The mitigation of taste-based discrimination by algorithms on peer-to-peer platforms

Erasmus University Rotterdam
Erasmus School of Economics

Master thesis Financial Economics

Supervisor: Zhu, H.

Sabine Perquin
425700

Abstract

In this paper, the effect of being a woman with children on their lending success on peer-to-peer platforms is studied by looking at the interest rate spread, loan amount and the default rate. It is expected that the algorithms implemented on peer-to-peer platforms cannot mitigate discrimination against this group of borrowers, which occurs because of female stereotypes that society holds. This would imply that mothers on the lending platform are charged worse lending terms than other borrowers. Using a loan dataset from the European peer-to-peer platform Bondora, the key findings are that mothers are not discriminated against on the platform but are charged more favorable loan terms. These results are driven by their better financial background and higher status and provides evidence against taste-based discrimination. This suggests that the female stereotype does not hold within the lending market and shows the potential for algorithms in mitigating discrimination within financial markets.

Table of Content

1. Introduction

Innovative technologies have led to more equal treatment of individuals within financial markets (Gonzales Martínez et al., 2020; Morse & Pence, 2020). Nevertheless, cultural models of gender and family life continue to disadvantage women with children in financial opportunities (Robinson, 2002). This raises concerns within the lending market as minority groups are discriminated against and do not have access to the same markets and opportunities as their equivalents (Gonzales Martínez et al., 2020). The last couple of years, the lending market has been evolving which has led to the rise of peer-to-peer platforms. On these platforms, personal, face-to-face originated loans have made way for automated, algorithmic originated loans and the interactions between borrower and lender are performed anonymously. This has made financial intermediaries abundant and unnecessary (Barasinska & Schäfer, 2014). These peer-to-peer platforms provide new opportunities within financial markets, as borrowers have an additional source for funding and lenders an additional product to invest in. However, the question arises whether these peer-to-peer platforms enhance equality within financial markets and mitigate discrimination because of their technology driven decision making. Therefore, this research examines whether unfavourable selection of groups occurs on peer-to-peer platforms based on characteristics that are not necessarily linked to their creditworthiness.

Behavioural biases can be promoted on peer-to-peer platforms as the financial intermediary is removed and financial experts are no longer involved in the transaction. Instead of employing a financial professional to assess the creditworthiness of the loan applicant, most peer-to-peer platforms make use of algorithms to determine a credit score. By implementing algorithms, more objective decision-making may be performed and biases in human decision-making can be mitigated (Morse & Pence, 2020). This suggests that technology has the potential to reduce discrimination within the lending market. Nevertheless, decreasing discrimination in financial decision-making is not preordained when using statistical techniques, and algorithms are often, mistakenly, seen as unbiased and credible (Bartlett et al., 2017). One should not ignore the role of humans completely when making algorithmic decisions because algorithms can incorporate biases through the datasets that are used for training the algorithms and through the biases of the development team (Morse & Pence, 2020). Hence, whether equality within financial markets can be enhanced by using algorithms is ambiguous.

The matter of whether technology mitigates discrimination is especially relevant in the lending market, as discriminative behaviour breaches fair lending and civil rights law (Morse & Pence, 2020). Previous literature has researched which personal characteristics of borrowers have unjustly influenced the lending decision in the lending market. In traditional finance, characteristics such as gender and race are of influence in the lending decision, even after controlling for variables that can affect the creditworthiness (Alesina et al., 2013). Similar results are found in peer-to-peer studies, where there is evidence in favour of discrimination based on gender (Chen et al., 2013, 2017; Pope & Sydnor, 2017). Robinson (2002) shows that in traditional mortgage lending, discrimination is present by activation of female stereotypes. The traditional female stereotype holds the belief that mothers stay at home and are responsible for the child rearing. Thus, this involves discrimination based on gender and the number of children the person has, which is referred to as the familial status. However, the research of Robinson (2002) uses data from 1992 and female stereotypes have been evolving since. Recent literature lacks on examining female stereotypes and their impact in financial markets (Huber et al., 2010). Additionally, whether algorithms on peer-to-peer platforms can mitigate discrimination on gender and familial status has not been examined yet. Therefore, this research aims to fill this literature gap by examining how female stereotypes affect the lending success on peer-to-peer platforms and the following research question is proposed:

To what extent are borrowers discriminated based on female stereotypes when lending decisions are performed by machine learning algorithms on peer-to-peer platforms?

This study measures discrimination by determining the effect of gender and familial status on lending success with data from the European peer-to-peer platform Bondora. Lending success is measured by three outcome variables: the interest rate spread, the loan amount and the default rate. Based on previous studies, it is expected that algorithms are not resistant against human biases and that discrimination against mothers is present on the peer-to-peer platform. If there is discrimination on the platform, then mothers are irrationally seen as riskier by the lenders. This translates into unjustly charging mothers higher costs of credit and originating them lower loan amounts whereas their default rates are lower than other borrowers. Nevertheless, the findings of the analyses performed in this research suggest the contrary, as women with children are charged lower interest rates, higher loan amounts and have a lower default rate. This implies that mothers are less risky than other borrowers, which is explained by their significantly higher creditworthiness and better status. Moreover, the results entail that the machine learning algorithm performed on Bondora do not promote discriminative decision-making. This shows that algorithms have the potential to reduce inequality within financial markets.

This research paper is structured in the following way. The next section discusses behavioural literature on discrimination and elaborates on discrimination within economic markets based on gender stereotypes after which three hypotheses are formed. Then, the data and methodology are explained in Section 3 and Section 4, after which the results of the analyses are discussed and interpreted. Section 6 contains a conclusion and discussion.

2. Theoretical framework

Within financial markets, it is of great relevance to understand, measure and eliminate discrimination as it can lead to exclusion of groups of people from essential financial services and its opportunities (Gonzales Martínez et al., 2020). The European non-discrimination law, as formed in the European Convention on Human Rights, forbids discrimination across various contexts and on a range of grounds, such as colour, ethnicity, birth, and sex (European Commission, 2017). The attention to discrimination legislation and enforcement varies across European Union member states but lending decisions across all countries may not be based on personal characteristics that are irrelevant to the transaction (Morse & Pence, 2020). Nevertheless, traditional lenders exhibit patterns in providing financial services that are in line with discrimination against certain groups across several attributes, even when these decisions result in lower profits (Alesina et al., 2013). Therefore, this section elaborates on discrimination theory and discrimination within economic markets. Furthermore, mitigation of discrimination by technology is discussed whereafter three hypotheses are formed.

2.1 Discrimination theory

Systematic differences between groups of people have persisted across a range of areas, cultures, and periods. Existing research shows that, for example, minorities based on ethnicity have experienced worse school performance (Cohen et al., 2006), are more likely to have their cars searched for illegal contraband (Knowles et al., 2001) and receive harsher treatment in the justice system (Weitzer, 1996). Additionally, a gender-gap exists within the employment market as women are less likely to fill leadership positions than men (Goldman et al., 2006). The relations between different groups within financial markets can be explained by economic discrimination theory, which has shaped two different explanations: taste-based discrimination and statistical discrimination. Before diving deeper into the meaning of these two discrimination theories, the definition of discrimination is explained followed by theories that explain the motive behind discriminative behaviour.

According to psychological theory, discrimination is unjustified unequal treatment or unfair behaviour towards people based on their group membership of some arbitrary characteristic (Dion, 2002). This behaviour can be stimulated by stereotypes in society. Stereotypes are beliefs about particular social groups based on their characteristics that can be incorrect, over-generalized and, when presented with new information, reluctant to change. Thus, discriminative behaviour arises from stereotypes that people have in mind and leads to exclusion of these groups simply because they happen to be part of that category (Al Ramiah et al., 2010). In this research, discrimination is defined as unjustified behaviour that is based on stereotypes and is directed towards people that belong to a certain group, which is consequential for the financial opportunities for this group.

Why people, often unconsciously, exclude certain groups in their actions and exhibit discriminative behaviour is explained by Turner et al. (1979) based on the social identity perspective. They state that members of a group are motivated to attain a positive social identity to the extent that membership of a group becomes significant to their self-image, leading to ingroup favouritism. This means that agents rate members belonging to the group (ingroup) and members not belonging to the group (outgroup) and engage into a reward allocation task between the two groups. In this task, agents evaluate ingroup members more positively and prefer interacting with them, which leads to a biased and unfavourable treatment of the outgroup and discrimination against this group.

Within the field of behavioural economics, Tversky & Kahneman (1974) explain biased decision-making by three heuristic principles. When making decisions and assessing uncertain probabilities, people rely on these heuristic principles to reduce the complexity of the task. If the probability of an uncertain outcome must be judged, such as the probability of default of a loan application, people usually rely on the representativeness and availability heuristics. In the representativeness heuristic, people estimate the likelihood of an event by comparing it to an available paradigm in their minds. They base their decision on this existing paradigm, which they think is most relevant for the event, whereas this often leads to a biased and wrongful decision. Additionally, people tend to assess the probability of an event by searching their memories for available information which may not be correct when looking at the factual data, which is known as the availability heuristic. Thus, decisions that people make are often concerned with cognitive biases. Within lending markets, this means that agents tend to make biased decisions when assessing the probability of default of a loan because of these heuristics. This leads to unfavourable decision-making for minority groups and exclusion of those groups from the lending market.

Within economic markets, taste-based discrimination is introduced by Becker (1975). His theory states that employers get utility from satisfying their bias towards certain groups of people, which is based on prejudice, even when this is costly for the employer. When applying this theory to the lending market, this means that lenders are willing to sacrifice profits or pay higher costs to avoid originating loans to certain groups of people. This implies that if undesired groups want to participate in the lending market, they often must provide more favourable terms, such as higher interest rates on their loans. Taste-based discrimination as described by Becker cannot persist because other market participants, who do not base their lending decisions on prejudice, will accept the profitable loans that were denied by biased lenders. This way, they outperform prejudiced lenders and compete them away. However, in reality many financial markets are imperfect and taste-based discrimination can persist (Morse & Pence, 2020).

In the taste-based discrimination theory of Becker (1975), lenders stimulate discrimination based on irrational beliefs. However, a second theory states that discrimination can also be stimulated by rational choice, which is referred to as statistical discrimination. This is a theory developed by Arrow (1973) and Phelps (1972) and states that groups of people are discriminated against because of imperfect information. In other words, statistical discrimination is not stimulated by biases but is based on incomplete information and is economically efficient and profit maximizing for the decision maker. Thus, the main difference between taste-based and statistical discrimination is that with the latter, agents are rational and profit-maximizing, whereas with taste-based discrimination a distinction is made on personal characteristics of the applicant that are irrelevant to the transaction.

2.2 Discrimination within economic markets based on gender stereotypes

Within economic markets, unequal treatment of men and women occurs because of widely held gender stereotypes. These gender stereotypes arise from social norms, cultural beliefs, and traditional patterns which people are exposed to from an early age. People are influenced by these well-known stereotypes without any reminder of it and lenders' differential treatment may be attributable to these stereotypes (Gupta et al., 2008). As mentioned above, decision-making based on stereotypes can stimulate discrimination and gender stereotypes have led to biased evaluations against women because of ingroup favouritism (Al Ramiah et al., 2010). Managerial positions in firms, for example, are associated with masculine characteristics and younger men are disproportionately represented in these positions as opposed to women (Tresh et al., 2019). Moreover, discrimination against females occurs within financial markets, as women are less likely to get credit and pay higher costs (Alesina et al., 2013). In the research of Alesina et al. (2013), the credit market within Italy is examined and they find that women pay almost 28 basis points more for credit than men. Even after controlling for risk factors, women pay more for credit than men, although female-owned businesses have a better credit history and are less likely to go bankrupt. This implies that financial agents are susceptible to biases in line with taste-based discrimination.

However, Stefani & Vacca (2013) their results contradict these findings and are in compliance with non-discrimination laws within financial markets. They find that financial markets do not exhibit discrimination based on gender by assessing European small and medium-sized enterprises' access to credit. They distinguish women-led firms and men-led firms from one another by defining women-led firms as firms whose CEO or owner is a woman. Their findings suggest that female firms experience higher rejection rates than firms that are owned by males. However, their econometric analysis suggests that this is almost completely because of structural differences between male and female firms, such as firm size, age, and sector. This means that the credit constraints that female firms face are hardly because of gender and implies that discrimination does not occur within the financial market.

Gender stereotypes go beyond differences in gender only, and literature elaborates on the interaction between gender and familial status and their effects in economic markets. Women are not only discriminated against because of their gender, their familial status can also impede their opportunities, which is researched extensively within the labour market. Society stereotypes mothers as the one in charge of child-rearing whereas fathers are expected to increase their participation in the labour-market when having children. Hence, a popular argument for explaining discrimination against women with children is that future incomes of mothers will decrease whereas that of fathers will remain stable or increase, which is based on the traditional female stereotype that society holds (Robinson, 2002). Nevertheless, when controlling for creditworthiness factors and when examining wages before men and women start forming families, the gender gap already exists and mothers are discriminated against more than fathers within the employment market (Combet & Oesch, 2019; Cukrowska-Torzewska & Lovasz, 2020). This implies that mothers are disadvantaged within the employment market as opposed to childless-women and fathers, solely because of irrelevant personal characteristics and beliefs based on stereotypes.

Within financial markets, less research is conducted but similar results are found. Robinson (2002) examines discrimination of women across race lines and finds that women, especially from racial minorities and women with children, have been excluded from the mortgage market and can therefore not utilize its services and opportunities. The results indicate that minority mothers were discriminated against when staying at home, whereas white mothers were not. Additionally, single women with children were disadvantaged when applying for a mortgage as opposed to single men with children. Lenders assume, based on female stereotypes, that women will leave a firm once she expects or has children, whereas men will stay. This results in a higher expected probability of default for mothers, despite the illegality of this belief. Additionally, having children can negatively influence the lender's beliefs as the expenses of taking care of children may contribute to larger debt obligations. Nevertheless, children can also enhance financial stability and if children are reason for a higher probability of default, then both fathers and mothers should be disadvantaged within the mortgage market and not only mothers.

2.3 Discrimination in Financial Technology

Behaviour in peer-to-peer markets deviates from behaviour in traditional credit markets because the two markets are different from one another. First, borrowers and lenders do not have personal contact on lending platforms and therefore do not build relationships that can influence the lending success (Barasinska & Schäfer, 2014). Furthermore, it is common that multiple lenders fund one loan request, as opposed to traditional lending where a loan is usually originated by one institution. Lastly, many Fintech platforms make use of algorithms, for example for assessing the credit worthiness, which are believed to make more objective decisions (Bartlett et al., 2017). Therefore, taste-based discrimination is expected to be less of a problem on peer-to-peer platforms. Nevertheless, investors on peer-to-peer platforms often lack financial expertise and experience and information on the borrowers is available and transparent. This raises the possibility that irrelevant personal characteristics could guide lenders on peer-to-peer platforms more by stereotypes as opposed to traditional lenders and increases taste-based discrimination (Herzenstein & Andrews, 2008).

Bartlett et al. (2017) examine algorithmic credit scoring among Fintech lenders to identify discrimination in the United States. Their results show that these lenders discriminate around one-third less than traditional mortgage lenders. This shows that there is prospective for Fintech platforms in reducing discrimination. Nevertheless, they still find evidence for taste-based discrimination through pricing strategies within Fintech applications. This implies that even though Fintech lenders discriminate less, they do discriminate against minorities through their prices, even after controlling for variables related to the creditworthiness of the borrower.

Pope & Sydnor (2017) have researched discrimination on peer-to-peer platforms by examining the characteristics of the picture attached to the loan application. They use data from one of the leaders in peer-to-peer lending in the United States, Prosper.com, where borrowers have the option to add unverified personal information in their applications in the form of text and pictures. They find that lenders on the Prosper platform respond to signals about personal characteristics of the lender more than to information about the borrower's credit profile. Blacks and overweight people are less likely to get funded based on their picture whereas women are more likely to be funded. This signals that lenders are irrationally discriminating applicants on personal characteristics, which implies the presence of taste-based discrimination. Nevertheless, it is important to note that in this context, lenders voluntarily choose the borrowers they wish to invest in instead of an algorithm, and thus taste-based discrimination is caused by human biases and not algorithmic imperfections.

Ravina (2007) finds similar results and shows that personal characteristics influence the funding likelihood and the interest rate that borrowers pay, even after controlling for their creditworthiness. She shows that beauty, weight and being overweight negatively affect the likelihood of getting a loan on the peer-to-peer platform. Additionally, black borrowers pay significantly higher interest rates on their loans. This implies that lenders are guided by stereotypes and prejudice in their lending decisions and provides evidence for taste-based discrimination. Moreover, Ravina (2007) analyses the role of similarity between lenders and borrowers. She finds that commonality has a strong, positive effect on the decision-making of the lender. If borrower and lender belong to the same race, gender or live close to each other, the funding success increases. This implies that lenders exhibit ingroup-favouritism on these platforms, again leading to taste-based discrimination.

Chen et al. (2017) find taste-based discrimination on a Chinese peer-to-peer platform as women are charged higher interest rates when applying for loans than men, even though they are less likely to default. However, additional results show that female borrowers are more likely to be funded, which can be explained by the finding that loans provided to women are less likely to default. This suggests that gender influences the funding success and the cost of credit and that the two types of discrimination, taste-based and statistical, co-exist in the online peer-to-peer lending context.

In summary, a shift from taste-based discrimination towards primarily statistical discrimination is expected on Fintech platforms because of algorithmic decision-making. Nevertheless, previous literature evidence for taste-based discrimination within the peer-to-peer lending market. Therefore, the question arises whether algorithms are suitable for mitigating discrimination on the platforms or whether discrimination in financial markets remains to exist despite innovative lending options.

2.4 Biases in algorithms

Most peer-to-peer systems classify borrowers within different risk profiles based on a user profile. These profiles include demographic information and variables related to their creditworthiness. Bondora even uses data form third parties, such as social media behaviour. All this data is put into algorithms to perform the credit-scoring and matching of borrowers and lenders which influences biases in the financial market (Bondora, 2021a). The effect of technology and algorithms in the credit market has been studied by several researchers.

Fuster et al. (2017) state that innovations in technology, such as machine learning algorithms used by Fintech lenders, have brought concerns about distributional impact across different groups because of personal characteristics. In their research, they show that predictions of who defaults and who not depends on the functional form and the distribution of characteristics across the applicants. This means that decision-making based on algorithms is very sensitive to the design of the algorithm and they find that disparity in credit market outcomes across different groups can be increased by technology.

The paper of Morse & Pence (2020) infers that technology can reduce discrimination inflicted by human discretion. In their paper they outline a framework consisting of five technological implementations that can promote discrimination instead of inhibiting it. These five “gateways for discrimination” are: human involvement in the design of the algorithm, biased embedded training datasets, the scoring of customers’ creditworthiness based on variables that proxy for membership of a minority group, statistical discrimination in profiling shopping behaviour and technology-facilitated advertising. Based on this framework, they conclude that taste-based discrimination is less of a factor when using algorithms. However, one should be aware of the statistical discrimination that can persist in financial decision-making.

Lastly, Bozdag (2013) shows that both human and technical biases are present in algorithms that are used in decision-making. Humans do not only impact the design of the algorithms, but they also affect the way that algorithms work by influencing the classification process manually. Furthermore, factors such as company policies, personal judgements and regulatory requirements will still induce bias as all these services are provided by humans. Hence, moving from human decision-making to algorithmic driven processes does not remove all human biases and is no guarantee for objective decision-making.

2.5 Measuring taste-based discrimination

The two discrimination theories result in different policy recommendations, and it is important to understand the extent to which the results are consistent with one of the two theories. In Becker’s theory, agents act irrational whereas in the statistical discrimination theory agents act rational but have imperfect information which causes them to exclude certain groups of people. Therefore, if agents correctly incorporate information when assessing the credit score of an applicant, statistical discrimination will result in accepted loans that have the same average net returns regardless of the characteristics of the applicant. However, if agents are discriminating on the basis of taste-based discrimination, then loans that are originated to minority groups, with higher interest rates for example, should have higher net returns than other loans (Pope & Sydnor, 2017). In this study, discrimination based on female stereotypes is examined by looking at the gender of the applicant and the number of children and their loan success. This is measured by the interest rate, loan amount and the default likelihood. Taste-based discrimination occurs if women with children are charged less favourable loan conditions but default less often than other groups of borrowers.

2.6 Hypotheses

Even though a strand of literature states that algorithms replace human decision-making with objective decision-making, it is expected that algorithms cannot mitigate taste-based discrimination based on the evidence in favour of human biases being present in algorithmic decision-making (Bozdag, 2013; Morse & Pence, 2020). According to the behavioural theory developed by Becker (1975), investors want to avoid lending credit to minorities based on stereotypes and prejudice. This means that, if algorithms cannot reduce human biases, women with children are discriminated against on peer-to-peer platforms because of gender stereotypes that society holds (Robinson, 2002). As mentioned, one of these gender stereotypes that influences agents' biases holds the belief that mothers will leave their job to take care of their children, whereas fathers will generate an income. This belief is irrational and shown to be wrongful, as the gender gap already commences before women start forming families (Combet & Oesch, 2019). Additionally, within modern society, traditional gender roles have changed and it is not standard anymore that women stay at home with the children (Cukrowska-Torzewska & Lovasz, 2020). A second belief for expecting higher risk for mothers is that having children bears high expenses which can increase the debt obligations. However, if this is the case, both mothers and fathers should face higher costs of credit because of their increased debt obligations, which is not what literature finds. Additionally, an argument against this belief is that having children can lead to a higher sense of responsibility and a more stable financial situation (Robinson, 2002; Taft et al., 2013). Thus, these counterarguments disprove the belief that mothers are economically weaker than other borrowers. Nevertheless, if algorithms cannot inhibit taste-based discrimination, this belief is passed on to the risk rating process on the peer-to-peer platform, and thus the platform unjustly sees women with children as riskier. As a result, mothers are charged higher costs of credit, even after controlling for their creditworthiness. To empirically test this expectation, the following hypothesis is developed:

H1: Being a woman with children increases the interest rate on the peer-to-peer platform

It is common on peer-to-peer platforms to have one loan funded by multiple lenders. Most lenders on these platforms bid small amounts on several loans to diversify their investment portfolio (Herzenstein & Andrews, 2008). On the Bondora platform, most bidding happens automatically but in a similar manner, and thus for higher loan amounts, a larger number of investors is typically needed. If women with children are perceived as riskier, the likelihood of finding investors for these loans is lower and it is harder for these women to borrow larger loan amounts. Furthermore, lenders' evaluation of the ability to repay loans of women with children is less favourable for higher loan amounts, also reducing the funding likelihood of these loans. Therefore, it is expected that mothers are assigned a lower loan amount as opposed to other borrowers and the second hypothesis is formulated:

H2: Being a woman with children decreases the loan amount on the peer-to-peer platform

If agents perceive women with children as riskier than other groups of people because of biases, this also means they belief their probability of default is higher. Nevertheless, as mentioned previously, the economic position of women with children is not economically weaker than that of other borrowers (Combet & Oesch, 2019; Robinson, 2002; Taft et al., 2013). This implies that the beliefs of the agents are irrational and that the actual default rate of women with children is lower than that of other borrowers if there is taste-based discrimination on the platform. Therefore, the last hypothesis is as follows:

H3: Being a woman with children decreases the default likelihood on the peer-to-peer platform

3. Data

3.1 Bondora platform

To empirically test the hypotheses, a dataset from the European peer-to-peer lending platform Bondora is used. Bondora is founded in 2008 and has been operating since 2009. The platform started offering consumer loans in Estonia, but in 2012 the loan marketplace opened to investors across whole Europe and in 2013 Finland and Spain could utilize the loan products on the platform. Ever since, the platform has grown and last year it achieved their third consecutive year of being profitable, with a net profit of €3.4M. No investing experience is required when investing on the platform as Bondora provides tools that automatically set up portfolios for lenders. In some of the tools they provide, lenders manually pick out the loans they want to invest in. However, most tools provide automatic matching of lenders and borrowers based on preferred characteristics that are set by the lender.

Besides automatically matching lenders and borrowers, Bondora rolled out their risk-based pricing model in January 2015, which replaced auction bidding-based interest-rate setting. In 2016 they updated their pricing model to also consider trends in the global business and economic climate. For their model, they use a proprietary credit scoring database containing the loan data they have amassed over the years. Their risk rating is predicted by using all the datapoints on the borrower which are collected during the screening process, which can include employment records, information on income, social media, and other non-traditional data. Then, through statistical analyses, they determine exactly what variables influence the risk rating and predict this rating using machine learning methods. The rating represents the combined likelihood of recovery and the default risk of the loan and is categorized into 8 different risk-categories ranging from AA (safest grade) to HR (riskiest grade). Based on these ratings and the expected return, an interest rate is determined with the lifetime cash flow model (Bondora, 2021a). Thus, as humans are still involved in the design of the risk rating algorithm and to some extent in the matching of borrowers and lenders, there is room for human biases to affect the lending decision.

3.2 Data

The dataset used in this analysis is collected from the website of Bondora, where they provide a public dataset on all loan data that is allowed by the data protection laws. The dataset contains daily loan data and is retrieved on the 7th of June 2021. The dataset provides 112 variables for a total of 172,863 loan applications within a timeframe of twelve years, from February 2009 to June 2021. The variables range from loan specific characteristics to borrower characteristics, such as income, liabilities, education, and occupation. All loan applications in the dataset have been approved by Bondora and listed on their platform.

In this research, the effect of personal characteristics on the loan success is examined, which is measured by the three variables: the interest rate spread, loan amount and the probability of default. The interest rate spread is calculated by subtracting the annual average EURIBOR rate of that year from the interest rate of the loan. Two loan specific variables are used in the analyses: the risk rating of the loan as determined by Bondora and the duration of the loan. Additionally, ten borrower specific variables are used: gender, marital status, children yes or no, age, employment, income, debt-to-income ratio, credit score, and the number and amount of previous loans. Gender and children are added to measure discrimination, where the variable children indicates the familial status of the applicant. Employment is added to account for the belief that mothers leave their jobs whereas fathers do not. Age and marital status are added to control for their effects on the loan success. Furthermore, to control for the creditworthiness of the borrower, income, debt-to-income, the personal credit score, the number and amount of loans previously obtained on the platform are added (Ravina, 2007). An overview of the definitions of the variables included in the analyses can be obtained in Table 1, Appendix A.

3.3 Descriptive statistics

After removing all observations with missing values for the variables: gender, marital status, number of children, credit score and risk rating, and after removing outliers from the dataset, 16,419 number of loan applications remain. An overview of the descriptive statistics of the variables can be obtained in Table 3.1. Out of the 16,419 loan applications, 44% of the loans have defaulted, which happens if the amount overdue is equal to or larger than 3 monthly payments (Bondora, 2021b). The loan amount on the platform ranges from 100 euros to 10,360 euros. The average loan has an amount of 2,585 euros and an interest rate spread of 25.46%. Additionally, the average loan has a duration of almost 44 months, which is equivalent to 3 years and 8 months, and most loans have a risk rating of C.

In total, 9,738 unique borrowers have applied for these loans, most of which are female (5,239). On average, 47% of the borrowers have children (4,549), out of which 2,071 of the borrowers are mothers and 2,478 are fathers. Most borrowers have a partner or are married (5,824). Furthermore, the average borrower is almost 37 years old, is employed and has an income of 1,002 euros per month. Regarding the creditworthiness of the borrowers, most borrowers have not had any payment problems (6,908) and the average debt-to-income ratio is 28.16%.

Table 3.1

Descriptive statistics on the variables used in the regression analysis

Variable	N	Mean	Sd	Min	Max
Loan-specific variables					
Interest rate spread	16,419	25.46	8.77	5.46	75.90
Loan amount	16,419	2,585	2,158	100	10,630
Default	16,419	0.44	0.50	0	1
Loan duration	16,419	43.83	17.76	1	60
Lender-specific variables					
Gender	9,738	0.46	0.50	0	1
Children	9,738	0.47	0.50	0	1
Marital Status	9,738	0.60	0.99	0	1
Age	9,738	36.81	11.42	21	75
Income	9,738	1,002	602.76	283	16,400
Employment	9,738	0.97	0.17	0	1
Credit score	9,738	0.29	0.45	0	1
Number of previous loans	9,738	0.21	0.78	0	21
Amount of previous loans	9,738	24.26	384.24	0	15,867
Debt-to-income	9,738	28.16	17.24	0	79.96

4. Methodology

This section elaborates on the linear regressions that are performed to test the relationship between the interest rate spread, loan amount, default rate and gender and familial status. Furthermore, it explains the robustness check that is done and the machine learning model that is performed to get a better understanding of how the risk ratings, and consequently the interest rates, are set on the Bondora platform.

4.1 Linear regressions

To examine the effect of gender and familial status on key variables related to the lending success, a multiple linear regression is performed. A linear regression fits the optimal linear relationship between the dependent and independent variables by minimizing the distances between the actual observations and the observations predicted by the regression. The goodness-of-fit across the linear regressions is assessed by the adjusted R_2 of the regression (James et al., 2000). This is a measure that evaluates the quality of the regression by determining how well the model fits the data whilst accounting for the number of predictors in the model. The adjusted R_2 can be interpreted as the percentage of variance in the dependent variable that is explained by the independent variables in the regression. Thus, a larger adjusted R_2 means that a larger proportion of the variance is explained by the independent variables and indicates higher quality of the model.

In this research, the linear regression model predicts a linear relationship between the dependent variables, interest rate spread, loan amount and default rate, and the twelve independent variables. The two independent variables of interest are gender and whether the borrower has children, and all other variables are added as control variables. For the variable gender, women are set as the reference variable. The interaction effect between gender and children is added to determine the effect of being a woman and having children on the lending success. The linear regression relies on some strong assumptions, such as the absence of multicollinearity among the independent variables. Therefore, the correlation between the independent variables is checked and from the correlation matrix in Table 2, Appendix A it can be observed that none of the variables have a high correlation, which occurs at a correlation coefficient of 0.7 or higher (Uyanık & Güler, 2013).

The first hypothesis states that it is expected that women with children experience higher interest rates on the Bondora platform. Therefore, the first linear regression that is performed measures the effect of gender and children on the interest rate spread (Formula 1).

$$Int\ Spread_{i,t} = \beta_0 + \beta_1 Gender_i + \beta_2\ Children_i + \beta_3\ Gender * Children_i + \beta_6\ Controls_{i,t} + Y_t + C_i + \epsilon_{i,t} \tag{1}$$

In this equation, β_0 is the intercept with the y-axis and β_p is the slope coefficient for each independent variable and indicates the direction and size of the relationship. Y_t controls for time-fixed effects, C_i for country fixed effects and $\epsilon_{i,t}$ is the error of the regression. The natural logarithm of the interest rate spread is taken because the interest rate spread is not normally distributed, which conflicts with a second assumption of the linear regression, which states that all numerical variables should have a normal distribution. For the same reason, the natural logarithm is taken of all other numerical variables except for the debt-to-income ratio, which is already close to a normal distribution without taking the logarithm. It is expected that the coefficient of the interaction variable is positive, which implies that women with children pay higher interest rates than other borrowers.

The second linear regression examines the relationship between the loan amount originated and the personal characteristics of the borrower to test the second hypothesis (Formula 2).

$$Amount_{i,t} = \beta_0 + \beta_1 Gender_i + \beta_2\, Children_i + \beta_3\, Gender * Children_i + \beta_6\, Controls_{i,t} + Y_t\ + C_i + \ \epsilon_{i,t} \tag{2}$$

Again, the natural logarithm for the dependent variable loan amount is taken because the distribution is not normal. It is expected that the coefficient of the interaction variable is negative, which indicates that women with children are originated lower loan amounts than other groups of borrowers.

To test the third and last hypothesis, a linear regression is performed with default as dependent variable, which is a dummy variable with category 1 if the loan has defaulted and 0 otherwise (Formula 3).

$$Default_{i,t} = \beta_0 + \beta_1 Gender_i + \beta_2\, Children_i + \beta_3\, Gender * Children_i + \beta_6\, Controls_{i,t} + Y_t\ + C_i + \ \epsilon_{i,t} \tag{3}$$

If there is taste-based discrimination on the platform, the interaction coefficient of gender and children is negative which means that women with children are less likely to default on their loans than other borrowers.

4.2 Random forest to predict the risk rating

The Bondora platform uses machine learning algorithms to calculate the risk rating of the loan applications, which is an important component in determining the interest rate of the loans (Bondora, 2021a). To get better insights in what variables drive the predictions of the risk rating, a machine learning method is performed after which the variable importance is assessed. By getting a better understanding of which factors drive the risk rating predictions, the pricing of the loans can also be better understood. If the predictions of the risk rating are mostly driven by the variables gender and children, which are deemed irrelevant to the creditworthiness of the borrower, then this also means that the interest rate is largely determined by these characteristics, which would support the expectation of finding taste-based discrimination on the peer-to-peer platform.

Bondora does not disclose the machine learning method they use in their risk rating model, and thus the exact method cannot be simulated. However, a simplified model for predicting the risk rating is performed with a random forest. This is an ensemble method which combines the predictions of numerous other methods to get more accurate results. Furthermore, a random forest is used because this method is robust to overfitting (James et al., 2000). The random forest is performed with the predictors gender and children and ten control variables. These control variables are age, marital status, employment status, the credit score, income, debt-to-income ratio, number of previous loans, amount of previous loans, loan duration and loan amount. The response variable in this random forest is the risk rating of the loan. This response variable is transformed into a binary variable to be able to perform a random forest on a binary classification problem, which is done to make the random forest model more reliable and improve the model performance (James et al., 2000). Thus, the loans that have a rating within the top four risk categories (AA, A, B and C) are categorized as non-risky loans and the loans with other ratings (D, E, F, and HR) are categorized as risky loans.

A random forest combines the predictions of B decision trees into one prediction. When building the decision trees, the random forest adds a force of randomness by only considering m predictors at each split in the decision trees. Thereby, the method prevents the forest from being dominated by one strong predictor variable and thereby reduces overfitting of the model. The final prediction of the dependent variable is assigned to the class with the majority votes of the B trees (James et al., 2000). Before performing a random forest and predicting the risk rating of the loans, the splitting criteria in each tree should be determined. In this classification problem, the splits in the trees are made by minimizing the Gini index, which measures the purity of the node. The Gini index determines the probability of a variable being incorrectly classified when selected randomly. Thus, the Gini index can range from 0 to 1, where 0 indicates that all the predictions belong to a certain class and a value of 1 indicates that there is a random distribution of the variables across the different classes. Hence, a lower Gini value indicates a better classification (Breiman et al., 1984). Furthermore, two parameters are tuned by performing five-fold cross validation. In five-fold cross validation, the data is randomly divided into five folds. Four of the folds are used as a training set and the other fold as a test set, where each fold is used as a test set once. Then, the model is trained on the training set and fitted on the test set. The optimal value is retrieved at the lowest cross validation error. The two parameters that are tuned are the number of predictors, *m*, and the number of trees, *B*. The hyper parameter *m* controls the balance between the predictive strength of the decision trees and the decorrelation of the trees (Breiman, 2001). The parameter *B* determines how large the forest is, where a larger number of trees increases the stability of the random forest. However, computations can get very expensive if *B* is set very large. Thus, choosing the optimal value for *B* is a trade-off between stable estimates and computational time (James et al., 2000). Once the random forest is performed with the optimal parameters, the performance of the method is assessed. This is done by determining the accuracy of the model, which indicates how many predictions are correctly made by the random forest (James et al., 2000).

The variable importance in the predictions of the risk rating is analysed to get a better understanding of how the risk rating is determined. This is done with the permutation method. In the permutation method, one variable is permuted, and predictions are done on the test set with this permuted dataset. The accuracy of the permuted predictions is compared to the accuracy of the original predictions. The predictor variables are ordered from most important to least important by the change in accuracy. Thus, if the permutation of a variable results in a high decrease in accuracy, this variable is considered as important (Breiman, 2001).

5. Results

5.1 Results of the linear regression on the interest rate spread

To test the first hypothesis, the interest rate spread is regressed on three regression models that include the interaction variable between gender and children. When looking at the interaction coefficient, we generally find that the results of all three regression models indicate that women with children do not pay higher interest rates than other borrowers (Table 5.1). The interaction coefficient of women with children is negative and significant in all three regression models, and therefore we can conclude that mothers pay lower interest rates than fathers and women without children. Model 1 is the regression model without time fixed effects, country fixed effects and any other control variables. In this model, women with children pay 8.0% lower interest rates than other borrowers. The effect of the interaction variable is approximately similar in model 2, which includes time fixed effects. The adjusted R_2 of model 3 is the highest and thus this model fits the data best. From this model it can be obtained that, after controlling for other personal characteristics such as age, marital status, age, employment, and variables related to the creditworthiness, such as the debt-to-income ratio, total income, credit score and loan duration, mothers are charged 0.1% lower interest rates than other borrowers. Because this model controls for differences in creditworthiness, this shows that mothers do not pay lower interest rates because of financial differences but because they have a better social status on the peer-to-peer platform than other borrowers. When looking only at gender and children, we find contradicting results. Being a woman or having children leads to significantly higher interest rates in models 1 and 2. When controlling for the creditworthiness of these two groups in model 3, the effects become smaller and insignificant. Thus, the coefficients in models 1 and 2 are partially capturing the financial background of the groups of borrowers which implies that, besides having a higher social status, mothers also have a stronger economic position than women and borrowers with children.

When diving into the characteristics of these groups of borrowers, the stronger economic position of mothers is affirmed (Table 3.1, Table 3, Appendix A). The descriptive statistics of the different groups indicate that the average income of mothers is 1,265 euros per month, which is higher than the average income of women (1,136 euros), borrowers with children (1,050 euros) and all borrowers (1,002 euros). Furthermore, the debt-to-income ratio of mothers is the lowest. To empirically test the difference in creditworthiness between the groups of borrowers, two Kruskall-Wallis tests are performed in section 5.4.

Based on the significant interaction coefficients of the regression models, the first hypothesis can be rejected which means that being a woman with children does not increase the interest rate on peer-to-peer platforms. Rather, mothers pay lower interest rates than other borrowers on the peer-to-peer platform Bondora.

Table 5.1

Results of the regression models with the interest rate spread

Variables	(1)	(2)	(3)
Woman	0.062***	0.058***	0.003
Children	0.023***	0.019**	0.005
Woman * Children	-0.080***	-0.079***	-0.001*
Year fixed effects	No	Yes	Yes
Country fixed effects	No	No	Yes
Controls	No	No	Yes
Adjusted R_2	0.004	0.127	0.747

*Coefficients of the predictor variables included in the linear regression where *** indicates significance at the 1% level, **indicates significance at the 5% level, *indicates significance at the 10% level.*

5.2 Results of the linear regression on the loan amount

The second set of regression models are performed to test the effect of gender and children on the loan amount. Again, three different models are performed of which the results can be obtained in Table 5.2. The interaction coefficient in the first model is positive and significant and resembles an increase of 13.8% in the loan amount if the borrower is a woman and has children. When adding year fixed effects, the increase in the loan amount is slightly larger for women with children (model 2). Model 3 has the highest adjusted R_2 and 38.2% of the variance in the loan amount can be explained by the independent variables. However, the coefficient for the interaction variable is not significant. In this set of regression models, the variables gender and children generally show similar results as the interaction coefficients. For example, the significant coefficient in model 3 for the variable gender indicates that the loan amount of women is 4.8% higher than that of men.

The second hypothesis states that being a woman with children decreases the loan amount. Based on the significant interaction coefficients of model 1 and 2, this hypothesis can be rejected. Thus, being a mother increases the loan amount on the peer-to-peer platform Bondora.

Table 5.2

Results of the regression models with the loan amount

Variables	(1)	(2)	(3)
Woman	0.030	0.028	0.048***
Children	0.004	0.002	-0.012
Woman * Children	0.138***	0.143***	0.000
Year fixed effects	No	Yes	Yes
Country fixed effects	No	No	Yes
Controls	No	No	Yes
Adjusted R_2	0.006	0.034	0.382

*Coefficients of the predictor variables included in the linear regression where *** indicates significance at the 1% level, **indicates significance at the 5% level, *indicates significance at the 10% level.*

5.3 Results of the linear regression on the default likelihood

From the third set of regression models, we find that women with children have a lower probability of default (Table 5.3). These results are used to test the third hypothesis, which states that women with children are expected to have lower default rates than other borrowers. All three regression models have negative and significant coefficients for the interaction variable. From the interaction coefficient of model 3, which has the highest adjusted R_2, it can be obtained that being a woman with children decreases the default likelihood with 4.5%. However, contradicting results are found for the separate variables gender and children. Whereas being a mother decreases the default rate, being a woman or having children increases the default likelihood significantly with approximately the same amount. When combining this with the findings that women and borrowers with children pay higher interest rates and the general impression that the descriptive statistics give on the income and debt-to-income ratio of these groups, this supports the idea of these groups having a lower creditworthiness than mothers.

Based on the significant interaction coefficient in all three models, the third hypothesis can be accepted which implies that women with children have a lower probability of default than other borrowers on the peer-to-peer platform.

Table 5.3

Results of the regression models with the default likelihood

Variables	(1)	(2)	(3)
Woman	0.057***	0.056***	0.058***
Children	0.061***	0.060***	0.051***
Woman * Children	-0.070***	-0.067***	-0.045**
Year fixed effects	No	Yes	Yes
Country fixed effects	No	No	Yes
Controls	No	No	Yes
Adjusted R_2	0.002	0.009	0.112

*Coefficients of the predictor variables included in the linear regression where *** indicates significance at the 1% level, **indicates significance at the 5% level, *indicates significance at the 10% level.*

5.4 Kruskall-Wallis test to determine differences in creditworthiness

The results of the regression models consistently show that mothers perform better on Bondora as they default less, are originated higher loan amounts, and pay lower interest rates on the peer-to-peer platform. This contradicts existing literature but can be explained by the variables related to the financial background of the borrower. The descriptive statistics on the income and debt-to-income ratio of the different groups of borrowers suggest that mothers have the highest creditworthiness. To statistically test this, a Kruskall-Wallis test is performed between the four groups: all borrowers, women, borrowers with children and mothers, on the two variables related to their financial background: income and debt-to-income.

The Kruskall-Wallis test is a non-parametric test that determines whether the means of these groups are significantly different from each other. This test is preferred over ANOVA because the data does not meet the normality assumption of ANOVA as can be seen in Figures 1 and 2 in Appendix A (Glass et al., 1972). The Kruskall-Wallis test ranks the variables and calculates the average rank per group. If the mean averages of all groups are similar, then there is no statistical significant difference between the groups (Vargha & Delaney, 1998). To avoid the presence of overlapping observations, the group of mothers is compared against the complement of the other three groups of borrowers. This means that it can be assumed that all four groups are independent, which is needed to perform a Kruskall-Wallis test (Breslow, 1970). Thus, the means of the four groups: men without children, women without children, fathers and mothers are compared to determine whether the financial background of mothers is significantly different than that of the other groups of borrowers. The boxplots of these four groups across the two variables income and debt-to-income suggest that mothers have the highest income and the lowest debt-to-income ratio of the four groups as suggested in the previous section (Figure 3-4, Appendix A).

First, a Kruskall-Wallis test is performed on the independent variable income to test for differences in the income of the four groups. Thus, the null hypothesis for this test is that the mean income of the different groups of borrowers are the same. The p-value for this test is 2.20e-16. As this p-value is less than the significance level of 0.05, the null hypothesis can be rejected. This means that there are significant differences in income between the groups. The second Kruskall-Wallis test is performed on the debt-to-income ratio, and thus the null hypothesis states that the means for the different groups of borrowers for the debt-to-income ratio are the same. The p-value for this test is 2.07e-10 which is smaller than 0.05. Therefore, the null hypothesis can be rejected, and it can be concluded that the debt-to-income ratios of the groups are significantly different. This shows that mothers have a significantly, stronger economic position than other borrowers on the peer-to-peer platform and implies that the better loan conditions for mothers come forth from profit-maximizing and rational behavior. These findings are in line with statistical discrimination rather than taste-based discrimination.

Reason for the stronger economic position of mothers can be explained by the economic stability that is needed to raise a child. Robinson (2002) argues that, whereas children lead to higher debt obligations, children also lead to a greater commitment towards the other parent and their home situation and enhances stability. Thus, from a borrower's perspective, having children increases the sense of responsibility and drive to get the finances straight. This is affirmed by Warren (2006), who finds that the life course, familial status and marital status of a person determines its economic stability and wealth and she shows that couples with school-children are among the most wealthy groups of people within the United Kingdom. Taft et al. (2013) find similar results and state that married people are more financially literate, which leads to a better financial situation and less financial concerns. Supporting this idea of having a stabler home-situation when having children, the average likelihood of mothers being married is 0.87 versus a likelihood of 0.60 for all borrowers in general. This implies that the stronger economic position of mothers can be explained by their higher sense of responsibility and steadier home situation. From a lender's perspective, this updated belief of mothers having a strong economic position because of the responsibilities to their children gives them a higher status on the peer-to-peer platform, as implied by the main results after controlling for the creditworthiness of the borrower.

A second explanation for mothers being more successful on the peer-to-peer platform is that this group of borrowers anticipate a rejection when applying for a loan (Stefani & Vacca, 2013). Thus, to prevent their loan application from being rejected, women, and especially the one with children, only apply for a loan when they are certain of a strong economic position. This implies that women with a low creditworthiness do not even try to get a loan and that the women that are accepted on the platform are more successful due to their strong economic position. Whether this explanation holds on Bondora cannot be tested in this research as Bondora does not disclose information on rejected loans.

Besides showing that mothers are economically strong on the peer-to-peer platforms, the results also entail that the risk rating model used on the Bondora platform is not prone to promoting taste-based discrimination. This in line with the research of Morse & Pence (2020) that states that taste-based discrimination is less of a concern when using machine learning models. Nevertheless, they state that statistical discrimination is likely to sustain, which is in line with the main results of this research. Women are rationally charged lower interest rates and higher low amounts, which indicates that this type of discrimination comes forth from maximizing profits on the peer-to-peer platform, and thus implies a form of statistical discrimination.

5.5 Robustness check

From the previous analyses, it can be obtained that the results of the linear regression analyses remain consistent after adding control variables, which indicates that the results are robust. However, one additional check is performed to further test the robustness of the main results.

Bondora uses a pricing model to determine the risk rating which is used for determining the level of interest rates. In January 2016, they replaced the first version of this model with a second version to improve the rating performance. If women with children on the Bondora platform have a stronger economic position as suggested by the main results, then the version of the rating model should not matter in their loan success. Therefore, the dataset is split up in two subsets to test whether the results remain consistent across the different versions of the pricing model. One subset includes the loan applications before January 2016 and the other subset includes the loan applications after January 2016. Besides testing the robustness of the results if a different rating model is used, splitting the dataset this way also checks for the robustness of the results across two different time periods.

Generally, the results of the robustness check indicate that the version of the pricing model and the time period do not influence the treatment of women with children on the peer-to-peer platform (Table 4-6, Appendix A). The results of both samples generally show that women with children are charged lower interest rates, higher loan amounts and have a lower probability of default on Bondora. These findings indicate that there is no taste-based discrimination when using either one of the two models which supports the main results. Furthermore, both versions of the risk rating model do not promote biases in the loan success based on irrelevant personal characteristics. This supports the main findings and the belief that machine learning algorithms can diminish human biases and mitigate discrimination within financial markets.

5.6 Random forest results

Based on existing literature, it is expected that the algorithms on peer-to-peer platforms are prone to taste-based discrimination against women with children. Therefore, it is presumed that the personal characteristics gender and children are of great importance in the predictions of the risk rating, even though these variables should be unrelated to the risk of the loan. To test this expectation, a random forest is performed on the variables gender and children and ten control variables. The dataset is split into 70% training data, 15% validation data and 15% test data.

Two hyperparameters are tuned before performing the final random forest model is performed by using the validation set. These two parameters are the number of predictors, m, that are used in the decision trees and the number of trees used in the forest. The optimal value for the number of predictors, m, is 4 and the optimal number of trees, B, is 750 trees (Table 7 and 8, Appendix A). Thereafter, the random forest is trained on the training dataset and predictions are made on the test dataset. The accuracy of the random forest model is 75.23%, which means that around three-quarters of the predictions made by the model are correct (Table 5.4).

Table 5.4
Confusion matrix with the prediction results of the random forest

		Actual	
		Non-risky	**Risky**
Predictions	**Non-risky**	1,069	391
	Risky	219	784

Confusion matrix of the random forest where the actual and predicted values of the two classes are displayed. The accuracy of the model is 75.23%.

After determining the accuracy of the model, the variable importance is computed with the permutation method to test whether gender and children dominantly influence the risk rating of the loan. All variables are ranked according to their mean increase in accuracy scaled by the standard deviation (Figure 5, Appendix A). Contrary to the expectation that the variables gender and children are of great importance in the predictions of the risk rating, these two variables are among the least important predictors. When permuting the variable gender, the accuracy of the random forest decreases by 0.163% (Table 9, Appendix A). This means that if the variable gender is eliminated from the model, the accuracy of the model becomes 75.07% instead of 75.23%. Furthermore, when removing the variable children in the risk rating predictions, the accuracy decreases with 0.253%. Thus, the variables gender and children do not have a large, statistical impact on the predictions of the risk rating. This implies that these variables are not influencing the level of the interest rate much. Rather, most of the important variables in the predictions of the risk rating of the loan are related to the creditworthiness of the borrower, such as their personal credit score, debt-to-income ratio, and their total income. This suggests that these variables are dominantly influencing the level of interest rates and contradicts the expectation of finding taste-based discrimination. The only personal characteristic that is of great importance in the risk rating predictions is the age of the borrower. If the variable age is permuted in the dataset, the accuracy of the model decreases with 7.22%. As the age of the borrower should be irrelevant to the risk rating when controlling for the creditworthiness, this could imply that borrowers are discriminated against based on age. However, the importance of age can also be explained by the strong, positive correlation between age and financial literacy and financial wellbeing (Taft et al., 2013).

6. Conclusion

This study contributes to the growing literature on unequal treatment of groups within financial markets and on the potential for algorithms in financial decision-making by determining whether discrimination against women with children exists on peer-to-peer platforms. A dataset from the European peer-to-peer platform Bondora is used to determine the loan success of this group of borrowers based on three variables: the interest rate, the loan amount, and the default likelihood. The significant findings imply that women with children are not treated adversely on the platform because of taste-based discrimination. The analyses show that women with children pay significantly lower interest rates as opposed to other borrowers, are originated higher loan amounts, and have a lower probability of default which signals that there is statistical discrimination on the platform.

These findings contradict most existing literature on taste-based discrimination against women in financial markets and rather suggest that the platform provides more favorable loan terms to women with children because they have a stronger economic position and higher status. This suggestion is supported by the data and the Kruskall-Wallis tests that show that the creditworthiness of women with children is significantly higher than that of other borrowers. This can be explained by the mother's ambition to have a strong economic position because of the increased drive towards a financial and stable environment once women have children (Robinson, 2002). Additionally, the traditional female stereotype does not necessarily hold nowadays as mothers often keep working when having children and do not always take care of the child-rearing anymore (Cukrowska-Torzewska & Lovasz, 2020). Thus, within modern society, lenders do not perceive mothers as economically weak. This implies that the pricing strategies that are pursued on the Bondora platform come from a profit-maximizing rationale, which is in line with statistical discrimination rather than taste-based discrimination.

The outcome of the analyses also adds a new aspect to literature on the use of machine learning algorithms within financial markets. Previous literature is ambiguous on the potential for algorithms in financial decision-making as both human and technical biases influence their performance. This study looks at the bias in machine learning models by examining whether discrimination in the lending market is present and by determining the important variables in the risk rating predictions. Both methods indicate that machine learning algorithms do not promote discriminative biases. This means that borrowers are not discriminated against based on female stereotypes when lending decisions are performed by algorithms and shows the potential for replacing human decision-making with algorithmic decision-making within financial markets.

Due to the limited data made available by Bondora, this study faced two limitations. First, Bondora only publishes data on the loans that are accepted by the platform. Therefore, rejected loans were not included in the dataset and the funding likelihood of mothers could not be examined or linked to their creditworthiness. Even though discrimination does not play a role in the pricing strategy and loan amount on the platform, discriminative biases are possibly present in the approval process of the loan applications on the platform. Therefore, a suggestion for future research includes an analysis on borrower's characteristics that dominate the loan approval decision and determining the funding likelihood of women with children. Secondly, Bondora calculates the risk rating of a loan by using all the data points they have on the loan applicant. This includes externally validated data that is received from third parties, which cannot be shared because of data protection laws. This implies that even though a good indication of the risk rating prediction is done by simulating this model with the data that is publicly available, the actual risk rating depends on many more variables. Thus, the findings in this research provide a good basis for understanding the pricing model on peer-to-peer platforms, but a recommendation for future research is to control for a lot more factors if these are made available.

References

Al Ramiah, A., Hewstone, M., Dovidio, J. F., & Penner, L. A. (2010). The social psychology of discrimination: Theory, measurement, and consequences. In *Making equality count: Irish and international approaches to measuring discrimination*.

Alesina, A. F., Lotti, F., & Mistrulli, P. E. (2013). Do women pay more for credit? evidence from italy. *Journal of the European Economic Association*, *11*(SUPPL. 1), 45–66. https://doi.org/10.1111/j.1542-4774.2012.01100.x

Barasinska, N., & Schäfer, D. (2014). Is Crowdfunding Different? Evidence on the Relation between Gender and Funding Success from a German Peer-to-Peer Lending Platform. *German Economic Review*. https://doi.org/10.1111/geer.12052

Bartlett, R. P., Morse, A., Stanton, R., & Wallace, N. (2017). Consumer Lending Discrimination in the FinTech Era. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3063448

Bondora (2021a). How are Bondora risk rating calculated? Retrieved on 30 June, 2021 from: https://support.bondora.com/en/how-are-bondora-risk-ratings-calculated/

Bondora (2021b). What is Bondora's 3-step Collection & Recovery process? Retrieved on 28 June, 2021 from: https://support.bondora.com/en/what-is-bondora-s-3-step-collection-recovery-process/

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-013-9321-6

Breiman, L. (2001). Random forests. *Machine Learning*. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. In *Classification and Regression Trees*. https://doi.org/10.1201/9781315139470

Breslow, N. (1970). A generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship. *Biometrika*. https://doi.org/10.1093/biomet/57.3.579

Chen, D., Hao, L., & Xu, H. (2013). Gender Discrimination towards Borrowers in Online P2PLending. *Whiceb*, 55.

Chen, D., Li, X., & Lai, F. (2017). Gender discrimination in online peer-to-peer credit lending: evidence from a lending platform in China. *Electronic Commerce Research*. https://doi.org/10.1007/s10660-016-9247-2

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*. https://doi.org/10.1126/science.1128317

Combet, B., & Oesch, D. (2019). The Gender Wage Gap Opens Long before Motherhood. Panel

Evidence on Early Careers in Switzerland. *European Sociological Review*. https://doi.org/10.1093/esr/jcz009

Cukrowska-Torzewska, E., & Lovasz, A. (2020). The role of parenthood in shaping the gender wage gap – A comparative analysis of 26 European countries. *Social Science Research*. https://doi.org/10.1016/j.ssresearch.2019.102355

Dion, K. L. (2002). The social psychology of perceived prejudice and discrimination. *Canadian Psychology*. https://doi.org/10.1037/h0086899

European Comssion. (2017). Retrieved on June 4, 2021, from: https://ec.europa.eu/info/aid-development-cooperation-fundamental-rights/your-rights-eu/know-your-rights/equality/non-discrimination_en

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2017). Predictably Unequal? The Effects of Machine Learning on Credit Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3072038

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research*. https://doi.org/10.3102/00346543042003237

Goldman, B. M., Gutek, B. A., Stein, J. H., & Lewis, K. (2006). Employment discrimination in organizations: Antecedents and consequences. In *Journal of Management*. https://doi.org/10.1177/0149206306293544

Gonzales Martínez, R., Aguilera-Lizarazu, G., Rojas-Hosse, A., & Aranda Blanco, P. (2020). The interaction effect of gender and ethnicity in loan approval: A Bayesian estimation with data from a laboratory field experiment. *Review of Development Economics*, *24*(3), 726–749. https://doi.org/10.1111/rode.12607

Gupta, V. K., Turban, D. B., & Bhawe, N. M. (2008). The Effect of Gender Stereotype Activation on Entrepreneurial Intentions. *Journal of Applied Psychology*. https://doi.org/10.1037/0021-9010.93.5.1053

Herzenstein, M., & 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*.

Huber, S., Bookstein, F. L., & Fieder, M. (2010). Socioeconomic status, Education, and reproduction in modern women: An evolutionary perspective. *American Journal of Human Biology*. https://doi.org/10.1002/ajhb.21048

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2000). An introduction to Statistical Learning. In *Current medicinal chemistry*. https://doi.org/10.1007/978-1-4614-7138-7

Kenneth J. Arrow. (1973). The Theory of Discrimination. In *Discrimination in Labor Markets*.

Knowles, J., Persico, N., & Todd, P. (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy*. https://doi.org/10.1086/318603

Morse, A., & Pence, K. M. (2020). Technological Innovation and Discrimination in Household Finance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3536506

Phelps, E. S. (1972). The Statistical theory of Racism and Sexism. *American Economic Review*.

Pope, D. G., & Sydnor, J. R. (2017). *What ' s in a Picture ?: Evidence of Discrimination from Prosper . com What ' s in a Picture ? Evidence of Discrimination from Prosper . com*. *46*(1), 53–92.

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

Robinson, J. K. (2002). Race, gender, and familial status: Discrimination in one US mortgage lending market. *Feminist Economics*. https://doi.org/10.1080/13545700210167323

Stefani, M. L., & Vacca, V. P. (2013). Credit Access for Female Firms: Evidence from a Survey on European SMEs. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2297789

Taft, M. K., Hosein, Z. Z., & Mehrizi, S. M. T. (2013). The Relation between Financial Literacy, Financial Wellbeing and Financial Concerns. *International Journal of Business and Management*. https://doi.org/10.5539/ijbm.v8n11p63

Tresh, F., Steeden, B., de Moura, G. R., Leite, A. C., Swift, H. J., & Player, A. (2019). Endorsing and reinforcing gender and age stereotypes: The negative effect on self-rated leadership potential for women and older workers. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2019.00688

Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2420090207

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty - Heuristics and Biases Tversky Kahneman 1974. *Science*.

Uyanık, G. K., & Güler, N. (2013). A Study on Multiple Linear Regression Analysis. *Procedia - Social and Behavioral Sciences*. https://doi.org/10.1016/j.sbspro.2013.12.027

Vargha, A., & Delaney, H. D. (1998). The Kruskal-Wallis Test and Stochastic Homogeneity. *Journal of Educational and Behavioral Statistics*. https://doi.org/10.3102/10769986023002170

Warren, T. (2006). Moving beyond the gender wealth gap: On gender, class, ethnicity, and wealth inequalities in the United Kingdom. In *Feminist Economics*.

https://doi.org/10.1080/13545700500508502

Weitzer, R. (1996). Racial discrimination in the criminal justice system: Findings and problems in the literature. *Journal of Criminal Justice*. https://doi.org/10.1016/0047-2352(96)00015-3

Appendix A

Table 1
Variable description of the variables used in the regression analyses.

Variable	Variable type	Explanation
Interest rate spread	Dependent	Interest rate on Bondora minus the average annual Euribor rate
Loan amount	Dependent	Amount the borrower received in euros
Default	Dependent	1) if the loan was defaulted, 0) if otherwise
Gender	Independent	1) if female, 0) if male
Children	Independent	1) if children, 0) if otherwise
Marital status	Control	1) if borrower has partner, 0) if otherwise
Age	Control	Age of the borrower at the time of the application
Employment	Control	1) if employed, 0) if otherwise
Income	Control	Borrower's total monthly income in euros
Credit score	Control	Payment behaviour of borrower: 1) if borrower had previous payment problems, 0) if no payment problems
Debt-to-income	Control	Ratio of income paid to loans over borrower's monthly gross income
Number of previous loans	Control	Number of previous loans
Amount of previous loans	Control	Amount of previous loans in euros
Risk rating	Control	Bondora rating per loan determined by their rating model. AA – A - B – C – D – E – F - HR
Loan duration	Control	Current loan duration in months

Table 2
Pearson correlation matrix of the numerical variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1) Age	1.000					
(2) Duration	0.018	1.000				
(3) Income	0.003	-0.035	1.000			
(4) Debt-to-income	0.027	0.104	-0.096	1.000		
(5) Number of previous loans	0.061	0.040	0.073	0.181	1.000	
(6) Amount of previous loans	0.002	0.015	0.022	0.006	0.404	1.000

Table 3
Descriptive statistics on women, borrowers with children and mothers

Variable	N	Mean	Sd	Min	Max
Women					
Marital status	5,239	0.45	0	0	1
Age	5,239	35.52	10.76	21	75
Income	5,239	1,136	684.32	283	16,400
Employment	5,239	0.98	0.12	0	1
Credit score	5,239	0.30	0.46	0	1
Number of previous loans	5,239	0.21	0.80	0	21
Amount of previous loans	5,239	26.75	436.58	0	15,867
Debt-to-income	5,239	27.05	1.70	0	69.99
Applicants with children					
Marital Status	4,549	0.82	0.39	0	1
Age	4,549	35.84	8.22	21	68
Income	4,549	1,050	623.52	300	10,000
Employment	4,549	0.98	0.13	0	1
Credit score	4,549	0.31	0.47	0	1
Number of previous loans	4,549	23.13	0.84	0	1
Amount of previous loans	4,549	30.71	470.62	0	15,867
Debt-to-income	4,549	28.40	17.18	0	79.96
Women with children					
Marital Status	2,071	0.86	0.35	0	1
Age	2,071	36.61	8.14	21	68
Income	2,071	1,265	744.44	300	10,000
Employment	2,071	0.99	0.07	0	1
Credit score	2,071	0.31	0.46	0	1
Number of previous loans	2,071	0.25	0.89	0	17
Amount of previous loans	2,071	38.02	593.25	0	15,867
Debt-to-income	2,071	27.00	16.78	0	69.99

Figure 1

Histogram of the variable income

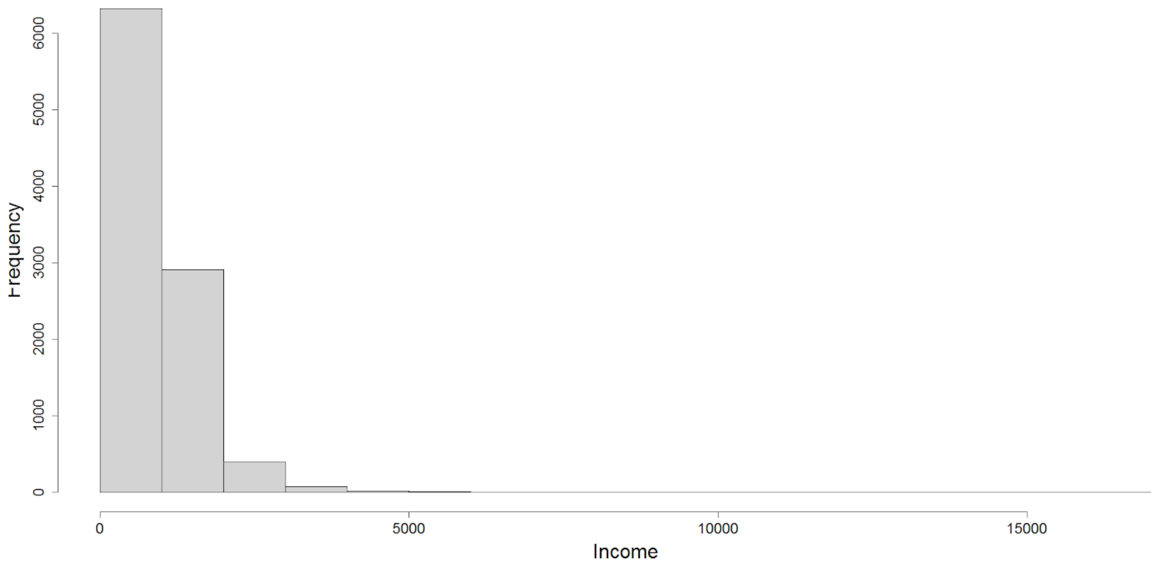

Figure 2

Histogram of the variable debt-to-income

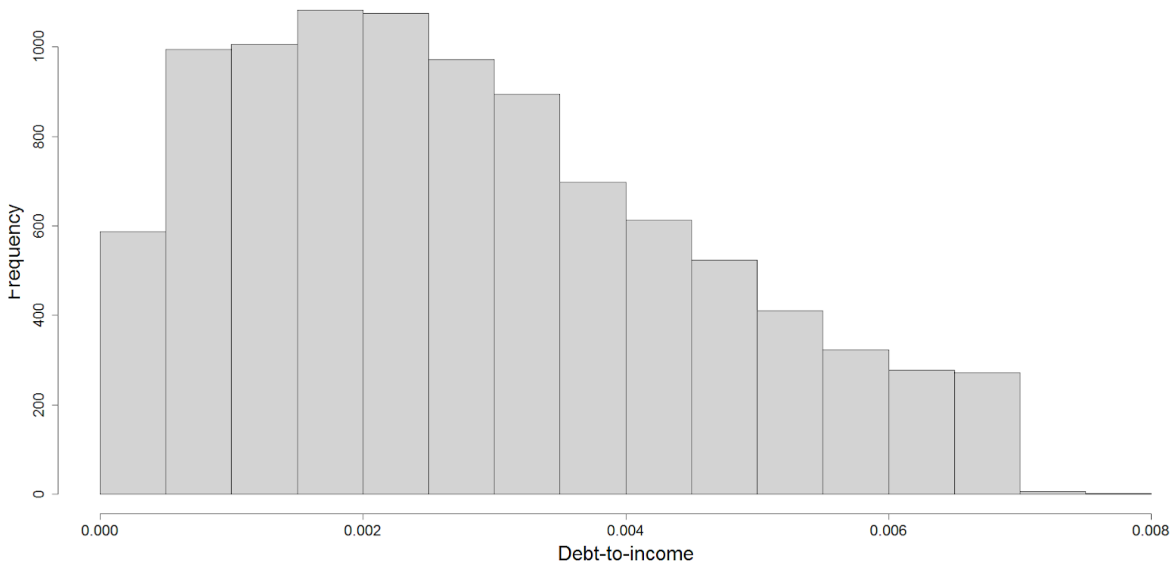

Figure 3
Boxplot that indicates the distribution of the variable income across the four groups

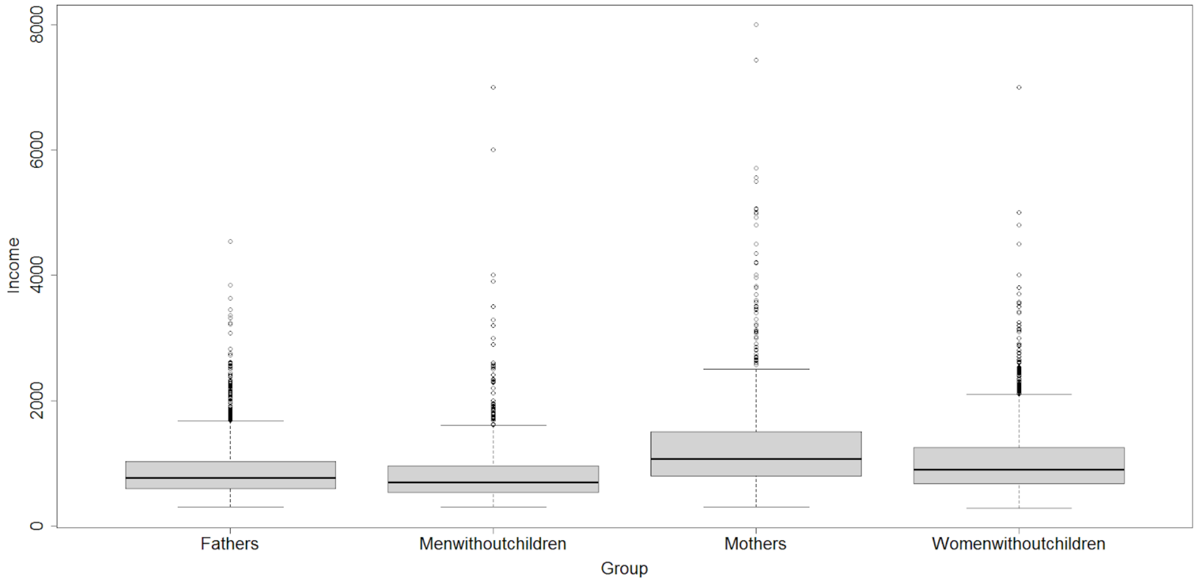

Figure 4
Boxplot that indicates the distribution of the variable debt-to-income across the four groups

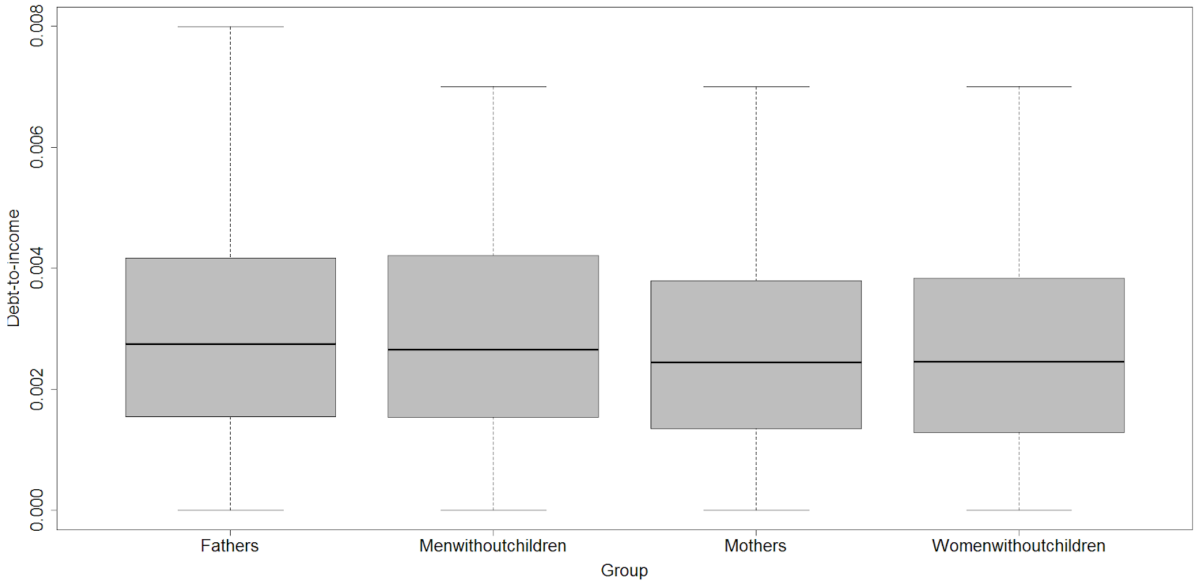

Table 4
Results of the linear regression on the interest rate spread for the two subsamples

	Before 2016			After 2016		
Variables	(1)	(2)	(3)	(1)	(2)	(3)
Woman	0.021*	0.022**	0.000	0.095***	0.085***	0.001
Children	0.019*	0.019*	0.018**	0.028**	0.019*	-0.001*
Woman * Children	-0.065***	-0.064***	-0.022**	-0.100***	-0.089***	-0.001
Year fixed effects	No	Yes	Yes	No	Yes	Yes
Country fixed effects	No	No	Yes	No	No	Yes
Controls	No	No	Yes	No	No	Yes
Adjusted R_2	0.003	0.049	0.507	0.009	0.092	0.935
Observations	6,863	6,863	6,863	9,556	9,556	9,556

Table 5
Results of the linear regression on the loan amount for the two subsamples

	Before 2016			After 2016		
Variables	(1)	(2)	(3)	(1)	(2)	(3)
Woman	0.046	0.043	0.031	0.015	0.017	0.054***
Children	0.014	0.011	-0.016	-0.006	-0.004	-0.012
Woman * Children	0.171***	0.174***	0.027	0.122***	0.119***	-0.022
Year fixed effects	No	Yes	Yes	No	Yes	Yes
Country fixed effects	No	No	Yes	No	No	Yes
Controls	No	No	Yes	No	No	Yes
Adjusted R_2	0.011	0.059	0.377	0.003	0.045	0.387
Observations	6,863	6,863	6,863	9,556	9,556	9,556

Table 6
Results of the linear regression on the default probability for the two subsamples

	Before 2015			After 2016		
Variables	(1)	(2)	(3)	(1)	(2)	(3)
Woman	0.362**	0.042**	0.051***	0.065***	0.066***	0.062***
Children	0.054***	0.052***	0.042**	0.066***	0.066***	0.057***
Woman * Children	-0.036	-0.033	-0.025*	-0.091***	-0.092***	-0.061***
Year fixed effects	No	Yes	Yes	No	Yes	Yes
Country fixed effects	No	No	Yes	No	No	Yes
Controls	No	No	Yes	No	No	Yes
Adjusted R_2	0.002	0.008	0.111	0.002	0.003	0.112
Observations	6,863	6,863	6,863	9,556	9,556	9,556

Table 7
Results of the five-fold cross validation to tune the parameter m

Number of predictors	Accuracy
1	72.15%
2	73.77%
3	73.49%
4	73.97%
5	73.93%
6	73.41%
7	73.00%
8	72.92%
9	73.37%
10	73.00%
11	72.96%
12	72.72%

Table 8

Results of the five-fold cross validation to tune the parameter B

Number of trees	Accuracy
250	72.80%
500	73.00%
750	73.16%
1000	73.08%

Figure 5

Variable importance of the predictors in the random forest with the scaled mean decrease in accuracy

Table 9
Mean decrease in accuracy in the random forest per variable

Variable	Mean decrease in accuracy
Credit score	7.325%
Age	7.220%
Debt-to-income	1.275%
Amount of previous loans	0.987%
Duration	0.832%
Loan amount	0.812 %
Income	0.777%
Number of previous loans	0.661%
Children	0.253%
Marital status	0.190%
Gender	0.163%
Employment	0.002%