paid each month by loan ID. The cash-flow table of Bondora downloaded on May 24, 2022, contains 4,343,194 rows, representing all payments during the lifetime of the platform. First, we ordered the payments according to loan ID. Consequently, we got the detailed cash-flow of 243,453 loan transactions. Then, we ordered the other details (rating, PD, status, etc.) based on lon ID from the portfolio table according to each loan.

Bondora assigns a status to each loan: current (transactions in progress), closed or unavailable (NA). For the latter category, other data is missing and therefore not suitable for analysis. As ex-post analysis is feasible only for closed transactions, so, first, we investigated the loans with a closed status. We calculated the IRR for each loan based on the cash-flow and the historical payment schedule. 62 537 transactions issued between 28 February 2009 and 14 March 2022 were analysed, excluding loans for which no payment was made other than at the time of disbursement.

Table 29 shows the ex-post performance of closed transactions according to rating categories. The rating was missing for 2,556 transactions. Thus, they are shown separately in the table. Average loan term refers to the difference between the last and the first payment, Sum CF is calculated by simply summing up all cash-flows.

29. Table: Ex-post performance of closed transactions

Rating	Number of loans	Default rate	Average loan amount	Average loan term (in days)	Average Sum CF	IRR mean	IRR St. Dev.	P (IRR < 0)
AA	2,672	5.01%	1,922	620	176	9.01%	10.54%	9.81%
A	3,166	6.16%	1,796	744	315	12.24%	12.32%	11.37%
B	8,205	6.79%	2,057	683	368	14.41%	11.67%	10.65%
C	11,246	8.94%	2,281	661	492	19.24%	14.52%	11.33%
D	11,099	13.86%	2,409	629	528	23.41%	19.46%	14.33%
E	1,0431	16.12%	2,597	557	533	25.61%	22.30%	15.31%
F	7,471	20.28%	3,025	501	728	44.57%	38.18%	15.13%
HR	5,691	31.15%	1,756	761	516	70.93%	127.33%	16.87%
NA	2,556	16.63%	642	712	166	28.36%	18.86%	2.27%
ALL	62,537	14.11%	2,261	635	484	27.72%	46.40%	12.96%

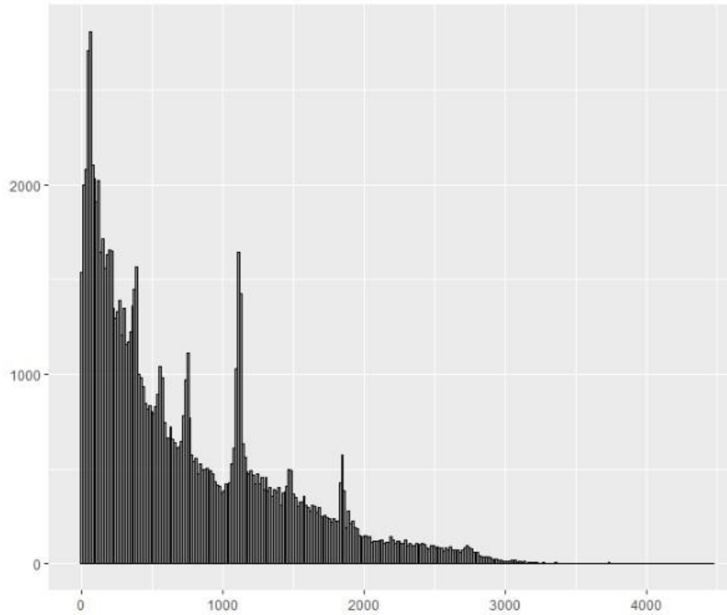
Source: Bondora, as of May 2022

The worst two categories offer really a high return with high volatility, but surprisingly, the effective loss (negative IRR) in those segments is not significantly higher than in the much better rating categories. Because of the definition of default, transactions with the defaulted flag do not necessarily represent an actual loss, and defaulting transactions that are past due for more than 90 days in the first year after origination may subsequently be recovered and all their obligations settled. Additionally, even non-defaulting transactions can result in losses if the default occurs outside the first annual period after the contract is signed. The average ex-post default rate denoted by the platform was 14.11% (see Table 29), much lower than the proportion of loans having non-performance status in both periods (see Figures 14 and 16). Investments with an effective loss—negative IRR—make up even less, specifically, 12.96% of the loans. On average, investors realizing negative IRR lost 23% of their initial investment. We can conclude that the majority of the defaulted closed transactions were recovered, ensuring an average IRR of 27.72%.

The picture is not that bright if a closer look is provided at transactions with a "current" status. Although we do not have the original repayment schedule of the loans, which could be indicative of non-payment and default problems, we also included in the analysis loans where the original maturity (calculated from the date of issuance and the original loan duration) is exceeded at the time of data collection or where no payment was made in the last one-year period. We assume here that these transactions can also be considered closed as there will be no further related payments. This represents 42,598 additional loans that are more problematic, 48% of which have received a default flag from the platform. In total, including closed status transactions, Bondora's total closed portfolio consists of 105,135 deals.

The distribution of actual loan term is presented on the histogram in Figure 19. The P2P loans are mainly short-term transactions, half of them were paid back below one year, while the average original loan term is around 3.5 years. We also found that the majority of the debtors (69%) prepaid the claim before the maturity date, as there is no prepayment charge to be paid by the borrower (Bondora, 2021).

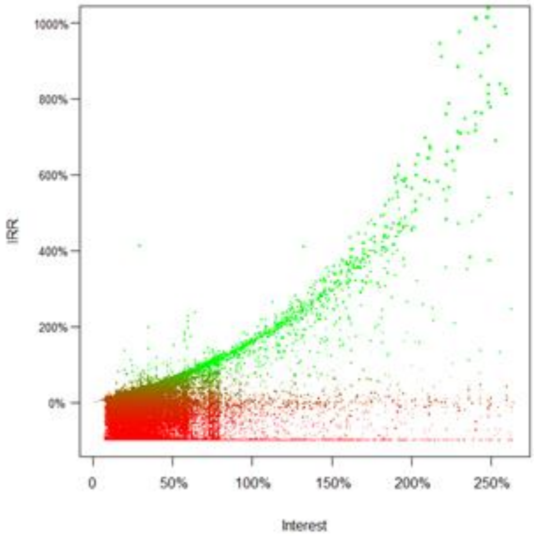
19. Figure: Distribution of actual loan term (in days)



Source: Bondora, as of May 2022

The relationship between the annual IRR of the loan and the interest rate priced initially by the platform is shown in Figure 20. The IRR distribution is wider for riskier, higher-level interest rates, and the correlation coefficient between interest rate and IRR is low, below 0.09. Although a substantial section of the high-risk borrowers performs well, offering above 100% return for the investors, a maximal, 100% loss can be realized at all interest levels.

20. Figure: IRR (internal rate of return) as a function of interest rate



Source: Bondora, as of May 2022

The ex-post performance of the extended closed category is shown in Table 30. The average IRR of the portfolio is negative, which means that investors on average not only do not receive compensation for the risk, but also make a loss on their original investment. The average IRR is negative even in the best rating categories, with only rating C, HR, and the unrated transactions resulting in a positive IRR. The IRR dispersion is high, but overall 41.63% of all transactions have a negative IRR and the realised loss is 55% of the amount invested (for transactions with a negative IRR, the nominal amount of payments received is on average 55% less than the initial investment). So, despite an average initial expected return of 9.58% - 15.52%, the average realised return is negative in most rating categories.

30. Table: Main Characteristics and Ex-Post Performance of the Extended Closed Dataset

Rating	Number of loans	Default rate	Average loan amount	Average loan term (in days)	Average Sum CF	IRR mean	IRR st. Dev.	P(IRR < 0)
AA	3,701	9.92%	1,843.50	764.40	-96.31	-4.00%	29.42%	30.96%
A	4,867	11.42%	1,693.45	900.82	18.30	-3.49%	31.14%	36.47%
B	12,292	12.41%	2,042.84	864.78	25.13	-1.15%	32.08%	33.25%
C	17,116	16.68%	2,349.80	834.13	9.38	0.29%	37.58%	33.46%
D	18,544	25.04%	2,574.12	783.91	-185.74	-3.86%	46.59%	39.70%
E	17,660	29.85%	2,753.38	644.83	-340.97	-9.12%	53.03%	44.51%
F	16,374	44.99%	2,943.98	529.52	-584.43	-17.29%	68.88%	56.99%
HR	11,880	53.10%	1,760.87	726.39	-228.66	4.29%	116.83%	53.21%
NA	2,701	19.29%	643.52	767.04	148.63	24.93%	25.64%	5.96%
ALL	**105,135**	**27.98%**	**2,355.22**	**736.35**	**-201.15**	**-4.17%**	**60.36%**	**41.63%**

Source: Bondora, as of May 2022

The negative IRR of platform loans is particularly striking when considering that the period 2012-2022 was a period of economic recovery and boom following the 2007-2009 crisis. The non-performing loan ratio ranged between 7.48% and 1.79% (Statista.com), with a monotonic downward trend after 2015. To assess the performance of P2P lending from an investor perspective, the 10-year performance of different asset classes is presented in Table 31.

31. Table: 10-years Performance of Different Investments

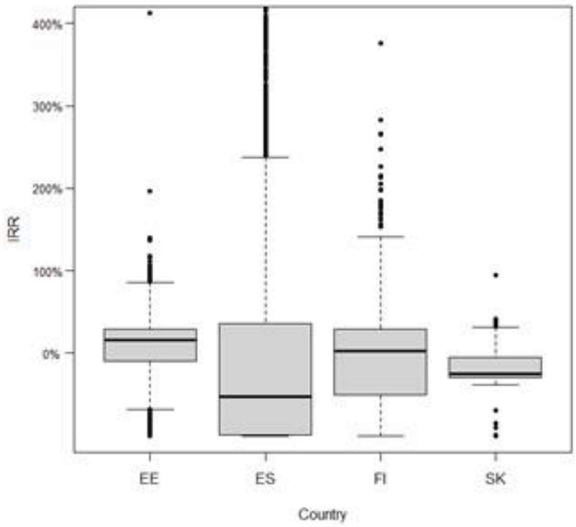
Investment type	10 year Annual return	Volatility
MSCI EUR High Yield Corporate Bond Index	3.21%	7.14%
MSCI Europe Large Cap Index (equity)	7.64%	13.80%
MSCI Euro Index (equity)	8.58%	16.43%
Bitcoin (EUR)	103.08%	97.26%
All transactions of Bondora	IRR: -4.17%	60%

Source: MSCI data as of 31th March 2023, Bitcoin statistics based on daily values available at investing.com

The IRR is not comparable to the ex-post returns of indices or individual assets, but it is clear that all asset classes performed well over the period and generated a positive risk premium. Thus, the negative performance of platform investments is not due to economic factors.

Breaking down the results by the country of the borrower, we found that loans for Estonian borrowers had a slightly positive average IRR of 2.93%, while the average IRR is negative for all other countries. Spain had the lowest IRR with an average of -22.57%, followed by Slovakia with -15% and Finland with -7.6%%. The distributions are shown in Figure 21. It seems that the willingness to pay reduces with physical distance.

21. Figure: IRR distribution according to country

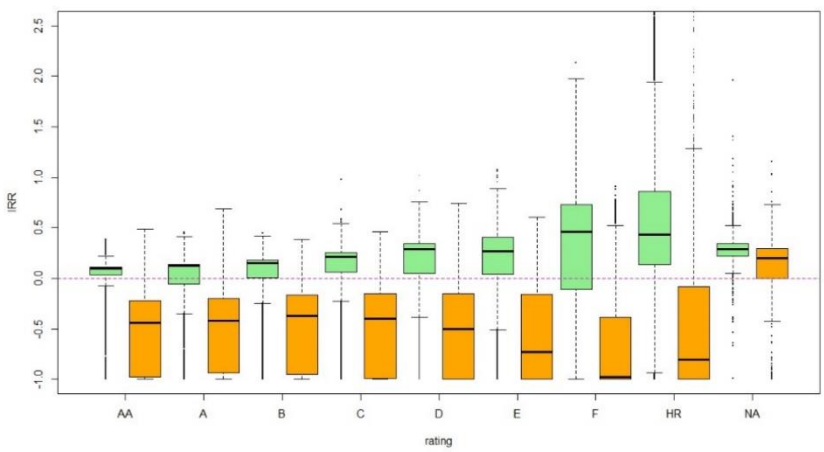


Based on the extended closed dataset of Bondora.

Source: Bondora, as of May 2022

The IRR distribution of the analyzed loans is shown in Figure 22. The green boxes show the distribution of the loans with a non-default status, while the orange boxes are the defaulted loans. The standard deviations increase as the rating worsens, but there is no significant difference in the standard deviation of the distributions within a given category. The IRR is positive for about 75% of the non-defaulted deals and negative for the majority of the loans with a default flag, except for the unrated category, in which even the loans with a default flag realized a positive IRR in 75% of the cases.

22. Figure: Distribution of IRR by rating category and default status



Source: Bondora as of May 2022

According to the analysis of the realized cash-flows and the IRR, investors do not seem to be compensated for the high risk undertaken by P2P loans. The ex-post return is negative on average, confirming the findings of Emekter et. al (2015). The negative IRR also casts doubts on the economic rationale of platform lending and raises the potential role of other motives such as altruism, as discussed in Berentsen and Markheim (2020). However, interest rates of up to 260% depending on creditworthiness do not reflect altruistic motives. It is also worth noting that, in addition to those analysed, only one cash-flow, the initial loan disbursement, was found for 4 997 loan IDs. We can consider them to be credit fraud, where the borrower had no intention to pay anything back. Fraud transactions make up more than 2% of all loan IDs.

5.4 Conclusion

This paper investigated the possible explanations for the rapid growth in P2P platforms' market share in the credit market. From a theoretical perspective, the strongest argument for the relevance of P2P lending is the reduction in information asymmetry through the use of alternative data and P2P platforms' information processing. We analyzed the loan-level data of an Estonian platform, Bondora, to find evidence on the benefit of using alternative information based on the performance of the credit risk model of the platform. We found that the grades assigned by the platform are in line with the default risk of the borrower, but the assigned default probabilities underestimated the real default rate in all segments. Our benchmark scoring model performed slightly better than the platform's model both on the in-sample and out-of-sample data; however, we used traditional explanatory variables: age, debt-to-income, home ownership, employment duration, country of origin, and existing liabilities. When looking at loans by country of origin, the performance of the models decreases significantly, indicating that both our model and the platform's model primarily capture different country-level creditworthiness. Therefore, alternative data may be needed to improve the models, but our results do not confirm that the platform can incorporate them.

By analyzing the ex-post performance of "closed" transactions, we found a significantly high average IRR of 27.72%. However, if we extend the analysis to transactions with a "current" status that are not expected to generate further cash flows, we find that the IRR is negative on average, with 41.63% of all transactions ending in a net loss (negative IRR).

There are huge differences in credit performance across countries. Foreign (non-Estonian) borrowers are not just lower rated, but their credit performance is significantly worse even in the same rating category, indicating the inefficiency of crossborder collection and higher information asymmetry. We conclude that the high credit risk reflected in the extremely high default rates is associated with net ex-post loss of the investment. Although the level of the interest rates excludes the altruistic motives of the investors, they are insufficient to compensate investors for the credit risk. The loan-sharking level interest rates on the other hand may lead to adverse selection on the borrower side.

Platform loans represent high-risk transactions that may not be acceptable for a traditional financial institution; thus, P2P lending complements traditional financial intermediation. However, the negative ex-post performance casts doubt on the rationale of P2P investments, even for market participants free of capital burden and reputation risk. Regulation is now better focused on the segment, but more transparency, disclosure of ex-post performance of loans and models and improvement of the crossborder collection would be needed to reduce information asymmetry and adverse selection.

Acknowledgement

This research was supported by National Research, Development and Innovation Office - NKFIH, K-138826

6. The resilience of the peer-to-peer market: a global perspective[7]

In recent years different alternative funding methods appeared on the financial market, supported by the Fintech revolution. One of this new form is peer-to-peer lending (P2P), which is a platform-based solution that links lenders and borrowers. There is an extensive literature focusing on various aspects of P2P lending during a time of economic prosperity. The purpose of this paper is to contribute to the understanding of the segment in the case of economic distress indicated by the pandemic. The data set covers 61 countries with a diverse economic background. Unsupervised machine learning approach was applied to classify countries based on their pre-COVID economic profile. Then the clustering performance was tested in terms of the P2P market reaction of each group after the pandemic outbreak. The outcome suggests that the model with three clusters is able to classify countries with respect to their P2P market behavior. In addition to that, the results imply that the growth rate of marketplace lending is in line with the economic instability of the country and the demand excessively increases from regions with weak banking system.

Keywords: Peer-to-peer lending, P2P, financial intermediation, COVID-19, pandemic

JEL Codes: G21, G29

[7] The paper in under consideration in Metroeconomica

6.1 Introduction

In previous years, a disintermediation tendency has started, meaning that the intermediary is eliminated from the transactions. This trend impacted several aspects of the economy, including the segment of consumer lending. The process is further supported by the emerging Fintech revolution and the introduction of digital innovations in the financial sector. As a result, different forms of alternative funding models appeared on the market in addition to conventional bank lending. One of the most extended solutions is peer-to-peer lending, which is an online platform linking lenders and borrowers. The platforms perform brokerage activity, and their initial aim was to provide more favorable conditions for both parties. The first platform was introduced in 2005, and since then the segment of marketplace lending showed a robust expansion, reaching $34.7 billion USD in the retail segment worldwide as of 2020 (CCAF, 2021). The blooming of the market is remarkable, having a strong presence in the leading economies.

The development of the segment was researched from several perspectives in the literature. However, most of the papers had the opportunity to investigate the emerging period of the platforms and utilized data from the time of economic prosperity, leaving space for further research focusing on the economic downturn. The outbreak of COVID-19 had a significant impact on the economy and the financial sector, and the long-term consequences of the crisis are still unpredictable. Currently, there is limited knowledge on the resilience of the segment and the behavior of its customers in case of an unexpected crisis. The purpose of this paper is to contribute to the existing literature by investigating the market's reaction from macro perspective. Unsupervised matching learning algorithm was applied, using pre-COVID economic and financial variables to classify countries regarding their P2P market behavior (lending volume change) after the outbreak. The following research questions are examined in this study:

1. Are pre-COVID economic variables applicable to classify countries according to their market reaction to the pandemic outbreak?
2. How does the P2P market respond to an external shock in the different types of economies?

The paper contributes to the existing literature on marketplace lending from several points of view. The data set covers a wide range of countries from diverse economic backgrounds and development level to gain a more comprehensive view on the market from a global perspective.

This approach enables us to explore trends that cannot be discovered on individual level. Previous studies focusing on the pandemic, used data from one economy only (Anh et al. 2021; Hidajat, 2021; Cumming et al. 2021), therefore the results reflect the specialties of the examined country and cannot be considered as a general conclusion on the whole market. In addition, papers using macroeconomic perspective examined the period of economic prosperity (Rubanov et al. 2019; Oh & Rosenkranz, 2020; Ramcharan & Crowe, 2013), therefore, there is limited information on the experiences during the economic downturn.

The research examines the volume change of platform lending at the country level. Overall, 61 economies are included in the analysis, covering diverse regions. In the first part of the analysis, different economic and financial indicators were selected before the time of the pandemic outbreak. All indicators were investigated in the previous P2P literature and were considered relevant from a platform perspective. Then k-means clustering is performed to identify similar patterns in the dataset and form groups based on that. As a robustness check, the analysis is conducted from the other way around. Countries are grouped based on their reaction to the pandemic and the economic background of each group is compared. In the second part of the research, two-sample t-tests are applied to compare the change in clusters, focusing on the P2P lending volume after the outbreak.

According to the results, three types of economies can be differentiated. The outcome suggests that the model using pre-COVID economic variables is applicable to classify countries regarding their P2P market behavior after the outbreak. In addition to that, the results imply that the growth rate of marketplace lending is in line with the economic instability of the country. In the case of developed countries with strong economic background and advanced financial institutions, the P2P lending volume showed a modest decline. Clients are less exposed in general, and due to the stable, accessible, and extended banking system, it is possible to manage those who are facing financial difficulties. On the other side, there was a robust growth in lending volume in case of the developing regions with low economic performance. The large demand is further boosted by the moderate efficiency of the financial sector. Besides that, the significant level of non-performing loans and the limited access to bank funding suggest that remarkable portion of borrowers are not bank eligible even during the times of economic prosperity. Therefore, it is reasonable that the demand notably rises for alternative funding options after the pandemic outbreak. This finding is in line with Cumming (2021) who highlights that during market turbulences more high-risk applicants attempt to request funding from the marketplaces. The results are further supported by the previous literature from the time of economic prosperity, where analysis was performed at the national level. Jagtiani and Lemieux (2018) found that in the US market, the lending volume increases in areas where the local economy has weaker performance. Polyzos et al. (2021) emphasizes that P2P expansion is in line with higher financial instability, unemployment, and lower GDP. Havrylchyk et al. (2017) explored that the lower level of bank network in the United States supports the spread of platform lending. Overall, the results justified the mentioned trend on global perspective during a time of economic distress.

The structure of the paper is as follows. Section 6.2 presents the related literature on peer-to-peer lending and summarizes the key findings. Section 6.3.1 provides an overview on the dataset, while Section 6.3.2. performs cluster analysis on the pre-pandemic period. Section 6.3.3. examines the differences between the clusters in terms of P2P lending volume. Section 4 concludes and provides possible practical implications.

6.2 Literature review

As mentioned, peer-to-peer lending is a relatively new form of financial intermediary, as the first marketplace was launched sixteen years ago. Despite its modest presence on the financial market, the segment experienced a remarkable development in the previous years and served as a basis for several research studies. The literature covers various aspects of platform lending, including their credit risk assessment (Emekter et al. 2015; Byanjankar et al. 2015) and scoring models (Serrano-Cinca & Gutiérrez-Nieto, 2016; Ye et al. 208; Wang et al. 2018), the determinants of successful funding (Herzestein, 2011; Michels, 2012; Yum et al. 2012) and the reasons behind the spread of the market (Milne & Parboteeah, 2016). However, most of the papers focus on the emerging period and use data sets from the time of economic prosperity, leaving space for further research on the investigation of economic downturn. In addition to that, the few papers that examine the pandemic are based on one economy only, although using a wide range of country-level data could bring further information on the resistance of the market.

As mentioned, this paper contributes to the growing literature on marketplace lending which examines the impact of pandemic on the P2P segment. Cumming et al. (2021) found that platform lending is proven to be more stable during the COVID-19 crisis compared to bank lending in the consumer segment of the United States, which is contrary to intuitive expectations. The marketplace loan volumes declined two months earlier than bank lending volumes; however, the drop in case of banks was two times larger. The resistance of the platforms is further supported by Najaf et al. (2022) who explored that the pandemic resulted in a remarkable change on the market and the form of P2P lending is considered as the most viable funding option for debtors. Nigmonov et al. (2020) analyzed the early stage of the crisis focusing on the secondary market of a leading Estonian platform. The analysis highlights that despite the market turbulences the probability of successful listings increased. In terms of loan performance, Anh et al. (2021) investigated the dataset of LendingClub, a market leader US platform and concluded that the pandemic had a positive impact on the probability of loan default. This result is supported by Hidajat (2021), who explored that loan default rates nearly doubled during the first three quarters of 2020 in the Indonesian market of peer-to-peer lending.

This paper also relates to the substantial part of the literature focusing on peer-to-peer lending from a macroeconomic perspective. Rubanov et al. (2019) examined the spread of alternative finance models with cluster analysis using a data set of 31 economies. According to the results, the volume of P2P consumer lending is the driver of the classification, shaping the market of alternative finance to the largest extent. Polyzos et al. (2021) built a theoretical framework to simulate the impact of platform lending in the economy under different scenarios. The outcome suggests that the extension of marketplace lending is in line with the lowering level of financial stability, the decreasing GDP, and higher unemployment. In terms of empirical studies, Jagtiani and Lemieux (2018) examined account-level US data and found that the presence of peer-to-peer lending is stronger in regions where customers are underbanked, specifically in highly concentrated markets and territories with lower bank branches. This finding is supported by Havrylchyk et al. (2017) who concluded that the segment of the platform lending is more extended in case of regions with lower density of branch network and bank concentration based on data from the two leading US platforms. Ramcharan and Crowe (2013) found that the decline in house prices on the state level has a negative impact on P2P conditions, leading to higher interest rates in the US. Foo et al. (2017) proved that macroeconomic indicators have a significant correlation with platform lending, especially with credit spreads. In addition to economic variables, Oh and Rosenkranz (2020) explored that financial literacy has a positive correlation with market expansion, using a dataset of 62 countries. Another interesting aspect is raised by Shao and Bo (2021), who found on the Chinese market that positive news from the media regarding platforms have a boosting effect on their lending activity.

In general, the literature investigating the impact of COVID-19 on the segment of peer-to-peer lending is still modest, as the outbreak of the pandemic and its influence on the market is a relatively new phenomenon. Current studies have only a few months of observation period, and the long-term effect of this economic shock is still unpredictable, leaving several open questions regarding the future of the segment. In addition to that, macroeconomic indicators proved to be relevant factors based on the literature, in order to better understanding on the spread of the market. Therefore, it is reasonable to further investigate them from a platform perspective.

6.3 Analysis of the data

There is an unfolding discussion on the resistance of marketplace lending during an external shock. To gain a comprehensive view on the behavior of the market, global-level data needs to be examined, covering various regions with different economic background. As a first step, different group of countries are identified that share similar profile, using pre-COVID data. Then, the change in the P2P market volume is analyzed in each group to explore how the market reacts to the shock in the different types of economies.

The first part of this section provides an overview of the data and presents general statistics on the market of peer-to-peer lending. The second part defines the different groups of countries using k-means clustering and performs robustness check. Finally, the third one estimates t-tests to explore the difference between the clusters and investigates the characteristics of each group.

6.3.1 Data

The data set used for the analysis comes from two sources. First, the macroeconomic and financial indicators - reflecting the pre-COVID period of the economy and banking system - were downloaded from IMF and the GlobalEconomy.com, which is a comprehensive datahub collecting several economic variables for a wide range of countries, using multiple official data sources, e.g., World Bank, World Economic Forum, etc. (TheGlobalEconomy.com, 2022). The time horizon of the data set covers the period of 2017-2019. The average of these three years was calculated in the analysis in order to properly capture the recent trends before the pandemic outbreak.

The second source is the Cambridge Centre for Alternative Finance (CCAF), from which the lending volume of the peer-to-peer segment was obtained on the country level. The Centre's aim is to investigate the development of the alternative finance market and to summarize the recent trends in its annual benchmark report. They developed a comprehensive data hub on the segment of alternative finance that is continuously extended (CCAF, 2021). The time horizon of the data includes the years of 2019 and 2020. The percentage change in the volume of marketplace lending was estimated for each county to reflect the behavior of the market during the shock.

The description of each variable is presented in Table 32.

32. Table: Variables included in the analysis

Variable	**Description**	**Time period**
Clustering		
GDP	GDP per capita in million USD	2017-2019
Cost of living	The price level paid by households and non-profit organizations. The value covers all goods and services in an average consumption bucket. The world average is considered to be 100.	2017-2019
Unemployment rate	The portion of the labor force not having employed status. The value is expressed as a percent.	2017-2019
Inflation	The annual percent change in the consumer price index. The value is expressed as a percent.	2017-2019
Nonperforming loans	The percentage of non-performing loans to the total loans in the bank portfolios. The value is expressed as a percent.	2017-2019
Financial Institutions Depth Index	The index covers aggregated information on the following: bank credit to the private sector as a percent of GDP, pension fund assets to GDP, mutual fund assets to GDP, insurance premiums to GDP. The value ranges between 0 and 1. The index is the part of the IMF's Financial Development Index Database.	2017-2019
Financial Institutions Access Index	The index covers aggregated information on the following: bank branches and ATMs per 100	2017-2019

	000 adults. The value ranges between 0 and 1. The index is the part of the IMF's Financial Development Index Database.	
Financial Institutions Efficiency Index	The index covers aggregated information on the following: banking sector net interest margin, lending deposit spread, non-interest income to total income, overhead cost to total assets, return on assets and equity. The value ranges between 0 and 1. The index is the part of the IMF's Financial Development Index Database.	2017-2019
T-tests		
P2P volume	The change in the volume of P2P lending from 2019 to 2020. The value is expressed as a percent.	2019-2020

Source: Author's collection from theglobaleconomy.com, IMF and CCAF, 2017-2020

The selected economic variables are related to the segment of marketplace lending, evidenced by the previous literature. He and Li (2021) analyzed regional platform failures using per capita GDP. Agarwal et al. (2020) examined household expenditures when estimating the performance of P2P claims. The role of unemployment rate was investigated by Yoon et al. (2019) as a default predictor on the P2P portfolio. Nigmonov and Alam (2022) examined the impact of inflation and interest rate on the default risk of the platforms. Jagtiani and Lemieux (2018) and Havrylchyk et al. (2017) analyzed the relationship between the extension of platform lending and bank concentration and branch network. De Roure et al. (2016) focused on the linkage with bank lending volume to identify the complementary or substitutional function of the platforms. Oh and Rosenkranz (2022) explored the impact of different financial development indexes (depth, access and efficiency) on P2P expansion.

The data set used for the analysis covers 61 countries from various regions. All countries were selected in the sample where lending volume and macroeconomic information is available. The dataset covers Europe, North and South America, Asia, Australia, and Africa. CCAF reports separately the volume of consumer lending, business lending and property lending. In the analysis, consumer lending and business lending were included, as they represent the most significant part of the P2P market, approximately 95% (CCAF, 2021). Besides that, property lending is secured with a collateral; therefore, the risk level is considerably lower, making it barely comparable with the other two lending products.

6.3.2 Analysis – Clustering

The purpose of this section is to identify similar patterns in the dataset and form a group of countries based on that. The economic variables used for clustering were all investigated in the P2P literature; therefore, it is assumed that they play significant role in the main trends in marketplace lending. The pre-COVID period is used between 2017 and 2019 to classify the countries. The average of these three years was calculated to properly reflect the state of the economy before the pandemic outbreak. The descriptive statistics of the 8 input variables are presented in Appendix VIII.

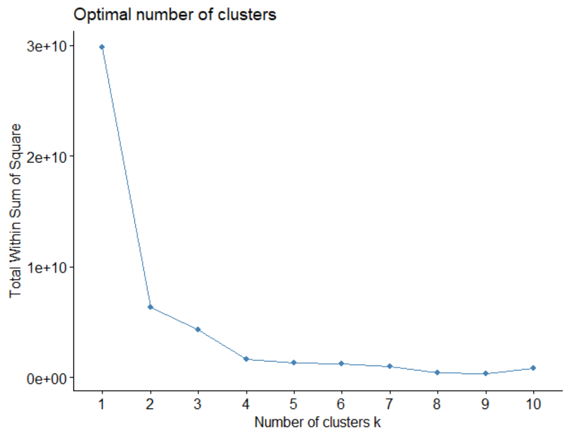
As a next step, clustering is performed to identify countries with a similar economic profile. K-means clustering is chosen which is one of the most commonly applied unsupervised machine learning algorithms. The aim of this method is to form groups where the variance is maximized between the clusters and minimized within them. It is a centroid-based algorithm, meaning that we define the number of clusters and set centroids that are the center of each cluster. Then we assign each point to the nearest centroid based on Euclidean distance. SSE is estimated, which is the sum of the squared distance between the centroid and each data point in order to evaluate the performance of the clustering. The iteration process aims to minimize SEE to define final clusters. The SSE is estimated using the following formula:

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} dist(c_i, x)^2 \qquad (10)$$

where K = number of clusters, c_i = centroid of each cluster; dist = Euclidean distance; x = observation point; C_i = clusters (Steinbach et al., 2005). This method is commonly applied in the literature, especially in the case of country-level analysis, as it derives easily interpretable results (Carrillo & Castillo, 2020; Rubanov et al., 2019; Kigerl, 2016).

In order to define the optimal number of clusters, the elbow function is estimated, where the possible number of clusters (K) are plotted with the related within-cluster sum of square (WSS) values. The aim is to minimize the WSS while selecting a K which retains sufficient information after clustering. When selecting the optimal K, it should be considered whether another additional cluster would lead to a significant reduction in WSS or not. Based on Figure 23, the three number of clusters seems optimal, as there is a steep drop between one and three clusters.

23. Figure: Optimal number of clusters based on the elbow method

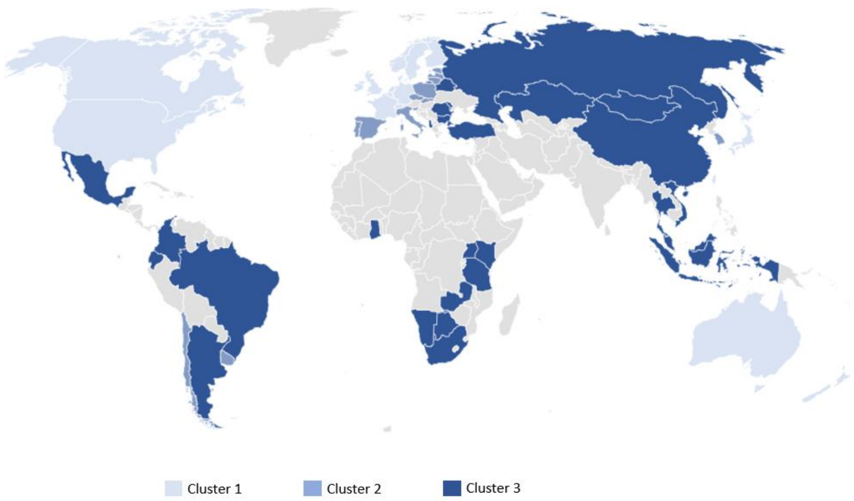


Source: Author's estimation based on theGlobalEconomy.com 2017-2019

To evaluate the separation performance of the clusters, the ratio of the total between sum of squares (BSS) and total sum of squares (TSS) is estimated, which reflects the compactness of the groups. The value of the ratio is 85.5%, meaning that 85.5% of the variance is explained by the clustering algorithm. Iterations were carried out with K = 3 to derive the final clusters, sharing similarities in their economic profile.

To gain a better understanding of the results of the classification, the final clusters are presented visually on Figure 24.

24. Figure: The final groups based on the clustering algorithm



Source: Author's estimation based on theGlobalEconomy.com 2017-2019

To further interpret the results, the mean of each variable is presented in Table 32. As it can be seen, the three clusters differ from each other in terms of all P2P related economic metrics. The 61 countries are classified into the following groups:

- Cluster 1 represents the most developed countries, having a strong economic background with robust per capita GDP, high living cost and low inflation. The financial institutions are accessible, reflecting extended branch and ATM network. The level of bank credit to the private sector is robust and the efficiency of the banking system is high. Furthermore, the banking portfolio has strong performance evidenced by the low NPL level.
- Cluster 2 presents moderately developed countries with restrained per capita GDP and significant unemployment rate. The banking portfolio has a good performance with an extended branch network; however, the depth of the financial institutions is low, leaving space for alternative financing opportunities on the market.
- Cluster 3 groups developing countries, having modest economic performance, evidenced by the relatively low GDP and high inflation rate. The lending activity of the banking sector is low, and the branch network is not extensive, coupled with moderate portfolio performance. The financial indexes imply that the access to credit is limited, leading to a significant portion of underbanked customers.

32. Table: Mean values of each cluster

Cluster	GDP per capita (USD)	Cost of living (100 = world average)	Unemployment rate (%)	Inflation (%)	NPL to all loans (%)	Financial Institutions Access Index	Financial Institutions Depth Index	Financial Institutions Efficiency Index
1	53 323	146	4.8	1.35	1.74	0.58	0.73	0.64
2	22 383	101	7.0	2.39	3.85	0.62	0.39	0.60
3	6 679	65	7.6	5.55	6.32	0.42	0.27	0.57

Source: Author's estimation based on theGlobalEconomy.com 2017-2019

6.3.3 Analysis – Two-sample t-test

In the second part of the analysis, two-sample t-test is performed in order to examine how the P2P market reacts during an external shock, in the case of each cluster. The reaction of the market is captured by the change in the lending volume from 2019 to 2020. The purpose of the test is to examine if there is a significant difference in the mean values across the clusters; therefore, pairwise combinations are analyzed. The results are presented in Table 33.

33. Table: The results of the t-tests

Combination	P-value
1-2	0.1457724
1-3	0.0605798 *
2-3	0.0545826 *

Source: Author's estimation based on CCAF 2020
** indicates significance at 10%.*

The t-tests imply that there is a significant difference between groups in case of clusters 1-3 and 2-3, as the p-value is below 0.1. It suggests that the three clusters based on pre-COVID variables have a good classification performance in terms of P2P market reaction. Table 34 summarizes the change in mean value of the P2P volume in the three clusters.

34. Table: The P2P volume change in each cluster

Cluster	Mean of P2P volume change
1	-9 %
2	20 %
3	865 %

Source: Author's estimation based on CCAF 2020

The trend of marketplace lending clearly differs in the three types of economies. Cluster 1, which represents the most developed countries, showed a moderate decline of -9% in its lending volume. First, it might be explained with the high living standard, as people are less exposed financially in case of a crisis. Furthermore, their stable and extended banking system has the ability to manage those who are facing financial difficulties. In addition to that, countries belonging to this cluster have quite developed P2P market in general, where the regulatory framework is constantly expanding, especially in the United Kingdom, Australia and the United States (Davis, 2016; Magee, 2011; FCA, 2019), providing a burden to finance clients with poor credit background.

In contrast, cluster 2 demonstrated a slight volume increase of 20%. The economic conditions are moderate in this class, people have lower level of income, and their financial background is more fragile. In addition to that, the depth of the financial institutions is limited, which might enhance the attractiveness of alternative funding opportunities, especially in case of an economic bottleneck. Finally, cluster 3 which covers developing countries showed an excessively strong growth rate of 865%. It has to be mentioned that this value is not driven by outliers, but it is indicated by the vast majority of the group. First, the remarkable demand might be implied by the modest economic conditions and the weak banking system. The access to funding and the efficiency of the banking system is underdeveloped, making platform lending more beneficial anyway. Besides that, the high level of non-performing loans suggest that significant portion of borrowers are not bank eligible even during the times of economic prosperity. Therefore, it is reasonable that the demand notably rises for alternative funding options after the pandemic outbreak.

As a robustness check, the same analysis is performed from the other way around. First, the countries are classified into three groups based on their P2P volume change after the pandemic. In case of a single variable, running a cluster analysis is not relevant, therefore the countries are divided based on distinct intervals. Economies with decline higher then 15% are Group 1, between 15% and -15% are Group 2 and above 15% are Group 3. As a next step, the economic and financial background of the groups is compared with the mean values. The results are presented in Table 35.

35. Table: Results of the robustness analysis

Cluster	GDP per capita (USD)	Cost of living (100 = world average)	Unemployment rate (%)	Inflation (%)	NPL to all loans (%)	Financial Institutions Access Index	Financial Institutions Depth Index	Financial Institutions Efficiency Index
1	32 749	112	6,0	2,32	3,80	0,51	0,49	0,62
2	27 594	101	7,5	2,53	3,85	0,58	0,52	0,60
3	14 358	81	6,7	5,62	5,19	0,47	0,33	0,57

Source: Author's estimation based on theGlobalEconomy.com 2017-2019

The results of the robustness test are in line with the analysis. The decline in marketplace lending could be observed in the most developed countries with stable banking system. The solid growth of the market was demonstrated in countries with instable economy and underdeveloped financial sector.

Overall, the results suggest that during economic distress, the growth rate of P2P lending is in line with the economic and financial instability of the country. This finding is supported by previous literature from the time of economic prosperity, which highlights that the extension of marketplace lending is stronger in underdeveloped regions. Jagtiani and Lemieux (2018) found that in the US market, the lending volume increases in areas where the local economy has weaker performance. Polyzos et al. (2021) emphasizes that P2P expansion is in line with higher financial instability, unemployment, and lower GDP. Furthermore, Havrylchyk et al. (2017) explored that the lower level of bank network in the US supports the spread of platform lending. Generally, it implies that in the period of prosperity, P2P lending is more extended in the less developed areas. However, the economic downturn amplified this trend, especially from a macro perspective. Countries, having the most developed P2P market, for example, UK, Australia, Latvia, Lithuania experienced a decline after the pandemic outbreak, while countries with moderate market share, for example, Tanzania, Turkey, Uganda showed robust growth. This finding is in line with Cumming (2021), who draws attention that platforms serve riskier debtors, meaning that the adverse selection cost is higher. This is more visible during market turbulences when more desperate bad debtors attempt to request funding from the marketplaces.

6.4 Conclusion

There is an extensive literature focusing on the segment of P2P lending, however, most of the studies investigated the market during the time of economic prosperity. The purpose of this paper is to examine the behavior of the market in case of an external shock, represented by the pandemic, which is a relatively new perspective, leaving a space for further research. The applied data set covers 61 countries various regions. Unsupervised machine learning algorithm was used to differentiate countries with the same economic profile before the COVID outbreak. Then the clustering performance was tested on the P2P market reaction of each cluster after the spread of the pandemic. The results suggest that the model is able to classify countries according to their P2P market behavior. In addition to that, the outcome implies that the growth rate of marketplace lending is in line with the instability of the economy. There was an excessive demand from developing countries with a weak banking system, where it is likely that most of the clients are not bank eligible even during normal economic conditions. On the other side, in countries with a strong financial background, a moderate decline was observed. The P2P market is also the most developed here, and there is an extensively growing regulatory framework on the segment which serves as a burden to finance high-risk clients. As a potential practical implication, it might be relevant for supervisory authorities to regularly monitor the segment of marketplace lending, especially in developing countries where the market showed a robust expansion in the last few months. These economies are already struggling with high default rates in their banking system, and due to interrelationships on the financial market, the high NPL ratios in the P2P segment might further deepen the lending and liquidity issues.

7. Related publications

Most of the results (and their earlier versions) included in this dissertation were presented in various conferences and published in selected journals. Please find below the details.

Publications:

- Dömötör, B., Illés, F., & Ölvedi, T. (2023). Peer-to-peer lending: Legal loan sharking or altruistic investment? Analyzing platform investments from a credit risk perspective. Journal of International Financial Markets, Institutions and Money, 86, 101801.
- Ölvedi, T. (2022). The Characteristics of Peer-To-Peer Applicants. The Journal of Alternative Investments, 25(2), 66-86.
- Dömötör, B., & Ölvedi, T. (2021). The Financial Intermediary Role of Peer-To-Peer Lenders. In *Innovations in Social Finance* (pp. 391-413). Palgrave Macmillan, Cham.
- Dömötör, B. M., & Ölvedi, T. (2021). A személyközi hitelezés létjogosultsága a pénzügyi közvetítésben. *Közgazdasági Szemle*, *68*(7-8), 773-793.
- Ölvedi, T. (2021). The liquidity aspects of peer-to-peer lending. Studies in Economics and Finance, vol 39./1 August 2021
- Ölvedi, T. (2020). An overview of peer-to-peer lending. *Economy and Finance: English Language Edition of Gazdaság és Pénzügy*, *7*(2), 218-232.

Conferences:

- World Finance & Banking Symposium (2021)
  Ölvedi Tímea: The characteristics of peer-to-peer applicants
- 12th Annual Financial Market Liquidity Conference Budapest (2021)
  Ölvedi Tímea: The characteristics of peer-to-peer applicants
- PRMIA Hungary Chapter Research Conference (2021)
  Ölvedi Tímea: The characteristics of peer-to-peer applicants
- 11th Annual Financial Market Liquidity Conference Budapest (2020)
  Dömötör Barbara - Illés Ferenc - Ölvedi Tímea: Information Processing of Peer-to-peer lending platforms
- Magyar Közgazdaságtudományi Egyesület (MKE) Conference (2020)
  Ölvedi Tímea: Peer-to-peer lending - Likviditási problémák
- PRMIA Hungary Chapter Research Conference (2020)
  Ölvedi Tímea: Peer-to-peer lending, Liquidity aspects

8. References

Agarwal, S., Alok, S., Ghosh, P., & Gupta, S. (2020). Financial inclusion and alternate credit scoring for the millennials: Role of big data and machine learning in fintech. Working Paper.

Ahelegbey, D. F., Giudici, P., & Hadji-Misheva, B. (2019). Latent factor models for credit scoring in P2P systems. *Physica A: Statistical Mechanics and its Applications*, *522*, 112-121.

Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235-251). Academic Press.

Allen, F., & Gale, D. (1995). A welfare comparison of intermediaries and financial markets in Germany and the US. european economic review, 39(2), 179-209.

Anh, N.T. T., Hanh, P. T. M., & Le Thu, V.T. (2021). DEFAULT IN THE US PEER-TO-PEER MARKET WITH COVID-19 PANDEMIC UPDATE: AN EMPIRICAL ANALYSIS FROM A LENDING CLUB PLATFORM. *International Journal of Entrepreneurship*, *25*(7), 1-19.

Arner, D. W., Barberis, J., & Buckley, R. P. (2015). The evolution of Fintech: A new post-crisis paradigm. Geo. J. Int'l L., 47, 1271.

Au, C. H., Tan, B., & Sun, Y. (2020). Developing a P2P lending platform: stages, strategies and platform configurations. *Internet Research*, *30*(4), 1229-1249.

Banai Á. (2014). A bankrendszer szerepe az üzleti ciklusokban. MNB szakmai cikk pp. 4.

Barasinska, N. (2011). *Does gender affect investors' appetite for risk? Evidence from peer-to-peer lending* (No. 1125). DIW Discussion Papers.

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.

Basel Committee on Banking Supervision, (2009). Enhancements to the Basel II Framework, July 2009.

Bassani, G., Marinelli, N., & Vismara, S. (2019). Crowdfunding in healthcare. The Journal of Technology Transfer, 44(4), 1290-1310.

Bazot, G. (2018). Financial consumption and the cost of finance: Measuring financial efficiency in Europe (1950–2007). *Journal of the European Economic Association*, *16*(1), 123-160. https://doi.org/10.1093/jeea/jvx008

Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. Journal of business venturing, 29(5), 585-609.

Berger, S. C., & Gleisner, F. (2009). Emergence of financial intermediaries in electronic markets: The case of online P2P lending. *BuR Business Research Journal*, *2*(1).

Berlinger, E., Lovas, A., & Juhász, P. (2017). State subsidy and moral hazard in corporate financing. Central European Journal of Operations Research, 25, 743-770.

Békés, G., & Kézdi, G. 2021. Data Analysis for Business, Economics, and Policy.

Bhuvaneswari, R., & Segalini, A. (2020). Determining secondary attributes for credit evaluation in P2P lending. arXiv preprint. arXiv:2006.13921.

Bondora. (2019). Background information about Bondora. Available at: https://support.bondora.com/hc/en-us/articles/212499589-Background-information-about-Bondora. Accessed: 05.10.2019.

Bondora (2020). Introducing the Bondora Rating. Available at: https://www.bondora.com/blog/introducing-bondora-rating/. Accessed: 10.05.2021.

Bondora (2020). Public reports. Available at: https://www.bondora.com/en/public-reports. Accessed: 05.10. 2019.

Bondora (2020). What is the secondary market? Available at: https://www.bondora.com/en/secondary-market. Accessed: 05.10.2020.

Bondora (2021). Portfolio Manager Guide. Available at: https://www.bondora.com/en/portfolio-manager. Accessed: 11.10.2021

Bondora (2021). What are the prepayment risks? Available: https://support.bondora.com/en/what-are-the-prepayment-risks

Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, *130*(3), 453-483.

Bureau of Consumer Financial Protection. (2019). The consumer credit card market. Available at: https://files.consumerfinance.gov/f/documents/cfpb_consumer-credit-card-market-report_2019.pdf. Accessed: 08.10.2020.

Bureau of Economic Analysis (2020). Regional data. Available at: https://apps.bea.gov/iTable/iTable.cfm?reqid=70&step=1&acrdn=8. Accessed: 17.10.2020.

Byanjankar, A., Heikkilä, M., & Mezei, J. (2015, December). Predicting credit risk in peer-to-peer lending: A neural network approach. In *2015 IEEE symposium series on computational intelligence* (pp. 719-725). IEEE.

Byanjankar, A., Mezei, J., & Wang, X. (2020). Analyzing peer-to-peer lending secondary market: What determines the successful trade of a loan note? In *World Conference on Information Systems and Technologies* (pp. 471-481). Springer, Cham.

Caglayan, M., Pham, T., Talavera, O., & Xiong, X. (2019). *Asset mispricing in loan secondary market* (No. 19-07). EconPapers.

Caldieraro, F., Zhang, J. Z., Cunha Jr, M., & Shulman, J. D. (2018). Strategic information transmission in peer-to-peer lending markets. *Journal of Marketing*, *82*(2), 42-63.

Cambridge Centre for Alternative Finance (2020): The global alternative finance market benchmarking report. Available at: https://jbs.cam.ac.uk/wp-content/uploads/2020/08/2020-04-22-ccaf-global-alternative-finance-market-benchmarking-report.pdf. Accessed: 15.07.2020

Cambridge Centre for Alternative Finance (2021). The 2nd Global Alternative Finance Market Benchmarking Report. Available at: https://www.jbs.cam.ac.uk/wp-content/uploads/2021/06/ccaf-2021-06-report-2nd-global-alternative-finance-benchmarking-study-report.pdf. Accessed: 09.09.2021

Carrillo-Larco, R. M. & Castillo-Cara, M. (2020). Using country-level variables to classify countries according to the number of confirmed cases of COVID-19: An unsupervised machine learning approach. *Wellcome open research*, *5*.

CCAF (2021). A global alternative finance data depository. Available at: https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/research/a-global-alternative-finance-data-depository/. Accessed: 09.12.2021

Chaffee, E. C., & Rapp, G. C. (2012). Regulating online peer-to-peer lending in the aftermath of Dodd-Frank: In search of an evolving regulatory regime for an evolving industry. *Wash. & Lee L. Rev.*, *69*, 485.

Chen, D., Lai, F., & Lin, Z. (2014). A trust model for online peer-to-peer lending: a lender's perspective. *Information Technology and Management*, *15*(4), 239-254.

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.

Chen, X., Qin, Y., Xiao, H., & Zhang, Y. (2019). Microfinancing and Home-purchase Restrictions: Evidence from China's Online Peer-to-Peer Lending. *Available at SSRN 3429030*.

Cole, R. A., Cumming, D. J., & Taylor, J. (2019). Does FinTech compete with or complement bank finance?. *Available at SSRN 3302975*.

Cornaggia, J., Wolfe, B., & Yoo, W. (2018). Crowding out banks: Credit substitution by peer-to-peer lending. *Available at SSRN 3000593*.

Croux, C., Jagtiani, J., Korivi, T., & Vulanovic, M. (2020). Important factors determining Fintech loan default: Evidence from a lendingclub consumer platform. *Journal of Economic Behavior & Organization*, *173*, 270-296. https://doi.org/10.1016/j.jebo.2020.03.016

Csóka, P., & Jean-Jacques Herings, P. (2018). Decentralized clearing in financial networks. *Management Science*, *64*(10), 4681-4699.

Cumming, D. J., Leboeuf, G., & Schwienbacher, A. (2020). Crowdfunding models: Keep-it-all vs. all-or-nothing. *Financial Management*, *49*(2), 331-360. https://doi.org/10.2139/ssrn.2447567

Cumming, D. J., Martinez-Salgueiro, A., Reardon, R. S. & Sewaid, A. (2021). COVID-19 bust, policy response, and rebound: equity crowdfunding and P2P versus banks. *The Journal of Technology Transfer*, 1-22.

Das, Sanjiv R. "The future of fintech." *Financial Management* 48, no. 4 (2019): 981-1007. https://doi.org/10.1111/fima.12297

Davis, K. (2016). Peer-to-peer lending: structures, risks and regulation. *JASSA*, (3), 37-44.

De Roure, C., Pelizzon, L., & Tasca, P. (2016). How does P2P lending fit into the consumer credit market?

De Roure, C., Pelizzon, L., & Thakor, A. V. (2021). P2P lenders versus banks: Cream skimming or bottom fishing?.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies*, *51*(3), 393-414.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, *91*(3), 401-419.

Dietrich, A., & Wernli, R. (2016). What drives the interest rates in the P2P consumer lending market? Empirical evidence from Switzerland. *SSRN Electronic Journal*, *10*.

Di Maggio, M., Kermani, A., Keys, B. J., Piskorski, T., Ramcharan, R., Seru, A., & Yao, V. (2017). Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging. *American Economic Review*, *107*(11), 3550-88.

Douglas, W. O., & Bates, G. E. (1933). The Federal Securities Act of 1933. Yale LJ, 43, 171.

Dowd, K. (2007). *Measuring market risk*. John Wiley & Sons.

Dömötör, B. M., & Ölvedi, T. (2021). A személyközi hitelezés létjogosultsága a pénzügyi közvetítésben. *Közgazdasági Szemle*, *68*(7-8), 773-793.

Dömötör, B., & Ölvedi, T. (2021). The Financial Intermediary Role of Peer-To-Peer Lenders. In *Innovations in Social Finance* (pp. 391-413). Palgrave Macmillan, Cham.

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.

Duffie, D., & Singleton, K. J. (2012). Credit risk: pricing, measurement, and management. Princeton university press.

Economic Inclusion 2019 (2019). Household survey results. Available at: https://economicinclusion.gov/surveys/2019household/. Accessed: 17.10.2020.

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.

Euro Area Statistics (2020). Bank interest rates - Deposit. Downloaded: 2020.12.02. Available: https://www.euro-area-statistics.org/bank-interest-rates-deposits?cr=eur&lg=en&page=0&charts=M..B.L22.L.R.A.2250.EUR.N+M..B.L22.F.R.A.2250.EUR.N+M..B.L22.K.R.A.2250.EUR.N&template=1

European Central Bank (2020). Statistics. Downloaded: 2020.12.02. Available: https://www.ecb.europa.eu/stats/html/index.en.html

European Commission (2020). Regulation on European Crowdfunding Service Providers (ECSP) for business. Available at: https://ec.europa.eu/info/business-economy-euro/growth-and-investment/financing-investment/crowdfunding_en. Accessed: 2021.04.15.

Experian (2020). What is a fair credit score? Available at: https://www.experian.com/blogs/ask-experian/what-is-a-fair-credit-score/. Accessed: 01.12.2020.

Experian (2020). What is a good credit score? Available at: https://www.experian.com/blogs/ask-experian/credit-education/score-basics/what-is-a-good-credit-score/. Accessed: 01.12.2020.

Experian (2020). What is the difference between FICO score and credit score? Available at: https://www.experian.com/blogs/ask-experian/what-is-the-difference-between-fico-score-and-credit-score/. Accessed: 03.12.2020.

FCA (2019) Loan-based (‘peer-to-peer’) and investment-based crowdfunding platforms: Feedback to CP18/20 and final rules. Available at :https://www.fca.org.uk/publication/policy/ps19-14.pdf. Accessed: 24.03.2021

FED (2021). Consumer Credit - G.19. Available at: https://www.federalreserve.gov/releases/g19/current/. Accessed: 14.10.2021

Federal Deposit Insurance Corporation (2020). Deposit market share reports—Summary of deposits. Available at: https://www7.fdic.gov/sod/sodMarketBank.asp?barItem=2. Accessed: 15.10.2020.

Federal Reserve (2020). State-level debt-to-income ratio, 1999 - 2020. Available at: https://www.federalreserve.gov/releases/z1/dataviz/household_debt/state/map/#year:2019. Accessed: 11.10.2020.

Federal Reserve Bank of New York (2021). Center for microeconomic data. Available at: https://www.newyorkfed.org/microeconomics/databank. Accessed: 13.10.2020.

FICO (2020). Learn About The FICO® Score and its Long History. Available at: https://www.fico.com/25years/. Accessed: 05.12.2020.

Foo, J., Lim, L. H., & Wong, K. S. W. (2017). Macroeconomics and fintech: Uncovering latent macroeconomic effects on peer-to-peer lending. arXiv preprint arXiv:1710.11283.

Frame, W. S., Wall, L. D., & White, L. J. (2018). Technological change and financial innovation in banking: Some implications for fintech.Freixas, X., & Rochet, J. C. (2008). *Microeconomics of banking*. MIT press.

Freedman, S., & Jin, G. Z. (2008). Do social networks solve information problems for peer-to-peer lending? Evidence from prosper.com. *NET Institute Working Paper*, No. 08-43.

Freedman, S., & Jin, G. Z. (2017). The information value of online social networks: lessons from peer-to-peer lending. *International Journal of Industrial Organization*, *51*, 185-222.

Freixas, X., & Rochet, J. C. (2008). *Microeconomics of Banking*. MIT Press.

Gavurova, B., Dujcak, M., Kovac, V., & Kotásková, A. (2018). Determinants of successful loan application at peer-to-peer lending market. *Economics & Sociology*, *11*(1), 85-99.

Ge, R., Feng, J., Gu, B., & Zhang, P. (2017). Predicting and deterring default with social media information in peer-to-peer lending. *Journal of Management Information Systems*, *34*(2), 401-424.

Gerschenkron, A. (1962). Economic backwardness in historical perspective (1962). *The Political Economy Reader: Markets as Institutions*, 211-228.

Goldstein, I., Jiang, W., & Karolyi, G. A. (2019). To FinTech and beyond. *The Review of Financial Studies*, *32*(5), 1647-1661.

Gonzalez, L., & Loureiro, Y. K. (2014). When can a photo increase credit? The impact of lender and borrower profiles on online peer-to-peer loans. *Journal of Behavioral and Experimental Finance*, *2*, 44-58.

Gosztonyi, M., & Havran, D. (2021). Highways to Hell? Paths Towards the Formal Financial Exclusion: Empirical Lessons of the Households from Northern Hungary. *The European Journal of Development Research*, 1-34. https://doi.org/10.1057/s41287-021-00434-9

Greenwood, J., & Jovanovic, B. (1990). Financial development, growth, and the distribution of income. *Journal of political Economy*, *98*(5, Part 1), 1076-1107.

Han, L., Xiao, J. J., & Su, Z. (2019). Financing knowledge, risk attitude and P2P borrowing in China. *International Journal of Consumer Studies*, *43*(2), 166-177.

Hanley, J. A. (1989). Receiver operating characteristic (ROC) methodology: The state-of-the-art. *Crit Rev Diagn Imaging*, *29*(3), 307-335.

Harkness, S. K. 2016. Discrimination in lending markets: Status and the intersections of gender and race. *Social Psychology Quarterly*, *79*(1), 81-93.

Harvey, S. (2018). Lending club's note trading platform facade: An examination of peer-to-peer (P2P) lending secondary market inefficiency. University of Dayton. Ohio. https://ecommons.udayton.edu/uhp_theses/199

Havrylchyk, O., & Verdier, M. (2018). The financial intermediation role of the P2P lending platforms. *Comparative Economic Studies*, 60(1), 115-130. https://doi.org/10.1057/s41294-017-0045-1

Havrylchyk, O., Mariotto, C., Rahim, T., & Verdier, M. (2017). What drives the expansion of the peer-to-peer lending? http://dx. doi. org/10.2139/ssrn, 2841316.

Havrylchyk, O., Mariotto, C., Rahim, T., & Verdier, M. (2020). The Expansion of Peer-to-Peer Lending. *Review of Network Economics*, *19*(3), 145-187.

He, Q., & Li, X. (2021). The failure of Chinese peer-to-peer lending platforms: finance and politics. *Journal of Corporate Finance*, *66*, 101852.

Hellwig, M. (1991). Banking, financial intermediation and corporate finance. *European financial integration*, *35*, 63.

Hemer, J. (2011). *A snapshot on crowdfunding* (No. R2/2011). Arbeitspapiere Unternehmen und Region.

Herzenstein, M., Andrews, R. L., Dholakia, U. M., & Lyandres, E. (2008). The democratization of personal consumer loans? Determinants of success in online peer-to-peer lending communities. *Boston University School of Management Research Paper*, *14*(6), 1-36.

Herzenstein, M., Sonenshein, S., & Dholakia, U. M. (2011). Tell me a good story and I may lend you money: The role of narratives in peer-to-peer lending decisions. *Journal of Marketing Research*, *48*(SPL), S138-S149.

Hidajat, T. (2021). The Relationship Between Debt Literacy and Peer-To-Peer Lending: A Case Study in Indonesia. The Journal of Asian Finance, Economics and Business, 8(5), 403-411.

Hidajat, T. (2021, March). Pandemic, Lender Risk, and Borrower Bargaining Power. In the 3rd *International Conference of Banking, Accounting, Management and Economics (ICOBAME 2020)* (pp. 43-45). Atlantis Press.

Hildebrand, T., Puri, M., & Rocholl, J. (2010). Skin in the game: Evidence from the online social lending market. *Group*.

Holmström, B., & Tirole, J. (1998). Private and public supply of liquidity. *Journal of political Economy*, *106*(1), 1-40.

IMF. (2023). Financial Development Index Database. Available at: https://data.imf.org/?sk=f8032e80-b36c-43b1-ac26-493c5b1cd33b. Accessed: 2023.09.26.

Jagtiani, J., and Lemieux, C. (2018). Do fintech lenders penetrate areas that are underserved by traditional banks? *Journal of Economics and Business*, June 2018.

Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: evidence from the LendingClub consumer platform. *Financial Management*, *48*(4), 1009-1029.

Johnson, R. A. & Wichern, D. W. (2014). Multivariate statistical analysis (Vol. 6). London, UK:: Pearson.

Jorgensen, T. (2018). Peer-to-Peer Lending-A New Digital Intermediary, New Legal Challenges. *NJCL*, 231.

Käfer, B. (2018). Peer-to-Peer lending–a (financial stability) risk perspective. *Review of Economics*, *69*(1), 1-25.

Kaiser Family Foundation (2020). Poverty rate by race/ethnicity. Available at: https://www.kff.org/other/state-indicator/poverty-rate-by-raceethnicity. Accessed: 16.10.2020.

Kigerl, A. (2016). Cyber Crime Nation Typologies: K-Means Clustering of Countries Based on Cyber Crime Rates. International Journal of Cyber Criminology, 10(2).

Kirby, E., & Worner, S. (2014). Crowd-funding: An infant industry growing fast. *IOSCO Research Department*, 1-62.

Klafft, M. (2008, July). Online peer-to-peer lending: a lenders' perspective. In *Proceedings of the international conference on E-learning, E-business, enterprise information systems, and E-government, EEE* (pp. 371-375).

Kuwabara, K., & Thébaud, S. (2017). When beauty doesn't pay: Gender and beauty biases in a peer-to-peer loan market. Social Forces, 95(4), 1371-1398.

Larrimore, L., Jiang, C., Larrimore, J., Markowitz, D., & Gorski, S. (2011). Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. *Journal of Applied Communication Research, 39*(1), 19-37.

LendingClub (2021). Important Updates to the LendingClub Notes Platform. Available at: https://help.lendingclub.com/hc/en-us/articles/360050574891-Important-Updates-to-the-LendingClub-Notes-Platform. Accessed: 12.10.2021

LendingClub (2021). Lending Club Prices Initial Public Offering. Available at: https://s24.q4cdn.com/758918714/files/doc_news/2014/12/1/Lending-Club-Prices-Initial-Public-Offering.pdf. Accessed: 12.10.2021

Lenz, R. (2016). Peer-to-peer lending: Opportunities and risks. *European Journal of Risk Regulation*, *7*(4), 688-700.

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.

Li, Y., Ning, Y., Liu, R., Wu, Y., & Hui Wang, W. (2020). Fairness of classification using users' social relationships in online peer-to-peer lending. In Companion Proceedings of the Web Conference 2020 (pp. 733-742).

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. *Western Finance Association 2009 Annual Meeting Paper.*

Liu, Z., Shang, J., Wu, S. Y., & Chen, P. Y. (2020). Social collateral, soft information and online peer-to-peer lending: A theoretical model. European Journal of Operational Research, 281(2), 428-438.

Macchiavello, E. (2014). Peer-to-Peer lending and the democratization of credit markets: Another financial innovation puzzling regulators. Colum. J. Eur. L., 21, 521

Magee, J. R. (2011). Peer-to-peer lending in the United States: surviving after Dodd-Frank. *NC Banking Inst.*, *15*, 139.

Maskara, P. K., Kuvvet, E., & Chen, G. (2021). The role of P2P platforms in enhancing financial inclusion in the United States: An analysis of peer-to-peer lending across the rural–urban divide. *Financial Management*. https://doi.org/10.1111/fima.12341

Merton, R. C., & Thakor, R. T. (2019). Customers and investors: a framework for understanding the evolution of financial institutions. *Journal of Financial Intermediation*, 39, 4-18.

Michels, J. (2012). Do unverifiable disclosures matter? Evidence from peer-to-peer lending. The Accounting Review, 87(4), 1385-1413.

Milne, A., & Parboteeah, P. (2016). The business models and economics of peer-to-peer lending. *European Credit Research Institute.*

Mintos (2020). Mintos applies for European investment firm and electronic money institution licenses. Available:https://www.mintos.com/blog/mintos-applies-for-european-investment-firm-and-electronic-money-institution-licenses/. Accessed: 13.03.2021

Molnár, J. (2018). What does financial intermediation theory tell us about fintechs?. *Vezetéstudomány-Budapest Management Review*, *49*(5), 38-46.

MoneyWeek (2020): Market crash exposes the risks in P2P lending sites. Available at: https://moneyweek.com/investments/alternative-finance/601218/market-crash-exposes-the-risks-in-p2p-lending-sites. Accessed: 18.7.2020

Morse, A. (2015). Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending. *Annual Review of Financial Economics*, 7, 463-482.

Najaf, K. Subramaniam, R. K., & Atayah, O. F. (2022). Understanding the implications of lending FinTech Peer-to-Peer (P2P) lending during the COVID-19 pandemic. Journal of Sustainable Finance & Investment, 12(1), 87-102.

Nikolaou, K. (2009). Liquidity (risk) concepts: definitions and interactions.

Nigmonov, A., Shams, S., & Alam, K. (2020). Born in Crisis: Early impact of COVID-19 Pandemic on P2P lending market. *Available at SSRN 3721406*.

Nigmonov, A., Shams, S., & Alam, K. (2022). Macroeconomic determinants of loan defaults: Evidence from the US peer-to-peer lending market. *Research in International Business and Finance*, *59*, 101516.

Oh, E. Y., & Rosenkranz, P. (2020). Determinants of peer-to-peer lending expansion: The roles of financial development and financial literacy. *Asian Development Bank Economics Working Paper Series*, (613).

Oh, E. Y., & Rosenkranz, P. (2022). Determinants of peer-to-peer lending expansion: The roles of financial development and financial literacy. The Journal of FinTech, 2250001.

P2P-Banking (2016): Prosper closes secondary market. Available at: https://www.p2p-banking.com/services/prosper-prosper-closes-secondary-market/. Accessed: 18.07.2020

P2PMarketData (2021). P2P lending platforms of the world. Available at: https://p2pmarketdata.com/p2p-lending-platforms-of-the-world. Accessed: 15.7.2020.

Philippon, T. (2014). Has the US finance industry become less efficient? On the theory and measurement of financial intermediation. *American Economic Review*, *105*(4), 1408-38.

Polasik, M., Huterska, A., Iftikhar, R., & Mikula, Š. (2020). The impact of Payment Services Directive 2 on the PayTech sector development in Europe. *Journal of Economic Behavior & Organization*, *178*. https://doi.org/10.1016/j.jebo.2020.07.010

Polyzos, S., Samitas, A., & Rubbaniy, G. (2021). The perfect bail-in: Financing without banks using Peer-To-Peer Lending. Available at SSRN 3916661.

Pope, D. G., & Sydnor, J. R. (2011). What's in a picture? Evidence of discrimination from Prosper.com. *Journal of Human resources*, *46*(1), 53-92.

Pringle, J. J. (1975). Bank capital and the performance of banks as financial intermediaries: comment. *journal of Money, Credit and Banking*, *7*(4), 545-550.

Prosper (2020). How does Prosper verify a borrower's income or employment? Available at: https://prosper.zendesk.com/hc/en-us/articles/208500386-How-does-Prosper-verify-a-borrower-s-income-or-employment-. Accessed: 05.10.2020.

Prosper (2020). Legal compliance. Available at: https://www.prosper.com/plp/legal/compliance/. Accessed: 08.10.2020.

Prosper (2020). Latest SEC filings. Available at: https://www.prosper.com/prospectus. Accessed: 03.10.2020.

Prosper (2020). Searching through Prosper loan listings – listings API. Available at: https://developers.prosper.com/docs/investor/listings-api/. Accessed: 01.10.2020.

Prosper (2020). What are the minimum criteria to borrow on Prosper? Available at: https://prosper.zendesk.com/hc/en-us/articles/210013963-What-are-the-minimum-criteria-to-borrow-on-Prosper-. Accessed: 05.12.2020.

Prosper (2021). Our Story. Available at: https://www.prosper.com/about. Accessed: 09.08.2021.

Prosper (2021). Searching through Prosper loan listings – listings API. Available at: https://developers.prosper.com/docs/investor/listings-api/. Accessed: 15.11.2021.

Prosper (2021) Available at: https://prosper.zendesk.com/hc/en-us/articles/210013913-What-is-a-loan-through-Prosper-and-who-originates-it-.Accessed: 12.10.2021

Prosper (2023). What happens if a borrower misses a payment? Available at: What happens if a borrower misses a payment? – Help is on the way. (prosper.com) Accessed: 30.09.2023.

Ramcharan, R., & Crowe, C. (2013). The impact of house prices on consumer credit: evidence from an internet bank. Journal of Money, Credit and Banking, 45(6), 1085-1115.

Ratesetter (2021). RateSetter announces that Metro Bank is purchasing the loan portfolio. Available at: https://www.ratesetter.com/blog/ratesetter-announces-that-metro-bank-is-purchasing-the-loan-portfolio. Accessed: 12.10.2021

Ravina, E. (2008). Love & loans: The effect of beauty and personal characteristics in credit markets. *Journal of Finance*.

Ravina, E. (2019). Love & loans: The effect of beauty and personal characteristics in credit markets. *Available at SSRN 1107307*.

Reher, M. (2014). Do de novo secondary markets affect primary market interest rates? A Case Study of Peer-to-Peer Lending. *Michigan Journal of Business, 7*(1).

Rogers, C., & Clarke, C. (2016). Mainstreaming social finance: The regulation of the peer-to-peer lending marketplace in the United Kingdom. The British Journal of Politics and International Relations, 18(4), 930-945.

Rubanov, P. M., Vasylieva, T. A., Lieonov, S. V., & Pokhylko, S. V. (2019). Cluster analysis of development of alternative finance models depending on the regional affiliation of countries.

SEC. (2008). Order instituting cease-and-desist proceedings pursuant to Section 8a 657 of the Securities Act of 1933, making findings, and imposing a cease-and-desist 658 order. Available at: https://www.sec.gov/litigation/admin/2008/33-8984.pdf. Accessed: 10.12.2021

Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. Decision Support Systems, 89, 113-122.

Scholfeld E. (2008). SEC Outlines Its Reasoning For Shutting Down P2P Lender Prosper. Available at: https://techcrunch.com/2008/11/26/sec-outlines-its-reasoning-for-shutting-down-p2p-lender-prosper/?guccounter=1. Accessed: 10.12.2021

Schwarcz, S. L. (2008). Systemic risk. Geo. Lj, 97, 193.

Shao, S., & Bo, H. (2021). Behavioural aspects of China's P2P lending. *The European Journal of Finance*, 1-16.

Steinbach, M., & Kumar, V. (2005). Cluster analysis: Basic concepts and algorithms. Ch. 8 in *Introduction to data mining, 1st edn. Pearson Addison Wesley*.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American economic review*, *71*(3), 393-410.

Szűcs, B., & Váradi, K. (2014). Measuring and managing liquidity risk in the Hungarian practice. *Society and Economy*, *36*(4), 543-563.

Tan, P., Steinbach, M. and Kumar, V. (2005) Cluster Analysis: Basic concepts and algorithms. In: Introduction to Data Mining, Addison-Wesley, Boston, MA.

Tang, H. (2019). Peer-to-peer lenders versus banks: substitutes or complements?. *The Review of Financial Studies*, *32*(5), 1900-1938.

Tirole, J. (2010). *The theory of corporate finance*. Princeton University Press.

Thakor, A. V. (2020). Fintech and banking: What do we know? *Journal of Financial Intermediation*, 41, 100833.

Theglobaleconomy.com. Latest data (2021). Available at https://www.theglobaleconomy.com/updates_history.php. Accessed: 11.14.2021

TransUnion (2021). What is a good Credit Score Range with TransUnion? Available at: https://www.transunion.com/article/what-is-a-good-credit-score. Accessed: 16.08.2021.

Twino. (2021). TWINO Secures Investment Brokerage License from the Latvian Financial Regulator. Available at: https://blog.twino.eu/twino-secures-investment-brokerage-license-from-the-latvian-financial-regulator/. Accessed: 24.03.2021

United States Census Bureau (2020). American community survey. Available at: https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/2014/. Accessed: 19.10.2020.

Van Liebergen, B. (2017). Machine learning: A revolution in risk management and compliance? *Journal of Financial Transformation*, *45*, 60-67.

Walter, G. (2019). Vállalatfinanszírozás a gyakorlatban – Lehetőségek és döntések a magyar piacon. Alinea Kiadó

Wang, H., Chen, K., Zhu, W., & Song, Z. (2015). A process model on P2P lending. *Financial Innovation*, *1*(1), 1-8.

Wang, Z., Jiang, C., Ding, Y., Lyu, X., & Liu, Y. (2018). A novel behavioral scoring model to estimate the probability of default over time in peer-to-peer lending. Electronic Commerce Research and Applications, 27, 74-82.

Wang, Y., & Ni, X. S. (2020, April). Risk Prediction of Peer-to-Peer Lending Market by a LSTM Model with Macroeconomic Factor. In *Proceedings of the 2020 ACM Southeast Conference* (pp. 181-187).

Weiss, G. N., Pelger, K., & Horsch, A. (2010). Mitigating adverse selection in p2p lending–Empirical evidence from prosper. com. *Available at SSRN 1650774*.

Wolfe, B., & Yoo, W. (2017). Crowding out banks: Credit substitution by peer-to-peer lending. *SSRN Electronic Journal. https://doi. org/10.2139/ssrn*, *3000593*.

Wooldridge, J. M. (2012). Introductory econometrics: A modern approach: Cengage Learning. *A Figures*, *18*.

Ye, X., Dong, L. A., & Ma, D. (2018). Loan evaluation in P2P lending based on a random forest optimized by a genetic algorithm with a profit score. Electronic Commerce Research and Applications, 32, 23-36.

Yellen, J. (2013). Interconnectedness and systemic risk: Lessons from the financial crisis and policy implications. *Board of Governors of the Federal Reserve System, Washington, DC*.

Yoon, Y. Li, Y., & Feng, Y. (2019). Factors affecting platform default risk in online peer-to-peer (P2P) lending business: An empirical study using Chinese online P2P platform data. *Electronic Commerce Research*, *19*(1), 131-158.

Yum, H., Lee, B., & Chae, M. (2012). From the wisdom of crowds to my own judgment on microfinance through online peer-to-peer lending platforms. Electronic Commerce Research and Applications, 11(5), 469-483.

Zopa. (2020). Zopa launches new digital bank, offering consumers a compelling alternative when they need it most. Available at: https://www.zopa.com/blog/article/zopa-launches-new-digital-bank-offering-consumers-a-compelling-alternative-when-they-need-it-most. Accessed: 10.12.2021

9. Appendix

9.1 Appendix I of Chapter 3

25. Figure: The snapshot of a Prosper listing report as of 2020.12.02

Borrower Payment Dependent Notes Series 11882331

The following information pertains to the borrower loan being requested, that corresponds to the series of Notes to be issued upon the funding of the borrower loan, in the event the listing receives commitments to purchase Notes in an amount that equals or exceeds the minimum amount required for the loan to fund.

Amount:	**$4,000.00**	Prosper Rating:	**B**	Listing Duration:	**14 days**
Term:	**36 months**	Initial Status:	**F**		
Investor yield:	**15.74%**	Borrower rate/APR:	**16.74% / 20.45%**	Monthly payment:	**$142.09**
Investor servicing fee:	**1.00%**	Historical Return*:	**3.73% - 7.62%**		

* Historical Return Range for Single or Multiple Prosper Ratings (e.g., Loan Listing, Auto Invest) Represents the Historical Return Range for the Prosper Rating or mix of Prosper Ratings (as applicable) as of November 30, 2020. Historical Returns are based on actual payments (other than principal) received by the investor net of fees and losses (including from charged-off loans). To be included in the historical return ("Historical Return") calculation, the loan must have originated (a) on or after July 1, 2009, and (b) at least 12 months prior to the calculation date. We calculate the Historical Return for loans originated through Prosper as follows. First, loans are separated into distinct "Groups" based on the specific month and year in which they were originated and their Prosper Rating or mix of Prosper Ratings (as applicable) at origination. For each Group, we calculate: (a) the sum of the interest paid, plus late fees, minus servicing fees, minus collection fees, in each case on active loans, plus net recoveries on charged-off or defaulted loans, plus net debt sale proceeds on sold loans, minus gross principal losses; divided by (b) the sum of the principal balances outstanding on active loans at the end of each day since origination. We then annualize the result to get the "Historical Return" for the Group. Once this calculation is performed for every Group, we compute the 10th and 90th cumulative-outstanding-principal-dollar-weighted percentiles of the Historical Returns of the Groups within each Prosper Rating or mix of Prosper Ratings (as applicable) to get the "Historical Return Range" of the relevant Prosper Rating or mix of Prosper Ratings. For purposes of this calculation, "active" means loans that are current in payments or delinquent less than 120 days; loans that have paid off, charged-off or are in default are not considered active. The Historical Return calculation (a) is updated monthly; and (b) excludes the impact of servicing related non-cash corrective adjustments that may modify the outstanding balance or status of a borrower loan. The actual return on any Note depends on the prepayment and delinquency pattern of the loan underlying each Note, which is highly uncertain. Individual results may vary. Historical performance is no guarantee of future results and the information presented is not intended to be investment advice or a guarantee about the performance of any Note.

Borrower's Credit Profile

Prosper score (1-11):	**6**	First credit line:	**Dec-2002**	Debt/Income ratio:	**12%**
TU FICO range:	**640-659 (Dec-2020)**	Inquiries last 6m:	**0**	Employment status:	**Employed**
Current delinquencies:	**0**	Current / open credit lines:	**6 / 6**	Length of status:	**20y 3m**
Amount delinquent:	**$0**	Total credit lines:	**14**	Occupation:	**Teacher**
Public records last 24m/10y:	**0 / 1**	Revolving credit balance:	**$2,616**	Stated income:	**$75,000-99,999**
Delinquencies in last 7y:	**5**	Bankcard utilization:	**57%**		
		Has Mortgage:	**No**	Borrower's state:	**Pennsylvania**

Credit information was obtained from borrower's credit report and displayed without having been verified. Employment and income was provided by borrower and displayed without having been verified.

"Not available" indicates the information could not be obtained by the credit bureau.

Description

Debt Consolidation

Source: prosper.com, 2020

9.2 Appendix II of Chapter 3

36. Table: Descriptive statistics of the local economic variables

Variable	Mean	Stdev	Min	Max	No. of observations
FICO score	704.84	16.13	666.00	739.00	336
Total debt	48200.03	12275.61	26770.00	90220.00	336
Mortgage debt	32553.04	11054.86	14340.00	67370.00	336
Auto debt	4401.19	773.12	2420.00	7000.00	336
Credit card debt	2963.45	535.16	1650.00	4440.00	336
Credit card delinquency	0.07	0.01	0.04	0.13	336
Mortgage delinquency	0.01	0.01	0.00	0.07	336
Race_White	0.78	0.13	0.31	0.97	336
Race_Black	0.13	0.11	0.01	0.51	336
Education_high school	0.89	0.03	0.82	0.94	336
Education_bsc	0.31	0.06	0.19	0.59	336
DTI	1.45	0.30	0.39	2.18	336
Poverty rate	0.13	0.03	0.07	0.22	336
Median sales price	217844.35	100143.22	100200.00	615300.00	336
Mean household income	79849.60	14437.17	54881.00	127890.00	336

Student loan	5165.36	1373.36	3010.00	13600.00	336
Student loan delinquency	0.11	0.03	0.04	0.18	336
Population	6506874.10	7231304.74	575251.00	39538223.00	336
Man	0.49	0.01	0.47	0.52	336
Race_Latino	0.12	0.10	0.01	0.49	336
Employed	0.63	0.04	0.50	0.72	336
Median age	38.00	2.33	29.90	44.70	336
Total expenditures	270117.23	318392.63	23022.00	1947590.00	336
Total expenditures per cap	40709.18	6560.67	27978.00	65352.00	336
Household expenditure	26284.82	5182.62	16564.00	48706.00	336
Health expenditure	7006.01	1211.36	4547.00	11337.00	336
Financial services expenditure	3267.82	913.20	1764.00	7101.00	336
Number of branches	1787.77	1607.05	118.00	7264.00	336

Source: Author's estimation based on Prosper listing data 2014-2020.

9.3 Appendix III of Chapter 3

37. Table: Descriptive statistics of the variables related to the Prosper applicants

Variable	Mean	Std.Dev.	Min	Max	No. of observations
Amount	12 287.32	7 544.91	2 000.00	40 000.00	100 000
Current_deliquencies	0.20	0.79	0	36.00	100 000
Amount_deliquent	65.62	2 029.78	0	360 362.00	100 000
Deliquencies_last_7y	2.61	7.16	0	99.00	100 000
Total_credit_lines	24.22	12.81	2.00	147.00	100 000
Revolving_credit_balance	17 021.96	27 020.13	0	1 135 216.00	100 000
Bankcard_utilization	0.52	0.28	0	1.61	100 000
Monthly_payment	391.22	229.28	38.70	1 392.33	100 000
DTI	0.27	0.13	0	3.65	100 000
interest_average	18.46	8.00	5.48	33.96	100 000
fico_average	702.85	41.65	649.50	849.50	100 000

Source: Author's estimation based on Prosper listing data 2014-2020.

9.4 Appendix IV of Chapter 3

38. Table: Correlation matrix of the variables related to the Prosper applicants

	Revolving_credit_balance	Current_deliquencies	fico_average	Amount	Amount_deliquent	Deliquencies_last_7y	Total_credit_lines	Bankcard_utilization	Monthly_payment	DTI	interest_average
Revolving_credit_balance	1.000	-0.086	0.011	0.317	-0.055	-0.148	0.374	0.409	0.341	0.310	0.028
Current_deliquencies	-0.086	1.000	-0.217	-0.048	0.699	0.198	0.066	-0.013	-0.033	-0.050	0.098

fico_average	0.011	-0.217	1.000	0.219	-0.138	-0.400	0.083	-0.491	0.103	0.054	-0.609
Amount	0.317	-0.048	0.219	1.000	-0.024	-0.091	0.206	-0.047	0.948	0.169	-0.256
Amount_deliquent	-0.055	0.699	-0.138	-0.024	1.000	0.036	0.038	-0.014	-0.022	-0.020	0.044
Deliquencies_last_7y	-0.148	0.198	-0.400	-0.091	0.036	1.000	0.171	0.069	-0.047	-0.085	0.216
Total_credit_lines	0.374	0.066	0.083	0.206	0.038	0.171	1.000	-0.027	0.192	0.277	-0.071
Bankcard_utilization	0.409	-0.013	-0.491	-0.047	-0.014	0.069	-0.027	1.000	0.033	0.153	0.418
Monthly_payment	0.341	-0.033	0.103	0.948	-0.022	-0.047	0.192	0.033	1.000	0.237	-0.112
DTI	0.310	-0.050	0.054	0.169	-0.020	-0.085	0.277	0.153	0.237	1.000	0.227
interest_average	0.028	0.098	-0.609	-0.256	0.044	0.216	-0.071	0.418	-0.112	0.227	1.000

Source: Author's estimation based on Prosper listing data 2014-2020.

9.5 Appendix V of Chapter 4

39. Table: Number of loans used for the liquidity dimensions

DPD	A	AA	B	C	D	E	F	HR
0	93 536	82 057	266 450	448 810	509 116	485 985	232 185	20 582
1-30	6 036	3 486	18 373	33 656	38 730	40 133	27 493	2 564
31-60	1 171	798	4 482	8 850	12 250	14 267	13 113	1 479
61-90	676	374	2 419	4 830	7 832	9 262	9 609	1 055
91-120	505	239	1 885	3 029	4 634	6 408	7 156	753
120+	9 954	4 250	47 090	90 561	108 345	106 480	119 727	43 936

Source: Author's estimation based on the Bondora secondary market database as of August 2020

9.6 Appendix VI of Chapter 5

40. Table: List of All Variables in the Raw Portfolio Table

ReportAsOfEOD	Age	HomeOwnershipType
LoanId	DateOfBirth	IncomeFromPrincipalEmployer
LoanNumber	Gender	IncomeFromPension
ListedOnUTC	Country	IncomeFromFamilyAllowance
BiddingStartedOn	AppliedAmount	IncomeFromSocialWelfare

BidsPortfolioManager	Amount	IncomeFromLeavePay
BidsApi	Interest	IncomeFromChildSupport
BidsManual	LoanDuration	IncomeOther
UserName	MonthlyPayment	IncomeTotal
NewCreditCustomer	County	ExistingLiabilities
LoanApplicationStartedDate	City	LiabilitiesTotal
LoanDate	UseOfLoan	RefinanceLiabilities
ContractEndDate	Education	DebtToIncome
FirstPaymentDate	MaritalStatus	FreeCash
MaturityDate_Original	NrOfDependants	MonthlyPaymentDay
MaturityDate_Last	EmploymentStatus	ActiveScheduleFirstPaymentReached
ApplicationSignedHour	EmploymentDurationCurrentEmployer	PlannedPrincipalTillDate
ApplicationSignedWeekday	EmploymentPosition	PlannedInterestTillDate
VerificationType	WorkExperience	LastPaymentOn
LanguageCode	OccupationArea	CurrentDebtDaysPrimary
PreviousRepaymentsBeforeLoan	PrincipalDebtServicingCost	NextPaymentNr
PreviousEarlyRepaymentsBefoleLoan	InterestAndPenaltyDebtServicingCost	NrOfScheduledPayments
PreviousEarlyRepaymentsCountBeforeLoan	ActiveLateLastPaymentCategory	ReScheduledOn
DebtOccuredOn	EAD1	Rating_V0
CurrentDebtDaysSecondary	EAD2	EL_V1
DebtOccuredOnForSecondary	PrincipalRecovery	Rating_V1
ExpectedLoss	InterestRecovery	Rating_V2
LossGivenDefault	RecoveryStage	Status
ExpectedReturn	StageActiveSince	Restructured
ProbabilityOfDefault	ModelVersion	ActiveLateCategory
DefaultDate	Rating	WorseLateCategory

PrincipalOverdueBySchedule	EL_V0	CreditScoreEsMicroL
PlannedPrincipalPostDefault	GracePeriodStart	CreditScoreEsEquifaxRisk
PlannedInterestPostDefault	GracePeriodEnd	CreditScoreFiAsiakasTietoRiskGrade
InterestAndPenaltyPaymentsMade	NextPaymentDate	CreditScoreEeMini
PrincipalWriteOffs	PrincipalBalance	PrincipalPaymentsMade
InterestAndPenaltyWriteOffs	InterestAndPenaltyBalance	AmountOfPreviousLoansBeforeLoan
NoOfPreviousLoansBeforeLoan		

Source: Bondora webpage as of 2020 October

9.7 Appendix VII of Chapter 5

41. Table: Descriptive Statistics for the Variables in the Scoring Model

	Age	**IncomeTotal**	**DebtToIncome**	**Country**	**Employment duration**	**Education**	**Existing Liabilities**	**LiabilitiesTotal**	**NoOfPrevious Loans Before Loan**	**AmountOfPrevious Loans Before Loan**	**PreviousRepaymentsBeforeLoan**	**HomeOwnershipType**
Average	41	1 576	0.48	-	-	-	3	501	1	2 840	950	-
Standard deviation	12	5 178	19.83	-	-	-	4	1 046	2	4 431	1 852	-
Minimum	18	0	0	-	-	-	0	0	0	0	0	-
Maximum	75	1 012 019	4 607.82	-	-	-	40	145 042	25	44 417	34 077	-
Missing values	0	0	0	0	0	0	0	0	0	0	0	0
Type of variable	Continuous	Continuous	Continuous	Category	Category	Category	Continuous	Continuous	Continuous	Continuous	Continuous	Category

Source: Bondora webpage as of 2020 October

9.8 Appendix VIII of Chapter 6

42. Table: Descriptive statistics by clusters

Variable	1	2	3

	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max
GDP per capita (USD)	53 323	13 254	39 826	85 025	22 383	5 927	15 022	33 525	6 679	3 514	772	12 124
Inflation rate	1,35	0,61	0,53	2,30	2,39	1,56	0,90	7,23	5,55	6,77	0,17	37,84
Unemployment rate	4,81	1,61	2,35	8,95	6,97	3,22	2,38	15,52	7,57	6,42	0,77	27,47
Non-performing loans as percent of all bank loans	1,74	1,66	0,28	6,85	3,85	2,85	0,28	9,84	6,32	4,07	1,51	17,91
Cost of living (world average=100)	0,73	0,23	0,10	0,99	0,39	0,19	0,19	0,80	0,27	0,21	0,06	0,80
Financial Institutions Depth Index	0,58	0,23	0,18	0,90	0,62	0,20	0,33	0,98	0,42	0,23	0,05	0,91
Financial Institutions Access Index	0,64	0,10	0,47	0,77	0,60	0,07	0,50	0,73	0,57	0,10	0,37	0,73
Financial Institutions Efficiency Index	146	26	78	198	101	17	70	124	65	14	46	105

Source: Author’s estimation based on theGlobalEconomy.com and IMF 2017-2019