Unwarranted Gender Disparity in Online P2P Lending: Evidence of Affirmative Action

Xudong Shen
xudong.shen@u.nus.edu
NUS Graduate School, ISEP Programme
National University of Singapore

Tuan Q. Phan
tphan@hku.hk
HKU Business School
The University of Hong Kong

Tianhui Tan	
tant@nus.edu.sg
NUS Business School
National University of Singapore

Jussi Keppo
keppo@nus.edu.sg
NUS Business School
National University of Singapore

ABSTRACT
Closing the gender gap in financial access is important. Most research tends to empirically uncover the direct effect of gender on decisions. Yet, this view overlooks other indirect channels of gender discrimination, leading to systemic bias in identifying the overall discrimination effect. In this work, by collaborating with one of the largest online P2P lending platforms in China, we estimate a broadened discrimination notion called unwarranted gender disparity (UGD). UGD recognizes any disparate lending decisions that do not commensurate with the loan’s return rate, encompassing direct, indirect, and proxy discrimination. We develop a two-stage predictor substitution (2SPS) approach to estimate UGD. Somewhat surprisingly, we find significant female favoritism at almost all return rate levels. On average, female borrowers are 3.97% more likely to be funded than male borrowers with identical return rates. We further decompose and find at least 37.1% of UGD is indeed indirect or proxy discrimination. However, we also identify the observed UGD favoring female can be completely attributed to an adverse effect on the protected group. The defendant, then, must prove the disparate impact can be explained by factors that have a "manifest relation" to its business [1]. It overlooks indirect and proxy discrimination, such as discrimination against Black-sounding names [14] and redlining [28], when the econometrician controls for mediators that mediate the effect of gender on decisions and proxy variables that correlate with gender. This issue is prominent. For one thing, discrimination is often covert and indirect in the digital world. For another, many fields do adopt a broader view of discrimination, including law [1], stratification economics [26], and machine learning [43, 57]. As powerfully argued in Darity et al. [26], overlooking indirect discrimination “absolves the social system and privileged groups from criticism for their role in perpetuating the condition of the dispossessed”.

Our Contributions We consider a broadened discrimination notion called unwarranted gender disparity (UGD) [3, 7, 13], which aligns with the legal doctrine of disparate impact. Under disparate impact, the plaintiff can initiate a discrimination case by showing an adverse effect on the protected group. The defendant, then, must prove the disparate impact can be explained by factors that have a “manifest relation” to its business [1]. In P2P lending, the individual investors can determine their own investment amount to every loan. Therefore, the investors’ legitimate discriminant factor solely refers to the loan’s return rate 1. Correspondingly, we define UGD as the disparate funding of loans that differ in borrower gender but have identical return rates. UGD captures any disparate impact of equally “qualified” loans that differ in gender, encompassing direct, indirect, and proxy discrimination (illustrated in Fig. 1b).

We propose a two-stage predictor substitution (2SPS) approach to estimate UGD from observational data, illustrated in Fig. 2.

1 We define a loan’s return rate as the ratio of the borrower’s actual repayment amount to the principal. A return rate of 1 corresponds to a net rate of return 0. A return rate of 0 corresponds to a net rate of return -1, i.e., a complete loss of principal and interest.
loan’s return rate can be expressed as the product between the repayment ratio\(^2\) and \((1 + \text{the interest rate})\). Yet, the repayment ratio is only observed given the loan is successfully funded. Thus, in the first stage of 2SPS, we develop a survival model to predict the repayment ratio as a function of loan and borrower characteristics. In the second stage, using the predicted repayment ratio for the unsuccessful loans, we directly estimate UGD. We bootstrap both stages for 500 times to obtain confidence intervals. We also analyze the asymptotic properties of the 2SPS approach and show its robustness.

We obtain anonymized backend data from one of the largest P2P lending platforms in China. By applying the 2SPS approach, we find significant female favoritism at almost all return rates (except those \(\leq 0.1\)). The evidence of female favoritism is especially strong for borrowers who do not default, with return rates \(> 1.1\). On average, female borrowers are 3.97% more likely to be funded than male borrowers, given identical return rates. Using a novel decomposition technique, we estimate at least 37.1% of UGD is indirect or proxy discrimination. Corroboratively, the traditional view of only measuring direct gender discrimination will underestimate the female favoritism by 44.6%, equivalent to saying that female is 2.15% more likely to be funded than male, \textit{ceteris paribus}. We also conduct extensive diagnostics and robustness checks, which support the consistency of our results.

We further investigate a decision model where the investors primarily engage in accurate statistical discrimination \([3, 8, 51]\) but may also have gender animus \([11]\)—two workhorse frameworks of discrimination. We find that the observed female favoritism can be completely attributed to \textit{accurate statistical discrimination}. Namely, the UGD can be explained as the investors recognize the fact that women are less likely to default on their P2P loans, and are able to accurately predict the \textit{expected} return rate. Next, this model estimates the gender animus component is still against women. Namely, in order to be funded, women need to yield an higher expected return rate of 1.099 compared to men with 1.079. By simulating counterfactuals, we find that female borrowers would receive a more favorable treatment with UGD between 7.73% ~ 11.83%, if the investors only have accurate statistical discrimination but no gender animus.

Our findings suggest online P2P lending is a \textit{market-based affirmative action} \([23]\). We contend that gender gap persists \([29]\) and women are still disadvantaged in many other credit markets \([18, 20, 64]\). P2P lending, therefore, complements traditional bank lending by providing an alternative credit market where the affirmative action to support women is driven by the market forces rather than bureaucratic rules, is win-win, and can arise naturally from the rational crowd.

This work’s contribution is twofold. First, we are among the first to quantify a broadened notion of discrimination in online platforms such as P2P lending, joining the concurrent efforts in other domains such as bail decision \([7]\) and employment \([17]\). Second, we provide novel insights to whether and how P2P lending practice can help close the gender gap in financial access.

2 RELATED WORK

P2P Lending Uncovering the importance of FinTech enabled financial inclusion, there is an increasing number of research on online P2P lending. One stream of research tries to identify various determinants of funding success and borrower default \([38, 42]\), including appearance \([32, 53]\), social ties \([34, 45]\), and textual descriptions \([44]\). Herzenstein et al. \([39]\), Zhang and Liu \([65]\) studies the online investors’ herding behavior. Berger and Gleisner \([12]\) looks into the role of group leaders in crowdfunding.

Gender Discrimination in P2P Lending Although female is often discriminated against in traditional bank lending \([4, 15, 48]\), findings from P2P lending tend to show otherwise. For example, Barasinska and Schäfer \([10]\) finds no effect of gender on funding success on a German P2P lending platform. Pope and Sydnor \([52]\), Ravina \([53]\) find women are favorably treated on an American P2P lending platform \textit{Prosper.com}. Notably, Chen, Li, and Lai \([21]\) and Chen, Huang, and Ye \([22]\) study P2P lending platforms in China. While Chen, Huang, and Ye \([22]\) finds no significant gender effect on funding success, Chen, Li, and Lai \([21]\) finds women are more likely to be funded than men.

Our work differs from the existing research in two ways. First, the prior research focuses on the direct effect of gender on decisions. We consider a broader view that also encompasses indirect and proxy discrimination. Second, the prior research focuses on testing the existence of gender discrimination, while we are interested in further quantifying its extent.

Unwarranted Gender Disparity The idea of recognizing any unwarranted disparity that does not commensurate with the true qualifiedness traces back to Aigner and Cain \([3]\). Stratification economics \([26, 27]\) holds a similar broader view of discrimination.

---

\(^2\)We define a loan’s repayment ratio as the ratio of the borrower’s successful repayment amount to the total amount the borrower should repay, including principal and interest.
where indirect and proxy discrimination are interpreted as mechanisms that entrench the privileged groups’ socio-economical standing. Bohren et al. [17] distinguishes between direct discrimination and systemic discrimination, the latter of which is exactly the indirect and proxy discrimination considered in this work. Recently, Arnold et al. [7] developed a quasi-experimental tool to estimate unwarranted racial disparity (also called disparate impact therein) in bail decisions. Their illuminating approach, however, utilizes the quasi-random assignment of judges, and thus is inapplicable in many other contexts including ours. In contrast, the 2SPS estimate developed in this work is prediction-based and has a greater applicability.

3 BACKGROUND

3.1 The P2P Lending Platform

We collaborate with one of the largest online P2P lending platforms in China. The platform connects individual investors to individual borrowers across the country and offers secured personal loans. When a borrower submits a loan application, a loan listing containing borrowing amount, interest rate, loan term, borrower gender, a short description, etc., is generated on the platform. The loans are open for investors to subscribe for a fixed period of time. The investors may subscribe a small amount for every loan to diversify the risk. The platform operates an All-or-Nothing crowdfunding policy. A loan is considered successful if and only if the borrowing amount is fully reached. Otherwise, the loan is unsuccessful and the borrower will not receive any money. For the successful loans, the borrowers make repayments in equated monthly installments (EMI) according to the loan term.

3.2 Data

Our data consists of 1,006,161 loan listings with 12-month installment plans between January 1, 2016 and June 30, 2016. In consistent with the platform’s operational definition, an installment is considered defaulted if the payment is late for more than 90 days.

We conduct the following pre-processing steps. First, we drop the loans whose payments are non-conventional for ease of analysis, including (1) those who have some installment partially paid (N=2,806) and (2) those who first default some installment(s) but then pay again (N=3,301). Second, we winsorize the loans whose amount, age, # past ontime payments, # past late payments, # past failed borrowings, and # past aborted borrowings fall in the top or bottom 0.5% quantile to eliminate outliers (N=29,985). Lastly, we drop loans whose interest rates are lower than 16% (N=293,596), and we drop loans whose interest rates are lower than 16%. The final sample consists of 676,473 loans. Table 1 reports the descriptive statistics.

3.3 Notations

The loans are indexed by a subscript $i \in [I]$. We use $G_i \in \{m, f\}$ to denote the borrower’s gender: $m$ means male and $f$ means female. $R_i > 0$ is the interest rate. $X_i$ is a vector of loan characteristics, including borrower gender $G_i$, interest rate $R_i$, borrowing amount, and other borrower information. $D_i = 1$ (or 0) denotes the loan is successful (or unsuccessful). We use $\lambda_i \in [0, 1]$ to denote the loan’s repayment ratio, which is defined as the ratio of the borrower’s successful repayment amount to the total amount the borrower should repay. $\lambda_i < 1$ means the borrower defaults and $\lambda_i = 1$ means the borrower does not default. We use $Y_i \geq 0$ to denote the loan’s return rate, which is defined as the ratio of the borrower’s successful repayment amount to the principal. A return rate $Y_i = 1$ correspond to a net rate of return 0. In P2P lending, the loan’s return rate $Y_i$ can be expressed as follows:

$$Y_i = \lambda_i \times (1 + R_i).$$

3.4 Unwarranted Gender Disparity

DEFINITION 1. Unwarranted gender disparity (UGD) at a return rate level $y$ is defined as follows:

$$\Delta(y) = \mathbb{E}[D \mid G = m, Y = y] - \mathbb{E}[D \mid G = f, Y = y].$$

The average level of UGD is given by:

$$\Delta = \mathbb{E}_y[\Delta(y)],$$

where the expectation is taken over the population distribution of the return rate $y$. 

---

Tab. 1: This table reports the mean statistics of our data.
UGD measures any disparate treatment of the loans that differ in borrower gender but have identical return rates. Fig. 1 provides an illustration.

4 TWO-STAGE PREDICTOR SUBSTITUTION

We propose a two-stage predictor substitution approach (illustrated in Fig. 2) to estimate UGD from observational data. The difficulty in estimating UGD lies in that the repayment ratio is only observed when the loans are successfully funded—a fundamental missing data problem. Therefore, in the first stage of 2SPS, we develop a survival model to predict repayment ratio as a function of loan characteristics. In the second stage, we estimate UGD using the predicted repayment ratio for the unsuccessful loans.

4.1 First Stage: Survival Model

Survival model has shown success in financial default analysis [30, 49, 58, 60]. It is particularly suited for modelling repayment ratio because (i) it explicitly characterizes the risk of default over the loan term and (ii) since the borrower makes repayments in equated monthly installments, the sufficient statistics for repayment ratio is exactly at which month the borrower starts to default.

The survival model focuses on $T \in \mathbb{Z}$, the random variable that represents the number of months before the occurrence of default. We call $T$ default time for short. In our context, $T = 0$ means the borrower defaults the 1st month’s installment (and all subsequent months), and has repayment ratio $\lambda = 0$. $T = 1$ means the borrower defaults at the 2nd month (and all subsequent months), and has repayment ratio $\lambda = 1/12$. Similar is true for $T \in \{1, 2, \ldots, 11\}$. $T \geq 12$ means the borrower does not default, and has repayment ratio $\lambda = 1$.

A survival model is characterized by the hazard function,

$$h(t) = P(T = t | T \geq t),$$

which defines the instantaneous rate of default given that the borrower has not defaulted so far. $h(t)$ is the hazard rate at time $t$.

We start with the classical Cox proportional hazard (PH) model [24], which assumes the following hazard function,

$$h(t | X, \beta) = h_0(t) \exp(\beta X).$$

The baseline hazard $h_0(t)$ is non-parametric and describes the effect of time for individuals with $X = 0$, who serves as a reference cell. The parametric component $\exp(\beta X)$, then, describes the relative increase or decrease of hazard associated with $X$. The Cox PH model imposes two strong assumptions—the log-linearity assumption and the proportional hazard assumption—that are unlikely to strictly hold in reality. We generalize the model to relax these assumptions.

The Cox PH model is a log-linear model, where the continuous covariates act exactly linearly on the log-hazard. Our first generalization is to allow non-linear relationships by applying natural spline transformation [63],

$$h(t | X, \beta) = h_0(t) \exp(\beta \cdot f_{ns}(X)), \quad (6)$$

where $f_{ns}$ denotes the natural spline transformation. In implementation the natural spline transformation is only applied on the continuous covariates.

The proportional hazard (PH) assumption is the distinguishing feature of the Cox PH model. It, nonetheless, restricts that all loans’ hazard rates over time are of a common shape, determined by $h_0(t)$, and loan characteristics affect the hazard time-independently. Our second generalization is to allow time-dependent effects by adding time interactions,

$$h(t | X, \beta) = h_0(t) \exp \left( \beta^{(1)} f_{ns}(X) + \beta^{(2)} f_{ns}(X) f_{ns}(t) \right), \quad (7)$$

where both the covariates and time are natural spline transformed.

We fit the survival model of Eq. 7 to the successful loans using the survival package [62] in R. Briefly, the data is first converted to the counting process formulation [2, 5, 9]. Then, a partial likelihood is derived using the Elfron approach [33]. Finally, we obtain the maximum partial likelihood estimate of the survival model parameters. Using the fitted survival model, we obtain predicted repayment ratio by sampling from the predicted hazard rate over the loan term, shown in Procedure 1.

4.2 Second Stage: Non-Parametric Estimate

Using the predicted repayment ratio for the unsuccessful loans as well as a small portion of successful loans (<3%) whose payments are partially unobserved due to data cutoff, UGD of Equation 2 and 3 can be directly non-parametrically estimated. Practically, we divide the return rate into small intervals and assume UGD is homogeneous within these intervals to circumvent the issue of small sample size at unique return rate values. We obtain confidence intervals by bootstrapping both stages for 500 times. Concretely, in every iteration, we resample the data that is used to fit the survival model and estimate UGD using the newly fitted survival model.

Fig. 2: An illustration of the 2SPS estimate for unwarranted gender disparity. Variables are defined in Section 3.3.
4.3 Analysis of Bias

Using the Bayes rule, UGD has the following expression,

$$\Delta(y) = \frac{P_m(D = 1)P_m(\lambda(1 + R) = y | D = 1)}{\sum_{d \in \{0,1\}} P_m(D = d)P_m(\lambda(1 + R) = y | D = d)} - \frac{P_f(D = 1)P_f(\lambda(1 + R) = y | D = 1)}{\sum_{d \in \{0,1\}} P_f(D = d)P_f(\lambda(1 + R) = y | D = d)},$$

(8)

where we use subscript $g$ to denote the distribution is condition on gender $G = g$. We use $\Delta(y)$ to denote our 2SPS estimate of $\Delta(y)$, which uses predicted repayment ratio $\lambda$ for the unsuccessful loans$^3$.

$$\Delta(y) = \frac{P_m(D = 1)P_m(\lambda(1 + R) = y | D = 1)}{P_m(D = 1)P_m(\lambda(1 + R) = y | D = 1) + P_m(D = 0)\sum_r P_m(R = r | D = 0)P_m(\lambda = y/r | D = 0, R = r)} - \frac{P_f(D = 1)P_f(\lambda(1 + R) = y | D = 1)}{P_f(D = 1)P_f(\lambda(1 + R) = y | D = 1) + P_f(D = 0)\sum_r P_f(R = r | D = 0)P_f(\lambda = y/r | D = 0, R = r)}. \tag{9}$$

We define bias from the first stage of 2SPS as $b_{g,r}(l)$,

$$b_{g,r}(l) = P_g(\lambda = I | D = 0, R = r) - P_g(\lambda = I | D = 0, R = r). \tag{10}$$

The first-stage bias $b_{g,r}(l)$ is measured at a specific repayment ratio level $l$ on the distributions of predicted repayment ratio $\lambda$ condition on borrower gender $g$, interest rate $r$, and loans being unsuccessful $D = 0$. We now can express the asymptotic bias for our 2SPS estimate of UGD in terms of the first-stage bias:

$$b_{2SPS}(y) = \hat{\Delta}(y) - \Delta(y)$$

average female first-stage bias $b_f$

$$P_f(D = 0)\sum_r P_f(R = r | D = 0)b_{f,r}(y/r)$$

average male first-stage bias $b_m$

$$P_m(D = 0)\sum_r P_m(R = r | D = 0)b_{m,r}(y/r)$$

We find the 2SPS estimate is unbiased when the average first-stage biases $b_f, b_m$ are zero. Two comments are worth highlighting. First, as defined in Eq. 10, the first-stage bias $b_{g,r}(l)$ is measured without conditioning on $X$. We only need the predicted repayment ratio $\lambda$ to induce a distribution unbiased to $P_f(\lambda | D = 0, R = r)$, which is a much weaker condition than unbiasedness of the predicted repayment ratio itself, $\mathbb{E}[\lambda | X] = \mathbb{E}[\lambda | X]$. Second, in $b_{2SPS}(y)$ the first-stage bias $b_{g,r}(y/r)$ is further marginalized over the interest rate $r$ (red terms). I.e., only the average first-stage biases affect the 2SPS estimate.

The 2SPS estimate is also robust to the average first-stage biases $b_f, b_m$ when they are non-zero. The reasons are as follows. First, $b_f$ and $b_m$ are both down-scaled in $b_{2SPS}(y)$: the coefficient in front of them (blue terms) are always smaller than 1. Second, When $b_f$ and $b_m$ have the same sign, they partially cancel each other in $b_{2SPS}(y)$. Intuitively, the 2SPS estimate is partially protected from systematic over- or underestimation of repayment ratio.

5 EMPIRICAL RESULTS

5.1 Survival Model

We report the survival model’s fitted coefficients in Tab. A1 and Fig. A1 in Appendix. Notably, male borrowers have an increased hazard rate, 1.503 (log-S.E.=0.021) times that of the female. We visualize the fitted hazard and survival curves in Fig. 3. For both male and female borrowers, the hazard rate is the highest at the 1st month, drops abruptly at the 2nd month, and then slightly increases till around the 8th month, and finally decreases to around 0 at the 12th month. Our interpretation is: the uncollateralized nature of online P2P lending invites a number of ill-intentioned borrowers who intent to default all installments. Fewer borrowers default at the last several installments because it is uneconomical: they could either default at an earlier time to earn a higher return, or do not default at all to maintain their credit. We observe female has a lower propensity to default at all months.

We report extensive model diagnostics in Appendix B, using the scaled Schoenfeld residual [35, 55], the Cox-Snell residual [25], the concordance index [36], and the rank plot. The evidence shows the survival model is well-specified, has goodness-of-fit, and is predictive of borrowers’ default.

For ease of analysis, we ignore the small fraction (< 3%) of successful loans whose payments are partially unobserved due to data cutoff.

Preprint — under review.
5.2 Repayment Ratio and Return Rate

Using the predicted repayment ratio for the unsuccessful loans, we plot the distributions of repayment ratio and return rate in Fig. 4. We find female borrowers’ repayment ratio distribution dominates the male borrowers’. A higher fraction of female borrowers—92.4% compared to 88.9% for male borrowers—do not default and have repayment ratio $\lambda = 1$. And fixing any repayment ratio level $\lambda \in \{0, 1/12, \ldots, 11/12\}$ where default does occur, there is always a lower fraction of female compared to male.

Nonetheless, the female borrower’s return rate distribution does not dominate the male borrower’s. Female’s return rates are more concentrated in an intermediate range, between 1.16 ~ 1.24. Male’s return rate has larger variation: their loans are more likely to yield very high ($> 1.24$) and very low ($< 1.16$) return rates. The reason is, as reported in Tab. 1, men on average borrow at higher interest rates, which partially compensate their higher default risk.

5.3 Unwarranted Gender Disparity

Fig. 5 reports the 2SPS estimate of UGD at different return rate levels. Somewhat surprisingly, we find significant UGD favoring female at all return rate levels except in the interval $[0, 0.1]$, where the 95% CI contains 0 and thus is inconclusive. Several comments are in order. First, a return rate $\leq 0.1$ is the lowest and indicates the borrower defaults almost all installments. But these loans only account for around 1% of our sample. Second, the 95% CIs are especially tight in regions where the return rate is higher than 1.1. This is because around 90% of our sample does not default and falls into this region. For these borrowers who are indeed trustworthy, we observe a very strong evidence of UGD favoring female with the magnitude of around 4%.

Averaged over the population distribution of return rate, we find female borrowers are 3.97% (95% CI: −3.98% ~ −3.95%) more likely to have their loans successfully funded, given identical return rates.

5.4 Analysis

**Indirect and Proxy Discrimination** Using a novel decomposition technique, we investigate how much of UGD arises from indirect and proxy discrimination. We first replace the second stage of 2SPS with an OLS regression, which yields an OLS estimate of UGD. Then, by additionally controlling for the observed loan characteristics in OLS regression, we measure the reduction of UGD, which identifies the indirect and proxy discrimination through the observed loan characteristics. This estimate gives a lower bound of the total indirect and proxy discrimination, since there might be unobserved mediators and proxy variables. In a concurrent work, Bohren et al. [17] developed a similar decomposition-based approach to estimate indirect and proxy discrimination.

The results are reported in Tab. 2. First, the OLS estimate finds female is 3.88% more likely to be funded after controlling for return rates, which is very close to and corroborates our 2SPS estimate. Second, additionally controlling for observed loan characteristics significantly reduces the UGD to 2.44%. It means at least 37.1% of the found UGD constitutes indirect or proxy discrimination. Lastly, we also report in Tab. 2 the traditional view of only estimating the direct effect of gender, which controls for loan characteristics but not return rate. It estimates women are 2.15% more likely to be funded than men, ceteris paribus—a further 44.6% underestimation compared to UGD.

**Data Disaggregation** To investigate how UGD varies in different cohorts of borrowers, we report in Tab. 3 the reestimated UGD in disaggregated data subsets. We find consistent UGD favoring
female in most subsets, with mild variation in its magnitude. The largest UGD is found for student borrowers, with a 62.9 percent point increase from the population level UGD. The UGD favoring female is of higher magnitude for loans with higher interest rates. One possible explanation is because these loans are more risky, the investors are more likely to employ the simple heuristics of favoring female—who are known to be more trustworthy—to reduce risk.

Notably, we find the magnitude of UGD favoring female decreases with higher education, and even becomes favoring male for borrowers with Master or Doctorate degrees. But these borrowers only constitute around 0.2% of our sample.

**Simulated Overestimation of the Repayment Ratio** The 2SPS approach is observational in nature and faces threat from the fundamental missing data problem. We fit the survival model to the successful loans and use it to predict the repayment ratio for the unsuccessful loans. The concern is we might systematically overestimate the repayment ratio—or underestimate the default risk—for the unsuccessful loans. Although Section 4.3 theoretically shows the 2SPS estimate is partially protected from systematic overestimation, we are still interested in empirically testing its robustness.

Thus, we simulate the scenario where there is an omitted fixed effect associated with the unsuccessful loans. A fixed effect in the survival model is a constant multiplicative factor to the hazard rate. We assume the omitted fixed effect is in the direction that the actual hazard rates are higher than our current predictions. We vary the strength of the omitted effect from $\{1.5, 2, 2.5, 3\}$. Concretely, the unsuccessful loans’ predicted hazard rates throughout the loan terms are multiplied with $1.5, 2, 2.5, \text{or } 3$.

The results are reported in Fig. 6. We observe a mild, almost linear decrease in the magnitude of UGD. The UGD favoring female is still more than 3% when the unsuccessful loans’ default risks are underestimated by three-fold. Applying a linear extrapolation, if the actual UGD is zero or is in the direction that favors male, the actual hazard rates for the unsuccessful loans will have to be at least 10 times our current predictions. Using Fig. 3 as a reference, it means their hazard rates are at least around 5 percentage at every month.

We argue this scenario is extremely unlikely. Thus, our estimate of UGD is robust to potential overestimation of repayment ratio.

### 6 UNDERSTANDING UGD

#### 6.1 Decision Model

Next, we investigate a decision model where the investors primarily engage in accurate statistical discrimination [3, 8, 51] but may also have gender animus [11]—two workhorse frameworks of discrimination. Concretely, we assume every loan $i$’s repayment ratio is sampled from a gender-specific Gaussian distribution, $\lambda_i \sim N(\mu_g, \sigma_{g,0}^2)$, with repayment ratio mean $\mu_g$ and repayment ratio variance $\sigma_{g,0}$. The subscript $g$ denotes the variable varies by gender. Every loan $i$ emits a noisy signal of repayment ratio, $\tilde{\lambda}_i = \lambda_i + \sigma_{g,1} \epsilon_i \sim N(0, 1)$, with noise variance $\sigma_{g,1}$. We assume the investors observe the noisy repayment ratio signal $\tilde{\lambda}_i$ from loan characteristics, and are able to accurately predict the expected repayment ratio $\hat{\lambda}_i$.

$$\hat{\lambda}_i \mid (G_i = g) = \mathbb{E}[\lambda_i \mid \tilde{\lambda}_i] = (1 - \gamma_g)\mu_g + \gamma_g(\lambda_i + \sigma_{g,1}\epsilon_i)$$

$$\gamma_g = (\sigma_{g,0})^{-2} + (\sigma_{g,1})^{-2}$$

where $\gamma_g$ is the signal reliability. Then, the investors compute the expected return rate by multiplying the expected repayment ratio with $(1 + \text{the interest rate})$. Finally, the funding success is determined by whether the expected return rate exceeds a gender-specific decision threshold $\pi_g$, $D_i \mid (G_i = g) = \mathbb{I}[\hat{\lambda}_i(1 + R_i) \geq \pi_g]$.

This model gives the investors the benefit of the doubt. Gender animus and statistical discrimination are hard to distinguish in practice [16]. We primarily explain the observed UGD as arising from accurate statistical discrimination, where the investors accurately predicting the expected repayment ratio and the expected return rate. Any residual UGD is then attributed to disparate decision thresholds $\pi_m \neq \pi_f$, which is gender animus.

---

Preprint — under review.
We estimate parameters from the above decision model using a similar 2SPS approach. The first stage is the same. A survival model is fitted to predict the repayment ratio for the unsuccessful loans. This directly allows us to estimate the repayment ratio mean $\mu_g$ and variance $\sigma_{g,0}$. In the second stage, using the predicted repayment ratio and the estimated $\mu_g$ and $\sigma_{g,0}$, we use Bayesian inference to infer the posteriors of the repayment ratio noise variance $\sigma_{g,1}$, the signal reliability $y_g$, and the decision threshold $\pi_g$. We bootstrap both stages for 500 times to obtain confidence intervals. The Bayesian inference step is elaborated in Appendix C.

### 6.2 Analysis

Fig. 7 reports the estimated parameters. Female borrowers have higher repayment ratio mean $\mu_g$, lower repayment ratio variance $\sigma_{g,0}$, and lower repayment ratio noise variance $\sigma_{g,1}$. As the combined effect of $\sigma_{g,0}$ and $\sigma_{g,1}$, the signal reliability $y_g$ is also lower for female. These estimates are consistent with Fig. 4. Importantly, we find after first attributing UGD to accurate statistical discrimination, the residual gender animus component disfavors female. In order to be funded, female borrowers need to yield 2 percent higher expected return rate than male borrowers.

The above findings provide insights into the observed female favoritism. First, the considerable female favoritism can be completely attributed to accurate statistical discrimination, which is not rooted in animus but the fact that female is less likely to default on the studied P2P lending platform. Second, as illustrated from this model, the observed UGD favoring female does not necessarily reject the existence of gender animus against women. Many drivers may co-exist and contribute to the total UGD. One possibility that this model identifies is the investors are accurately predicting the expected return rate, which leads to female favoritism, but also hold gender animus against women, which reduces the magnitude of the female favoritism resulted from accurate statistical discrimination.

### 6.3 Counterfactual Analysis

We are interested in further quantifying the following counterfactual value: how large would the UGD be if the investors only engage in accurate statistical discrimination and do not have gender animus. We estimate these counterfactual values by simulating from the decision model of Sec. 6.1 and setting the male and female borrowers’ decision thresholds $\pi_m, \pi_f$ to some common value. The results are reported in Fig. 8 and Tab. 4.

We find accurate statistical discrimination alone may lead to less female favoritism for loans with low ($\leq 0.6$) and exceptional high ($>1.3$) return rates. But it leads to larger female favoritism for loans with return rates in (0.6, 1.3), where most of the loans are from (see Fig. 4). If the investors do not have gender animus, the average UGD favoring female will be of larger magnitude, between 7.73% and 11.83% depending on the values of decision threshold.

### 7 DISCUSSION

Our empirical findings suggest online P2P lending is a market-based affirmative action [23], which is arguably more efficient than bureaucratic rules in closing the gender gap in financial access. The P2P lending market functions much like a perfectly competitive market because (i) the size of P2P loans are typically small and (ii) there is a large number of investors. Via the standard demand-supply analysis [23], the market competition will impose the cost of gender animus—in the form of lower expected return rates—upon the investors who demand it. Therefore, competition reduces gender animus. Competition, however, forces the investors to use borrower gender to more accurately predict the expected return rate—i.e., engage in accurate statistical discrimination—because the investors can not perfectly observe the return rate. In sum, because women are less likely to default on their P2P loans, the market forces

---

**Fig. 7: Estimated decision model parameters.**

**Fig. 8: Counterfactual UGD by return rate when changing decision thresholds to $\pi_m = \pi_f = \pi \in \{1.06, 1.08, 1.10, 1.12\}$. Each red line connects four estimates, corresponding to $\pi = 1.06, 1.08, 1.10, 1.12$. The actual estimate is the same as Fig. 5.**

**Tab. 4: Counterfactual UGD when changing decision thresholds to $\pi_m = \pi_f = \pi \in \{1.06, 1.08, 1.10, 1.12\}$.**

| Decision threshold $\pi$ | Average UGD | Lower 95% CI | Upper 95% CI |
|--------------------------|-------------|--------------|--------------|
| 1.06                     | -0.0773     | -0.0799      | -0.0749      |
| 1.08                     | -0.0954     | -0.0984      | -0.0921      |
| 1.10                     | -0.1099     | -0.1134      | -0.1063      |
| 1.12                     | -0.1183     | -0.1222      | -0.1146      |
will drive the investors to rationally engage in accurate statistical discrimination favoring female.

In addition, affirmative action resulted from accurate statistical discrimination leads to an important narrative shift. The female borrowers’ favorable treatment becomes incentivized by their own expected return rates rather than bestowed upon because of their disadvantaged condition. Particularly, it generates a viable defence to the attack on affirmative action that less qualified female borrowers are selected over more qualified male borrowers [19]. In fact, more qualified female borrowers—in terms of expected return rates—are selected by the investors.

8 CONCLUSION

This work investigates a broadened notion of gender discrimination on one of the largest P2P lending platforms in China. We find empirical evidence of female favoritism, and show that it can be completely attributed to accurate statistical discrimination. We conclude with a discussion that online P2P lending can serve as a market-based affirmative action to support women.

REFERENCES

[1] 1971. Griggs v. Duke Power Company. 401 U.S. 424 (1971).
[2] Odd Aalen. 1978. Nonparametric inference for a family of counting processes. The Annals of Statistics (1978), 701–726.
[3] Dennis J Aigner and Glen G Cain. 1977. Statistical theories of discrimination in labor markets. J R Review 30, 2 (1977), 175–187.
[4] Alberto F Alessina, Francesca Lotti, and Paolo Emilio Mistrulli. 2013. Do women pay more for credit? Evidence from Italy. Journal of the European Economic Association 11, suppl. 1 (2013), 45–66.
[5] Per Kragh Andersen and Richard D Gill. 1982. Cox's regression model for counting processes. Springer.
[6] David G Blanchflower, Phillip B Levine, and David J Zimmerman. 2003. Discrimination leads to an important narrative shift. The female discrimination favoring female.
[7] Richard Berk, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. 2020. Gender, small firm ownership, and credit access: some insights from India. Management Science 66, 3 (2020), 1165–1181.
[8] Dongyu Chen, Xiaolin Li, and Fujun Lai. 2017. Gender discrimination in online peer-to-peer credit lending: evidence from a lending platform in China. Electronic Commerce Research 17, 4 (2017), 553–583.
[9] Xiao Chen, Bihong Huang, and Dezhu Ye. 2020. Gender gap in peer-to-peer lending: Evidence from China. Journal of Banking & Finance 112 (2020), 105533.
[10] Robert Cooter. 1994. Market affirmative action. San Diego L Rev. 31 (1994), 133.
[11] David R Cox. 1972. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological) 34, 2 (1972), 187–202.
[12] David R Cox and E Joyce Snell. 1968. A general definition of residuals. Journal of the Royal Statistical Society: Series B (Methodological) 30, 2 (1968), 248–265.
[13] William Darity et al. 2005. Stratification economics: the role of intergroup inequality. Journal of Economics and finance 29, 2 (2005), 144.
[14] William A Darity and Patrick L Mason. 1998. Evidence on discrimination in employment: Codes of color, codes of gender. Journal of Economic Perspectives 12, 2 (1998), 63–90.
[15] Bill Dedman. 1988. The color of Money. Atlanta Journal Constitution (1988), 1–4.
[16] Asli Demirgüç-Kunt, Leora Klapper, Dorothee Schaller, and Sanjaya Anand. 2022. The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19.
[17] Aislinn Bohren, Peter Hull, and Alex Imas. 2022. Preprint — under review.
[18] Aislinn Bohren, Kareem Haggag, Alex Imas, and Devin G Pope. 2019. Technical Report. National statistical discrimination: An identification problem

Preprint — under review.
[51] Edmund S Phelps. 1972. The statistical theory of racism and sexism. *The American Economic Review* 62, 4 (1972), 659–661.

[52] Devin G Pope and Justin R Sydnor. 2011. What’s in a Picture? Evidence of Discrimination from Prosper.com. *Journal of Human Resources* 46, 1 (2011), 53–92.

[53] Enrichetta Ravina. 2019. Love & loans: The effect of beauty and personal characteristics in credit markets. Available at SSRN 1107307 (2019).

[54] John Salvatier, Thomas V Wiecki, and Christopher Fonnesbeck. 2016. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* 2 (2016), e55.

[55] David Schoenfeld. 1982. Partial residuals for the proportional hazards regression model. *Biometrika* 69, 1 (1982), 239–241.

[56] Steven L Scott, Alexander W Blocker, Fernando V Bonassi, Hugh A Chipman, Edward I George, and Robert E McCulloch. 2016. Bayes and big data: The consensus Monte Carlo algorithm. *International Journal of Management Science and Engineering Management* 11, 2 (2016), 78–88.

[57] Xudong Shen, Yongkang Wong, and Mohan Kankanhalli. 2022. Fair Representation: Guaranteeing Approximate Multiple Group Fairness for Unknown Tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022), 1–1. https://doi.org/10.1109/TPAMI.2022.3148905

[58] Tyler Shumway. 2001. Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business* 74, 1 (2001), 101–124.

[59] Camelia Simoiu, Sam Corbett-Davies, and Sharad Goel. 2017. The problem of infra-marginality in outcome tests for discrimination. *The Annals of Applied Statistics* 11, 3 (2017), 1193–1216.

[60] Maria Stepanova and Lyn Thomas. 2002. Survival analysis methods for personal loan data. *Operations Research* 50, 2 (2002), 277–289.

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

[62] Terry M Therneau. 2021. A Package for Survival Analysis in R. https://CRAN.R-project.org/package=survival R package version 3.2-13.

[63] Terry M Therneau and Patricia M Grambsch. 2000. The cox model. In *Modeling survival data: extending the Cox model*. Springer, 39–77.

[64] Nirosha Hewa Wellalage and Sujani Thrikawala. 2021. Bank credit, microfinance and female ownership: Are women more disadvantaged than men? *Finance Research Letters* 42 (2021), 101929.

[65] Juanjuan Zhang and Peng Liu. 2012. Rational herding in microloan markets. *Management Science* 58, 5 (2012), 892–912.

*Preprint — under review.*
A FITTED SURVIVAL MODEL

|                | Coef. | Robust S.E. | Exp(Coef.) | p-value |
|----------------|-------|-------------|------------|---------|
| Male           | 0.408 | 0.021       | 1.503      | <.001   |
| Married        | -0.239| 0.024       | 0.787      | <.001   |
| APP Channel    | 0.170 | 0.034       | 1.186      | <.001   |
| Express Loan   | -0.284| 0.074       | 1.328      | <.001   |
| Repeated Borrower | -0.169| 0.039       | 0.844      | <.001   |
| Employment: 1  | -0.196| 0.049       | 0.822      | <.001   |
| Employment: 2  | 0.165 | 0.021       | 1.179      | <.001   |
| Employment: 3  | 0.027 | 0.028       | 1.027      | .342    |
| Employment: 4  | -0.175| 0.059       | 0.840      | .003    |
| Education: 1   | -0.513| 0.056       | 0.599      | <.001   |
| Education: 2   | -0.475| 0.028       | 0.622      | <.001   |
| Education: 3   | -0.673| 0.026       | 0.510      | <.001   |
| Education: 4   | -1.383| 0.278       | 0.251      | <.001   |
| ID Province: 1 | 76.711| 28 <0.001   |
| ID Province: 2 | 1.698 | 4 0.791 |
| ID Province: 3 | 2.286 | 4 0.683 |
| ID Province: 4 | 5.626 | 1 0.018 |
| ID Province: 5 | 5.466 | 1 0.019 |

Tab. A1: This table reports the main effect of the categorical covariates in the fitted survival model. The exponentiated coefficient is the multiplicative increase / decrease of hazard compared to the baseline hazard ($b_0(t)$ in Eq. 7).

Fig. A1: The figures plot the log-risk (with robust S.E.) of the continuous covariates in the fitted survival model. Exponentiated log-risk is the multiplicative increase / decrease of hazard compared to the baseline hazard ($b_0(t)$ in Eq. 7).

B SURVIVAL MODEL DIAGNOSTICS

Survival Model is Well-Specified. We first assess whether the proportional hazard (PH) assumption is satisfied in our survival model, using the scaled Schoenfeld residual [35, 55]. The Scaled Schoenfeld residual $\epsilon_i^{(j)}$ is measured for a specific loan $i$ and a specific covariate, such as age or education, which we index using the superscript $(j)$. Grambsch and Therneau [35] shows that, assuming an alternative survival model with time-varying coefficients $\beta^{(j)}(t)$ for the $j$-th covariate, if $\hat{\beta}^{(j)}$ is the $j$-th covariate’s fitted coefficient from the model of constant coefficients, we have

$$E[\epsilon_i^{(j)} + \hat{\beta}^{(j)}] \approx \beta^{(j)}(\tau_i),$$

(12)

where $\tau_i$ is the observation time of the loan $i$. I.e., the expected value of $\epsilon_i^{(j)} + \hat{\beta}^{(j)}$ approximates the actual, potentially time-varying coefficient $\beta^{(j)}(\tau_i)$. Using this result, two tests—the $\chi^2$ test and the visual test—can be developed to test for potential violations of the PH assumption.

Tab. B2 reports the $\chi^2$ test, which is a test of zero slope in the regression line fitted between the scaled Schoenfeld residual and observation time has zero slope.

|                | $\chi^2$ | df | p-value |
|----------------|----------|----|---------|
| Male           | 0.521    | 1  | 0.470   |
| Married        | 5.268    | 1  | 0.022   |
| APP Channel    | 3.630    | 1  | 0.057   |
| Express Loan   | 5.466    | 1  | 0.019   |
| Repeated Borrower | 5.626  | 1  | 0.018   |
| Employment     | 2.286    | 4  | 0.683   |
| Education      | 1.698    | 4  | 0.791   |
| ID Province    | 76.711   | 28 | <0.001  |
| $f_{ni}(Age)$  | 11.596   | 6  | 0.072   |
| $f_{ni}(# Past Failed Borrowings)$ | 12.942 | 4 | 0.012 |
| $f_{ni}(# Past Aborted Borrowings)$ | 3.746 | 4 | 0.442 |
| $f_{ni}(# Past Ontime Payments)$ | 59.121 | 9 | <0.001 |
| $f_{ni}(# Past Late Payments)$ | 32.308 | 5 | <0.001 |
| $f_{ni}(Amount)$ | 1.903   | 6  | 0.928   |
| $f_{ni}(Interest Rate)$ | 136.280 | 4 | <0.001 |

Tab. B2: The $\chi^2$ test for PH assumption. The null hypothesis is the regression line fitted between the scaled Schoenfeld residual and the observation time has zero slope.

To further investigate the degree of these violations, we conduct the visual test for these four covariates, which is shown in Fig. B2. The visual test fits a smooth curve to the scaled Schoenfeld residual and compares it against a horizontal fit, which represents a model of constant coefficient. For these four covariates, the visual test shows the violations are not economically meaningful. Constant coefficients still provide satisfactory approximations to the smooth curve, and always lie within the 95% CI. In fact, although we do not report due to the limited space, the approximation is visually no worse than the other covariates for which the $\chi^2$ test is unable to reject the null. Combining evidence from both the $\chi^2$ test and the visual test, we show the PH assumption is reasonably satisfied.

Preprint — under review.
Cumulative hazard of the Cox–Snell residual \( \epsilon \) (S.E.=0.001) [36]. It means the fitted model predicts which loan then means the loan has the lowest predicted default probability. From 0 to 1. A rank of 1 means the loan has the highest predicted default probability at its default month from that month’s risk set, which includes all loans that have not defaulted so far. A rank of 0 then means the loan has the lowest predicted default probability.

**C BAYESIAN INFERENCE**

The decision model explained in Sec. 6.1 can be written as follows:

\[
D_i | (G_i = g) \sim \text{Bern}(p = \Phi(\frac{1}{\sigma_{g,1}} \lambda_i - \frac{1}{\sigma_{g,1}} (\pi_g/\gamma_g) \times \frac{1}{1 + R_i} + \frac{1}{\sigma_{g,1}} \times ((1/\gamma_g - 1)\mu_g))),
\]

\[
y_g = \frac{(\sigma_{g,1})^{-2}}{(\sigma_{g,0})^{-2} + (\sigma_{g,1})^{-2}}.
\]

This model can be directly converted to a Bayesian latent variable model with the following unknown parameters: the repayment ratio noise variance \( \sigma_{f,1}, \sigma_{m,1} \) and the decision threshold \( \pi_f, \pi_m \). To complete the Bayesian model specification, we put the following weakly informative priors:

\[
1/\sigma_{f,1}, 1/\sigma_{m,1} \sim N^+(0, 2),
\]

\[
\pi_m, \pi_f \sim N^+(0, 2).
\]

We use \( N^+ \) to denote the half-Normal distribution. We put priors on the the inverse of \( \sigma_{f,1} \) and \( \sigma_{m,1} \) because we find it stabilizes the sampling procedure.

Performing Bayesian inference directly with this model, however, is compute-intense because the number of loans is more than half a million. We leverage the fact that both interest rate and repayment ratio only take some discrete values, and reframe the model as a Binomial variable at unique combinations of interest rate and repayment ratio, similar to Simoiu et al. [59]. After reformulation, the number of data is reduced to less than 500. The reformulation from a Bernoulli random variable to a Binomial random variable is without any loss and dramatically speeds-up inference.

We estimate the posterior distribution of \( \{\sigma_{g,1}, \pi_g \}_{g \in \{m,f\}} \) using Hamiltonian Monte Carlo (HMC) [31, 50], a kind of Markov chain Monte Carlo (MCMC) sampling approach [37, 47]. Specifically, we use the No-U-Turn sampler (NUTS) [40] implemented in the PyMC3 package [54] to sample from the posterior distribution. We set the target acceptance rate to 0.98, use 5000 draws for both warmup and estimation, and run 4 chains in parallel.

To yield confidence interval estimates, we similarly bootstrap both stages of 2SPS. The difference here is the second stage returns an estimated posterior distribution rather than a point estimate. We combine posteriors from different iterations by uniform averaging. After refactoring, the uncertainty of the predicted repayment ratios.

**Survival Model Has Goodness-of-Fit.** We use the Cox-Snell residual [25] to assess the model’s goodness-of-fit. If the model is correct and well-fitted, the Cox-Snell residual \( \epsilon_{CS} \) should resemble a censored sample from an unit exponential distribution. And if \( \epsilon_{CS} \) follows a censored unit exponential, the cumulative hazard of \( \epsilon_{CS} \) should be a straight line with zero intercept and unit slope. Fig. B3 implements this test. The estimated cumulative hazard of \( \epsilon_{CS} \) closely matches the 45° line and thus the survival model is well-fitted.

**Survival Model is Predictive.** The fitted survival model is predictive of the borrower’s default with a concordance index 0.638 (S.E.=0.001) [36]. It means the fitted model predicts which loan has higher repayment ratio—or is more trustworthy—with 63.8% accuracy, for all possible pairs of comparable loans. In Fig. B4, we plot the defaulted loans’ ranks as predicted by the survival model against their default months. Each defaulted loan’s rank ranges from 0 to 1. A rank of 1 means the loan has the highest predicted default probability at its default month from that month’s risk set, which includes all loans that have not defaulted so far. A rank of 0 then means the loan has the lowest predicted default probability.