

CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Bachelor thesis

2020

Martina Juračková

CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies



Martina Juračková

Analysis of Investor's Portfolio Diversification on the Czech
Peer-to-peer Market

Bachelor thesis

Prague 2020

Author: Martina Juračková

Supervisor: Mgr. Petr Polák, M.Sc.

Academic Year: 2019/2020

Bibliographic note

JURAČKOVÁ, Martina. *Analysis of Investor's Portfolio Diversification on the Czech Peer-to-peer Market*. 63 p. Bachelor thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies Supervisor Mgr. Petr Polák, M.Sc.

Abstract

This thesis deals with the topic of peer-to-peer lending. Using the data provided by the Czech platform Zonky, we study the behaviour of investors on the peer-to-peer lending market. This paper aims to determine whether the demographic factors of investor and the structure of his portfolio have a significant impact on portfolio performance. Based on the quantity invested, a sample of “small” investors (with the quantity invested from 10 000 CZK to 50 000 CZK) and “big” investors (with the invested volume exceeding 1 000 000 CZK) was chosen. Those contrasting samples were analysed. It was discovered that several factors affect portfolio performance. In terms of the portfolio’s structure, the size of the share invested in loans with very low and medium risk level and to borrowers from Prague was determined as significant. This is a common feature of both samples. On the other hand, compared to “big” investors, in the case of “small” investors, more variables were estimated as significant. This implies that the performance of a portfolio with lower amount invested is probably more sensitive to various loan’s characteristics. These findings are partially consistent with the results of similar papers conducted in the Czech Republic or abroad.

Abstrakt

V této práci se zabýváme oblastí peer-to-peer půjček. Na datech poskytnutých českou platformou Zonky zkoumáme chování investorů na trhu s peer-to-peer půjčkami. Hlavním cílem práce je určit, zda mají demografické charakteristiky investora a struktura jeho portfolia významný vliv na výkonnost investorova portfolia. Na základě informací o celkově investovaných částkách jednotlivých investorů byl vybrán vzorek „malých“ investorů s investovanou částkou v rozmezí 10 000 Kč až 50 000 Kč a „velkých“ investorů s investovanou částkou větší než 1 000 000 Kč. Tyto rozdílné vzorky byly následně analyzovány. V první řadě bylo zjištěno, že výkonnost portfolia je ovlivněna hned několika faktory. Z hlediska struktury portfolia byly jako důležité určeny podíly investované do půjček s velmi nízkým a středním rizikem a dále pak klientům s trvalým bydlištěm v Praze. Tento závěr je společný pro oba dva vzorky. Na druhou stranu se ukázalo, že v porovnání s velkými investory je počet statisticky signifikantních proměnných u malých investorů větší. Z toho vyplývá, že výkonnost portfolií s malými investovanými objemy je více citlivá na jejich složení. Výsledky této práce jsou do značné míry ve shodě se závěry podobných studií provedených v České republice i v zahraničí.

Keywords

Peer-to-peer lending, peer-to-peer platform, investment portfolio, interest rate, portfolio performance, Zonky

Klíčová slova

Peer-to-peer lending, peer-to-peer platforma, investiční portfolio, úrok, výkonnost portfolia, Zonky

Declaration of Authorship

I hereby declare that I compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

I grant a permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 31, 2020

Signature

Acknowledgment

I would like to express sincere gratitude to my thesis supervisor Mgr. Petr Polák, M.Sc. for provision of valuable advice on thesis' improvement and guidance. Further, I thank to Aleš Černý from Zonky for providing me with the necessary data and information that are not publicly available. Last but not least, I am grateful to my family and friends for the support and encouragement during my studies.

Bachelor's Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague



Author's name and surname: Martina Juračková

Supervisor's name: Petr Polák

Proposed Topic:

Analysis of investor's portfolio diversification on the Czech peer-to-peer market

Preliminary scope of work:

Research question and motivation

After the financial crisis in 2008, banks changed their lending conditions (made them tougher). As a consequence, some individuals and firms could not borrow money from banks because they were not able to fulfil the new conditions, and those who could borrow money from banks often struggled with high interest and inability to pay back. This unpleasant situation together with vast development of information technology and shared economy resulted in increase of popularity of online lending on the peer-to-peer lending market and peer-to-peer lending platforms. Online peer-to-peer lending, based on the idea of providing a loan without involving the financial institution, is beneficial for both parties – lender (investor) and borrower. In exchange for financing unsecured loans, lenders receive higher interest than the one offered in a bank. For borrowers, the advantage of peer-to-peer lending consists in obtaining a loan with better conditions than in a bank.

However, the fact that borrower seeks for a loan elsewhere than in a bank may indicate that for some reason a bank does not want to lend money to that particular individual. Such loans are riskier to invest in, also a situation in which the individuals are not able to pay back properly and in time might arise. All these risks are covered in the interest that is charged to borrowers, according to which investors make their decisions and create the portfolio. But this is probably not the only factor that influences the structure of the portfolio, personal characteristics of investor and his previous experience with investing in p2p loans might matter to. So, the research question of this thesis is: Do personal characteristics of investor have some impact on his portfolio diversification?

Contribution

There are many analyses concerned with peer-to-peer lending, most of them using data from the Lending Club which is the largest peer-to-peer platform on the US market. Recently, also many studies focused on Chinese p2p market occurred. Despite the fact that the first peer-to-peer platform (Zopa) was established in 2005, the peer-to-peer market in the Czech Republic is still quite undeveloped. Although it has experienced significant development in the last few years, not many analyses dealing with the question of peer-to-peer lending in the Czech Republic can be found. Also to the author's best knowledge, no analysis of portfolio diversification on Czech p2p market exists. Thus, the aim of this bachelor thesis is to shed light on this rather unexplored topic by selecting particular Czech peer-to-peer platform and analysing the relationship between investor's portfolio diversification and his personal characteristics.

Methodology

Choosing a Czech peer-to-peer platform and using its available data to conduct an analysis on the micro level. Specifically, analysing portfolio diversification of investors according to not only their personal

characteristics such as age, gender, occupation etc. but also their behavior on p2p platform described for example by the amount of money invested or number of investments during one month. The data will be used to run a regression model which will be estimated. The results will be compared with existing findings from abroad.

Outline

1. Introduction
2. History of P2P lending
3. Key characteristics
4. P2P lending in the Czech Republic
5. Literature review and hypothesis
6. Methodology
7. Empirical part and results
8. Conclusion

List of academic literature:

Bibliography

- CHEN, D. a C. HAN (2012). A comparative study of online P2P lending in the USA and China. *Journal of Internet Banking and Commerce*, 17(2), 1 – 15
- HU, Rongcai, Meng LIU, Pingping HE a Yong MA (2019). Can Investors on P2P Lending Platforms Identify Default Risk?. *International Journal of Electronic Commerce*, 23(1), 63-84
- LI, Zhihong, Lanteng WU a Hongting TANG (2018). Optimizing the Borrowing Limit and Interest Rate in P2P System: From Borrowers' Perspective. *Scientific Programming*, 1-14
- POLENA, M. a T. REGNER (2018). Determinants of borrowers' default in P2P lending under consideration of the loan risk class. *Games*, 9(4)
- SERRANO-CINCA, Carlos, Begoña GUTIÉRREZ-NIETO a Luz LÓPEZ-PALACIOS (2015). Determinants of Default in P2P Lending. *PLoS ONE*, 10(10), 1-22

List of Acronyms

AMT Amount

CNT Count

CZK The Czech Crown

GBP The British Pound

P2P Peer-to-peer

SE Sharing economy

UK The United Kingdom

US The United States of America

USD The United States Dollar

Contens

1. INTRODUCTION.....	1
2. THEORETICAL BACKGROUND OF PEER-TO-PEER LENDING.....	3
2.1 SHARING ECONOMY	3
2.2 P2P LENDING	5
2.3 P2P LENDING IN THE CZECH REPUBLIC.....	12
2.4 ZONKY.....	13
2.4.1 <i>Investing</i>	16
2.4.2 <i>Borrowing</i>	20
2.4.3 <i>Zonky Rentier</i>	22
3. LITERATURE REVIEW	23
3.1 INFORMATION ASYMMETRY AND PROPOSED SOLUTIONS	23
3.2 BORROWER'S CHARACTERISTICS AND FUNDING SUCCESS	24
3.3 CREDIT RISK	26
3.4 ZONKY.....	26
4. METHODOLOGY.....	28
4.1 PORTFOLIO.....	28
4.2 PORTFOLIO PERFORMANCE.....	28
4.3 THE APPROACH OF THIS THESIS	30
4.4 DESCRIPTION OF DATA	30
5. EMPIRICAL PART.....	41
5.1 ESTIMATED MODELS	42
5.1.1 <i>Estimation results of model AMT</i>	43
5.1.2 <i>Estimation results of model CNT</i>	46
6. CONCLUSION	49
REFERENCES	51
LIST OF APPENDICES	56
APPENDICES	57

1. Introduction

The term peer-to-peer (P2P) lending refers to providing a loan with the help of an online peer-to-peer platform instead of the traditional financial institution (typically a bank). The online peer-to-peer lending platform mediates the interaction between the lender (investor) and the borrower while the whole process is transferred into the internet. The concept of P2P lending belongs to the sharing economy (Galloway et al., 2009).

During the last decade and a half, the area of peer-to-peer lending has experienced a significant development mainly due to the increasing popularity of online communities and the development of the digital economy. The first online peer-to-peer lending platform Zopa was established in 2005 in the UK. Since then, many platforms around the world such as Prosper, Funding Circle, Lending Club, Mintos or Bondora have been founded. As Vysušil (2019) explains, online peer-to-peer lending has quickly become a severe competitor to the traditional ways of financing. Bankerat, the first P2P platform in the Czech Republic, was launched in 2010. Among the numerous platforms operating on the Czech peer-to-peer lending market (e.g. Bankerat, Prestito, Banking online, Bondster, Žlutý meloun, Mintos, Bondora), Zonky can be considered as the current leader.

The primary research focus of this thesis is to determine the main factors among the demographic features of investor (and borrower) and the loan's characteristics having a significant impact on the portfolio performance of the investors on the Czech P2P platform Zonky. And consequently, partly uncover the decision-making process of the investor on the Czech peer-to-peer market. Hypothetically speaking, this study could also help investors to decide whether to do or do not invest in a loan with particular characteristics.

On the field of the Czech P2P market, somewhat similar research has already been done by Hudcová (2017) with the conclusion that firstly, in the environment of the peer-to-peer lending platform, investors behave riskier than they from the theoretical point of view should and secondly, she identified that high interest rate has the greatest impact on the investor's expected return and the final composition of the lender's portfolio. This paper will elaborate on her work by using more detailed information about the investor and his portfolio provided directly by the P2P platform. The

regression analyses will be performed on two samples of investors (“small” investors whose amount invested ranges between 10 000 CZK and 50 000 CZK and investors with the quantity invested greater than 1 000 000 CZK) and the results of those groups will be compared. The hypothesis that the interest rate is not the only determinant affecting the portfolio performance significantly will be tested.

The main contribution of this thesis lies in the fact that the dataset used for the analyses is provided by Zonky itself. Thus, it does not contain only information about the lender’s portfolio (such as the volumes lender has invested in specific loans) but also characteristics of the lender and the borrower. Therefore, the real behaviour of investors can be examined more closely. Compared to Hudcová (2017) who gained data with the help of online questionnaire posted on the Facebook group “Investor on Zonky” and various webpages frequented by Zonky investors, our dataset provides us with the unbiased sample. Moreover, this work also lays possible foundations for future research examining the portfolio performance on the Czech peer-to-peer lending market.

The thesis is structured as follows: First, Chapter 2 situates peer-to-peer lending in the context of the sharing economy and continues with the explanation of general principles and risks in peer-to-peer lending, including a brief overview of peer-to-peer platforms around the world. This chapter also maps the peer-to-peer market in the Czech Republic and discusses platform Zonky in more detail. Chapter 3 brings a review of the literature relevant to the topic of this thesis. In Chapter 4, methodology, data and the variables used are described. Chapter 5 presents the results of the analyses conducted. Chapter 6 concludes.

2. Theoretical background of peer-to-peer lending

As peer-to-peer lending is a concept that belongs to the sharing economy, this section will be first dealing with the explanation of sharing economy. Second, peer-to-peer lending in general will be described proceeding with its main features and the most significant platforms around the world. Third, peer-to-peer lending in the Czech Republic will be analysed with the emphasis on currently the most significant Czech peer-to-peer platform – Zonky.

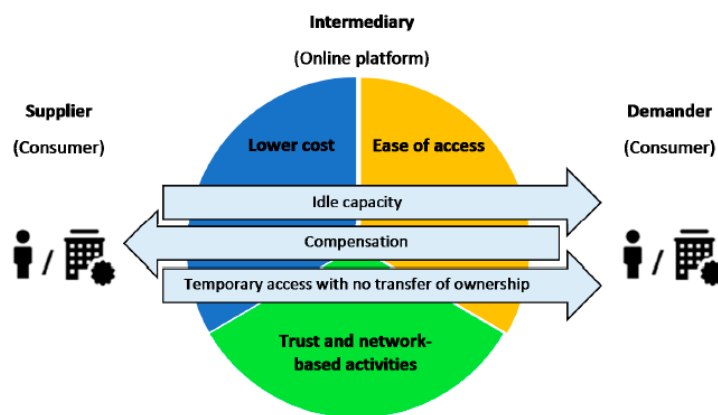
2.1 *Sharing economy*

Sharing economy (SE), sometimes also referred to as a collaborative economy, circular economy, peer-to-peer economy, access economy or pooling economy (Deloitte, 2017), has experienced rather significant development over the last few years. Such an expansion can be mainly attributed to the development of the digital economy. However, researchers in the field of sharing economy (Belk 2010, Sundararajan 2016, Albors, Ramos, and Hervás 2008) remind that the act of sharing as such is not a new concept, since barter and communal ways of life are as old as mankind. Nevertheless, until recently, there were certain limitations in sharing of physical goods. Specifically, a low level of trust between lender and borrower as well as the difficulties in matching supply and demand (Ranjbari, Morales-Alonso, and Carrasco-Gallego, 2018). However, the vast development of digital platforms and other large-scale mediating technologies managed to cover these gaps and resulted in the excitement over the sharing economy and collaborative consumption (Sutherland and Jarrahi, 2018).

Even though sharing economy is quite popular these days, until recently there has been no generally accepted definition for this term. Partly because this economic principle keeps evolving constantly but also due to disagreement among researchers concerning the interpretation for word sharing. The main dispute lies in the fact whether the provision of the access to an asset (i.e. sharing) brings some financial benefit to at least one of the parties involved. If that is the case, research conducted by Frenken and Schor (2017) suggests excluding such an exchange from sharing economy. To differentiate between these two approaches, Belk (2014) provides the terms true-sharing (no revenue is considered) and pseudo-sharing (monetary benefit is generated).

As can be seen, different views and frameworks exist for the sharing economy, implying also a variety of interpretations for the SE concept. Ranjbari, Morales-Alonso and Carrasco-Gallego (2018) tried to shed some light on the ambiguity in the interpretation of the sharing economy (the framework for SE they proposed is presented in Figure 1). In their paper a systematic literature review was conducted, 67 SE definitions from among 193 papers were collected and analysed. They define the SE as *an economic system, whose intermediary companies utilize online platforms to facilitate and lower the cost of the for-profit transactions of giving temporary access – without the transfer of ownership – to idle resources of consumers in the peer-to-peer networks that it has created, because of the trust built among its members, who may be individuals or businesses.* The authors additionally explain that online platforms connect consumers on both sides - the supply and demand sides, and also emphasize that for-profit activities mentioned in their definition do not necessarily involve a monetary transaction. Instead, they can sometimes refer to the exchange of different types of tangible or intangible assets.

Figure 1: The proposed framework for sharing



Source: Ranjbari, Morales-Alonso and Carrasco-Gallego (2018)

2.2 P2P lending

General information

The increasing popularity of online communities in the past decade in combination with the development of the digital economy has established a new way of loan provision – online peer-to-peer lending. Peer-to-peer lending is based on the idea of providing a loan without involving the financial institution, a principle that is not revolutionary since from the perspective of individuals, providing private loans without any mediation can be seen as a rather traditional way of borrowing money. What makes online P2P lending special is the fact that interaction between lender (investor) and borrower is mediated with the help of an online peer-to-peer lending platform (i.e. is transferred into the internet). Such an arrangement is beneficial for all the parties involved – lender, borrower and the platform. For borrowers, it is viewed as a way of obtaining a loan with better conditions than in a bank. From the point of investors, it can be regarded as an exceptional investment opportunity with not only an appealing interest rate offered but also some sort of social overtone included. Peer-to-peer platform profits from charging fees to both types of its clients (i.e. borrower and lender) for providing its intermediary services and realizing operations (Galloway et al., 2009).

Each P2P platform is different, but the core principles of online peer-to-peer lending remain similar for all of them. Initially, both the borrower and the lender need to register on the platform and provide necessary information about themselves to verify their identity. Some platforms might require additional personal information such as education level or occupation. Loan applicants are also required to report their income and credit history. Using several data sources (banking and non-banking registries, credit bureaus, social networks etc.) platform confirms the validity of acquired pieces of information and assigns borrower's loan request with a rating (or credit score). Each rating is matched with a given level of interest rate identifying the likelihood of a loan's default. Once loan request is accepted, transfer of money can be instant or take a few days depending on platform's approach towards fund provision (P2P platform can either provide loan amount before funding by investors begins or postpone the transfer until the whole amount is fully funded). Fees charged to the borrower and the investor vary among different platforms. The borrower repays the debt in monthly payments.

Based on the lender's expectation regarding the purpose of his investment, P2P lending platforms can be categorized as non-commercial (also referred to as micro-lending) and commercial (Ashta and Assadi, 2009). Whereas a lender on the commercial platform is focused solely on earning a profit, an investor on the non-commercial platform is rather interested in development. The main goal of the microlender consists in supporting the business development of small entrepreneurs or projects in economically undeveloped regions whose other sources of funding are very limited, while the role of profit is not that vital. The social overtone is further supported by the possibility of coaching the borrower on business management provided by micro-lending organization that at the same time helps to ensure the business will generate profit and borrower will be able to pay back the loan.

Risks

Like any other investment, peer-to-peer lending is as well accompanied by several risks. Probably the best known and the most discussed one is credit risk that arises when the borrower is unable or unwilling to repay the debt. Investing in secured loans can be seen as a convenient way of reducing credit risk. It does not obviously influence the borrower's ability to pay back. However, if the borrower defaults, investors can get the remaining amount of money by claiming the collateral provided by the borrower. There are some P2P lending platforms offering secured loans (LandlordInvest, CrowdProperty, Bridgecrowd, Crowdestate) but since the majority of P2P lending platforms provide unsecured loans the other way of dealing with credit risk is sound portfolio diversification.

The portfolio should be diversified not only in terms of different loans with various characteristics but ideally also across different P2P lending platforms. That leads us to the other type of risk – platform failure. If the platform itself gets into trouble, theoretically investors should not be concerned since typically there is a deal arranged with some other entity concerning the management of necessary administration. Such a scenario is not that uncommon. In the UK, Collateral and Lendy collapsed in 2018 and 2019, respectively. Chinese P2P market experienced a great shock in 2016 when one of its largest P2P lending platforms Ezubao was revealed to be fraud (Ponzi scheme). Lack of regulation from the side of the People's Bank of China

and search for alternative ways of loan provision led to a rapid growth of the P2P market in China. Instead of the quality of risk management, the only signal of trustworthiness appeared to be the speed and scale at which the platform was able to grow. Since Ezubao collapsed a wave of scandals and regulatory changes followed and the number of platforms on the Chinese P2P market has decreased considerably.

Liquidity risk in general occurs when an individual is unable to convert asset or security quickly into cash with minimized loss. Liquidity risk in P2P lending refers to the investor's inability to sell his loan participation and get money back. Peer-to-peer platforms handle this type of risk by establishing a secondary market. Similarly as any other market, the peer-to-peer lending market is also based on trust. Loss of confidence in any platform may evolve in a situation known as "bank run" – many investors want to get their money back at the same time. This sudden increase in the supply of loan participations on the secondary market however cannot be matched with a corresponding shift in demand for loan shares.

During a recession higher unemployment rates can be expected, followed by an increase in the likelihood of loan's default and consequently decrease in investor's return. Providers of loans secured by collateral in the form of property are affected too since the property prices are likely to decrease. Many P2P platforms have not yet experienced an economic downturn, nevertheless they are not indifferent towards this type of risk. In order to model what would happen in various scenarios, some of them opt to conduct stress tests of their loan book. For instance, platform Zonky has decided to estimate how would the number of defaulted loans (and thus return of investor's portfolio) change under parameters resembling the financial crisis in 2008. Results have shown that investors whose invested amount is allocated evenly into loans with various risk levels (i.e. sufficiently diversified¹) may anticipate on average 1 percentage point decrease in return of their portfolio.

¹ Zonky recommends its investors to invest the same amount of money in at least 120 loans from different risk categories

P2P platforms around the world

Prosper and Zopa, both established in 2005, are considered to be the founders of the P2P industry in the US and UK, respectively. Since then peer-to-peer lending has been experiencing exponential growth. The volume of loans issued by peer-to-peer platforms around the world in 2014 amounted to 23,7 billion USD (Deloitte, 2016). The study *Global Peer to Peer (P2P) Lending Market Size by Type (Business Lending and Consumer Lending), by Region (North America, Europe, Asia Pacific, Latin America, and Middle-East & Africa) and Forecast 2018 to 2025* by Adroit Market Research (business analytics and consulting company) estimated the value of P2P lending market in 2017 to be more than 231 billion USD and predicts it to exceed 820 billion USD by 2025 (under the assumption of compound annual growth rate (CAGR) above 25,4 %).

In the US with 59 billion USD funded amount and market share exceeding 70 percent, Lending Club has replaced Prosper in the role of the most significant P2P platform. Prosper follows Lending Club with 17 billion USD funded amount and market share slightly above 20 percent (*USA Peer-to-Peer Lending & Equity Crowdfunding Statistics (USD)*, no date).

In the UK, currently leading platform is Funding Circle (35 % market share and above 6 billion GBP funded amount), followed by Zopa (30 % market share and above 5 billion GBP funded amount) and Ratesetter with above 22 % market share and almost 4 billion GBP funded amount (*UK Peer-to-Peer Lending & Equity Funding Volumes | P2PMarketData*, n.d.). Other significant European platforms can be found in the Baltic area, namely Mintos in Latvia (more than 5 billion EUR funded amount) and Bondora in Estonia with 370 million EUR funded amount (*EU Peer-to-Peer Lending & Equity Funding Volumes | P2PMarketData*, n.d.). One remarkable feature of Mintos is that loans are brought to the platform by partner lending companies instead of being issued by Mintos. A brief overview of the platforms discussed in this section in more detail can be found in Table 1.

As liquidity (the degree to which asset can be converted quickly into cash without affecting substantially its price) is one of the cornerstones in any economy, innovations in peer-to-peer lending models have been very often incentivized by liquidity improvement. Some principles on which peer-to-peer lending was based in its

infancy do not hold anymore. Both investors and borrowers value their time greatly. Apart from liquidity, lesser demand on their time is another powerful stimulus for the platform's advancement. One example of such an improvement can be found on the US platform Prosper. In 2010, it shifted from a rather lengthy auction process² used in determining the interest rate for borrowers to setting the interest rate with the help of an internally developed model.

Another good illustration of P2P platforms evolution is the usage of auto-bids (automatic bidding mechanisms). Platforms like Bondora or Funding Circle do not allow investors to choose manually loans they would like to invest in anymore since 2016 and 2017, respectively. Bondora decided to implement this step because it has become clear that almost all loans have been fully funded by automatized mechanisms before manual investing could take part in. According to Bondora statistics before removing the possibility of manual investing, around 80 percent of investments were conducted with the help of an automated portfolio manager, whereas the share of manual investing was only 10 percent (*Auto Bidding | P2P-Banking*, n.d.). Nowadays, investors are left with the possibility to either use an automated portfolio manager that makes investment decisions based on preferred investment strategy or create their own unique investment strategy using API (Application Interface). Funding Circle abandoned manual investing because it perceived it as unfair (similarly as on Bondora, before lenders using manual investing could access the most beneficial loans, they were already gone) and not sufficiently diversified. Under the new investment strategy, investors can choose between two lending options – a balanced portfolio or a conservative portfolio that are set to yield 7.5 percent and 4.8 percent net annual return, respectively.

In terms of better platform liquidity, Zonky does not stay behind. Since August 2017, investors on this platform are enabled to sell their investments to other investors on the secondary market. Providing the lenders with the opportunity to gain liquidity before the loan's maturity is not an extraordinary thought, many platforms use that. However, compared to other platforms, this move was done rather quickly. Zonky

² In online auction borrowers set a maximum level of interest rate they are willing to pay. During a limited period of time investors can make their bids and even beat other investors by offering a lower interest rate. The winners of the auction are the investors offering the lowest interest rate.

launched its secondary market two years after the platform was founded, while making progress in this direction had taken Prosper or Zopa four to five years. Nevertheless, as pioneers on P2P markets in the US and UK the mentioned platforms were founded 10 years before Zonky (in 2005). Therefore, in contrast with Prosper and Zopa, in making improvements Zonky could draw inspiration from already well established and successful competitors.

What has also changed throughout Zonky's existence is the platform's approach towards the provision of money to borrowers. Originally, the borrower received money only after the loan request was fully funded by investors. In other words, if there were not enough investors willing to contribute in the loan, the financial transaction was not realized (loan request was rejected). Currently, the initial transfer of money is provided by Zonky and the funding phase follows as the second step. This can serve as another attempt of P2P platforms to satisfy the time demands of their clients by making various procedures quicker and less time-consuming.

In conclusion, as has been shown above the area of peer-to-peer lending keeps changing and evolving. Therefore, the results of studies conducted several years ago may not hold anymore. A good illustration of this fact is paper by Bachmann et al. (2011) that identifies the auction process as one of the means of setting the interest rate for borrowers. However, no platform mentioned in Table 1 uses this approach nowadays.

Table 1: Comparison of selected P2P platforms

	Prosper	Zopa	Lending Club	Bondora	Funding Circle	Mintos	Zonky
Founded	2005	2005	2007	2008	2010	2015	2015
Country of origin	US	UK	US	Estonia	UK	Latvia	Czech Republic
Manual investing	yes	no	yes	no	no	yes	yes
Auto-bid	yes	yes	yes	yes	yes	yes	no ³
Secondary market	no ⁴	yes	yes	yes	yes	yes	yes
Smallest possible bid	\$25 or \$5000 ⁵	£1000	\$25	€1	£10	€10	200 CZK
Loan term	3 or 5 years	1 to 5 years	3 or 5 years	3 to 60 months	6 months to 5 years	1 day to more than 72 months	6 months to 7 years

Source: websites of respective platforms

³ Even though Zonky does not provide investors with auto-invest tools (platform only offers service “Zonky Rentier“ that is set to allocate invested amount in various loans in order to yield 4.4 percentage net annual return, however, the investor has no impact on the selection of loans his money will be invested in), some websites like www.zotify.cz or www.zonkios.cz do.

⁴ Secondary market on Prosper was closed down in 2016.

⁵ Depending on the type of account – for individual Investment Accounts, the minimum amount is \$25, for Individual Retirement Account (IRA) minimum investment is \$5000.

2.3 P2P lending in the Czech Republic

Apart from Zonky (nowadays the most significant P2P platform in the Czech Republic), there are other Czech platforms such as Bankerat, Prestito, Banking online or Bondster operating on the market (these will be described later) as well as some platforms from abroad (Žlutý meloun, Mintos, Bondora). The establishment of the P2P lending market in the Czech Republic is associated with the foundation of Bankerat in 2010. Due to an amendment of the Consumer Credit Act in December 2016, some of those entities suspended conducting their business activities in the field of peer-to-peer lending (for example platform Benefi has limited its functioning only to the collection of payments from borrowers, it does not provide new loans anymore).

Bankerat, founded in 2010 and currently with more than 57 000 clients, has already brokered loans amounting to almost 1 billion CZK. The Czech National Bank did not categorize this platform as a non-banking provider of consumer credit loans. Thus, Bankerat mediates money provision in the form of a contract concluded between investor and borrower. Loan funding is based on the auction process. Platform charges a 5% fee of the principal sum to the borrower (*Poplatek Dlužníka | Bankerat, n.d.*) and a 1% fee of remaining debt to the investor (*Poplatek Věřitele | Bankerat, n.d.*). If the payment from the borrower is delayed, instead of a one-off money penalty, the borrower is charged a penalty interest rate. Average yields on Bankerat can exceed the return on investment on other platforms multiple times but naturally, high profitability goes hand in hand with a greater level of risk. In contrast to Zonky, Bankerat is indifferent to the composition of the investor's portfolio, it does not actively encourage investors to diversify their portfolio.

The other platform Banking online uses the auction approach as well. Borrowers are enabled to take loans amounting to 500 000 CZK with duration from 6 to 60 months. They are charged a 1.5% fee from the borrowed amount and investors have to pay a 0.8% fee from every payment received from the debtor.

On Bondster loans offered on the marketplace are issued by a non-bank loan providing company (third party) that also determines the interest rate for the borrower. What makes investing on this platform appealing is the possibility of investing in secured loans and also the fact that some of the loans are provided with a guarantee of buying back the loan by the loan issuer in case the borrower ceases to pay back. The

minimum investment is 100 CZK and the investor's fee is 1% p.a. of the invested amount.

The last platform Prestito does not provide intermediation services. Instead of that, it only creates an online environment where lenders and borrowers can be connected. The interest rate is set with the help of auction system. Borrowers advertise their loan request amounting from 10 000 CZK to 1 000 000 CZK (including maximum interest rate borrowers are willing to pay) on the website of Prestito and for a limited amount of time, lenders can make their bids. If the borrower finds some of the offers especially appealing, the auction can be prematurely ended. The minimum investment is 5 000 CZK and fees charged are one-off.

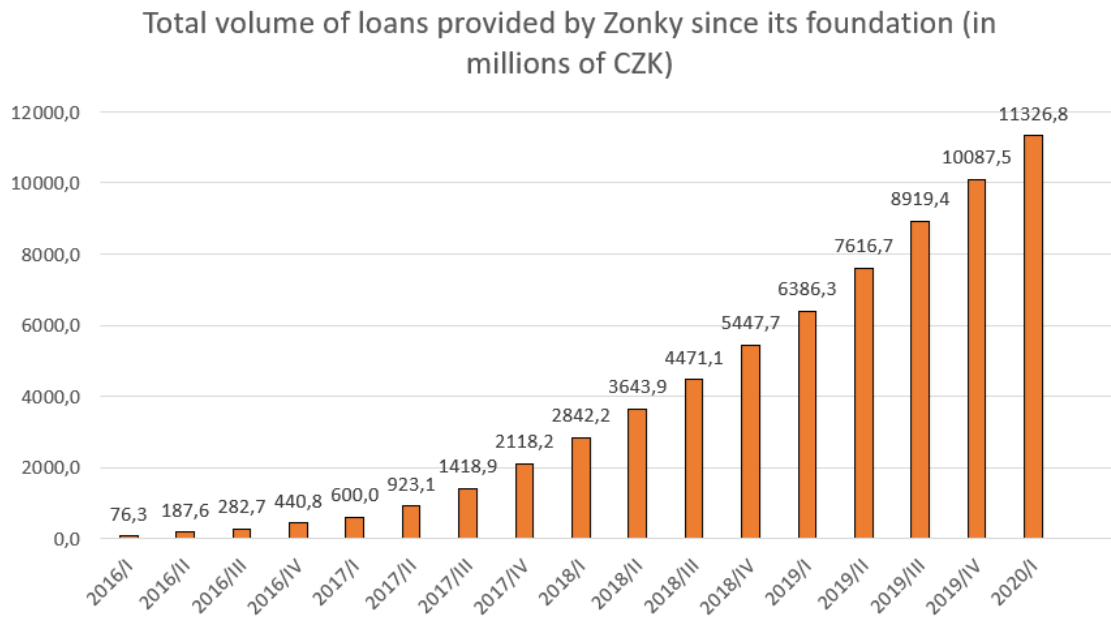
2.4 Zonky

Since the research focus of the thesis is aimed at Zonky, history, key characteristics, concepts and working principles of this Czech peer-to-peer platform should be introduced.

Zonky is part of the Home Credit Group belonging to the international investment group PPF. Therefore, it is not surprising that as well as its fund provider (Home Credit), Zonky is a non-banking provider of credit loans. The start-up project was legally founded in November 2014 (*Veřejný rejstřík a Sběrka Listin - Ministerstvo Spravedlnosti České Republiky*, n.d.), the platform itself started operating in June 2015. Considering the entrance on other than Czech (i.e. international) peer-to-peer market in the near future, the name Zonky was changed to Benxy in November 2019 (*Klient: Často Kladené Otázky | Zonky*, n.d.). However, the participants of the Czech peer-to-peer market did not perceive this change as severe since the platforms' name "Zonky" as well as its functioning in the Czech environment remained unchanged. The platform is designed primarily to operate online (for its clients can log in using a personal computer or mobile phone) but it also runs several local branches. Even though Zonky is categorized as a non-banking provider of credit loans, due to an amendment of the Consumer Credit Act, it is supervised by the Czech National Bank since December 2016.

The platform prides itself on the human approach regarding both *considering the borrowers' loan requests* (Hudcová, 2017) and borrowers' problems with repaying the loan. According to the survey "Trustworthy brand"⁶ conducted among thousands of Czech customers, this mild approach is appreciated by the public since 42 percent of the respondents have chosen Zonky as "The most trustworthy brand in the area of non-banking loans in 2019"⁷ (*Zonky Je Nejdůvěryhodnější Značka v Oblasti Nebankovních Půjček | ZonkyTimes.Cz – Finance, Investice, Rozhovory*, n.d.). Also, the advertising slogan "People lend to people"⁸ seems to work for the public (at least subconsciously) judging by the fact that in December 2019, the total sum of money Zonky borrowed since its foundation exceeded 10 billion Czech crowns as shown by Figure 2. Figure 3 presents the number of loans provided each quarter since Zonky's foundation.

Figure 2: Total volume of loans Zonky has provided



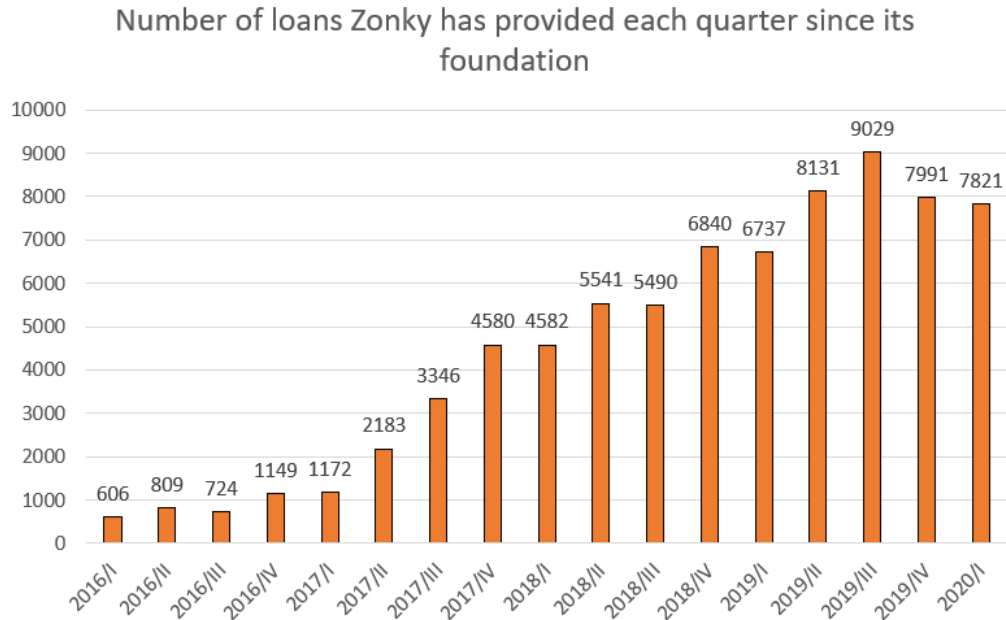
Source: author's own computations using Zonky's publicly available loanbook

⁶ „Důvěryhodná značka“

⁷ „Nejdůvěryhodnější značka v oblasti dodavatelů mimobankovních půjček za rok 2019“

⁸ „Lidé půjčují lidem“

Figure 3: The number of loans Zonky has provided

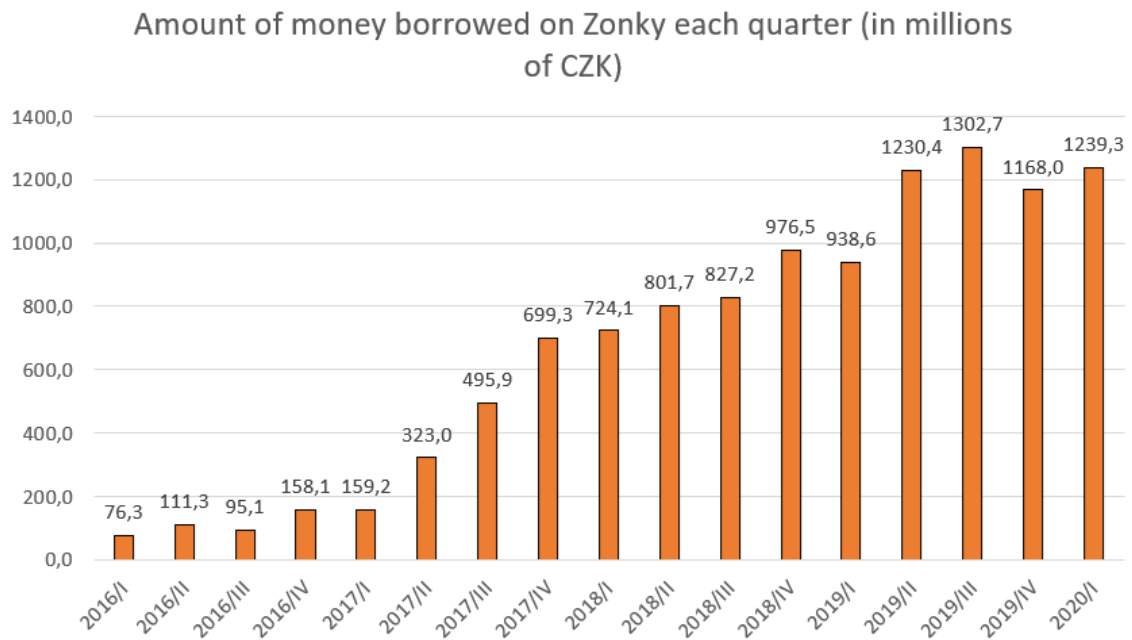


Source: authors' own computations using Zonky's publicly available loanbook

According to the Financial Stability Report 2018/2019 provided by the Czech National Bank (*Financial Stability Report 2018/2019 - Czech National Bank*, n.d.), the total volume of loans provided by non-bank providers of financial assets⁹ in 2018 amounted to 307.2 billion CZK. As can be seen from Figure 4, in 2018 Zonky has mediated loans of 3.33 billion CZK. Recently, Zonky's share on the consumer credit market has increased to approximately 5 percent. However, the platform's long-term target on this particular market is 10 percent.

⁹ Non-bank providers of financial assets expand the range of credit products traditionally provided by banks.

Figure 4: Amount of money borrowed on Zonky



Source: authors' own computations using Zonky's publicly available loanbook

2.4.1 Investing

As already mentioned above, online peer-to-peer lending is beneficial for both lenders (investors) and borrowers. In exchange for financing unsecured loans, lenders receive higher interest than the one offered in a bank. For borrowers, the advantage of peer-to-peer lending consists in obtaining a loan with better conditions than in a bank.

Registration process

Any person 18 years or older who has a Czech current account, uploads identification documents and passes verification checks can become a lender on platform Zonky. Interestingly, Zonky does not demand the Czech citizenship of investors. The contract between Zonky and investor can be signed online on the website via copying a special code obtained in the text message. Activation of investor's personal account on Zonky (referred to as "investor's wallet") and verification of

personal information obtained from the registration form can take a few days. The registration is considered as completed the moment the lender can invest in loans displayed on the “marketplace”.¹⁰

Marketplace

Each approved loan is placed on the marketplace (also referred to as the primary market) for an interval of 2 days. During this time individual investors (registered lenders) or the institutional investor¹¹ can contribute to the funding of a given loan. Each loan on the marketplace is provided with several pieces of information upon which the investor can decide whether to invest in or not. These features may be divided into two categories:

- i) general facts about the loan - interest rate (e.g. rating), interest rate after the deduction of fee, the duration of the loan in months (maturity), the proportion of a loan that has already been funded and by how many investors, the exact amount of regular monthly instalment and how much longer be the loan will available on the marketplace
- ii) information about borrower – nickname (since Zonky does not require their clients to reveal their real names on the marketplace), ID number in Zonky system, source of income (employee, entrepreneur, student, self-employed, pension, etc.), area of living (region), the purpose of loan (what it will be used for), verification of borrower’s income and identity in registries, whether the borrower is insured against the inability of paying back, whether has the borrower ever before (and how many times) borrowed money using Zonky

The borrower is also given the opportunity to elaborate on the personal story behind his loan in the form of a short paragraph displayed together with loans’

¹⁰ „Tržiště“

¹¹ Institutional investor is a legal person (entity) that works as a “market creator” – creates its own portfolio by investing in various loans (in particular in those that are either less attractive to other investors or would require more time than 2 days to be fully funded) which helps to maintain both good reputation of the platform and general interest in peer-to-peer lending by keeping sufficient supply of loans for other investors

characteristics. In case yet further information is needed, investors can ask additional questions. To get the idea of how the marketplace and detail of loan look like, see Appendix 1 and Appendix 2.

The minimum amount the investor can lend is 200 CZK. The maximum amount that can be invested is dependent on the number of loans that the lender has already invested in - see Table 2 below (*Dokumenty a Smlouvy, Parametry částek pro investování* | Zonky, n.d.). Zonky regards such limitations necessary in order to keep the investor's portfolio diversified to ensure minimization of risks in case of loans' default. This sort of "safety measure" is further supported by the fact that each loan must be funded by several investors (e.g. it is not permitted to an investor to fund a loan all by himself).

Table 2: The scale of possible amount invested on Zonky

The minimum amount of investments	The minimum amount that can be invested in CZK	The maximum amount that can be invested in CZK
0	200	5 000
101	200	10 000
201	200	20 000

Source: Zonky

Zonky assigns each loan with interest rate ranging from 2.99% p.a. to 19.99% p.a. corresponding to the risk lender is exposed to in the form of borrower's inability to repay (risk categories are marked in colours in Table 3 based on categorization on the platform's website). The risk category also determines the investment fee charged by the platform for its mediation services as a proportion of the amount of money invested in a loan with the respective rating. As shown in Table 3, fee increases with higher risk. In case payment from the borrower is more than 36 days past due, the investor does not pay the investment fee anymore until borrower resumes repaying again. Every monthly instalment received from the borrower is redistributed

by Zonky to respective investors. It is important to emphasize that any profit made by investors is a subject of 15% income tax.

Table 3: Interest rate, investment fee and expected yield

Interest rate (p.a.)	Investment fee	Expected yield
2.99%	0.2%	2.34%
3.99%	0.2%	3.34%
4.99%	0.5%	3.93%
5.99%	1.0%	4.22%
6.99%	1.5%	4.38%
8.49%	2.2%	4.64%
9.49%	2.5%	4.78%
10.99%	3.0%	5.01%
13.49%	3.5%	5.88%
15.49%	4.0%	6.26%
19.99%	5.0%	6.84%

Note: very low level of risk is marked blue (interest rate ranging from 2.99% to 4.99%), a low level of risk is marked green (interest rate from 5.99% to 10.99%), medium risk level is marked orange (interest rate 13.49%), high level of risk is marked red (interest rate of 15.49% and 19.99%)

Source: Zonky

Secondary market

Since August 2017, investors are enabled to use the other kind of market on the Zonky platform – the secondary market. This one is used by investors to sell investments they already possess to other investors which is usually done in order to gain liquidity immediately (before maturity). Unlike the primary market (e.g. marketplace) that mediates interactions between borrowers and lenders, secondary market connects only investors. Since its launch, the secondary market on Zonky has undoubtedly experienced some development. Originally, to guarantee that offered loan participation was of a decent quality (meaning interesting investment opportunity that was safe to invest in) and could be placed on the secondary market, two conditions regarding the borrower's payment history had to be met – first, payment from the borrower has never been more than one day past due and second, Zonky has already

received at least one instalment (*Spouštíme Sekundární Trh | ZonkyTimes.Cz – Finance, Investice, Rozhovory*, n.d.). Nowadays, it is possible to sell either investments that have been historically at maximum 89 days past due and presently exhibit no repayment problems (this loan participation is sold with 0% to 50 % discount) or investments that are currently up to 90 days past due and not yet labelled as defaulted (*Vylepšené Sekundární Tržiště 3.0 - Jak to Funguje? / Sekundární Trh - Zonky Forum*, n.d.). Such loan participation is again sold with discount. This improvement of the secondary market should serve as another way of increasing the liquidity of the investor's portfolio. In a case investment is sold on the secondary market within 12 months from its original purchase, the investor is obliged to pay a fee of 1.5% of the amount the investor has received from selling his loan participation (*Dokumenty a Smlouvy, Sazebník investora | Zonky*, n.d.).

2.4.2 Borrowing

Loan application

On Zonky, a loan in the range from 20 000 CZK to 750 000 CZK with duration from 6 months to 7 years (84 months) can be requested. The application process shares some similar features with the lender's registration process mentioned already above in section 2.4.1. As well as the lender, the borrower needs to be at least 18 years old, be in possession of a Czech bank account registered under his own name and finally upload two personal documents for identity verification. Moreover, the acknowledgement of the level of income (salary) in the form of an account statement is necessary. Any citizen of the European Union or Ukraine with permanent residence in the Czech Republic can become a loan applicant on Zonky.

After the loan application is completed, Zonky uses banking and non-banking registries to gain as much information about the borrower's payment morale as possible. Based on this "research", Zonky assigns the loan with the interest rate that is in line with Zonky's estimated probability of the borrower's potential inability of paying back and the offer is presented to the applicant who can either reject or accept.

The borrower repays the loan regularly each month. The platform charges the borrower a one-off fee of 2% of the total loan amount which is paid for in the form of

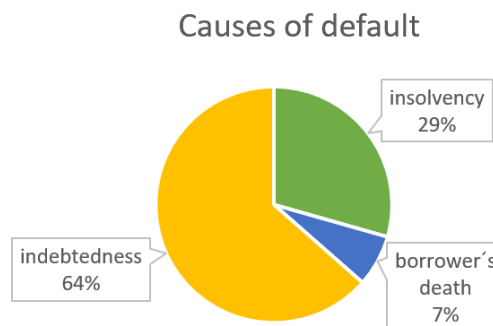
the first few instalments (usually two or three). Nonetheless, there are no restrictions concerning early debt repayment.

Default

In case regular monthly payment from the borrower has not been received, text message and email are automatically sent off to inform him about the delay. Furthermore, the platform attempts to reach the borrower at the telephone number listed in the registration form. Lack of communication from the side of the borrower can be punished by 500 CZK penalty. If the situation proceeds to an undesirable outcome of one payment being more than three months overdue or two or more instalments being past due, extraordinary measures are adopted. As the last resort, Zonky demands to restore remaining debt including not yet paid interest and other contractual penalties at once. In that case, the loan is labelled as defaulted and the borrower can be charged yet additional fees. It is even possible to claim the payment by legal means. Zonky can then keep the sum amounting all the expenses associated with requesting borrower's debt, 30% of the exacted amount at maximum (*Dokumenty a Smlouvy, Sazebník investora | Zonky, n.d.*). The rest is received by investors.

From the total number of 82 131 loans Zonky has ever provided (until June 2020), 1 864 of them were defaulted (which is approximately 2.27 %). Out of these 1 864 loans, 548 of them were defaulted due to insolvency reasons and 132 loans were defaulted because of the borrower's death. In the case of 1184 loans, the default was caused by indebtedness. The proportion of reasons for defaulting is depicted in Figure 5.

Figure 5: Causes of default



Source: Zonky

2.4.3 Zonky Rentier

An issue that has been bothering the platform for some time was the fact demand for borrowing has been outnumbering demand for investing (lending). According to Pavel Novák (CEO of Zonky) people on average borrow 150 000 CZK, whereas average lender invests only 40 000 CZK (*Zonky Chce Nalákat Větší Investory. Spouští Službu, Která Bude Investovat Za Ně, a Plánuje Vstoupit Do Španělska - CzechCrunch*, n.d.). Thus, the remaining demand needed to be covered by an institutional investor. Novák implies that Zonky Rentier, a service launched in November 2019, might help to attract new investors and consequently pull demand and supply closer to equilibrium.

This investment instrument uses a special algorithm developed by Zonky to allocate the amount provided by the investor (minimal investment is 50 000 CZK) in hundreds of loan offers with various ratings and maturities in order to achieve a gross yield of 7.36% p.a. As stated by Zonky, net yield after deduction of fees (investment fee 1.9% (*Dokumenty a Smlouvy, Sazebník investora | Zonky*, n.d.) of each sum invested and the fee of 1.51% p.a. of the total amount invested charged for managing investor's portfolio by the platform) is 4.4% p.a. Any profit made (paid out monthly) is furthermore a subject of 15% income tax.

To be able to participate in Zonky Rentier, the investor is required to meet the same criteria as (any other investor) mentioned in section 2.4.1. Once the money is invested, the algorithm initiates the portfolio creation process. Taking into account that the portfolio consists of at least 490 investments, this procedure takes from 10 to 60 days depending on the relative size of the investment. After 6 months from the initial transfer of money on the Zonky Rentier account, the return on an investment after deduction of fees is regarded reasonable and the investor is permitted to sell his portfolio or some part of it.

This service is especially attractive for investors who do not intend to spend time by creating their own investment strategy and investing manually. Zonky Rentier offers a convenient way of creating a diversified portfolio with a stable and reasonable yield without much demand on investor's time.

3. Literature review

Klafft (2008) and Everett (2015) claim that the environment of peer-to-peer lending does not vary significantly from the traditional banking system, both of those areas are facing similar challenges. Studying a case of Zopa (P2P platform in the United Kingdom), Hulme and Wright (2006) suggest that the concept of online peer-to-peer lending emerging from the urgent search of new forms of relationship in the field of finance in the forthcoming information age can become a severe competitor to the traditional banking sector. But Käfer (2018) perceive P2P lending as riskier compared to the traditional banking. That can be in a way confirmed by de Roure et al. (2016) who reveal that peer-to-peer loans are particularly appealing to borrowers whose loan application has not been approved of by common means (i.e. their loan application has been rejected by traditional financial institutions). However, Klafft (2008) argues that if borrowers with poor credit-rating have not been able to get their loan application accepted by the traditional banking system, they will be also very likely rejected by the P2P platform.

3.1 *Information asymmetry and proposed solutions*

As Bachmann et al. (2011) points out, issue of information asymmetry (a situation in which party on one side of an economic transaction is in possession of greater material knowledge than the party on the other side, typically occurs when the seller (or buyer) of a good or service is endowed with greater knowledge than the other party) is an elementary problem in peer-to-peer lending process. In order to make an informed decision about the investment, lenders want to be provided with the greatest variety of information about the borrower that is possible, while borrowers might be tempted to hide some aspects about themselves or report false data to get a loan with more favourable conditions (especially lower interest rate) (Polák, 2017).

To deal with this challenge, P2P platforms require borrowers to provide truthful information. The validity of those can be rather easily and quickly verified with the help of registries and external agencies. Apart from that, many platforms enable the borrowers to report not only other demographical characteristics such as age, race or

gender but also social information (borrower's hobbies, "friends" (virtually connected people in P2P lending platform), a story describing the purpose of the loan, photos, etc.). Every additional piece of information influences the likelihood of funding success. As described by Herzenstein and Andrews (2008), the borrower's effort put in providing personal information is the most influential factor for funding to become a success.

3.2 Borrower's characteristics and funding success

Research conducted by Bachmann et al. (2011) has shown that among all demographic characteristics, discrimination based on race is the one having the most significant impact on the likelihood of funding. Pope and Sydnor (2008) and Herzenstein and Andrews (2008) suggest that in comparison with other races, African Americans have a lower chance to get their loan funded (according to Pope and Sydnor (2008), their chance is 25 to 34 percent smaller). But Ravina (2007) argues that rather than in the likelihood of loan funding the racial discrimination can be more likely observed in the interest rate borrowers have to pay. Her study claims that in comparison with white borrowers, African Americans pay between 1.39 to 1.46 percent higher interest. On the other hand, Pope and Sydnor (2008) reveal that loan of African Americans are more probable to default and even the higher interest rate charged cannot serve as a sufficient compensation for this potential risk.

Pope and Sydnor (2008) moreover study the relationship between borrower's age and funding success. Treating people aged 35 to 60 as a base group, their findings reveal that individuals who are at least 60 years old are between 110 and 230 basis points (i.e. 1.1 to 2.3 percentage points) less likely to get their loan funded, whereas those under 35 years old appear to have 0.4 to 0.9 percentage points higher probability of successful funding.

Apart from demographic factors (sometimes referred to as hard factors), soft factors can influence the decision of the lender too, however, their validity cannot be verified. Lin (2009) describes soft credit information as *the information about borrowers' riskiness generated by his or her social network in the peer-to-peer lending community*. Based on the definition, photos, friends or group membership on the P2P

platform can be assigned to this category. The division of soft and hard information is also adopted by Chen and Han (2012). As implied by Greiner and Wang (2009), social capital may have in general positive impact on loan funding or interest rate, but it is not beneficial from the point of view of lender's investment decision making as social capital cannot be treated as a good indicator of loan payment. This can be partially confirmed by Klafft (2008) who argues that attaching a photo to a loan request affects predominantly likelihood of loan funding, while the interest rate determination is indifferent to this particular piece of soft information.

Membership in groups (special communities formed on P2P platforms) and virtual connection with other people ("friends") on the P2P platform may also prove as helpful in the lending process. As regards the former, Herrero-Lopez (2009) shows that mere membership in a trustworthy group doubles the likelihood of loan's full funding and according to Berger and Gleisner (2009) and Greiner and Wang (2009) also reduces the interest rate borrower is required to pay. In addition to that, Greiner and Wang (2009), Klein (2008) and Everett (2008) observe that unlike borrowers outside the group, borrowers that are members of a certain community exhibit a somewhat lower default rate. The latter can be supported by findings of Freedman and Jin (2008). They point out that if a loan is funded by borrower's friends, the number of late payments decreases and higher returns on investment are recognized.

But the process of lending might not be determined solely by the borrower's features mentioned above, characteristics of investor matter too. The study by Barasinska (2009) indicates that female lenders in comparison with male lenders are more probable to invest in loans with lower interest rates and lower credit rankings. In other words, female lenders are less risk-averse. For the author herself, this finding was surprising but she suggests the explanation may be found in the altruistic nature of women and purely profit-seeking behaviour of men. Other gender differences have been identified by Pope and Sydnor (2008) – single women are charged a 0.4 percent lower interest rate than men. At the top of that, Ravina (2007) discovers that similarities shared between lender and borrower (such as living in the same city or being of the same gender or ethnicity) have a favourable impact on the likelihood of loan funding.

Interesting finding is provided by Hu et al. (2019). With the use of data from Chinese P2P platform Renrendai, they discovered that instead of paying attention to

information about the borrower (e.g. age, education or income), investors rather focus on loan amount, loan term and borrower's creditworthiness since these were identified as key factors in determining the level of default risk.

3.3 Credit risk

Firms operating in various areas face different risks. As denoted by Aleš Černý (risk director of Czech peer-to-peer platform Zonky) on publicly held workshop organized by Zonky (Zonky, 2017), in case of peer-to-peer platforms, credit risk (the risk of default on a debt when a borrower is not able to pay back his debt) is the major one.

Freedman and Jin (2008) point out that if a loan is funded by a lender whose connection with the borrower on the P2P platform can be classified as friendship, the number of late payments decreases. Being borrower's friend in this virtual environment helps the investor to adequately estimate risks and trustworthiness of debtor and at the same time works as an incentive for loan repayment since investors can monitor borrowers. Greiner and Wang (2009), Klein (2008) and Everett (2008) observe that group membership can also be a factor influencing the timing of loan repayment. Unlike borrowers outside the group, borrowers that are members of a certain community exhibit somewhat lower default rate. Applying logistic regression on data from Lending Club, Serrano-Cinca et al. (2015) find that the purpose of the loan and annual income have a significant impact on the default rate. Polena and Regner (2018) deepen this study by involving various loan risk classes into the regression. Their findings imply that across all risk classes only a few variables can be identified as significant, the significance of the majority of them is associated with the risk class of loan.

3.4 Zonky

Even though the Czech peer-to-peer platform Zonky was founded in 2015, several researchers have already been interested in analysing this platform from many different perspectives.

Hudcová (2017) examines whether the relationship between the risk degree of P2P lending investors and their real behaviour exists. Considering the risk degree of

investors, she finds out that their behaviour deviates considerably from what the theory predicts. Lenders' behaviour is much riskier and furthermore, the only factor that influences an investor's investment decision is apparently the interest rate. This implies that in contrast with findings of lenders on other than Czech peer-to-peer markets, lenders on Zonky are not very much interested in other qualities of a borrower.

Using investment strategy inspired by the modified Markowitz portfolio model, Bock and Tichý (2017) demonstrate that it is possible to make a satisfactory profit (3,4 % p.a.) by investing in peer-to-peer loans on Zonky while keeping the investment portfolio conservative at the same time.

Čermáková (2018) investigates the possible existence of the relationship between personal characteristics of borrowers on Zonky and the probability of default. She concludes that factors having a positive impact on debt repayment are high education level, living in their own house or flat, age over 31 and number of children. Čermáková (2018) also remarks that not paying back the debt is less common in a situation with 2 children at maximum than when the borrower has 3 to 4 children. But it is important to bear in mind that the sociodemographic data acquired from Zonky the author uses are not available to investors. Therefore, these particular results serve better to the P2P platform to recognize "good" and "bad" loans than to help investors estimate the likelihood the loan he has invested in will default.

Based on the Modern portfolio theory (MPT) and rating-based approach, Jonáš (2019) studies a possible way of finding the optimal portfolio on Zonky. Using three different scenarios about the performance of defaulted loans ranging from "no more payments will be received" across "a bit more optimistic scenario" to "high rate of success in recovery operations", he applies portfolio optimization principles and discovers that while in case of "high rate of success in recovery operations" correlation between the interest rate and the expected rate of return is positive, "no more payments will be received" and "a bit more optimistic" scenarios exhibit that expected return of riskier loans is lower than expected return of less risky loans. Apart from confirming that diversification is important, he also supports the view that loans with lower interest rates provide a better risk-return relationship.

4. Methodology

4.1 Portfolio

An investment portfolio can be imagined like a pie that is divided into pieces of diverse sizes, depicting various asset classes and combination of risk levels. In general, a portfolio can consist of only one type of financial asset as well as of combination of many types of assets with various characteristics. Stocks, bonds, and cash are generally considered as basic elements but publicly non-tradable securities like real estate or art can be also included. The process of portfolio creation should be conducted in line with individuals' risk tolerance and investing objectives.

Risk aversion characteristics of each investor are different, the trade-off between additional risk and additional expected return is not the same for all the investors. Therefore, indifference curves showing the same utility level for diverse combinations of risk and return vary among investors. Risk attitude of an individual can be also expressed by the shape of utility function depicting the level of utility as a function of the amount of good consumed. In case an investor is risk-averse, the utility function is strictly concave. For a risk-neutral person, the utility function is linear and eventually, risk-seeking individual is assigned with a strictly convex utility function.

In this thesis, the investment portfolio refers to a combination of loans with various characteristics lender has invested in. The focus of this work does not consist in finding an optimal portfolio with the aim of maximizing expected utility and minimizing costs in the form of financial risk (i.e. the process of portfolio optimization). Nevertheless, some basic terms from portfolio theory need to be explained since they will be henceforth used.

4.2 Portfolio performance

The rational investor requires to be well informed about his portfolio's anticipated performance and consequently its profit or loss. The computation of the expected return can serve as one of the tools used for this purpose. Expected return is a profit or loss investor expects to yield on investment knowing the rates of return.

However, being based on historical data and not taking into account various currently present risks, the ex-ante estimate is only average of the historical returns and thus the anticipated expected return is not guaranteed to be in correspondence with the actual return on investment.

Generally, in order to gain the portfolio's expected return, the summation of the weighted averages of each security's anticipated rates of return (RoR) in the investment portfolio needs to be done. To be more precise, for an investor i , the expected return of his portfolio ER_i can be computed as a weighted average of all his investments in various securities (or loans in our case), where w_i is the weight of given security in the investor's portfolio and r_i is the expected return of the respective security:

$$ER_i = \sum_{i=1}^n w_i * r_i$$

Nevertheless, it is important to emphasize that the assessment of the success and profitability of the investor's portfolio based solely on the expected return estimation is rather unreasonable approach. Performance of two portfolios with identical expected returns but very diverse risk levels will be different. Taking into consideration the risk associated with achieving a particular return is always essential as is mentioned in the study of Marhfor (2016).

To create a list of all existing performance measures would be very challenging if not impossible since new methods of portfolio performance measurement keep evolving. Marhfor (2016) takes the most commonly used measures of financial performance and divides them into two categories – traditional (unconditional) performance measures and conditional performance measures. The traditional measures include Sharpe Ratio, Information Ratio, Modigliani and Modigliani Ratio, Sortino Ratio, Treynor Ratio, Value at Risk Measure, Jensen's Alpha, Treynor & Mazuy Measure and Henriksson & Merton Measure. All these ratios, influenced by the Capital Asset Pricing Model (CAPM), presume that the risk level stays the same over time. Unlike the traditional measures, the conditional measures assume that based on the state of the economy the portfolio risk and market premium changes over the evaluation period. Marhfor (2016) places Fama-French Three-Factor Model, Carhart Four-Factor

Model, Conditional Jensen's Alpha and Conditional GARCH Volatility and Performance Evaluation in opposition to the traditional measures.

4.3 The approach of this thesis

In this thesis, the portfolio performance assessment will be based on the formula the platform Zonky uses to inform the investors about their portfolio performance. The formula is as follows:

$$\left(\sum_{n=0}^N \frac{CF_n}{(1+r)^{\frac{n}{365}}} \right) + \frac{RP_N - EL_N * (1 - RR_N)}{(1+r)^{\frac{N}{365}}} = 0 ,$$

where CF means cash flow, n is the number of days between the cash flow and the beginning of the investment, N is the number of days between the beginning of the investment and the date on which the yield is computed, RP stands for the remaining principal at the end of the period, EL expresses the expected loss at the end of the period, RR is the expected recovery rate and r denotes yield of the investor.

4.4 Description of data

The dataset used in this thesis is of cross-sectional nature (i.e. each observation represents an individual investor and his portfolio). In total, there are 32 870 observations, each carrying information about demographic factors of given investor and his portfolio (ranging from the amount invested and number of investments in loans with various characteristics to average remaining term and portfolio's performance).

In more than 19 000 cases, the value of the portfolio's performance (our dependent variable) was missing. This can be attributed to the Zonky's criteria of portfolio performance computation (the formula was mentioned in the Methodology section) – as long as the investor's portfolio does not contain at least 50 active investment participations being held for 90 days at minimum, portfolio performance is not calculated. Investors without this figure stated were excluded from the dataset and further analysis. Observations with incomplete information on the lender's region of

living, age or gender have been removed as well. As a result, the size of the dataset shrank to 13 353.

The negotiation process with Zonky regarding the data provision was rather lengthy and due to Zonky clients' privacy protection, the scope of the dataset provided was different than expected. Therefore, the initially intended approach needed to be modified.

Characteristics of the investor

In order to protect the privacy of their clients, Zonky has modified the provided dataset in a way that it is impossible to decipher the real identity of any investor. Nevertheless, some basic characteristics of the investor such as age, gender and region of living have been provided. In the model, the variable *age* is measured in years. Then, there are several dummy variables – *male_dummy*, *borrower* and *region_dummy*. *Male_dummy* is a binary variable equal to 1 in case investor is a man and 0 otherwise, *borrower* is a dummy variable equal to 1 if the investor is also the borrower (i.e. investor has borrowed some amount of money on the Zonky platform) and 0 otherwise and *region_dummy* is equal to 1 if the investor's region of living is the Capital of Prague (Prague) and 0 otherwise.

Characteristics of the portfolio

For each investor in the dataset, the portfolio performance (which will be used as a dependent variable) estimated by the formula mentioned above as well as the total number of investments and the total amount invested (invested volume) is known. Furthermore, the dataset contains information about the exact number of investments and the volume invested into loans with various characteristics according to the borrower's source of income, borrower's region of living, the purpose of the loan and the interest rate.

Each of these four characteristics distinguishes between many subcategories. For example, we know that out of 26 investments made in total, the investor has lent 18 times to borrowers in an employment relationship, 7 times to borrowers who are self-employed and 1 time to a borrower whose source of income is his own entrepreneurship and at the same time purchasing a car was denoted as the purpose of

the loan 5 times, household 4 times, project 4 times, refinance 11 times and another purpose 2 times (similarly spread distribution of those 26 investments can be found for the respective subcategories of the borrower's region of living and the interest rate).

To simplify, only the characteristics considered as meaningful and interesting were used directly in the model. This step is based on thorough deliberation of what risky and safety features of a loan might be (the explanation is contained in Table 5,7,9,11 and respective paragraphs below them) in a combination with author's own computation of shares of the respective subcategories in the dataset and sample (those with the share around 10% on average were taken into further consideration with the only exception of pension - its share is lower but because this characteristic is denoted as important it was involved in the model nevertheless). The procedure is presented below (Table 4 to Table 11).

Table 4: Loan distribution in the dataset by the source of income

Source of income	Number of loans	Percent	Amount invested	Percent
Employee	4 330 188	76.17	1 203 918 497	76.50
Self-employed	767 059	13.49	209 881 825	13.34
Pension	264 582	4.65	69 196 348	4.40
Parental allowance	139 424	2.45	39 490 468	2.51
Entrepreneur	119 837	2.11	33 007 105	2.10
Another	42 501	0.75	11 851 746	0.75
Liberal profession	13 445	0.24	4 172 456	0.27
Student	7 348	0.13	2 026 234	0.13
Unemployed	773	0.01	236 525	0.02
Grand Total	5 685 157	100	1 573 781 252	100

Source: author's own computations

Table 4 provides the distribution of loans in the dataset according to the borrower's source of income. Majority of borrowers (76%) whose loans have been successfully funded are in the employment relationship, followed by those who are self-employed (13%) and pensioners (4%).

Table 5: Subcategories of the borrower's source of income

Subcategories	The ones used in the model	In the model represented as
employee, self-employed, pension, entrepreneur, parental allowance, student, unemployed, liberal profession, another source of income	employee	<i>employee</i>
	self-employed	<i>self_empl</i>
	pension	<i>retirement</i>

Based on the research done on the Facebook group “*Investor on Zonky*”, investors themselves consider the fact whether a borrower is in an employment relationship, self-employed or pensioner to some extent play a role in their decision-making process. Therefore, each of these parameters was included in the model as a separate variable. However, due to identifying the multicollinearity problem, only variables *self_empl* and *retirement* were used in the model eventually.

Table 6: Loan distribution in the dataset by the loan purpose

Purpose	Number of loans	Percent	Amount invested	Percent
Household	1 525 522	26.83	427 217 200	27.15
Car	1 375 293	24.19	362 211 150	23.02
Refinance	1 283 759	22.58	381 502 428	24.24
Another	885 885	15.58	240 036 288	15.25
Project	380 358	6.69	105 315 353	6.69
Health	70 476	1.24	17 704 836	1.12
Education	58 047	1.02	14 902 570	0.95
Electronics	54 774	0.96	12 689 601	0.81
Travelling	51 043	0.90	12 201 718	0.78
Grand Total	5 685 157	100	1 573 781 252	100

Source: author's own computations

Table 6 presents an overview of possibilities borrowers have stated as the purpose of the loan. As can be seen, loans for a car, household and refinance are assigned with the greatest weight and their respective proportions sum up to almost 75%

share in the dataset. On the opposite side of the scale, loans used for travelling, educational purposes or electronics with the share around 1% can be found.

Table 7: Subcategories of the loan's purpose

Subcategories	The ones used in the model	In the model represented as
car, household, travelling refinance, electronics, education, health, project, another purpose	car	<i>car_purp</i>
	household	<i>household_purp</i>
	refinance	<i>ref_purp</i>

In the case of the loan's purpose, loans for car, electronics or travelling were chosen as the representatives of the risky type of investment that are probable to decrease the portfolio's quality and subsequently its performance, while the loans with household improvements or refinancing purpose can be assumed to be somewhat safe to invest in. Given the fact that the share of loans for travelling and electronics is on average lower than 1%, these two subcategories were not further taken into account and only variables *car_purp*, *household_purp* and *ref_purp* were included in the model.

Table 8: Loan distribution in the dataset by the interest rate

Interest rate	Number of loans	Percent	Amount invested	Percent
8.49%	961 485	16.91	269 859 128	17.15
10.99%	906 435	15.94	264 832 265	16.83
5.99%	900 511	15.84	221 756 630	14.09
4.99%	659 827	11.61	162 650 447	10.34
13.49%	531 225	9.34	167 266 752	10.63
6.99%	512 506	9.01	137 344 810	8.73
9.49%	483 498	8.50	139 555 600	8.87
15.49%	318 801	5.61	100 190 634	6.37
3.99%	220 112	3.87	51 369 996	3.26
19.99%	146 806	2.58	48 312 704	3.07
2.99%	43 951	0.77	10 642 392	0.68
Grand Total	5 685 157	100	1 573 781 252	100

Source: author's own computations

In Table 8, the portion of loans with various interest rates is shown. The share of loans assigned with the interest rate from 5.99% to 10.99 % (i.e. low risk level) is major - 65%. What is more, the top three values in Table 8 belong to this risk level category as well. On the other hand, the proportion of loans with very low risk level and high risk level of interest rates 2.99% and 19.99% respectively can be found at the bottom of the table.

Table 9: Subcategories of the loan's risk level

Subcategories	The ones used in the model	In the model represented as
very low, low, medium, high	very low	<i>very_low</i>
	medium	<i>medium</i>
	high	<i>high</i>

The whole spectrum of various interest rates can be categorized into several groups. Table 3 in section 2.4.1 shows that Zonky distinguishes between very low, low, medium and high risk level subcategories. We decided to maintain this division and apply it in the following way. As was already mentioned above, the share of low risk level loans in the dataset is major. Thus, this subcategory was chosen as our base group and the remaining subcategories of very low, medium and high risk level were included in the model, denoted as *very_low*, *medium* and *high*, respectively.

Table 10: Loan distribution in the dataset by the borrower's region of living

Region of living	Number of loans	Percent	Amount invested	Percent
Central Bohemian	884 014	15.55	243 705 163	15.49
Moravian-Silesian	657 019	11.56	182 808 806	11.62
Prague	643 608	11.32	173 613 758	11.03
South Moravian	586 025	10.31	161 774 330	10.28
Ústí nad Labem	505 848	8.90	142 299 354	9.04
Olomouc Region	332 548	5.85	91 613 197	5.82
Plzeň	300 839	5.29	84 537 908	5.37
South Bohemian	293 539	5.16	82 192 407	5.22
Zlín	279 206	4.91	75 931 492	4.82
Liberec	267 322	4.70	74 862 851	4.76
Hradec Králové	263 396	4.63	73 283 363	4.66
Pardubice	263 079	4.63	73 746 501	4.69
Vysočina	222 330	3.91	61 812 469	3.93
Karlovy Vary	186 384	3.28	51 599 382	3.28
Grand Total	5 685 157	100	1 573 781 252	100

Source: author's own computations

In Table 10, the distribution of loans according to the borrower's region of living is displayed. As can be seen, the greatest proportion in the dataset belongs to borrowers from the Central Bohemian Region (15%), followed by Moravian-Silesian Region and Prague (both with 11% portion). But the share of the majority of regions does not exceed 6%.

Table 11: Subcategories of the borrower's region of living

Subcategories	The ones used in the model	In the model represented as
Capital of Prague (Prague), Ústí nad Labem Region, South Bohemian Region, Central Bohemian Region, Plzeň Region, Karlovy Vary Region, Liberec Region, Hradec Králové Region, Pardubice Region, Vysočina Region, South Moravian Region, Olomouc Region, Moravian-Silesian Region, Zlín Region	Capital of Prague (Prague)	<i>prague_region</i>
	Ústí nad Labem Region	<i>usti_region</i>

The intuition behind this arrangement is the following – we expect the borrower who lives in the region with a low level of unemployment, high average wage and high quality of life index will be less likely to experience troubles with repaying the debt regularly and in time, whereas the borrower residing in the region that is generally known for its high unemployment rate, low average wage and poor quality of life index is anticipated to be more troublesome. As typical examples of these two opposing groups were chosen Prague (considered as a region that usually denotes a loan that is safe to invest in) and Ústí nad Labem Region (marked as a risky feature of the loan that can possibly decrease the investor's portfolio performance).

Comparison of investors

The focus of this thesis consists in identifying the effect of a lender's demographic factors and the amount invested in loans with various characteristics on the portfolio's performance. However, the investment portfolio of various investors differs. Not only in its composition but also by the quantity totally invested. Therefore, to be able to compare among investors, we examine the structure of their portfolio and describe it with via shares (i.e. the proportion of the portfolio invested in loans with specific characteristics). The level of investment is chosen solely by the investor. For that reason, these portions are expressed primarily in terms of the volume invested. Since the number of investments and the invested amount of any investor is known for each subcategory used in the model, we were able to compute the shares of the

respective subcategories in the investor's portfolio in terms of percent. This was done from two perspectives – firstly, the share of the amount invested in a particular subcategory out of the total invested amount (marked as *amt_sh*) and secondly, the share of the number of investments in a given subcategory out of the total number of investments made (denoted as *cnt_sh*). As was demonstrated in the tables above (Table 4,6,8 and 10), the distribution of the number of specific loans and the amount invested into them is similar. Thus, we will use a model with the share of the number of investments as a robustness check of our results.

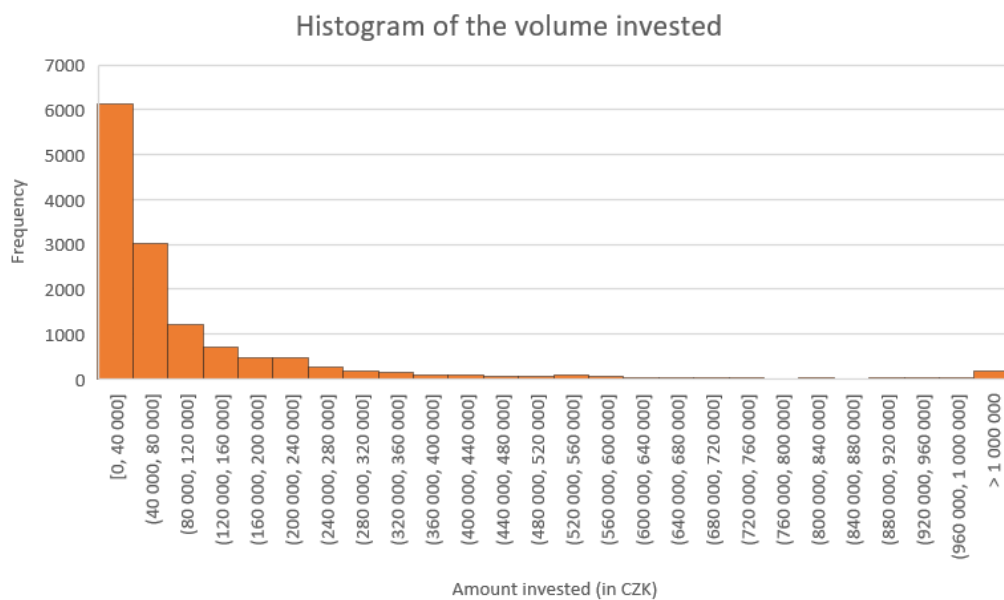
Table 12 summarizes the reflections above and presents all the variables included in the regression in their “shares” form, at which they were used in the model.

Table 12: Description of the regression variables

Variable	Description
Characteristics of the investor	
<i>age</i>	investor's age in years
<i>male_dummy</i>	=1 if investor is a man
<i>region_dummy</i>	=1 if investor's area of residence is Prague
<i>borrower</i>	=1 if investor is a borrower either
Characteristics of the portfolio	
<i>self_empl_amt_sh / self_empl_cnt_sh</i>	the share in the portfolio invested into loans of self-employed borrowers in terms of the quantity invested/number of investments
<i>retirement_amt_sh / retirement_cnt_sh</i>	the share in the portfolio invested into loans of borrowers who are pensioners in terms of the quantity invested/number of investments
<i>prague_region_amt_sh / prague_region_cnt_sh</i>	the share in the portfolio invested into loans of borrowers with Prague denoted as their region of living in terms of the quantity invested/number of investments
<i>usti_region_amt_sh / usti_region_cnt_sh</i>	the share in the portfolio invested into loans of borrowers with Ústí nad Labem Region denoted as their region of living in terms of the quantity invested/number of investments
<i>household_purp_amt_sh / household_purp_cnt_sh</i>	the share in the portfolio invested into loans with the purpose of household improvements in terms of the quantity invested/number of investments
<i>ref_purp_amt_sh / ref_purp_cnt_sh</i>	the share in the portfolio invested into loans with the purpose of refinancing in terms of the quantity invested/number of investments
<i>car_purp_amt_sh / car_purp_cnt_sh</i>	the share in the portfolio invested into loans for car in terms of the quantity invested/number of investments
<i>very_low_amt_sh / very_low_cnt_sh</i>	the share in the portfolio invested into loans assigned to very low risk level subcategory in terms of the quantity invested/number of investments
<i>medium_amt_sh / medium_cnt_sh</i>	the share in the portfolio invested into loans assigned to medium risk level subcategory in terms of the quantity invested/number of investments
<i>high_amt_sh / high_cnt_sh</i>	the share in the portfolio invested into loans assigned to high risk level subcategory in terms of the quantity invested/number of investments

Next, two samples were drawn from the dataset using the criteria of the total invested amount. The samples used are from the opposing sites of the scale – the former (Sample 1) contains investors whose invested amount ranges from 10 000 CZK to 50 000 CZK (i.e. we could call them “small investors”), the latter (Sample 2) is formed of investors with invested volume greater than 1 000 000 CZK. Figure 6 represents the histogram of the invested quantity in the dataset. The threshold 1 000 000 CZK for Sample 2 was chosen deliberately in order to make the difference in the aggregate invested amount as great as possible and subsequently compare samples that are contrasting. Specimen consisting of top 10 percent of investors with the highest amount invested was considered as an alternative of Sample 2. However, the lowest amount invested in this sample is 239 244 CZK and thus the criteria of great contrast in terms of the quantity invested is not met that well. The regression results of this sample are attached in Appendix 4.

Figure 6: Histogram of the quantity investors in the dataset invested



Source: author's own computations

5. Empirical part

Given the nature of the dependent variable (portfolio performance in percent), the Ordinary Least Squares (OLS) model was used in the thesis. Subsequently, CLM assumptions declared by Wooldridge (2015) have been verified. The results of the procedure performed can be found in Appendix 3.

The following regressions estimate the relationship between the investor's portfolio performance and the demographic characteristics of the investor together with various features of the loan. The former tests the effect of a lender's demographic factors and the structure of the investment portfolio in terms of amount invested in loans with various characteristics on the portfolio's performance (model AMT). The latter investigates whether the demographic factors of a lender and the composition of the investment portfolio in terms of the number of investments in various loans have a significant impact on the portfolio's performance (model CNT).

In both models, characteristics of loan are expressed as shares (e.g. the proportion of the investments (invested amount) in respective subcategories out of the total number of investments (totally invested amount), both in percent). It should be also mentioned that other arrangements of loan's characteristics were taken into consideration, added in the model and estimated but due to either multicollinearity or their low proportion in the overall sample, they were eventually excluded in order to produce relevant outcomes.

MODEL AMT (AMT = amount, i.e. quantity invested)

$$\begin{aligned}
 portfolio_perf = & \beta_0 + \beta_1 age + \beta_2 male_dummy + \beta_3 region_dummy + \beta_4 borrower \\
 & + \beta_5 self_empl_amt_sh + \beta_6 retirement_amt_sh \\
 & + \beta_7 prague_region_amt_sh + \beta_8 usti_region_amt_sh \\
 & + \beta_9 household_purp_amt_sh + \beta_{10} ref_purp_amt_sh + \beta_{11} car_purp_amt_sh \\
 & + \beta_{12} very_low_amt_sh + \beta_{13} medium_amt_sh + \beta_{14} high_amt_sh + u
 \end{aligned}$$

MODEL CNT (CNT = count, i.e. number of investments)

$$\begin{aligned}
 portfolio_perf = & \beta_0 + \beta_1 age + \beta_2 male_dummy + \beta_3 region_dummy + \beta_4 borrower \\
 & + \beta_5 self_empl_cnt_sh + \beta_6 retirement_cnt_sh \\
 & + \beta_7 prague_region_cnt_sh + \beta_8 usti_region_cnt_sh \\
 & + \beta_9 household_purp_cnt_sh + \beta_{10} ref_purp_cnt_sh + \beta_{11} car_purp_cnt_sh \\
 & + \beta_{12} very_low_cnt_sh + \beta_{13} medium_cnt_sh + \beta_{14} high_cnt_sh + u
 \end{aligned}$$

Since we wanted to compare the behaviour of two dissimilar groups (described in Chapter 4) and find out whether the factors affecting the portfolio performance of those samples are different, the regressions were run twice – each time for a different sample.

5.1 *Estimated models*

This section presents coefficients estimated for the regression models introduced above. The tables below summarize the relationship between the dependent variable *portfolio_perf* (portfolio performance in percent) and various independent variables (representing the demographic features of the investor and the characteristics of his portfolio) as estimated by the OLS regression. At this point, it should be reminded that any profit made by an investor is a subject of 15% income tax. However, in our analysis, we will relax this requirement and assume that no tax burdens are imposed on the investor's portfolio.

By simple comparison of the results for both samples, it can be seen that firstly, in case of Sample 1 more variables were estimated as significant (and very often at higher significance level) than in respective regressions for Sample 2 and secondly, unlike for Sample 1 the purpose of the investment seems to have no impact on the portfolio performance of the investors from Sample 2. The common feature can be found in the fact that the demographic factors of the investor have been identified as insignificant in all the models except for the statistically significant gender variable *male_dummy* for Sample 2. Moreover, in all the respective regressions, variables *prague_region*, *very_low* and *medium* are labelled as significant.

5.1.1 Estimation results of model AMT

Table 13 captures the estimation results of the two samples using the model with the amount invested.

Table 13: OLS regression - model AMT

	Sample 1		Sample 2	
	Estimate	Standard error	Estimate	Standard error
<i>(intercept)</i>	4.0457 ***	0.2434	4.9711 ***	1.4081
<i>age</i>	-0.0014	0.0018	0.0001	0.0048
<i>male_dummy</i>	0.0704	0.0455	0.4880 **	0.1835
<i>region_dummy</i>	0.0397	0.0384	-0.0062	0.0912
<i>borrower</i>	-0.0640	0.0731	-0.2078	0.1933
<i>self_empl</i>	0.0049	0.0036	-0.0153	0.0176
<i>retirement</i>	0.0386 ***	0.0053	0.0351	0.0231
<i>prague_region</i>	0.0203 ***	0.0047	0.1046 **	0.0365
<i>usti_region</i>	-0.0098 *	0.0046	-0.0276	0.0216
<i>household</i>	0.0012	0.0030	-0.0169	0.0169
<i>ref_purp</i>	0.0025	0.0025	-0.0184	0.0136
<i>car_purp</i>	0.0164 ***	0.0029	-0.0046	0.0158
<i>very_low</i>	-0.0144 ***	0.0013	-0.0121 **	0.0038
<i>medium</i>	0.0055 *	0.0024	0.0256 ***	0.0049
<i>high</i>	0.0194 ***	0.0016	-0.0014	0.0047
F statistic	(df=14; 6986) = 61.67		(df=14; 175) = 8.204	
p-value (F)	< 2.2e-16		2.265e-13	
Residual standard error	(df = 6986) = 1.374		(df = 175) = 0.5727	
N	7001		190	
R ²	0.11		0.3963	
Adjusted R ²	0.1082		0.348	

Note: *** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05; . p-value < 0.1;

The analysis shows, that the differences in the portfolio performance of the “small” investors (Sample 1) can be partially explained using variables *retirement*, *prague_region*, *car_purp*, *very_low*, *high* significant at the 1% level and *usti_region*, *medium* significant at the 5% level.

The interpretation of those is following. In general, raising the share of the amount invested in a loan with given characteristic by 1 percentage point ceteris paribus increases or decreases (depending on the sign of the estimate) the portfolio performance on average by one times the estimated coefficient of characteristic in question percentage points. For example, in case of *retirement*, holding other factors fixed, the rise in the share of the amount lent to borrowers whose source of income is pension by 1 percentage point increases the portfolio performance on average by 0.0386 percentage points. On the other hand, an investor can experience a decrease in his portfolio performance as well. Namely, increasing the share of the amount invested in either loans from very low risk level subcategory or loans with borrower’s area of residence being the Ústí nad Labem Region by 1 percentage point ceteris paribus lowers *portfolio_perf* on average by 0.0144 and 0.0098 percentage points, respectively.

The negative sign of *very_low* may be a bit surprising at first. However, it could be explained by the fact that while the average portfolio performance of the investors in Sample 1 is 4.9%, the expected yield of investing in a loan from very low risk subcategory ranges from 2.34% to 3.93% (as shows Table 3). So, investing in such loans pulls the portfolio performance lower than it on average should be.

In case of Sample 2 (excluding constant), *medium* is statistically highly significant, meaning investors who increase the share of the amount invested in loans from medium risk level category by 1 percentage point will ceteris paribus on average experience 0.0256 percentage point increase in *portfolio_perf*. The important role of the level of risk is further supported by the significance of *very_low*. But compared to the estimate of *medium*, the coefficient of *very_low* implies that, holding other factors fixed, investors who increase the share of the amount invested in loans assigned with very low risk level by 1 percentage point can anticipate on average 0.0121 percentage point decrease in their portfolio performance.

Male_dummy is the only representative of the demographic features of investors in Sample 2 evaluated as having an impact on the investor's portfolio performance. It is significant at the 1% level, and also causes the biggest nominal change in the dependent variable. The estimate shows that holding other factors fixed, portfolio performance of men should be on average 0.488 percentage points higher compared to women. Variable *prague_region* is significant at the 1% level as well with the following interpretation – holding other factors fixed, the rise in the share of the quantity lent to borrowers whose region of living is Prague by 1 percentage point improves the portfolio performance by 0.1046 percentage points on average.

Variable *high* was also expected to be estimated as important but its statistical insignificance and negative sign implies that for investors with great quantity invested, increasing the share of the amount invested in high risk level loans is not worth the risk.

5.1.2 Estimation results of model CNT

Table 14 presents the estimation results of Sample 1 and Sample 2 using the model with the number of investments.

Table 14: OLS regression - model CNT

	Sample 1		Sample 2	
	Estimate	Standard error	Estimate	Standard error
<i>(intercept)</i>	4.2037 ***	0.2532	3.9217 *	1.7360
<i>age</i>	-0.0011	0.0018	-0.0007	0.0049
<i>male_dummy</i>	0.0597	0.0454	0.4137 *	0.1877
<i>region_dummy</i>	0.0365	0.0383	0.0043	0.0943
<i>borrower</i>	-0.0621	0.0729	-0.2176	0.1959
<i>self_empl</i>	-0.0003	0.0040	-0.0370 .	0.0215
<i>retirement</i>	0.0472 ***	0.0055	0.0470 .	0.0261
<i>prague_region</i>	0.0281 ***	0.0053	0.1475 **	0.0490
<i>usti_region</i>	-0.0142 **	0.0053	-0.0055	0.0273
<i>household_purp</i>	-0.0018	0.0032	-0.0028	0.0220
<i>ref_purp</i>	0.0020	0.0027	-0.0150	0.0159
<i>car_purp</i>	0.0147 ***	0.0031	-0.0006	0.0192
<i>very_low</i>	-0.0157 ***	0.0014	-0.0129 **	0.0045
<i>medium</i>	0.0054 *	0.0026	0.0208 **	0.0066
<i>high</i>	0.0192 ***	0.0017	0.0005	0.0061
F statistic	F statistic (df=14; 6986) = 65.53		(df=14; 175) = 6.848	
p-value (F)	< 2.2e-16		4.42e-11	
Residual standard error	(df=14; 6986) = 1.369		(df = 175) = 0.5925	
N	7001		190	
R ²	0.1161		0.3539	
Adjusted R ²	0.1143		0.3022	

Note: *** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05; . p-value < 0.1;

In term of the significance levels, regression results of Sample 1 in model CNT match with those of model AMT. The only exception is *usti_region*, its significance shifts upwards from the 5% level to the 1% level. The estimated coefficients are occasionally slightly different, but the signs stay unchanged.

Compared to the model AMT, the interpretation of significant variables (*retirement*, *prague_region*, *car_purp*, *very_low*, *high*, *usti_region*, *medium*; intercept not included) is a bit different – holding other factors fixed, the growth in the share of the number of investments in a loan with given characteristic by 1 percentage point increases or decreases (depending on the sign of the estimate) the portfolio performance on average by one times the estimated coefficient of characteristic in question percentage points. For instance, the coefficient of *prague_region* implies that 1 percentage point increase in the share of the number of investments in loans with borrower's region of living being Prague ceteris paribus raises the portfolio performance on average by 0.0281 percentage points. Even though the coefficient of *borrower* is estimated to have the greatest nominal impact on the dependent variable, this variable is insignificant. Choosing from the significant variables (not including intercept), the greatest change in *portfolio_perf* can be attributed to *retirement*.

As regards Sample 2 variables *very_low*, *medium* and *prague_region*, remain significant at the 1% level. The negative estimate of the former means that investors who increase the share of the number of investments to very low risk level subcategory by 1 percentage point will ceteris paribus experience on average 0.0129 percentage point decrease in *portfolio_perf*. The coefficients of the latter can be interpreted in the following way – holding other factors fixed, 1 percentage point growth in either the share of the number of investments to medium risk level subcategory or in the share of the number of investments to borrowers with region of living being Prague will increase on average portfolio performance by 0.0208 and 0.1475 percentage points, respectively.

Compared to the model AMT, the significance level of *male_dummy* changes from the 1% level to the 5% level. The coefficient of *male_dummy* is also the bearer of the greatest nominal change in the dependent variable among significant variables. Compared to women, portfolio performance of men should be on average ceteris paribus 0.4137 percentage points higher.

Retirement and *self_empl* are both significant at the 10% level, suggesting the borrower's source of income might play a role as well. Similarly as in model AMT, even though in the case of Sample 2 variable *high* was anticipated to be estimated as important, it was determined as insignificant. The same explanation applies - investors with high volume invested do not perceive increasing the share of the amount invested in high risk level loans worthy the level of risk.

The negative signs of *very_low* in the regression results for both samples can be a bit unintuitive at first, but the same reasoning as in the case of model AMT can be provided.

Majority of the investors in Sample 2 (around 95 percent) have invested up to the total amount of 6 million CZK. However, a few outlying values were identified. Since outliers can have a serious effect on the OLS estimates, those observations were removed and the respective regressions were estimated again. The results do not differ considerably from the ones presented in Table 13 and Table 14. Variable *male_dummy* in model AMT was assigned with lower significance level and *retirement* in model CNT became insignificant but otherwise, the estimated coefficients were almost identical and their signs did not change at all.

6. Conclusion

This thesis examined which of the demographic features of investor and loan's characteristics have a significant impact on the portfolio performance. The analysis was performed on two different samples – “small” investors with the quantity invested ranging from 10 000 CZK to 50 000 CZK (Sample 1) and lenders with the total invested amount exceeding 1 000 000 CZK (Sample 2). The following findings can be concluded.

Firstly, it was discovered that for both groups the fact whether investors lend to a borrower whose region of living is Prague or invest in a loan with medium and very low risk level subcategory are statistically significant (the first and the second increasing the portfolio performance while the third lowering it). Whereas in Sample 1, these as well as four other factors (borrower being a pensioner, car denoted as the purpose of the loan, borrower's area of residence being the Ústí nad Labem Region and high risk level subcategory) were assigned with high significance level, for Sample 2 fewer additional variables were estimated as significant and their significance level is almost exclusively lower. This implies the second key finding - compared to the investors with the greatest invested volumes, portfolio performance of “small” investors is likely more sensitive to the various loan's characteristics.

Thirdly, the regression results of Sample 2 reveal that portfolio performance of investors with quantity invested greater than 1 000 000 CZK is not affected by the purpose of the loan at all. Moreover in general, demographic characteristics of investors were denoted as insignificant with the only exception of gender dummy marked as significant in both regressions for Sample 2.

Hudcová (2017) who investigated on the determinants influencing the expected return of the investors on the Czech P2P platform Zonky concludes that high interest rate is the only determinant of the final image of the lender's portfolio. The results of this thesis show that the level of the interest rate is indeed highly influential when it comes to the portfolio performance. But unlike Hudcová (2017), this study furthermore suggests that investors also pay attention to other factors such as the borrower's qualities and the characteristics of a loan and is therefore in partial agreement with the finding from abroad as well (e.g. Bachmann et al. (2011)).

The valuable contribution of this paper is in the fact that the dataset used for the analysis was provided directly by the Czech peer-to-peer platform Zonky. Conducting an online survey was not necessary. Thus, our sample of investors is unbiased.

Since the results imply that the factors influencing the portfolio performance of investors with different invested quantity are not identical, it could be interesting to divide the investors into several groups according to the total amount invested and compare the highly significant characteristics among them. Future research can also focus on the robustness of our results by using one of the variety of ratios evaluating the portfolio performance mentioned in the Methodology section as a dependent variable and challenge the findings of this thesis. Last but not least, as generally believed (and on the field of the Czech P2P lending market confirmed by Hudcová (2017) and also by this study), the loan's interest rate affects the image of the lender's portfolio greatly. Thus, incorporating particular interest rates together with only some additional characteristics of a borrower (for example only the source of income) in the model could help to reveal the attitude of investors towards the loan's interest rate even more.

References

- Adroit Market Research (2017). *Global Peer to Peer (P2P) Lending Market Size by Type (Business Lending and Consumer Lending), by Region (North America, Europe, Asia Pacific, Latin America, and Middle-East & Africa) and Forecast 2018 to 2025*.
- Albors, J., Ramos, J. C. and Hervas, J. L. (2008) 'New learning network paradigms: Communities of objectives, crowdsourcing, wikis and open source', *International Journal of Information Management*, 28(3), pp. 194–202. doi: 10.1016/j.ijinfomgt.2007.09.006.
- Ashta, A. and Assadi, D. (2009) 'An Analysis of European Online micro-lending Websites', *Cahiers du CEREN*, 29(June), pp. 147–160. doi: citeulike-article-id:12156499.
- auto bidding | P2P-Banking*. (n.d.). Retrieved July 20, 2020, from <https://www.p2p-banking.com/tag/auto-bidding/>
- Bachmann, A., Becker, A., Buerckner, D., and Hilker, M. K. (2011) 'Online peer-to-peer lending - a literature review', *Journal of Internet Banking and Commerce*, 16(2), pp. 1–18. doi: 10.1007/978-3-531-92534-9_12.
- Barasinska, N. (2009). 'The Role of Gender in Lending Business: Evidence from an Online Market for Peer-to-Peer Lending'. *The New York Times*. Berlin.
- Belk, R. (2010). 'Sharing'. *Journal of Consumer Research*. 36(5), pp. 715-734.
- Belk, R. (2014) 'Sharing versus pseudo-sharing in web 2.0', *Anthropologist*, 18(1), pp. 7–23. doi: 10.1080/09720073.2014.11891518.
- Berger, S. C. and Gleisner, F. (2009) 'Emergence of Financial Intermediaries in Electronic Markets: The Case of Online P2P Lending BuR - Business Research', *Official Open Acces Journal of VHB*, 2(1), pp. 39–65.
- Bock, M., T. J. (2017) 'Assessment of Investor's Portfolio of P2P Loans', *ACTA VSFS*, 11(2), pp. 121–143.
- Chen, D. and Han, C. (2012) 'A comparative study of online P2P lending in the USA and China', *Journal of Internet Banking and Commerce*, 17(2), pp. 1 – 15
- Čermáková, J. (2018). *Credit Risk of P2P Lending on the Czech Market*. B.A. Thesis. Charles University.

Deloitte (2016). *A temporary phenomenon? Marketplace lending. An analysis of the UK market.*

Deloitte (2017). *Sdílená ekonomika. Bohatství bez vlastnictví.*

Dokumenty a smlouvy, Parametry částek pro investování | Zonky. (n.d.). Retrieved April 5, 2020, from <https://zonky.cz/dokumenty#investori>

Dokumenty a smlouvy, Sazebník investora | Zonky. (n.d.). Retrieved April 14, 2020, from <https://zonky.cz/dokumenty#investori>

EU Peer-to-Peer Lending & Equity Funding Volumes | P2PMarketData. (n.d.). Retrieved July 20, 2020, from <https://p2pmarketdata.com/p2p-lending-funding-volume-eu/>

Everett, C. R. (2008) 'Group membership, relationship banking and loan default risk: The case of online social lending', *Banking and Finance Review*, 7(2), pp. 15–54.

Everett, C. R. (2015) 'Group membership, relationship banking and loan default risk: The case of online social lending', *Banking and Finance Review*, 7(2), pp. 15–54. doi: 10.2139/ssrn.1114428.

Financial Stability Report 2018/2019 - Czech National Bank. (n.d.). Retrieved April 6, 2020, from <https://www.cnb.cz/en/financial-stability/fs-reports/financial-stability-report-2018-2019/>

Freedman, S. and Jin, G. Z. (2008) *Do Social Networks Solve Information Problems for Peer-to-Peer Lending? Evidence from Prosper.com. Technical report., SSRN Electronic Journal.* doi: 10.2139/ssrn.1936057.

Frenken, K. and Schor, J. (2017) 'Putting the sharing economy into perspective', in *Environmental Innovation and Societal Transitions.* Elsevier B.V., pp. 3–10. doi: 10.1016/j.eist.2017.01.003.

Galloway, I. (2009) 'Peer-to-peer lending and community development finance', *Community Investments*, 21(3), pp. 19–23.

Greiner, M. E., and Wang, H. (2009) 'The Role of Social Capital in People-to-People Lending Marketplaces', *Thirtieth International Conference on Information Systems*, Phoenix, Arizona, USA, 15 December to 18 December 2009.

Herrero-Lopez, S. (2009) 'Social Interactions in P2P Lending', *The 3rd Workshop on Social Network Mining and Analysis*, Paris, France, 28 June 2009.

Herzenstein, M. and Andrews, R. L. (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), pp. 1–45. doi: 10.1002/nml.

- Hu, R., Liu, M., He, P. and MA, Y. (2019) 'Can Investors on P2P Lending Platforms Identify Default Risk? ', *International Journal of Electronic Commerce*, 23(1), pp. 63-84
- Hudcová, T. (2017). *Analysis of Czech P2P lending investor' behaviour drivers*. B.A. Thesis. Charles University.
- Hulme, M. and Wright, C. (2006) 'Internet based social lending: Past, present and future', *Social Futures Observatory*, 11(10), pp. 1–115. Available at: <http://karz.googlecode.com/svn-history/r105/trunk/reference/internetbasedsociallending.pdf>.
- Islam, M. R. (2018) 'Sample size and its role in Central Limit Theorem (CLT) ', *International Journal of Physics & Mathematics*, 1(1), pp. 37-47.
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An introduction to statistical learning: with applications in R*. New York: Springer
- Jonáš, F. (2019). *Portfolio optimization for an P2P investor on Zonky*. B.A. Thesis. Charles University.
- Käfer, B. (2018) *Peer-to-Peer Lending - A (Financial Stability) Risk Perspective*, *Review of Economics*. doi: 10.1515/roe-2017-0020.
- Klauff, M. (2008). *Peer to peer lending: auctioning microcredits over the internet*.
- Klein, T. (2008). *Performance in Online Lending Platforms*.
- Klient: často kladené otázky | Zonky. (n.d.). Retrieved April 5, 2020, from <https://zonky.cz/otazky-a-odpovedi#odpoved31>
- Li, Z., Wu, L. and Tang, H. (2018) 'Optimizing the Borrowing Limit and Interest Rate in P2P System: From Borrowers' Perspective', *Scientific Programming*, pp. 1-14
- Lin, M. (2009) 'Peer-to-Peer lending: An Empirical Study', *15th Americas Conference on Information Systems*, San Francisco, CA, USA, 6 August to 8 August 2009.
- Marhfor, A. (2016) 'Portfolio Performance Measurement: Review of Literature and Avenues of Future Research', *American Journal of Industrial and Business Management*, 06(04), pp. 432–438. doi: 10.4236/ajibm.2016.64039.
- Polák, P. (2017). *Portfolio diversification on p2p loan markets*. M.A. Thesis. Charles University.
- Polena, M. and Regner, T. (2018) 'Determinants of Default in P2P Lending under Consideration of the Loan Risk Class', *Games*, 9(4), pp. 1–82.
- Pope, D. G. and Sydnor, J. R. (2008) 'What's in a Picture? Evidence of Discrimination from Prosper.com', *Journal of Human Resources*, 46(1), pp. 53–92.

Poplatek dlužníka | Bankerat. (n.d.). Retrieved April 10, 2020, from <https://www.bankerat.cz/pujcka/poplatek-dluznika/>

Poplatek věřitele | Bankerat. (n.d.). Retrieved April 10, 2020, from <https://www.bankerat.cz/investice/poplatek-veritele/>

Ranjbari, M., Morales-Alonso, G. and Carrasco-Gallego, R. (2018) ‘Conceptualizing the sharing economy through presenting a comprehensive framework’, *Sustainability (Switzerland)*, 10(7). doi: 10.3390/su10072336.

Ravina, E. (2007) ‘Beauty, Personal Characteristics, and Trust in Credit Markets’, *SSRN Electronic Journal*. doi: 10.2139/ssrn.972801.

de Roure, C., Pelizzon, L. and Tasca, P. (2016) ‘How Does P2P Lending Fit into the Consumer Credit Market?’, *SSRN Electronic Journal*, 3(30). doi: 10.2139/ssrn.2756191.

Serrano-Cinca, C., Gutierrez-Nieto, B. and Lopez-Palacios, L. (2015) ‘Determinants of Default in P2P Lending’, *PLoS ONE*, 10(10), pp. 1–22.

Spouštíme sekundární trh | ZonkyTimes.cz – finance, investice, rozhovory. (n.d.). Retrieved April 12, 2020, from <https://www.zonky.cz/zonkytimes/spoustime-sekundarni-trh/>

Sundararajan, A. (2016). *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. Cambridge, Massachusetts: MIT Press

Sutherland, W. and Jarrahi, M. H. (2018) ‘The sharing economy and digital platforms: A review and research agenda’, *International Journal of Information Management*, 43(September), pp. 328–341. doi: 10.1016/j.ijinfomgt.2018.07.004.

UK Peer-to-Peer Lending & Equity Funding Volumes | P2PMarketData. (n.d.). Retrieved July 20, 2020, from <https://p2pmarketdata.com/p2p-lending-funding-volume-uk/>

USA Peer-to-Peer Lending & Equity Crowdfunding Statistics (USD). (n.d.). Retrieved July 20, 2020, from <https://p2pmarketdata.com/p2p-lending-funding-volume-usa/>

Veřejný rejstřík a Sbirka listin - Ministerstvo spravedlnosti České republiky. (n.d.). Retrieved April 7, 2020, from <https://or.justice.cz/ias/ui/rejstrik-firma.vysledky?subjektId=880547&typ=UPLNY>

Vylepšené sekundární tržiště 3.0 - Jak to funguje? / Sekundární trh - Zonky Forum. (n.d.). Retrieved April 13, 2020, from <https://forum.zonky.cz/t/vylepsene-sekundarni-trziste-3-0/606>

Vysušil, T. (2019). *Peer-to-peer lending as a substitute to conventional investment products*. B.A. Thesis. Charles University.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. 6th edition. Mason, OH: Thomson/South-Western

Zonky (2017) *Riskový workshop Zonky: jak chytré investovat*. Available at: <https://www.youtube.com/watch?v=WLtrAnC4aN0> (Accessed: 10 April 2020)

Zonky chce nalákat větší investory. Spouští službu, která bude investovat za ně, a plánuje vstoupit do Španělska - CzechCrunch. (n.d.). Retrieved April 16, 2020, from <https://www.czechcrunch.cz/2019/11/zonky-chce-nalakat-vetsi-investory-spousti-sluzbu-ktera-bude-investovat-za-ne-a-planuje-vstoupit-do-spanelska>

Zonky je nejdůvěryhodnější značka v oblasti nebankovních půjček | ZonkyTimes.cz – finance, investice, rozhovory. (n.d.). Retrieved April 6, 2020, from <https://zonky.cz/zonkytimes/zonky-je-nejduveryhodnejsi-znacka-v-oblasti-nebankovnich-pujcek/>

List of appendices

Appendix 1: Marketplace on Zonky


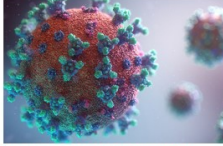




Appendix 2: Detail of loan from the marketplace from Appendix 1

Appendix 3: CLM assumptions verification

Appendix 4: Regression results of top 10 percent of investors

Appendices

Appendix 1: The marketplace on Zonky

<p>Investovat můžete ještě 1 den 20 h 15 m</p>  <p>Refinancování nevýhodné půjčky</p> <p>Č. klienta 1000268250 Požádáno</p> <p>Chci si snížit úroky z nevýhodné půjčky a splatit kreditní kartu.</p> <p>Investováno 24 % z 90 000 Kč</p> <table border="0"> <tr> <td>Roční úrok</td> <td>Po odečtení poplatku</td> <td>Délka splácení</td> </tr> <tr> <td>2,99 %</td> <td>2,79 %</td> <td>84 měsíců</td> </tr> </table>	Roční úrok	Po odečtení poplatku	Délka splácení	2,99 %	2,79 %	84 měsíců	<p>Investovat můžete ještě 1 den 20 h 10 m</p>  <p>Splacení kreditek s velkým úrokem.</p> <p>Č. klienta 1000110963</p> <p>Splatím kredity s velkým úrokem + refinancuji jeden zonky úvěr. Snížím si...</p> <p>Investováno 20 % z 200 000 Kč</p> <table border="0"> <tr> <td>Roční úrok</td> <td>Po odečtení poplatku</td> <td>Délka splácení</td> </tr> <tr> <td>2,99 %</td> <td>2,79 %</td> <td>84 měsíců</td> </tr> </table>	Roční úrok	Po odečtení poplatku	Délka splácení	2,99 %	2,79 %	84 měsíců	<p>Investovat můžete ještě 09:00:37</p>  <p>Investici do bytu pro dceru</p> <p>Č. klienta 1000075260</p> <p>Půjčku využiji k nákupu bytu pro dceru. Je jí 8 let, takže nejbližších 10-15 let bu...</p> <p>Investováno 12 % z 750 000 Kč</p> <table border="0"> <tr> <td>Roční úrok</td> <td>Po odečtení poplatku</td> <td>Délka splácení</td> </tr> <tr> <td>2,99 %</td> <td>2,79 %</td> <td>84 měsíců</td> </tr> </table>	Roční úrok	Po odečtení poplatku	Délka splácení	2,99 %	2,79 %	84 měsíců
Roční úrok	Po odečtení poplatku	Délka splácení																		
2,99 %	2,79 %	84 měsíců																		
Roční úrok	Po odečtení poplatku	Délka splácení																		
2,99 %	2,79 %	84 měsíců																		
Roční úrok	Po odečtení poplatku	Délka splácení																		
2,99 %	2,79 %	84 měsíců																		
<p>Investovat můžete ještě 1 den 23 h 44 m</p>  <p>Muzikoterapie</p> <p>Č. klienta 1000274628</p> <p>Kvalitní nástroje pro muzikoterapii stojí velké částky, a pokud mám možnost vy...</p> <p>Investováno 21 % z 100 000 Kč</p> <table border="0"> <tr> <td>Roční úrok</td> <td>Po odečtení poplatku</td> <td>Délka splácení</td> </tr> <tr> <td>2,99 %</td> <td>2,79 %</td> <td>84 měsíců</td> </tr> </table>	Roční úrok	Po odečtení poplatku	Délka splácení	2,99 %	2,79 %	84 měsíců	<p>Investovat můžete ještě 08:51:50</p>  <p>Rodinná investice</p> <p>Č. klienta 1000047992</p> <p>Občas přijde příležitost</p> <p>Investováno 12 % z 660 000 Kč</p> <table border="0"> <tr> <td>Roční úrok</td> <td>Po odečtení poplatku</td> <td>Délka splácení</td> </tr> <tr> <td>2,99 %</td> <td>2,79 %</td> <td>84 měsíců</td> </tr> </table>	Roční úrok	Po odečtení poplatku	Délka splácení	2,99 %	2,79 %	84 měsíců	<p>Investování jen přes rezervace 1 den 15 h 11 m</p>  <p>Dodelání prací na domě</p> <p>Č. klienta 1000348136</p> <p>Marmolit na soklu, okapový chodník, základy pro terasu, skleník a přístřeš...</p> <p>Investováno 82 % z 50 000 Kč</p> <table border="0"> <tr> <td>Roční úrok</td> <td>Po odečtení poplatku</td> <td>Délka splácení</td> </tr> <tr> <td>3,99 %</td> <td>3,79 %</td> <td>66 měsíců</td> </tr> </table>	Roční úrok	Po odečtení poplatku	Délka splácení	3,99 %	3,79 %	66 měsíců
Roční úrok	Po odečtení poplatku	Délka splácení																		
2,99 %	2,79 %	84 měsíců																		
Roční úrok	Po odečtení poplatku	Délka splácení																		
2,99 %	2,79 %	84 měsíců																		
Roční úrok	Po odečtení poplatku	Délka splácení																		
3,99 %	3,79 %	66 měsíců																		

Appendix 2: Detail of loan from marketplace from Appendix 1

Investorů	Investováno	Roční úrok (rating)	Po odečtení poplatku	Délka splácení	Výše splátky	Zbývá hodin
22	24 % z 90 000 Kč	2,99 %	2,79 %	84 měs.	1 320 Kč/měs.	44:16:20



REFINANCOVÁNÍ NEVÝHODNÉ PŮJČKY

č. klienta 1000268250
Zaměstnanec Královéhradecký kraj

2. půjčka na Zonky Ověřený v registrech Ověřený příjem Pojištěno

Chci si snížit úroky z nevýhodné půjčky a splatit kreditní kartu.

Dotazy (1) Investoři (22)

[Půjčte si na refinancování](#) [Pro investování se přihlaste](#)

Appendix 3: CLM model assumptions verification

Multiple Linear Regression (MLR) assumptions:

- 1) **Linearity in parameters** – Dependent variable (in our case portfolio performance) can be expressed as a linear combination of the independent variables and the error term u .
- 2) **Random sampling** – We have a random sample of size n following the population model.
- 3) **No perfect collinearity** – In the sample (as well as in the population), none of the independent variables is constant and there are no exact linear relationships among the independent variables.

To be able to eliminate the presence of multicollinearity between independent variables, VIF (Variance inflation factor) test was conducted. According to James et al. (2013), the value of VIF greater than 5 suggests multicollinearity may be present. As can be seen for both samples from the tables below, the critical value of 5 is not exceeded in any model and thus multicollinearity problem is not an issue.

Sample 1 model AMT			
Variable	VIF	Variable	VIF
age	1.026538	usti_region_amt_sh	1.124514
male_dummy	0.020522	household_purp_amt_sh	2.343379
region_dummy	1.010835	ref_purp_amt_sh	2.433148
borrower	1.008445	car_purp_amt_sh	1.920024
self_empl_amt_sh	1.198475	very_low	1.624015
retirement_amt_sh	1.182586	medium	1.911943
prague_region_amt_sh	1.145577	high	1.666707

Sample 1 model CNT			
Variable	VIF	Variable	VIF
age	1.024867	usti_region_cnt_sh	1.153661
male_dummy	1.019805	household_purp_cnt_sh	2.451210
region_dummy	1.011317	ref_purp_cnt_sh	2.499507
borrower	1.008381	car_purp_cnt_sh	2.014237
self_empl_amt_sh	1.240271	very_low	1.668653
retirement_amt_sh	1.218114	medium	2.077654
prague_region_amt_sh	1.191365	high	1.783348

Sample 2 model AMT			
Variable	VIF	Variable	VIF
age	1.087938	usti_region_amt_sh	1.576336
male_dummy	1.064456	household_purp_amt_sh	3.799474
region_dummy	1.101677	ref_purp_amt_sh	4.482515
borrower	1.079477	car_purp_amt_sh	4.018103
self_empl_amt_sh	1.395085	very_low	1.673550
retirement_amt_sh	1.840307	medium	2.096932
prague_region_amt_sh	1.714808	high	2.042840

Sample 2 model CNT			
Variable	VIF	Variable	VIF
age	1.077787	usti_region_cnt_sh	1.795243
male_dummy	1.041050	household_purp_cnt_sh	4.180001
region_dummy	1.098967	ref_purp_cnt_sh	4.950254
borrower	1.036346	car_purp_cnt_sh	4.036563
self_empl_cnt_sh	1.575020	very_low	1.893101
retirement_cnt_sh	2.046162	medium	3.018915
prague_region_cnt_sh	2.226822	high	2.736459

- 4) **Zero conditional mean** – The error u has an expected value of zero given any values of the explanatory variables, i.e. $E(u|X) = 0$.

- 5) **Homoskedasticity** – The error u has the same variance given any values of the explanatory variables, i.e. $\text{Var}(u|X) = \sigma^2$ (or in other words the variance of errors is constant).

To verify this assumption, the Breusch-Pagan test with the null hypothesis of homoskedasticity (H_0 : homoskedasticity vs. H_A : heteroskedasticity) was performed. In each respective model for both samples, the p-value was sufficiently small and therefore, we were able to reject the null hypothesis (in academic research, the threshold p-value is typically 0.05 but we were able to reject the null hypothesis even at 0.01 significance level). Since the violation of the homoskedasticity assumption causes the confidence intervals, t-statistic, F-statistics and LM statistic to be invalid, White's standard errors have been applied in order to correct for the heteroskedasticity problem.

- 6) **Normality** – The population error u is independent of the explanatory variables x_1, \dots, x_k and is normally distributed with zero mean and variance σ^2 , i.e. $u \sim N(0, \sigma^2)$.

To check the normality assumption, Kolmogorov-Smirnov test, Shapiro-Wilk test, Anderson-Darling test, Jarque-Bera test and many others with the H_0 : sample distribution is normal vs H_A : the distribution is non-normal can be conducted. Even in a case the null hypothesis of normality is rejected, the violation of the normality assumption should not cause major problems, as long as some criteria regarding the sample size are met.

According to the Central Limit Theorem, regardless of the initial shape of the data if the sample is large enough the sampling distribution will be normal. The “magical” threshold of the sample is typically considered to be greater than 30. However, as Islam (2018) points out to become fixated to this number of observations is not reasonable because in fact there is very little evidence supporting the claim that the sample size of 30 is the magical line separating normal and non-normal distributions. But on the other hand, he concludes that “*most statisticians agree on that if the parent distribution is symmetric and relatively short-tailed, then the sample mean reaches*

approximate normality for smaller samples than if the parent population is skewed or long-tailed”.

As the size of our respective samples is hundreds and even thousands of observations, we will apply the Central Limit Theorem and assume that normality is met.

Appendix 4: Regression results of top 10 percent of investors

	Model AMT		Model CNT	
	Estimate	Standard error	Estimate	Standard error
<i>(intercept)</i>	4.5262 ***	0.5415	4.9225 ***	0.6081
<i>age</i>	-0.0040 .	0.0024	-0.0047 .	0.0024
<i>male_dummy</i>	0.0726	0.0880	0.0611	0.0882
<i>region_dummy</i>	0.0102	0.0479	0.0181	0.0481
<i>borrower</i>	-0.1394	0.1120	-0.1513	0.1125
<i>self_empl</i>	0.0102	0.0079	0.0036	0.0093
<i>retirement</i>	0.0207 *	0.0101	0.0215 *	0.0107
<i>prague_region</i>	0.0481 ***	0.0128	0.0558 ***	0.0156
<i>usti_region</i>	-0.0247 *	0.0101	-0.0238 *	0.0114
<i>household_purp</i>	-0.0041	0.0067	-0.0083	0.0076
<i>ref_pur</i>	-0.0009	0.0055	-0.0066	0.0061
<i>car_purp</i>	0.0177 **	0.0064	0.0131 .	0.0071
<i>very_low</i>	-0.0168 ***	0.0018	-0.0177 ***	0.0019
<i>medium</i>	0.0097 ***	0.0026	0.0093 **	0.0028
<i>high</i>	0.0141 ***	0.0019	0.0137 ***	0.0020
F statistic	(df=14; 1344) = 30.05		(df=14; 1344) = 29.48	
p-value (F)	< 2.2e-16		< 2.2e-16	
Residual standard error	(df = 1344) = 0.8284		(df = 1344) = 0.8302	
N	1359		1359	
R ²	0.2384		0.235	
Adjusted R ²	0.2304		0.227	

Note: *** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05; . p-value < 0.1;