Bivariate Tests of Credit Market Discrimination with an Application to Intersectionality

Anastasia Cozarenco*
Montpellier Business School and Centre for European Research in Microfinance (CERMi)

Ariane Szafarz
Université Libre de Bruxelles (ULB), CERMi, and Robert F. Wagner Graduate School of Public Service, New York University

November 2021

Abstract

This paper develops a bivariate test for discrimination in lending which is applicable regardless of the profit orientation of the lender. The originality of the testing approach comes from combining denial rates and recovery rates. We also extend the test to address intersectional discrimination. In our French microcredit dataset, the tests reveal that the positive—and socially consistent—intersectional bias toward migrant women hides the striking fact that European Union women's loan applications are handled more harshly than those of their male counterparts, which suggests that pro-social lenders are not immune to discriminatory attitudes stemming from entrenched gender stereotypes.

***Keywords**:* Discrimination in lending; Intersectional discrimination; Credit market; Migrants; Gender; Ethnicity; Recovery rate.

***JEL Classifications**:* C12, C44, J15, J16, G21, D63

***Acknowledgements**:* This study was carried out within the framework of the Chair in Social and Sustainable Finance founded by Montpellier Business School in partnership with the Caisse d'Epargne Languedoc-Roussillon. The authors thank Cécile Abramowicz, Moez Bennouri, Patrick Sentis, and Ilan Tojerow for their helpful comments and Roxanne Powell for excellent copy-editing.

* Corresponding author. Address: Montpellier Business School, 2300, avenue des Moulins, 34185, Montpellier cedex 4; France. Email: a.cozarenco@montpellier-bs.com. Phone: +33.4.67.10.26.35.

1. Introduction

According to Black et al. (1978, p. 186), "discrimination [in lending] occurs if a loan application is rejected due to personal rather than economic characteristics of the borrower." Yet different groups of borrowers can have different economic characteristics, so that differences in denials rates across segments are insufficient to prove discrimination in lending. This crucial point, made by Ferguson and Peters (1995), has received surprisingly little attention. Recently, however, a new stream of papers developed so-called "outcomes tests" for discrimination in lending (Dobbie et al., 2020; Simoiu et al., 2017). These tests scrutinizing the lenders' profitability of marginal applicants have brought a major achievement to the literature. Their implementation brings new practical challenges, such as the need for taking a proper account of endogenous differences in charged interest (Butler et al., 2021). Also, the profitability outcome is unsuitable for tracking biased loan allocation by lenders with a non-profit orientation, such as providers of social credit and microfinance institutions. By going back to the original model of Ferguson and Peters (1995), based on observed defaults rather than profits, this paper offers a new approach agnostic on the utility function of the lender. It applies this approach to a hand-collected dataset from a European microfinance institution providing consumption loans to modest households.

Extending the framework to prosocial lending helps understand how biased loan allocation can go beyond the supply-side perspective according to which discrimination appears as a cost to the prejudiced lender. The same is true in markets with credit rationing (Stiglitz & Weiss, 1981), where unfair denial based on taste discrimination can hurt applicants without affecting the lender's profit[1] and therefore fly under the radar of tests based on profitability measures. By contrast, our test design combines the two sides of the credit market, namely it uses both loan denial experienced by the applicants and default occurrences experienced by the lender. Highlighting the demand-side

[1] Coffman et al (2021) provide a labor-market example of "costless discrimination" (Méon and Szafarz, 2011), where an employer prefers male over female workers with identical resumes.

dimension of biased loan allocation is important because discrimination is objectionable on ethical grounds and, from a philosophical point of view, ethics alone is sufficient motivation to identify biases in any economic market. Discriminatory practices that incur no profit losses to the lender can be as harmful to denied applicants. Arguably, such practices can even be more deeply ingrained when they are not financially penalized.

When it comes to prejudices, the credit market has been far less scrutinized than the labor market (Bertrand & Duflo, 2017).[2] Yet fairness aside, discrimination in lending is economically important for at least two reasons. First, access to credit has been, and still is, recognized as a challenging barrier for entrepreneurs belonging to vulnerable groups who are potentially subject to biases motivated by gender, race, religion, sexual orientation, or ethnic background. Second, there is an urgent need for methodological innovation in the field of discrimination in lending, where the design of existing tests is still controversial (Ross, 2002; Dymski, 2006) and data limitations often lead researchers to use ill-suited methods. Typical data limitations include inaccessible information on unsuccessful applicants, unavailability of the amount requested by applicants (as opposed to loan size), and delay/default in repayment.

The available tests for discrimination in lending with administrative data can be split into two groups. The tests in the first group (Ladd, 1998) assess disparate treatment, i.e., whether loan applications by a given category of applicants may be treated more harshly. From a theoretical standpoint, these tests rely on the assumption that, *all else being equal*, applicants belonging to different categories (say, men and women) are equally creditworthy. But this working assumption may or may not be true. For instance, the microfinance literature has uncovered evidence that women entrepreneurs tend to reimburse their loans more reliably than men (Armendariz & Morduch, 2010; Agier & Szafarz, 2013a). Let us assume for now that this observation is not due to a bias against women causing female applicants to be more thoroughly vetted to begin with. In

[2] Regarding the financial industry, Egan et al. (2017) find a gender gap in misconduct punishment.

that case, the facts would suggest that women are better credit risks and, therefore, equally probable loan approvals for men and women may be discriminatory. In addition, the *all-else-equal* condition exposes the analysis to the issue of omitted variables: Missing characteristics correlated with credit risk may help detect spurious discrimination (Guryan & Charles, 2013). Recent experimental evidence collected by Agan and Starr (2018) shows that withholding potentially relevant information, such as criminal records, may trigger racial discrimination. Moreover, if loan allocation has been automated through computerized decision-making, missing variables may result in so-called algorithmic discrimination (Williams et al., 2018; Bartlett et al., 2019).[3] Together, the required assumption of creditworthiness and the omitted-variable issue make a significant case against solely using loan allocation records to assess discrimination in the credit market.

Second, the outcome tests are based on Becker's (1993) argument that harsher treatment of minority loan applicants should lead them to a lower occurrence of default. The tests in this group require identifying the outcome targeted by the lender, which in turn depends on its utility function. The main difficulty in implementing outcome tests stems from the infra-marginality problem (Ayres, 2002) consisting in having two groups of applicants with different risk distributions, which may could lead to a biased average outcome (Arnold et al., 2018). To address this issue, Dobbie et al. (2020) use a 2SLS method where loan take-up is instrumented by the leniency level of the loan officer in charge. The authors identify discrimination by scrutinizing how the long-term profit of the lender is affected by the probability of taking-up a loan across groups of applicants characterized by citizenship, gender, and age. A significantly positive coefficient for a group of

[3] To reduce the risk of omitting relevant variables, scholars tend to include as many control variables as possible in probit regressions, which in turn may create multicollinearity. Additional issues go beyond the concern of omitted variables (Ross, 2000). Loan granting is more than a binary, approval vs. denial decision. Typical loan proposals made by banks come with loan conditions such as loan size, charged interest rate, loan duration, collateralization requests, and even additional arrangements involving financial products, such as a life insurance or the opening of a bank account. Discrimination may be absent from the approval process but present in differentiated loan conditions (Agier & Szafarz, 2013b; Bayer et al., 2018).

applicants is considered evidence of bias against that group. Regardless of the nature of the outcome variable, valid discrimination tests based on outcomes require accounting for the endogeneity associated with the selection of borrowers and their loan conditions. For instance, default might be more frequent for higher-interest loans, which are presumably riskier (Stiglitz & Weiss, 1981). Endogeneity prevents ex post outcomes from being straightforward explanatory variables for loan conditions. Any practical implementation of outcome tests must therefore address endogeneity properly.

This paper tackles discrimination in the credit market by combining the advantages of both groups while circumventing their limitations. First, we go beyond the approval process by using reimbursement records as well. Doing so mimics the banking practice of basing approval decisions on credit scoring models (Boyes et al., 1989; Roszbach, 2004; Robb & Robinson, 2018) and thereby reduces the likelihood that captured biases are implicit or outside of the discriminator's awareness, to use Bertrand et al.'s terms (2005). Second, to address the endogeneity concern, we operationalize the theory proposed by Ferguson and Peters (1995) where credit discrimination is detected if an identified subset of the applicant pool suffers from both a lower (or equal) default rate and a higher (or equal) denial rate, provided that at least one inequality is strict. We address endogeneity concerns with instruments specific to the prosocial lending methodology.

Microfinance institutions and prosocial lenders tend to favor small loans and put reimbursement ahead of profit (Cozarenco & Szafarz, 2020). To reduce their operational costs, they typically grant loans with identical interest rate to all their borrowers. Therefore, the model of Ferguson and Peters (1995) offers an ideal theoretical context for developing our testing method. Note that Ferguson and Peters (1995, p. 744) implicitly acknowledge the infra-marginality problem since they claim that distinguishing between *marginal* and *average* borrowers is key and by subsequently defining discrimination in lending as "the use of different credit standards across the two components of the population, i.e. a policy that leads to the marginal borrower from each component of the population having a different credit score."[4] Importantly, their detection of lending discrimination based on observed loan denials or default rates is applicable to contexts with fixed interest rates. In addition, our estimation approach allows biases to be either negative or positive, which is particularly relevant for studying pro-social lending.

Section 2 presents the bivariate model combining approval and recovery rates. Repayment success is measured by the recovery rate, a ratio insensitive to loan size. Using the recovery rate as outcome variable has advantages and drawbacks. Its main advantage is that Becker's argument regarding the occurrence of defaults holds regardless of the (non)profit orientation of the lender, and so using the recovery rate maintains an agnostic view on the "true" utility function of the lender, which is a significant asset when dealing with socially oriented lenders. On the flip side, the recovery rate is a relatively rough indicator, which is certainly less informative than any smooth function like profit or interest rates. Subsequently, we derive the appropriate decision rule for the discrimination tests, with a special attention to intersectional discrimination.

To illustrate the merits of our suggested approach, Section 3 offers an application using microcredit data that combines two criteria: gender and citizenship (EU versus non-EU), where non-EU applicants are typically migrants from the South. The demand-side tests detect a positive bias in favor of female applicants (which makes perfect sense for microcredit) but no significant discrimination based on citizenship. Regarding intersectionality, we find that the globally positive bias toward women hides a divide between EU and non-EU women: Only non-EU women really benefit from the positive bias in loan allocation. Next, our bivariate methodology shows that the (seemingly) positive demand-side bias in favor of female applicants can be justified by higher recovery rates for the lender.

[4] This definition is close to the following statement by Ayres (2002, p. 135) "In the mortgage context, a test of disparate treatment would want to ask whether the least qualified whites to whom banks were willing to lend had a higher default rate than the least qualified minorities to whom banks were willing to lend."

In addition, bivariate intersectional tests reveal that, surprisingly, EU women suffer from negative bias compared with both EU men and non-EU women. While the bias in favor of non-EU women can be rationalized as the empowerment of migrants from non-EU countries (Aldén & Hammarstedt, 2016), there is no such argument explaining why EU women should be treated less favorably than male applicants from the same zone, suggesting that this is evidence of implicit (unintentional and likely unconscious) intersectional discrimination (De Andrés et al., 2020). In sum, our empirical design can help navigate the intricacies of lenders' attitudes toward demographic categories.

2. Testing for Discrimination in the Credit Market

Assessing discrimination in the credit market is hard for reasons pertaining to both economic theory and econometric issues. The problem stems from "the ambiguity of legal and theoretical definitions of discrimination" (Dimsky, 2006, p. 215), which conditions the way in which the question is framed econometrically. In the US, credit discrimination is a legal offense. The country's legal framework includes the 1968 Fair Housing Act, the 1974 Equal Credit Opportunity Act, and the 1975 Home Mortgage Disclosure Act. According to these laws, lenders must treat all borrowers equally with respect to protected characteristic such as race and gender. Since 1989, US lenders must report the race and ethnicity of their applicants. This general principle provides scholars with many testing possibilities.[5]

Economists typically present discrimination as a double standard that fails to be justified by an organization's profit maximization objective. In the credit market, lenders are said to exert discrimination if they take a harsher approach to granting loans to an identified category of applicants, such as women, without any economic justification. Discriminatory biases are not

[5] Recent contributions on discrimination in lending include, e.g., Cheng (2015), Beck et al. (2018), Cozarenco and Szafarz (2018), and Delis et al. (2020). Haselmann et al. (2018) show that social connections can lead to crony lending. Fisman et al. (2020) provide evidence on the impact of inter-group animosity on the credit market. Bayer et al. (2018), Bhutta and Hizmo (2021) and Ambrose et al. (2021) found mixed evidence of the impact of race and ethnicity on access to credit and loan conditions in the mortgage market.

necessarily intentional because stereotyping is a common human feature (Fein & Spencer, 1997). According to Buttner and Rosen (1988), women entrepreneurs suffer from gender stereotypes in terms of their perceived levels of leadership, autonomy, and emotionalism in running a business.

Lang and Nakamura (1993) stress that credit discrimination may arise because of information costs. If some variables potentially connected to creditworthiness are unobservable, lenders may be tempted to use gender as a proxy for credit risk, which leads to statistical discrimination (Arrow, 1971, 1998). Credit discrimination may be taste-based as well (Becker, 1971), meaning that the lender will willingly deny loans partly or totally to some groups (s)he dislikes. In fact, discrimination implies disparate treatment, but the reverse is not true: Disparate treatment may well be economically justified by objective credit risk characteristics. If female entrepreneurs had a higher credit risk than men, all else equal, then disparate treatment would be the lender's rational reaction and would not be considered gender discrimination. Since reimbursement records are needed to assess credit risk, researchers who fail to consider them are left with no other choice than to rely on the (strong) assumption of equal creditworthiness across all tested characteristics. In sum, the main problem plaguing empirical studies on credit discrimination is that the lender's assessment of creditworthiness is a black box for researchers (Cornée, 2019).

One way to identify biases in loan allocation would be to run an experiment holding loan application characteristics fixed while varying gender, as in Fay and Williams (1993), or exploit exogenous variation in loan allocation provided by an explicit staff rotation policy, as in Fisman et al. (2020). Another way would be to have access to the detailed decision process of the lender, and this is the approach we took. It is based on regressions explaining the lender's decision-making while controlling for as many covariates as possible. This approach is reliable if all the relevant variables taken into consideration by the lender are also considered by the econometrician.[6]

[6] This condition is fulfilled in the case of our application to microcredit in Section 3. We had access to the full application files. In addition, the loan officers do not meet their clients, so that we did not miss out on any "soft" information stemming from informal contacts between applicants and loan officers (see Iyer et al., 2016, on the link between soft information and biases in lending).

To explain our approach, let us now consider a rational risk-neutral lender deciding whether to issue a loan with fixed conditions. There are two options: Either the loan is denied, and the future return is zero, or the loan is approved, and the future cash flow depends on the outcome of the loan. If no default occurs, reimbursement is $LS(1+r)$, where *r* is the interest rate charged and *LS* is the loan size. A default, on the other hand, generates a loss for the lender. The resulting loss given default (LGD) corresponds to the debt write-off. Ultimately, the lender's optimization problem boils down to approving the loan if the present value of $E[(1+r)LS - LGD]$ is higher than LS, where LGD is the only random part of the future cash flows. To derive a simple decision-making rule that mimics the usual banking procedure, we use the relative measure known as the recovery rate:

$$Recovery\ rate = \frac{LS(1+r)-LGD}{LS} \tag{1}$$

Equation (4) yields the following decision-making rule:

$$Loan\ Approval = 1 \iff E[Recovery\ rate] \geq 1 \tag{2}$$

We assume that the expected recovery rate is rational, and therefore unbiased, which rules out any discriminatory loan allocation that might be attributed to statistical discrimination based on wrong preconceptions.

The recovery rate is a key variable in credit scoring. Banks typically address informational asymmetries between lenders and borrowers by assessing applicants' creditworthiness, using credit scoring techniques. In line with the Basel framework on credit risk management, scoring models are based on reimbursement records (Boyes et al., 1989). They rely on various statistical methods, including discriminant analysis, linear and logistic regressions, neural networks, and hybrid models. These models associate credit risk levels to clientele segments, so that loan applicants falling in categories with a lower estimated creditworthiness therefore face a higher probability of denial, while controlling for their personal credit history.

When the pool of borrowers is split into two observable categories, say women and men, and the administrative data are limited to applicants' characteristics and loan approval/denial, the common practice is testing for discrimination under the assumption of equal creditworthiness. The test is involves estimating the following probit model:

$$Approval_i = \mathbb{1}[\alpha_A F_i + \beta_A' \boldsymbol{Z_i} + v_i > 0] \quad (3)$$

where $\mathbb{1}[Y > 0]$ is the dummy variable taking the value of 1 if Y is positive, and 0 otherwise. Index i refers to loan applicants; F_i takes the value of 1 if applicant i is a woman, and 0 otherwise; $\boldsymbol{Z_i}$ is a vector of control variables that includes the applicant's characteristics; v_i is the error term. Discrimination is assessed by running a bilateral test on α_A in Equation (3), with a fixed significance level.

Dropping the equal-creditworthiness assumption requires additional data, such as observations of recovery rates. From an economic standpoint, assessing creditworthiness *per se* is motivated by risk management and cannot be considered discriminatory. In contrast, disparate treatment of a given group would result from group members exhibiting both higher (or equal) creditworthiness and higher (or equal) probability of denial. The empirical setting that we suggest relies on the theory developed by Ferguson and Peters[7] (1995) while using recovery rates, as proposed by Shaffer (1996). Using recovery rates rather than default rates is a way to make sign comparisons intuitive and therefore easy to interpret. The tests exploit probabilities of default estimated from both individual credit histories and group-based recovery rates (Altman et al., 2004). The originality of our bivariate approach stems from combining the basic idea of outcome

[7] Similarly, favorable bias would mean group members exhibiting both lower (or equal) creditworthiness and lower (or equal) probability of denial. To motivate their model, Ferguson and Peters (1995, p. 740) compare two existing claims on discrimination in lending: "The first claim is that differences in denial rates are likely due to differences in average credit quality between white and minority applicants. The second claim is that equal default rates indicate that minorities and whites are being held to the same credit standard (…). [B]oth of the claims (…) cannot be true simultaneously. In fact, if the first claim is correct, then contrary to the second claim, equal default rates imply that minorities are being discriminated against." Note that the model of Ferguson and Peters (1995) is a necessary, but not sufficient, condition for discrimination since it fails to identify discrimination if a group has both a higher approval rate and a higher creditworthiness but the difference regarding approval rates strongly exceeds the difference regarding creditworthiness.

tests (Ayres, 2002; Dobbie et al., 2020) and the routine practice of banks (Butler et al., 2021). In addition, using recovery rate as outcome variable is robust to contextual features, such as loan characteristics (duration, collateralization, and interest rate) and lender's legal status (for-profit, nonprofit, or hybrid). In sum, Equation (2) follows the banking routine,[8] which dictates that any applicant's characteristic that constitutes a positive factor for the recovery rate ought to increase the probability of loan approval. Since recovery rate can only be observed for approved applicants, we need to address the concern of endogeneity stemming from this selection bias by applying the Heckman (1979) estimation method.

Our two-equation model used to test for the presence of biased credit granting writes:

$$Recovery\ rate_i = \alpha_R F_i + \beta_{\boldsymbol{R}}' \boldsymbol{X_i} + \varepsilon_i \tag{4}$$

$$Approval_i = \mathbb{1}[\alpha_A F_i + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i} + v_i > 0] \tag{5}$$

where $v \sim N(0,1)$ and $E(\varepsilon|v) \neq 0$. The control variables in vector $\boldsymbol{X}$ include the applicant's characteristics while vector $\mathbf{Z}$ is obtained by stacking $\boldsymbol{X}$ and a set of instruments affecting approval decision-making but not the recovery rate, as required by Heckman's estimation method. In practice, we estimate two equations—one for the recovery rate and the other for approval probability—and compare the signs of the coefficients of interest across the two equations.

The test procedure is described in Table 1. A bias either favoring or hindering women is observed where $\alpha_A \leq 0$ and $\alpha_R \geq 0$, with at least one strict inequality. Appendix A extends our test procedure to intersectionality. Further, we refer to strong or weak bias depending on the number of strict inequalities (at a given level of significance). Thus, strong bias means that both criteria are significant (for example, lower approval rate *and* higher recovery rate) while weak bias points to one significant inequality only (for example, higher denial rate with equal recovery rate).

[8] Actual recovery rates are observed after repayment. Using *ex post* recovery rates to assess the fairness of loan allocation relies on the assumption that the lender's expectations were formed rationally.

The classification in Table 1 detects both negative bias against women and positive bias favoring them thanks to the signs of the tested parameters, α_R and α_A.

Table 1. Detecting Simple Discrimination with Bivariate Estimation

	Higher approval rate for women: $\boldsymbol{\alpha_A} > 0$	**Insignificant difference between approval rates: $\boldsymbol{\alpha_A = 0}$**	**Lower approval rate for women: $\boldsymbol{\alpha_A} < 0$**
Higher recovery rate for women: $\boldsymbol{\alpha_R} > 0$	No bias detected	Weak negative bias	Strong negative bias
Insignificant difference between recovery rates: $\boldsymbol{\alpha_R = 0}$	Weak positive bias	No bias detected	Weak negative bias
Lower recovery rate for women: $\boldsymbol{\alpha_R} < 0$	Strong positive bias	Weak positive bias	No bias detected

Note: This table provides the decision rule of the test for discrimination based on Heckman estimation of the recovery rate and loan approval rate. α_A and α_R are the coefficients of the dummy variable that takes the value of 1 if the applicant is a woman, and 0 otherwise, in Equations (4) and (5), respectively.

The decision rule in Table 1 generalizes Figure 2 in Ferguson and Peters (1995). It is consistent with the intuition of how biased loan allocation works, namely, by making borrowing more difficult for applicants who display a characteristic that is visible to the lender, even though it has either no influence or a positive influence on objective creditworthiness, all else equal. Yet the proposed rule subsumes this specific situation since it deals with both positive and negative forms of discrimination. It also introduces a weak/strong dichotomy, which emphasizes the extent of potentially disparate treatment by lenders. The next section applies our proposed method for addressing intersectionality in the credit market to a case study that combines positive and negative biases.

3. An Application to Intersectional Discrimination in Pro-Social Credit

3.1. What Do we Known about Intersectional Biases?

Recently, discrimination in lending, both positive and negative, has gained renewed interest owing to the considerable success of microcredit, which targets vulnerable populations, with a special focus on poor women. Evidence shows that not all women are treated equally by microlenders (Agier & Szafarz, 2013a; Beck et al., 2018), confirming the view that stereotypical biases vary within demographic categories, such as gender (Hall et al., 2019; Harkness, 2016) and ultimately leading to the concept of intersectional discrimination (Ruwanpura, 2008). This application contributes to the literature by using our methodology for tracking intersectional biases in the pro-social credit market.

Thus far, identifying intersectional discrimination has raised two methodological challenges: The first relates to the multiplicity of reference groups, the second to the potential coexistence of positive and negative biases. Let us illustrate these two issues with an example involving the criteria of gender and attractiveness. Consider a lender who is prejudiced against female loan applicants and, separately, has gender-dependent opposite stereotypical views toward attractiveness (negative for women and positive for men). The first problem stems from interpreting the empirical results. The routine use of a single reference group, say plain-looking men, would constitute a strong limitation since it would prohibit comparisons between the fates of good-looking men and plain-looking women. Instead, our framework elicits multiple pairwise comparisons.

The second issue relates to the potential coexistence of negative bias against a segment of the borrower pool and positive bias toward another segment. Such a combination may mask meaningful intersectional effects. For instance, the lender who favors plain-looking women and good-looking men could well fly under the radar of tests for discrimination based on either gender or attractiveness. Researchers who fail to detect any significant simple bias (i.e., based on a single criterion) seldom further explore the lender's attitude toward subgroups. We address this issue with a testing design that makes explicit which bias is positive and which one is negative. This normative approach assumes that vulnerable groups are pre-defined. For instance, a lender favoring men is said to exert negative discrimination, whereas promoting female borrowers would be called affirmative action, even though, formally speaking, the two actions are identical. The potential legal consequences of this prerequisite are discussed in the conclusion.

3.2. Context

Microcredit in developed countries is a young industry filling a niche market (Cozarenco & Szafarz, 2019 & 2020b). Starting with the Arkansas pilot project (Taub, 1998), microfinance institutions (MFIs) in the US have typically replicated the Grameen model and focused on disadvantaged women. Likewise, Europe has witnessed the development of Grameen-style MFIs, such as the French ADIE, created in 1989. Bendig et al. (2014) mentioned the existence of over 500 MFIs in Europe. The largest survey to date (Diriker et al., 2018) points out that the legal structures of European MFIs contrast with those of standard financial intermediaries. These structures include NGOs (40%), non-bank financial institutions (29%), cooperatives and credit unions (19%), commercial banks (6%), and others (6%). European MFIs are typically subsidized, and their activity is limited to supplying standardized short-to-medium-term microcredit.

Our hand-collected dataset was provided by a subsidized French NGO set up in 2007 that grants microcredit to economically fragile individuals rejected by traditional banks owing to low income, unemployment, insecure employment, over-indebtedness, and/or bad credit history.[9] At the time, the NGO was mainly funded by the *Caisse d'Epargne Bretagne Pays de Loire*, a regional branch of the French Federation of Savings Banks. Between 2007 and 2014, the lender received 6,237 loan applications, of which 3,709 were approved. These figures put the overall approval rate at 59%. The full-period average loan size was EUR 2,231, the average duration 33 months, and the

[9] For more information about the lender, see its website: http://www.parcoursconfiance-bpl.fr/about-us/.

annual interest rate matched the return on the state-controlled *Livret A*,[10] which fluctuated between 1% and 4% during the investigation period. By microcredit standards, this is a particularly favorable rate for borrowers. The interest rate is determined irrespective of borrower characteristics, which is a common practice in microfinance. This feature of microcredit makes our setting, where discrimination is identified regardless of loan pricing,[11] particularly suitable for our dataset.

Poor women and migrants are two of the impoverished groups typically targeted by European microfinance, along with the unemployed (Bendig et al., 2014). Therefore, we will use *gender* and *citizenship* as characteristics to illustrate the test methodology. We will check whether the lender makes a (positively or negatively) biased loan allocation toward women and/or non-EU citizens. We did not observe the precise nationalities of non-EU applicants, but informal contacts with the NGO's staff revealed that most non-EU applicants were from developing countries.

The screening of loan applicants took place remotely, based on paper files that were filled in by partner NGOs, local public authorities, and charitable organizations such as the Red Cross, *Fondation Abbé Pierre* or *Restaurants du Coeur*. All these organizations serve disadvantaged populations and help them apply for loans if necessary. Loan officers received paper applications and uploaded the relevant information in the lender's system, which has the notable advantage of keeping track of all applications regardless of whether they were successful or not. Each application was analyzed by a loan officer who shortlisted applicants to be presented to the credit committee, composed of two or three loan officers and the NGO's director. The final decision was made by the committee on the basis of discussions and did not involve any algorithmic inputs.

Our data was retrieved from the paper applications. We can thus reasonably assume that we had access to the same information set as the lender. In addition, we observed the repayment

[10] The *Livret A* is a regulated, riskless passbook savings account. Its interest rate serves as a reference for financial agreements.
[11] See Delis and Papadopoulos (2019) about discrimination in loan pricing.

conduct of all borrowers as of November 2016. We therefore cleaned the sample by removing loans with a later maturity date.

Our final sample comprises 5,789 applicants and 3,262 borrowers, of whom 421 (13%) defaulted. For defaulted loans, the dataset records both the default date and the monetary loss (in EUR), that is, the outstanding debt at the time of default. In sum, we followed each application file throughout its entire journey, whether it ended in rejection, default, or successful repayment.

3.3. Data

Table 2 shows summary statistics concerning applicants, borrowers, and defaulters. The average recovery rate among defaulters is 41.49%, meaning that the average defaulter creates a loss for the lender of 58.51% of the initial loan amount. In the microcredit niche market, all selected applicants do take out the loan. Hence, there is no bias associated with post-offer loan rejection by applicants (Bai & Lu, 2020).

Table 2. Descriptive Statistics for Applicants and Borrowers

	Panel I: Applicants				**Panel II: Borrowers**		
	Full Sample	**Rejected**	**Borrowers**	**Difference**	**Repaying**	**Defaulting**	**Difference**
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Percentage		43.646	56.354		87.094	12.906	
Recovery rate (%)			92.447		100	41.492	58.508***
Female (dummy)	0.485	0.450	0.511	-0.061***	0.517	0.473	0.045*
Non-EU (dummy)	0.086	0.089	0.083	0.005	0.082	0.090	-0.007
Age (years)	38.671	38.330	38.936	-0.006*	39.217	36.999	0.022***
Single (dummy)	0.590	0.553	0.619	-0.065***	0.621	0.601	0.020
# of children	0.934	0.981	0.898	0 .082**	0.890	0.952	-0.063
HH income (EUR)	963.254	987.626	944.338	43.288**	958.152	851.328	0.107***
Unemployed (dummy)	0.404	0.383	0.420	-0.038***	0.413	0.470	-0.057**
Bad credit history (dummy)	0.332	0.386	0.289	0.097***	0.266	0.447	-0.180***
Observations	5,792	2,530	3,262		2,841	421	

Note: This table reports the descriptive statistics for the variables collected by the MFI. Panel I and Panel II show the mean values of the observations for applicants and borrowers, respectively, as well as the results of t-tests for equal means for rejected applicants and borrowers, and for repaying and defaulting borrowers, respectively. *** p<0.01, ** p<0.05, * p<0.1

The test for intersectional discrimination focuses on gender and citizenship. Table 2 shows that the percentage of women is higher among borrowers than among applicants (51% versus 45%), which implies that women enjoy a higher approval rate. Accordingly, women are less likely to default: Only 47% of defaulters are female. Table 2 indicates no significant difference in approval rates between EU and non-EU applicants. Likewise, EU and non-EU borrowers are as likely to default.

Other characteristics recorded by the lender include age, marital status, number of children, household monthly income, and unemployment. These variables will serve as controls in the regressions. On average, borrowers are slightly older than rejected applicants (39 years of age against 38). We observe the same trend among repaying borrowers, whose average age, 39, exceeds the defaulters' average age of 37. Marital status, number of children, and household income interact with loan approval but not with default. Unemployed applicants are more likely to obtain a loan than other applicants. This may be due to affirmative action since unemployed borrowers have a higher default rate according to Panel II. Unsurprisingly, applicants with a bad credit history are rejected more frequently than other applicants, and they also have a higher likelihood of default. In our dataset, bad credit history is a dummy variable taking the value of 1 if the applicant is registered in the French National Bank's register of household credit repayment, which lists the individuals who have missed repayments on a loan and those who have filed an over-indebtedness application.[12]

3.4. Probit Estimation

Following the testing procedure described in Appendix A (Table A1), we first estimate a probit equation to check for simple discrimination toward women and non-EU citizens. The approval equation includes year fixed effects to account for both a potential time trend and the MFI's life cycle (Bogan, 2012). Like Butler et al. (2021), we use standard errors that are two-way clustered at the county (*département*) level and by year of application.

First, in Table 3, column (1), we disregard potential intersectional biases. The average marginal effects indicate that women's loan applications are 2.8% more likely to be approved than

[12] https://particuliers.banque-france.fr/sites/default/files/media/2018/11/23/654ta18_818235_depliant_ficp_web-en.pdf

men’s. Such a level of affirmative action seems reasonable given the pro-women bias often claimed by the microfinance industry. In contrast, the coefficient pertaining to non-EU citizenship is insignificant, suggesting that the lender has a neutral attitude toward non-EU citizens versus EU ones.

Next, we factor in intersectional discrimination by adding the interaction term standing for non-EU women in Table 3, column (2). Panel I shows the estimation results and Panel II features the total marginal effects used to assess intersectional biases and their standard errors, to be interpreted according to Table A1. Owing to the non-linearity of the model, the effect of one variable on the approval probability depends on local values of all the variables (Berry et al., 2010). To address this issue, we compute local marginal effects for each individual applicant and the corresponding z-statistics. Figures B1–B5 in Appendix B confirm the results displayed in Table 3.

Overall, the results show that intersections do matter. The first three lines in Table 3, Panel II, compare non-EU women with other subgroups. The typical social mission of microfinance requires lenders to offer credit to disadvantaged groups. In our case, the most disadvantaged group is made up of female migrants, also referred to as non-EU women. Table 3 shows that non-EU women are indeed treated better than any other subgroup. Figures B1–B3 suggest that, all else being equal, non-EU women are 6.3% more likely to receive a loan than EU-women, 8.3% more likely than EU men, and 13.4% more likely than non-EU men. These facts corroborate laboratory evidence about microlending in Bolivia provided by Martinez et al. (2020), who uncovered positive discrimination in favor of non-indigenous women.

Table 3. Probit Estimation of Approval

	Simple discrimination	Intersectional discrimination
	(1)	(2)
Variables	Panel I: Average marginal effects	
Female ($\widehat{\alpha}$)	0.028**	0.020
	(0.014)	(0.015)
Non-EU ($\widehat{\theta}$)	-0.010	-0.050*
	(0.022)	(0.027)
Female*Non-EU ($\widehat{\mu}$)		0.114***
		(0.038)
Control variables		
Age (years/100)	0.130**	0.130**
	(0.062)	(0.061)
Single (dummy)	0.102***	0.102***
	(0.019)	(0.019)
# of children	0.003	0.002
	(0.006)	(0.006)
Ln (HH income in thousands of EUR)	0.188***	0.188***
	(0.037)	(0.037)
Unemployed (dummy)	-0.019	-0.020
	(0.015)	(0.014)
Bad credit history (dummy)	-0.102***	-0.102***
	(0.014)	(0.014)
Year FE	Yes	Yes
Unemployment rate in town of residence	0.016***	0.016***
	(0.004)	(0.004)
Unemployment rate at national level	-0.131***	-0.129***
	(0.048)	(0.047)
Pseudo R-squared	0.163	0.164
Number of observations	5,792	5,792
	Panel II: Total marginal effects	
Non-EU women vs. EU men		0.083**
		(0.033)
Non-EU women vs. EU women		0.063**
		(0.032)
Non-EU women vs. non-EU men		0.134***
		(0.038)
EU women vs. EU men		0.020
		(0.015)
Non-EU men vs. EU men		-0.050*
		(0.027)

Note: This table reports the test results for simple (column (1)) and intersectional (column (2)) forms of discrimination with a probit estimation of approval, using Equations (1) and (2), respectively. In Panel I, we report the average marginal effects. Panel II gives the total marginal effects for intersectional discrimination. Standard errors (SE) clustered by county-year are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

By contrast, EU women and EU men seem to have similar approval probabilities. The total marginal effect fluctuates between 0.1% and 2.4% (Figure B4a), with an average value of 2%

(Panel II). The corresponding z-statistics takes values below the 10% significance threshold (Figure B4b), suggesting that the MFI adopts homogenous screening for the two groups. However, non-EU men have lower approval probabilities than EU men. The marginal effect of approval varies between -6.1% and -0.4% (Figure B5a) with z-statistics scattered below the 10% significance threshold (Figure B5b) and only 9% of the computed values lying above the threshold. Taken together, the figures suggest that the marginal effect is significant at the 10% level, which may indicate borderline discrimination against non-EU men as compared to EU men.

3.3 Bivariate Estimation

As Sections 2.2 and 2.3 showed, using recovery rates to proxy clients' credit risk enriches the analysis by making the previous assumption of equal creditworthiness unnecessary. The econometric framework relevant for this approach includes the bivariate model presented in Appendix A and Heckman estimation to address the endogeneity issue stemming from borrower selection. Again, we proceed in two steps: First we consider simple discrimination only, and then we move to intersectional discrimination with several reference groups.

To check for simple discrimination in relation to women and non-EU citizens, taken separately, we run the regressions reported in Table 4: Column (1) provides the estimated coefficients of the recovery-rate equation and column (2) displays the marginal effects for the selection equation. We run a Heckman estimation and use the unemployment rates at the time of loan application (both in the applicant's place of residence and in France as a whole) as instruments. Table 4 shows that these two unemployment rates affect loan approval in opposite directions, which makes perfect sense given the business model of microcredit: Applicants from places with a higher unemployment rate are more likely to receive a loan, corroborating the social bottom line of the MFI, whereas an increase in the national unemployment rate tends to reduce approval rates. The correlation between the local and national unemployment rates is 0.40, well below the upper threshold of 0.9 suggested by Hair et al. (2010) for using two explanatory variables in the same regression. Following Delis et al. (2020), we argue that the two unemployment rates are good instruments because: a) they are predetermined and do not directly affect the recovery

Table 4. Bivariate Estimation of Simple and Intersectional Forms of Discrimination

	Simple discrimination		Intersectional discrimination	
	Recovery rate	Approval	Recovery rate	Approval
	(1)	(2)	(3)	(4)
Variables	Panel I: Coefficients and average marginal effects			
Female ($\hat{\alpha}$)	0.016**	0.028**	0.021***	0.020
	(0.007)	(0.014)	(0.007)	(0.015)
Non-EU ($\hat{\theta}$)	-0.016	-0.009	0.009	-0.050*
	(0.014)	(0.022)	(0.016)	(0.027)
Female*Non-EU ($\hat{\mu}$)			-0.055**	0.114***
			(0.027)	(0.038)
Control variables				
Age (years/100)	0.131***	0.131**	0.132***	0.130**
	(0.030)	(0.062)	(0.030)	(0.061)
Single (dummy)	-0.004	0.102***	-0.004	0.102***
	(0.007)	(0.019)	(0.007)	(0.019)
# of children	-0.003	0.003	-0.003	0.002
	(0.004)	(0.006)	(0.004)	(0.006)
Ln (HH income in thousands of EUR)	0.029***	0.188***	0.029***	0.188***
	(0.010)	(0.037)	(0.010)	(0.037)
Unemployed (dummy)	-0.017*	-0.019	-0.016*	-0.020
	(0.009)	(0.015)	(0.009)	(0.014)
Bad credit history (dummy)	-0.069***	-0.102***	-0.069***	-0.102***
	(0.011)	(0.014)	(0.011)	(0.014)
Year FE	Yes	Yes	Yes	Yes
Instruments				
Unemployment rate in town of residence		0.016***		0.016***
		(0.004)		(0.004)
Unemployment rate at national level		-0.131***		-0.129***
		(0.048)		(0.048)
Constant	0.905***		0.901***	
	(0.016)		(0.016)	
Rho ($\hat{\rho}$)	0.086***		0.089***	
	(0.035)		(0.034)	
Mills ratio	0.018***		0.019***	
	(0.008)		(0.008)	
Number of observations	3,262	5,792	3,262	5,792
	Panel II: Total marginal effects			
Non-EU women vs. EU men			-0.025	0.085**
			(0.025)	(0.033)
Non-EU women vs. EU women			-0.046**	0.065**
			(0.023)	(0.032)
Non-EU women vs. non-EU men			-0.034	0.134***
			(0.028)	(0.038)
EU women vs. EU men			0.021***	0.020
			(0.007)	(0.015)
Non-EU men vs. EU men			0.009	-0.050*
			(0.016)	(0.027)

Note: This table reports the test results for simple (columns (1) and (2)) and intersectional (columns (3) and (4)) forms of discrimination with a Heckman estimation. In Panel I, columns (1) and (3) show the estimated coefficients for the recovery rate in Equations (6) and (8), respectively. Columns (2) and (4) show the average marginal effects for the approval rate in Equations (7) and (9), respectively. Panel II gives the total marginal effects for intersectional discrimination. Standard errors (SE) clustered by county-year are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

rate (see Table C1, Appendix C), and b) they strongly correlate with approval.

The results in Table 4 show that, at the 95% level of confidence, women have higher recovery rates and are more likely to have their loans approved. By contrast, non-EU citizenship has no significant impact on either the recovery rate or the probability of approval. Overall, the bivariate estimation provides no evidence of simple discrimination based on gender or citizenship. The significance of the Mills ratio and the correlation between the error terms of the two equations (ρ) confirm the relevance of the Heckman approach to addressing the selection bias, which would otherwise distort the results.

The impacts of the control variables are in line with the summary statistics in Table 2. Age positively impacts the recovery rate and older applicants are more likely to be successful. Single applicants are more likely to secure a loan, but their creditworthiness does not differ significantly. According to Table 1, this result may be interpreted as a weak positive bias in favor of single applicants. The number of children has no significant impact. The approval process is in line with economic rationality as far as household income, unemployment, and credit history are concerned.

In Table 4, columns (3) and (4), we implement the tests for intersectional biases presented in Table A2 even though we failed to detect any simple discrimination. Panel I reports the marginal effects and robust standard errors clustered at the county-year level for the model combining Equations (8) and (9). Panel II provides the figures that are essential if we wish to assess intersectional discrimination by comparing groups' total marginal effects. Each line in Panel II corresponds to a specific test involving two intersectional groups. Column (3) checks whether there is any significant difference between recovery rates while column (4) pits one group's approval rates against the other's.

The econometric issues stemming from the non-linearity of the estimated loan-approval model are the same as in the previous section but adding intersectional discrimination to the picture imposes another layer of complexity. Therefore, instead of using the same individual marginal effects as in Appendix B, we opt for a simpler, yet more aggregate, representation of local marginal effects. Namely, Figure 1a uses polynomial approximations to draw curves representing the (non-linear) marginal effects of each intersectional group. By symmetry, Figure 1b draws the constant marginal effects estimated from the (linear) recovery-rate equation. In both figures, the effects associated with (at least) 5%-significant estimates are shown in black while less significant results are presented in light grey.

Figure 1 helps summarize the conclusions regarding intersectional discrimination as follows. First, non-EU women benefit from a weak positive bias compared with both EU and non-EU men. By the same token, we detect a strong positive bias in favor of non-EU women compared with EU women. Even though the recovery rate of non-EU women is significantly lower than that of EU women, their likelihood of receiving a loan is significantly higher. Arguably, this bias is in line

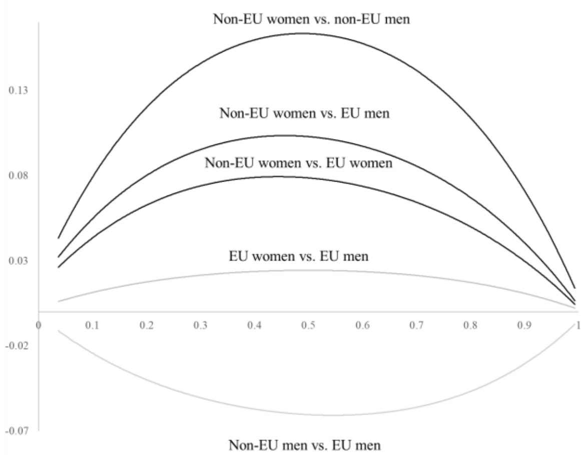


Figure 1a. Marginal effects for Approval

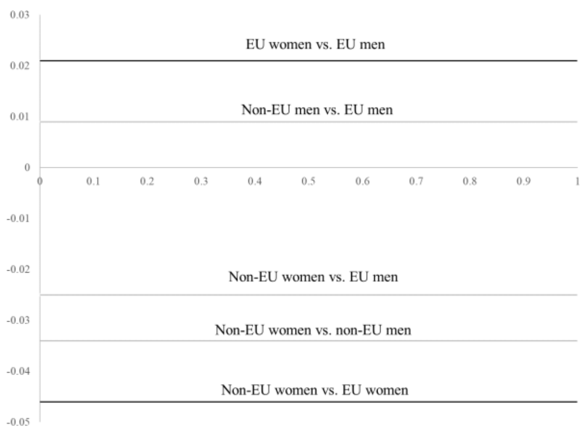


Figure 1b. Marginal effects for Recovery rate

with the social mission of helping vulnerable borrowers such as migrant women. Less expectedly, given the pro-women agenda of microfinance, Figure 1 suggests that the recovery rates of EU women are significantly higher than those of EU men, but this higher creditworthiness is not rewarded in terms of loan approval (insignificant marginal effect). The decision rule in Table A2 would therefore suggest that EU women are weakly discriminated against when compared with EU men. Last, Figure 1 reveals that non-EU and EU men have similar recovery rates, but non-EU men face a lower approval probability (albeit insignificant at the 5% level) that may be consistent with a negative bias against non-EU men compared with EU men.

Table 5. Summary of Test Results for Discrimination (at the 5% level)

Panel I: Simple Discrimination — Probit Estimation			
Comparing	with men		with EU applicants
women	Positive bias		
non-EU applicants			No bias detected
Panel II: Simple Discrimination — Bivariate Estimation			
Comparing	with men		with EU applicants
women	No bias detected		
non-EU applicants			No bias detected
Panel III: Intersectional Discrimination — Probit Estimation			
Comparing	with EU men	with EU women	with non-EU men
non-EU women	Positive bias	Positive bias	Positive bias
EU women	No bias detected		
non-EU men	No bias detected		
Panel IV: Intersectional Discrimination — Bivariate Estimation			
Comparing	with EU men	with EU women	with non-EU men
non-EU women	Weak positive bias	Strong positive bias	Weak positive bias
EU women	Weak negative bias		
non-EU men	No bias detected		

Note: This table summarizes the results for simple and intersectional forms of discrimination in Tables 3–4, based on the decision rules in Tables 1, A1, and A2.

Table 5 brings together all the test results using the 5% threshold. Panel I, based on a probit estimation, suggests a positive bias toward women and no bias toward non-EU citizens. Panel II adds recovery rates to the picture, suggesting that the positive bias toward women as regards loan approval may be rationalized by their higher creditworthiness. Integrating intersectionality leads to more nuanced conclusions. In Panel III, the only bias detected is in favor of non-EU women, who are more likely than any other group to be granted a loan. Last, the Heckman estimation (Panel IV) confirms the positive bias in favor of non-EU women compared with any reference group. This bias is weak with reference to men, regardless of their nationality, but strong with reference to the other (i.e., EU) women. But the most striking test result in Panel IV pertains to the comparison, overlooked by other estimations, between the fates of female and male EU citizens. The intersectional discrimination tests uncover a weakly disparate treatment of EU women compared with EU men.

In sum, the positive bias expressing affirmative action in favor of non-EU women is paradoxically stronger when they are compared to EU women rather than to any group of men. This difference in treatment between women of different nationalities is magnified by the harsher handling of loan applications made by EU women compared to their male counterparts. Overall, the winners are non-EU women and the losers are EU women.

The positive attitude of prosocial lenders toward the typically most disadvantaged group (i.e., non-EU women) makes perfect sense. The same type of preference would, however, point to the typically privileged group—EU men—as predictable losers, which contradicts our empirical findings. In contrast, our results are consistent with previous work highlighting the fact that gender stereotypes can be the driving force behind disparate treatment of female loan applicants, and these stereotypes can resist the barrier of the microfinance social mission (Garikipati et al., 2017). Our results confirm that intersectionality matters (Hall et al., 2019). From a methodological standpoint, focusing on a single interaction term in the approval equation fails to grasp the full complexity of intersectional biases. By contrast, our bivariate approach has delivered new insights showing how even pro-social lenders with a women's empowerment agenda can be affected by negative gender stereotypes.

4. Conclusion

The long history of discriminatory behavior on many markets indicates that competition fails to drive out discrimination (Sunstein, 1991) suggesting that at least some discriminatory biases have no impact on profits. Following this argument, we suggest testing for discrimination in lending by studying recovery rates and checking whether applicant groups would combine systematically higher, resp. lower, or equal loan denial rates and systematically higher, resp. lower, or equal recovery rates, with at least one strict inequality. This methodology has the merit of taking into consideration both demand-side and supply-side factors.

Regarding the current discussion on infra-marginality, the Ferguson and Peters (1995) models that has guided the development of our testing design is meant to capture situation where the *marginal* minority applicant is less successful than the *marginal* majority applicant taking into consideration potential differences in default probabilities. Therefore, this model accounted for infra-marginality even before the concept was defined in the literature.[13] In addition, the bivariate testing design we propose applies to any type of lender, be they for-profit, non-profit or even hybrid, i.e., combining social and financial objectives.

In several markets, such as pro-social lending and markets with credit rationing, discrimination can be costless, and has therefore little chances to be washed out by competition. We argue that, in some markets, costless discrimination appears to be the rule rather than the exception. These situations—and how to act on them—likely deserve more consideration than they have received so far. Recent evidence shows that economists tend to underestimate existing discriminatory behaviors plaguing not only economic markets, but also the academic profession, particularly economics scholars (Card et al., 2020; Dupas et al., 2021).

[13] Butler et al. (2021) indicate that biases attributable to infra-marginality in repayment tests tend to work against finding discrimination.

The method proposed in this paper may be subject to caveats. Three appreciable features concur to improve the data analysis proposed in this paper: 1) we exploited data on both loan granting and repayment, 2) the dataset included all variables collected by the lender itself and, 3) there was no face-to-face meeting between applicants and loan officers. Yet it is still conceivable that unobservables correlated with gender and/or ethnicity might have affected both loan approval and loan repayment. That, however, is unlikely to be the case. Another potential issue lies in the assumptions related to the lender's attitude toward credit risk. Recovery rates used in this paper might imperfectly capture the full credit risk structure by neglecting the impact of default likelihood at the time of loan issuance. In our setting, the lender is assumed to be risk-neutral and form rational expectations of recovery rates; its only decision is loan approval vs. denial. We also imposed the functional forms used in our estimates. Extending our setting to discriminatory biases associated with learning deficits, inattention, and wrong beliefs—like in Coffman et al. (2021)—is a fruitful avenue for further work on discrimination in behavioral economics.

The case study presented in this paper shows that our test design is relevant in the context of prosocial lending (Serrano-Cinca & Gutiérrez-Nieto, 2016) if we wish to assess the fairness of the loan granting process with respect to pre-defined reference groups. Microcredit in the North is, however, a niche market with special features, such as a flat interest rate and no collateralization. One may therefore question the external validity of our results. We contend that, like the instrument-based methods proposed by Butler et al. (2021) and Bayer et al. (2020) for detecting (simple) discrimination in lending, our *intersectional* discrimination approach is transposable to any outcome variable and therefore applicable to a wide range of lending institutions in which biased loan allocation is suspected.

A promising field of application pertains to algorithmic discrimination in bank lending. Through a systematically fine-tuned assessment of loan allocation to subcategories of applicants, our method pushes forward the research agenda on algorithmic-based detection of biases called for by Kleinberg et al. (2018). Theoretical refinements include both a probabilistic approach to credit risk and the setting of many loan conditions, such as maturity, loan size, interest rate, and collateralization (Melnik & Shy, 2015).[14] The credit market is characterized by a lack of transparency in the screening of loan applicants, and clear-cut econometric identification is often out of reach for scholars and courts wishing to assess the fairness of a lender. By using both loan approval and repayment records, our approach helps address an important ethical problem while reducing the impact of unobservables, which have plagued standard tests for discrimination in lending (Qi et al., 2018). Ultimately, the main limitation stems from data availability since the empirical design relies on high-quality, individual—and typically sensitive—data from a single lending institution (Delis et al., 2020).

In our setting, the standard dichotomy between taste-based and statistical discrimination was challenged by the possibility of affirmative action (i.e., favoring loan applicants who belong to typically disadvantaged groups). Depending on the context, this positive discrimination may be interpreted as taste-based, statistical, or both. Taste-based biases could stem from socially minded lenders willing to disregard recovery rates information and concentrate on group membership only. Prosocial lenders using statistical discrimination might prioritize groups with a lower creditworthiness, interpreted as a signal of economic hardship. Disentangling these two types of discriminatory behavior is often arduous as regards negative biases. If we acknowledge the possibility of both positive and negative forms of discrimination toward intersectional groups, then it is hazardous to assume that all detected biases will have the same origin. In sum, owing to the

[14] These aspects matter moderately for the microfinance industry, which typically grants standardized loans free of any collateral requirement, but they can play a more decisive role when studying commercial banks. To address this situation, one could assume joint decision-making about loan approval and interest rate, where differences in interest rates can, in turn, affect the probability of default.

prosocial dimension of positive discrimination, the identification of taste-based vs. statistical discrimination becomes even more complicated.

Our empirical results uncovered weak disparate treatment in one case (EU women compared with EU men) and affirmative action in favor of one group (non-EU women compared with any reference group). One explanation for these results could be that the prosocial MFI in our case study concentrated on the financial inclusion of poor non-EU women while paying less attention to its loan officers' stereotypical, hostile beliefs toward female borrowers. The social orientation of the lender makes it unlikely that this animus was deliberate. The evidence is, however, in line with previous studies showing that even some well-meaning lenders treat loan applications by women more harshly than those by men with similar characteristics (Alesina et al., 2013; Agier & Szafarz, 2013a).

Even though most discriminatory behaviors have significantly negative economic consequences, their most serious defect is ethical. Discrimination in lending is prohibited by law in many countries.[15] The US Equal Credit Opportunity Act (ECOA) enacted in October 1974 protects loan applicants from discriminatory conduct based on race, color, religion, national origin, sex, marital status, or age. Yet ECOA is silent on how the rule should be understood when it comes to intersectional biases. For instance, is discrimination against Black women unlawful if both Black men and White women are treated fairly? With the increase in scale and scope of credit market operations, further clarification on this legal issue would be welcome. Affirmative action in lending is another avenue that might practically address the ECOA's remaining loopholes. Paradoxically, the legal treatment of such a positive bias is hampered by the anti-discrimination principles that

[15] See Ongena and Popov (2016) about cross-country differences in cultural dimensions. Schiek and Lawson (2016) specifically address the European Union anti-discriminatory framework about intersectionality.

motivated the ECOA. For this reason, organizations willing to take affirmative action tend to frame it as “diversity management” (Kelly & Dobbin, 1998).

More than forty years after ECOA first came into effect, discrimination in lending is still an issue—calling for relevant tools to investigate the presence of bigotry. Our novel methodological approach is hopefully a step in that direction.

References

Agan, A. and S. Starr (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *Quarterly Journal of Economics*, 133(1), 191–235.

Agier, I. and A. Szafarz (2013a). Microfinance and gender: Is there a glass ceiling on loan size? *World Development*, 42, 165-181.

Agier, I. and A. Szafarz (2013b). Subjectivity in credit allocation to micro-entrepreneurs: Evidence from Brazil. *Small Business Economics*, 41, 263-275.

Ai, C. and E.C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(1), 123-129.

Aldén, L. and M. Hammarstedt (2016). Discrimination in the credit market? Access to financial capital among self-employed immigrants. *Kyklos*, 69(1), 3-31.

Alesina, A.F., F. Lotti, F., and P.E. Mistrulli (2013). Do women pay more for credit? Evidence from Italy. *Journal of the European Economic Association*, 11(suppl 1), 45-66.

Altman, E., A. Resti, and A. Sironi (2004). Default recovery rates in credit risk modelling: A review of the literature and empirical evidence. *Economic Notes*, 33, 183–208.

Ambrose, B.W, J.N. Conklin, and L.A Lopez (2021) Does borrower and broker race affect the cost of mortgage credit? *Review of Financial Studies*, 34(2), 790–826.

Armendariz, B. and J. Morduch (2010). *The Economics of Microfinance*. MIT Press.

Arnold, D., Dobbie, W., and C.S. Yang (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*, 133(4), 1885-1932.

Arrow, K.J. (1971). The theory of discrimination, Discussion paper, Princeton University, Department of Economics, Industrial Relations Section.

Arrow, K.J. (1998). What has economics to say about racial discrimination? *Journal of Economic Perspectives*, 12, 91-100.

Asiedu, E., J.A. Freeman, and A. Nti-Addae (2012). Access to credit by small businesses: How relevant are race, ethnicity, and gender? *American Economic Review*, 102(3), 532-537.

Atrey, S. (2019). Intersectional discrimination. Oxford, UK: Oxford University Press.

Ayres, I. (2002). Outcome tests of racial disparities in police practices. *Justice Research and Policy*, 4(1-2), 131-142.

Bai, Y. and L. Lu (2020). Households rejecting loan offers from banks. *Journal of Banking and Finance*, 119, 105339.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace. (2019). Consumer-lending discrimination in the fintech era, National Bureau of Economic Research, WP 25943.

Bayer, P., F. Ferreira, and S.L. Ross (2018). What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders. *Review of Financial Studies*, 31(1), 175–205.

Beck, T., P. Behr, and A. Madestam (2018). Sex and credit: Do gender interactions matter for credit market outcomes? *Journal of Banking and Finance*, 87, 380-396.

Becker, G.S. (1971). *The Economics of Discrimination*, second edition. Chicago, IL: University of Chicago Press.

Becker, G.S. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy*, 101, 385, 389.

Bendig, M., M. Unterberg, and B. Sarpong (2014). Overview of the microcredit sector in Europe 2012-2013. European Microfinance Network.

Berry, W.D., J.H. DeMeritt, and J. Esarey (2010). Testing for interaction in binary logit and probit models: Is a product term essential? *American Journal of Political Science*, 54(1), 248-266.

Bertrand, M., D. Chugh, D., and S. Mullainathan (2005). Implicit discrimination. *American Economic Review*, 95(2), 94-98.

Bertrand, M. and E. Duflo (2017). Field experiments on discrimination. In A.V. Banerjee and E Duflo (Eds.), *Handbook of Economic Field Experiments*, Vol. 1, North Holland, pp. 309-393.

Bhutta, N., and A. Hizmo, A. (2021). Do minorities pay more for mortgages? *Review of Financial Studies*, 34(2), 763-789.

Black, H., R.L. Schweitzer, R. L., and L. Mandell (1978). Discrimination in mortgage lending. *American Economic Review*, 68(2), 186-191.

Bogan, V. L. (2012). Capital structure and sustainability: An empirical study of microfinance institutions. *Review of Economics and Statistics*, 94(4), 1045-1058.

Boyes, W.J., Hoffman, D.L., and S.A. Low (1989). An econometric analysis of the bank credit scoring problem. *Journal of Econometrics*, 40(1), 3-14.

Butler, A.W., E.J. Mayer, and J.P. Weston (2021). Racial discrimination in the auto loan market. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3301009.

Buttner, E.H. and B. Rosen (1988). Bank loan officers' perceptions of the characteristics of men, women, and successful entrepreneurs. *Journal of Business Venturing*, 3, 249-258.

Card, D., S. DellaVigna, P. Funk, and N. Iriberri. 2020. Are referees and editors in economics gender neutral? *Quarterly Journal of Economics*, 135, 269–327.

Cheng, S. (2015). Potential lending discrimination? Insights from small business financing and new venture survival. *Journal of Small Business Management*, 53(4), 905-923.

Collins, P.H. and S. Bilge (2016). *Intersectionality*. Cambridge, UK: Polite Press.

Coffman, K.B., Exley, C.L., and M. Niederle (2021). The role of beliefs in driving gender discrimination. *Management Science*, 67, 3551-3569.

Cornée, S. (2019). The relevance of soft information for predicting small business credit default: Evidence from a social bank. *Journal of Small Business Management*, 57(3), 699-719.

Cozarenco, A. and A. Szafarz (2018). Gender biases in bank lending: Lessons from microcredit in France. *Journal of Business Ethics*, 147(3), 631–650.

Cozarenco, A. and A. Szafarz (2019). Microfinance in the North: Where do we stand? in Hudon, M. M. Labie, A. Szafarz (Eds.) *Research Agenda for Financial Inclusion and Microfinance*, Cheltenham: Edward Elgar Publishing, 125-137.

Cozarenco, A. and A. Szafarz (2020a). The Regulation of prosocial lending: Are loan ceilings effective?" *Journal of Banking and Finance*, 121, 105979.

Cozarenco, A. and A. Szafarz (2020b). Financial inclusion in developed countries: Gender gap or poverty trap? in Hartarska, V. and R. Cull (Eds.) *Handbook of Microfinance, Financial Inclusion and Development*, Edward Elgar Publishing, forthcoming.

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review,* 43(6), 1241-1299.

De Andrés, P., R. Gimeno, R., and R.M. de Cabo (2020). The gender gap in bank credit access. *Journal of Corporate Finance*, 101782.

Delis, M.D., I. Hasan, M. Iosifidi, and S. Ongena (2020). Gender, credit, and firm outcomes. *Journal of Financial and Quantitative Analysis*, forthcoming.

Delis, M.D. and P. Papadopoulos (2019). Mortgage lending discrimination across the US: New methodology and new evidence. *Journal of Financial Services Research*, 56(3), 341-368.

Diriker, D., P. Landoni, and N. Benaglio (2018). *Microfinance in Europe: Survey Report 2016-2017*. European Microfinance Network & Microfinance Centre.

Dobbie, W., A. Liberman, D. Paravisini, and V. Pathania. (2020). Measuring bias in consumer lending. *Review of Economic Studies,* forthcoming.

Dupas, P., A.S. Modestino, M. Niederle, M., and J. Wolfers (2021). Gender and the dynamics of economics seminars. National Bureau of Economic Research, WP 28494.

Dymski, G.A. (2006). Discrimination in the credit and housing markets: Findings and challenges, in W.M. Rodgers III (Ed.) *Handbook on the Economics of Discrimination*, Cheltenham: Edward Elgar Publishing, pp. 215-259.

Egan, M.L., G. Matvos, G., and A. Seru (2017). When Harry fired Sally: The double standard in punishing misconduct. National Bureau of Economic Research, WP 23242.

Fay, M. and L. Williams (1993). Gender bias and the availability of business loans. *Journal of Business Venturing*, 8, 363–376.

Fein, S. and S.J. Spencer (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology*, 73(1), 31.

Ferguson, M.F. and S.R. Peters (1995). What constitutes evidence of discrimination in lending? *Journal of Finance*, 50, 739-748.

Fisman, R., A. Sarkar, J. Skrastins, and V. Vig (2020). Experience of communal conflicts and inter-group lending. *Journal of Political Economy*, 128(9), 3346-3375.

Fisman, R., D. Paravisini, and V. Vig (2017). Cultural proximity and loan outcomes. *American Economic Review*, 107(2), 457-492.

Garikipati, S., S. Johnson, I. Guérin, and A. Szafarz (2017). Microfinance and gender: Issues, challenges and the road ahead. *Journal of Development Studies*, 53 (5), 641-648.

Greene, W.H. (2012). *Econometric Analysis, 7th Edition*. Pearson Education, Prentice Hall.

Guryan, J. and K.K. Charles (2013). Taste-based or statistical discrimination: The economics of discrimination returns to its roots. *Economic Journal*, 123(572), F417-F432.

Hair, J.F.J., W.C. Black, B.J. Babin, B.J., and R.E. Anderson (2010). *Multivariate Data Analysis* (Seventh Edition), Pearson Prentice Hall.

Hall, E.V., A.V. Hall, A.D. Galinsky, and K.W. Phillips (2019). Mosaic: A model of stereotyping through associated and intersectional categories. *Academy of Management Review*, 44(3), 643–672.

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

Haselmann, R., D. Schoenherr, and V. Vig (2018). Rent seeking in elite networks. *Journal of Political Economy*, 126(4), 1638-1690.

Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica*, 47, 153-61.

Iyer, R., A.L. Khwaja, E.F. Luttmer, E. F., and K. Shue (2016). Screening peers softly: Inferring the quality of small borrowers. *Management Science*, 62(6), 1554-1577.

Kelly, E. and F. Dobbin (1998). How affirmative action became diversity management: Employer response to antidiscrimination law, 1961 to 1996. *American Behavioral Scientist*, 41(7), 960-984.

Kleinberg, J., J. Ludwig, S. Mullainathan, and C.R. Sunstein (2018). Discrimination in the age of algorithms. *Journal of Legal Analysis*, 10, 113-174.

Ladd, H.F. (1998). Evidence on discrimination in mortgage lending. *Journal of Economic Perspectives*, 12, 41–62.

Lang, W.W. and L.I. Nakamura (1993). A model of redlining. *Journal of Urban Economics*, 33(2), 223-234.

Martinez, R.G., G. Aguilera-Lizarazu, A. Rojas-Hosse, and P. Aranda Blanco (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.

Melnik, A.Z. and O. Shy (2015). Exclusion, competition, and regulation in the retail loan market. *Journal of Banking and Finance*, 52, 189-198.

Méon, P.G. and A. Szafarz (2011). The modern corporation as a safe haven for taste-based discrimination: An agency model of hiring decisions. *Labour Economics*, 18, 487-497.

Norton, E. C., Wang, H., and C. Ai (2004). Computing interaction effects and standard errors in logit and probit models. *Stata Journal*, 4, 154-167.

Ongena, S. and A. Popov (2016). Gender bias and credit access. *Journal of Money, Credit and Banking*, 48(8), 1691-1724.

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

Puhani, P.A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear difference-in-differences models. *Economics Letters*, 115, 85-87.

Qi, S., S. Ongena, and H. Cheng (2018). When they work with women, do men get all the credit? Swiss Finance Institute Research Paper (18-01).

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

Robb, A. and D.T. Robinson (2018). Testing for racial bias in business credit scores. *Small Business Economics*, 50(3), 29–443.

Ross, S.L. (2000). Mortgage lending, sample selection and default. *Real Estate Economics*, 28(4), 581-621.

Ross, S.L. (2002). *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement*. Cambridge, MA: MIT Press.

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

Ruffle, B.J. and Z. Shtudiner (2014). Are good-looking people more employable? *Management Science*, 61(8), 1741–2011.

Ruwanpura, K.N. (2008). Multiple identities, multiple discrimination: A critical review. *Feminist Economics*, 14(3), 77-105.

Schiek, D. and A. Lawson (Eds.) (2016). *European Union non-discrimination law and intersectionality: Investigating the triangle of racial, gender and disability discrimination*. New York: Routledge.

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

Shaffer, S. (1996). Evidence of discrimination in lending: An extension. *Journal of Finance*, 51(4), 1551-1554.

Simoiu, C., S. Corbett-Davies, and S. Goel (2017). The problem of infra-marginality in outcome tests for discrimination. *Annals of Applied Statistics*, 11(3), 1193-1216.

Stiglitz, J.E. and A. Weiss (1981). Credit rationing in markets with imperfect information. *American Economic Review*, 71(3), 393-410.

Sunstein, C.R. (1991). Why markets don't stop discrimination. *Social Philosophy and Policy*, 8(2), 22-37.

Taub, R.P. (1998). Making the adaptation across cultures and societies: A report on an attempt to clone the Grameen Bank in southern Arkansas. *Journal of Developmental Entrepreneurship*, 3(1), 53.

Williams B.A., C.F. Brooks, and Y. Shmargad (2018). How algorithms discriminate based on data they lack: Challenges, solutions, and policy implications. *Journal of Information Policy*, 8, 78-115.

Appendix A. Extension to Intersectional Discrimination

We consider the following approval equation

$$Approval_i = \mathbb{1}[\alpha_A F_i + \theta_A D_i + \mu_A F_i \cdot D_i + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i} + v_i > 0] \tag{A1}$$

In addition to the *gender* dummy variable F, Equation (A1) includes *ethnicity*, represented by variable D (1 for non-white, 0 for white)[16] and its product with F. Intersectional discrimination is associated with any bias targeting the intersection $(F \cap D)$, namely, non-white women. The interaction term, $(F_i \cdot D_i)$ in Equation (A1) captures the intersectional effect. The expected approval of applicant i is given by:

$$E[Approval_i | F_i, D_i, \boldsymbol{Z_i}] = \Phi(\alpha_A F_i + \theta_A D_i + \mu_A F_i \cdot D_i + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i}) \tag{A2}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function (cdf). The corresponding decision rules are summarized in Table A1.

Table A1. Detecting Intersectional Discrimination using Probit Estimation

Panel I: Definition of coefficients			
Comparing	with white men	with white women	with non-white men
white women	$\delta_1 = \Phi(\alpha_A + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i}) - \Phi(\beta_{\boldsymbol{A}}' \boldsymbol{Z_i})$		
non-white men	$\delta_2 = \Phi(\theta_A + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i}) - \Phi(\beta_{\boldsymbol{A}}' \boldsymbol{Z_i})$		
non-white women	$\delta_3 = \Phi(\alpha_A + \theta_A + \mu_A + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i}) - \Phi(\beta_{\boldsymbol{A}}' \boldsymbol{Z_i})$	$\delta_4 = \Phi(\alpha_A + \theta_A + \mu_A + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i}) - \Phi(\alpha_A + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i})$	$\delta_5 = \Phi(\alpha_A + \theta_A + \mu_A + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i}) - \Phi(\theta_A + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i})$
Panel II: Test results and interpretation			
$\forall j = 1, \ldots, 5$	$\delta_j > 0$	$\delta_j = 0$	$\delta_j < 0$
	Positive bias	No bias	Negative bias

Note: This table describes the test for intersectional discrimination using the probit model of loan approval in Equation (A1). Panel I defines the coefficients and Panel II summarizes the decision rule. α_A is the coefficient of the dummy variable that takes the value of 1 if the applicant is a woman, and 0 otherwise; θ_A is the coefficient of the dummy variable that takes the value of 1 if the applicant is non-white, and 0 otherwise; and μ_A is the coefficient of the interaction term. Vector Z_i includes the control variables. $\Phi(\cdot)$ is the standard normal cumulative distribution function. The δ_j ($\forall j = 1, \ldots, 5$) are the test statistics for pairwise comparisons.

To address intersectional discrimination by using recovery rates and approval probabilities, we extend the model as follows:

[16] Gender and ethnicity are the two typical characteristics explored in the literature (Hall et al., 2019; Asiedu et al., 2012).

$$Recovery\ rate_i = \alpha_R F_i + \theta_R D_i + \mu_R F_i \cdot D_i + \beta_{\boldsymbol{R}}' \boldsymbol{X_i} + \varepsilon_i \tag{A3}$$

$$Approval_i = \mathbb{1}[\alpha_A F_i + \theta_A D_i + \mu_A F_i \cdot D_i + \beta_{\boldsymbol{A}}' \boldsymbol{Z_i} + v_i > 0] \tag{A4}$$

Table A2 summarizes the decision rule. To test for biases either favoring or hindering applicants belonging to two groups, we estimate the $\boldsymbol{\gamma_j}'$s from the recovery equation (A3) and the $\boldsymbol{\delta_j}'$s from the approval equation (A4). Then we use the Wald test and the delta method to assess the signs of $\boldsymbol{\gamma_j}'$s and $\boldsymbol{\delta_j}'$s, respectively. The decision rule mimics the rule indicated in Table A1, but it is applied to sums of coefficients and differences in cdfs rather than to single coefficients.

Table A2. Detecting Intersectional Discrimination with Bivariate Estimation

Panel I: Definition of coefficients

Comparing	with white men	with white women	with non-white men
white women	$\delta_1 = \Phi(\alpha_A + \beta_A'\boldsymbol{Z_i}) - \Phi(\beta_A'\boldsymbol{Z_i})$ $\gamma_1 = \alpha_R$		
non-white men	$\delta_2 = \Phi(\theta_A + \beta_A'\boldsymbol{Z_i}) - \Phi(\beta_A'\boldsymbol{Z_i})$ $\gamma_2 = \theta_R$		
non-white women	$\delta_3 = \Phi(\alpha_A + \theta_A + \mu_A + \beta_A'\boldsymbol{Z_i}) - \Phi(\beta_A'\boldsymbol{Z_i})$ $\gamma_3 = \alpha_R + \theta_R + \mu_R$	$\delta_4 = \Phi(\alpha_A + \theta_A + \mu_A + \beta_A'\boldsymbol{Z_i}) - \Phi(\alpha_A + \beta_A'\boldsymbol{Z_i})$ $\gamma_4 = \theta_R + \mu_R$	$\delta_5 = \Phi(\alpha_A + \theta_A + \mu_A + \beta_A'\boldsymbol{Z_i}) - \Phi(\theta_A + \beta_A'\boldsymbol{Z_i})$ $\gamma_5 = \alpha_R + \mu_R$

Panel II: Test results and interpretation

$\forall j = 1, \ldots, 5$	$\delta_j > 0$	$\delta_j = 0$	$\delta_j < 0$
$\gamma_j > 0$	No bias	Weak negative bias	Strong negative bias
$\gamma_j = 0$	Weak positive bias	No bias	Weak negative bias
$\gamma_j < 0$	Strong positive bias	Weak positive bias	No bias

Note: This table describes the test for intersectional discrimination using Heckman estimation of the recovery rate and loan approval rate. Panel I defines the coefficients and Panel II summarizes the decision rule. α_i (resp. θ_i) is the coefficient of the dummy variable that takes the value of 1 if the applicant is female (resp. non-white), and 0 otherwise, and μ_i is the coefficient of the interaction term, in Equation (A3) for $i = R$, and in Equation (A4) for $i = A$. $\Phi(\cdot)$ is the standard normal cumulative distribution function. The γ_j (resp. δ_j) ($\forall j = 1, \ldots, 5$) are the test statistics for pairwise comparisons in Equation (A3) (resp. Equation (A4)).

Appendix B. Intersectional Discrimination: Coefficients from Table 3

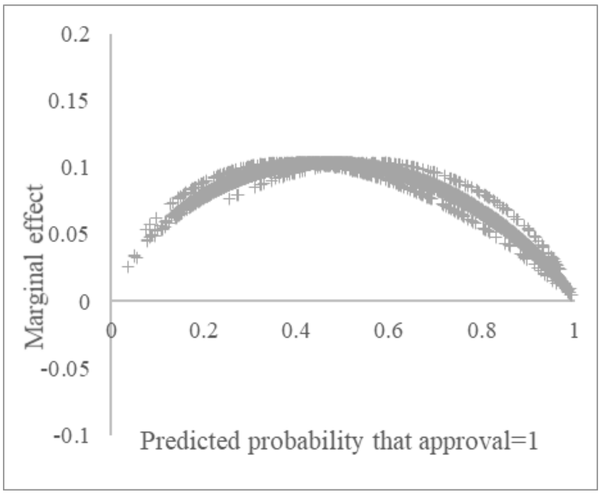


Figure B1a. Marginal Effect of Non-EU Women vs. EU Men

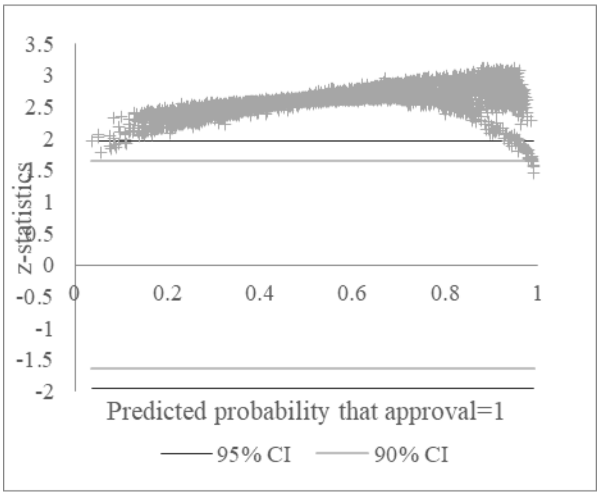


Figure B1b. Z-Statistics for Non-EU Women vs. EU Men

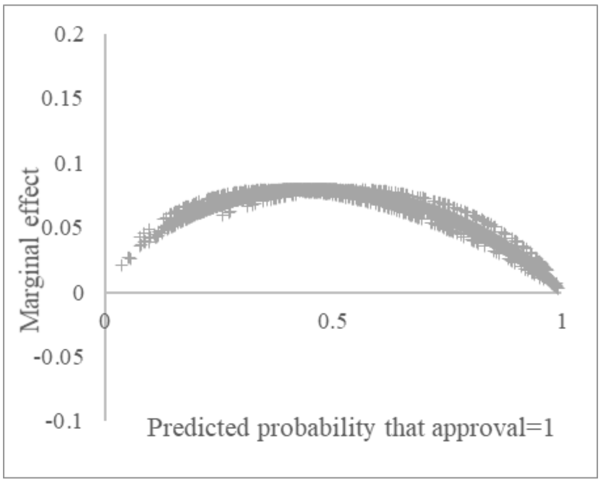


Figure B2a. Marginal Effect of Non-EU Women vs. EU Women

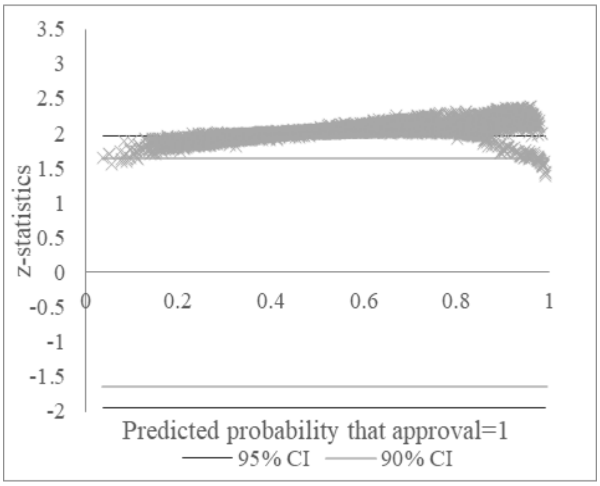
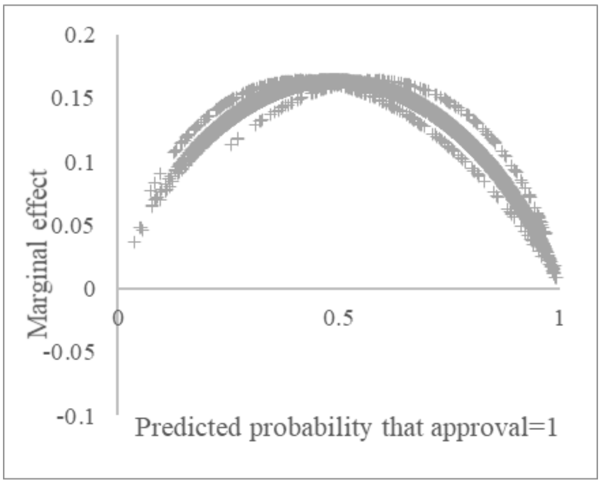


Figure B2b. Z-Statistics for Non-EU Women vs. EU Women



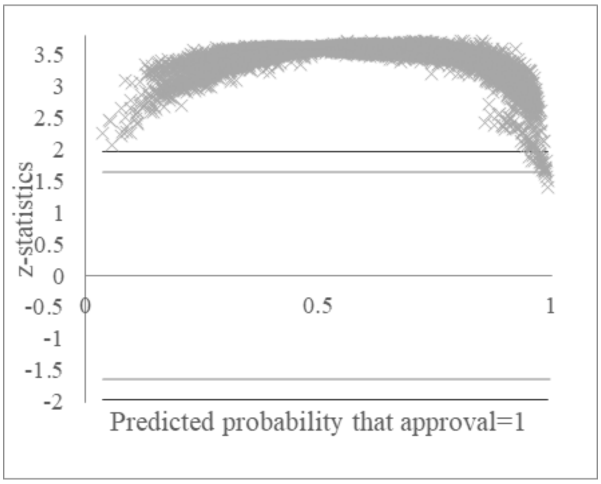
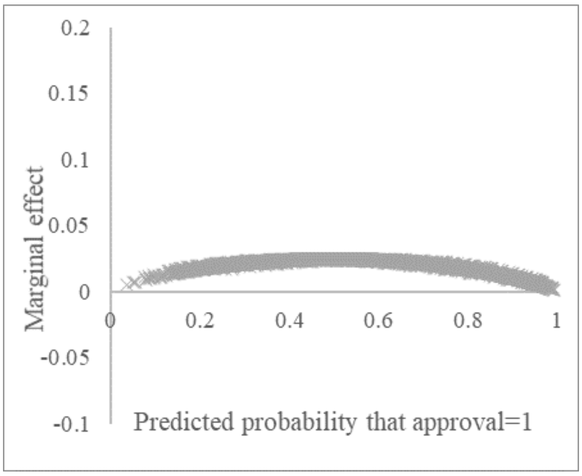


Figure B3a. Marginal Effect of Non-EU Women vs. Non-EU Men

Figure B3b. Z-Statistics for Non-EU Women vs. Non-EU Men



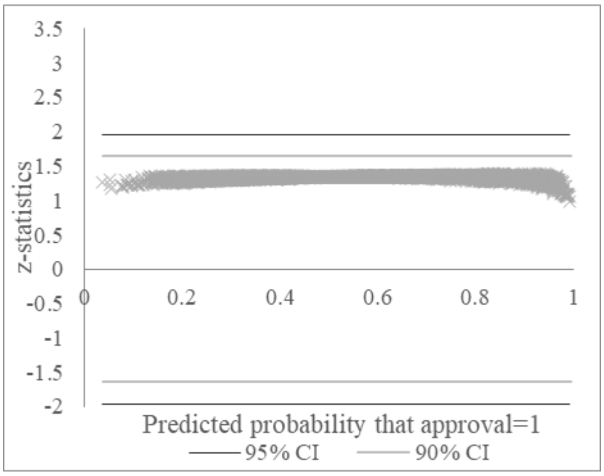
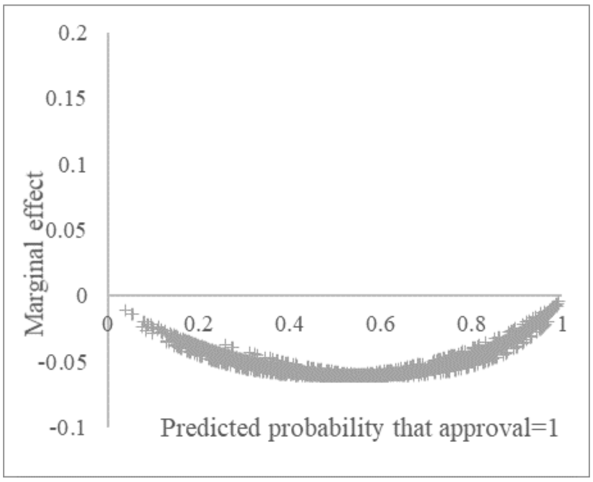


Figure B4a. Marginal Effect of EU Women vs. EU Men

Figure B4b. Z-Statistics for EU Women vs. EU Men



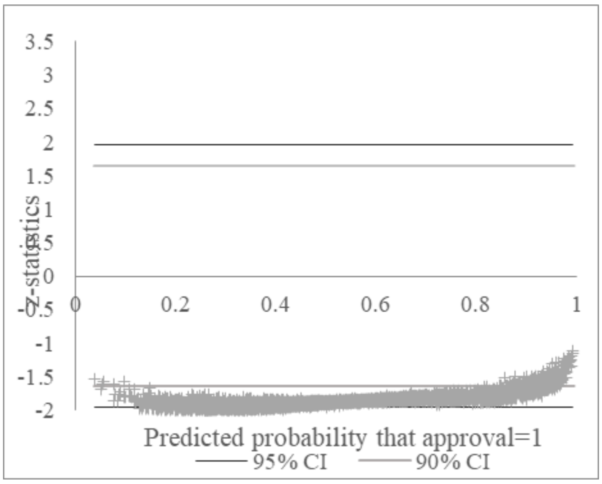


Figure B5a. Marginal Effect of Non-EU Men vs. EU Men

Figure B5b. Z-Statistics of Non-EU Men vs. EU Men

Appendix C. Recovery Rate: Controlling for Unemployment Rates

Table C1. Recovery Rate: Heckman and OLS Estimations

	Heckman		OLS
Variables	(1)	(2)	(3)
Female ($\hat{\alpha}$)	0.021***	0.021***	0.020***
	(0.007)	(0.007)	(0.007)
Non-EU ($\hat{\theta}$)	0.008	0.009	0.010
	(0.016)	(0.016)	(0.016)
Female*Non-EU ($\hat{\mu}$)	-0.054**	-0.055**	-0.058**
	(0.027)	(0.027)	(0.027)
Control variables			
Age (years/100)	0.132***	0.132***	0.128***
	(0.030)	(0.030)	(0.031)
Single (dummy)	-0.004	-0.004	-0.007
	(0.007)	(0.007)	(0.007)
# of children	-0.003	-0.003	-0.003
	(0.004)	(0.004)	(0.004)
Ln (HH income in thousands of EUR)	0.029***	0.029***	0.024**
	(0.010)	(0.010)	(0.010)
Unemployed (dummy)	-0.016*	-0.016*	-0.015
	(0.009)	(0.009)	(0.009)
Bad credit history (dummy)	-0.069***	-0.069***	-0.066***
	(0.011)	(0.011)	(0.011)
Year FE	Yes	Yes	Yes
Unemployment rate in town of residence	0.001		0.000
	(0.003)		(0.003)
Unemployment rate at national level		0.003	0.005
		(0.019)	(0.020)
Constant	0.894***	0.875***	0.864***
	(0.026)	(0.137)	(0.140)
Number of observations	3,262	3,262	3,262
R-squared			0.040

Note: This table reports specifications that replicate our baseline model in Equations (8) and (9) but also include unemployment rates in the town of residence and at the national level (our independent variables) directly in the model to show that this variable does not significantly correlate with the recovery rate. Specification 3 reports the results of an OLS estimation of the recovery rate including both independent variables. Standard errors (SE) clustered at the county-year level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1