

# Fair Allocation through Selective Information Acquisition

William Cai  
Stanford University  
willcai@stanford.edu

Nikhil Garg  
Stanford University  
nkgarg@stanford.edu

Johann Gaebler  
Stanford University  
jgaeb@stanford.edu

Sharad Goel  
Stanford University  
schoel@stanford.edu

## ABSTRACT

Public and private institutions must often allocate scarce resources under uncertainty. Banks, for example, extend credit to loan applicants based in part on their estimated likelihood of repaying a loan. But when the quality of information differs across candidates (e.g., if some applicants lack traditional credit histories), common lending strategies can lead to disparities across groups. Here we consider a setting in which decision makers—before allocating resources—can choose to spend some of their limited budget further screening select individuals. We present a computationally efficient algorithm for deciding whom to screen that maximizes a standard measure of social welfare. Intuitively, decision makers should screen candidates on the margin, for whom the additional information could plausibly alter the allocation. We formalize this idea by showing the problem can be reduced to solving a series of linear programs. Both on synthetic and real-world datasets, this strategy improves utility, illustrating the value of targeted information acquisition in such decisions. Further, when there is social value for distributing resources to groups for whom we have *a priori* poor information—like those without credit scores—our approach can substantially improve the allocation of limited assets.

## CCS CONCEPTS

• **Theory of computation** → **Mathematical optimization**; • **Computing methodologies** → **Machine learning**; • **Applied computing** → *Economics*.

## KEYWORDS

Algorithmic fairness, lending, information acquisition

### ACM Reference Format:

William Cai, Johann Gaebler, Nikhil Garg, and Sharad Goel. 2020. Fair Allocation through Selective Information Acquisition. In *Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)*, February 7–8, 2020, New York, NY, USA. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3375627.3375823>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

AIES '20, February 7–8, 2020, New York, NY, USA

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-7110-0/20/02...\$15.00

<https://doi.org/10.1145/3375627.3375823>

## 1 INTRODUCTION

Approximately one in seven U.S. households have unmet demand for small-dollar loans, and are often unable to secure credit from traditional financial institutions as they have little or no formal credit history [1]. However, in the majority of these households, individuals receive regular income and typically pay their bills on time, which suggests many in fact would have low risk of default [1]. One barrier to providing loans to this low-risk yet underserved sub-population is that it can be more expensive and time-consuming to screen individuals with non-traditional financial histories, limiting the inclusiveness of the banking system.

Motivated by this problem, we propose and analyze a strategy in which one can pay to acquire additional information on applicants. The task for a budget-constrained decision maker is thus to first select a set of candidates to screen and then, given the results of that screening process, determine to whom to allocate the remaining resources. In modeling this situation, we assume there is a fixed cost for screening each applicant, and that decision makers have prior knowledge of the distribution of information they would receive if they choose to screen an applicant. In practice, we note that such prior knowledge could be obtained by screening a small random sample of applicants to learn the resulting information distributions.

We derive an efficient algorithm for computing an optimal, utility-maximizing strategy for the general screening and allocation problem. To do so, we first show that once a set of candidates has been selected to screen, it is optimal to allocate the remaining resources according to a threshold rule, with assets distributed to those candidates having post-screening expected utility above a fixed threshold. Further, for any fixed threshold policy, we show that one can find the optimal set of candidates to screen (while satisfying the budget constraint) via a linear program. Intuitively, one should screen candidates near the margin, for whom the screening process could reveal information that could push a candidate across the threshold. But to do this rigorously, one also needs to account for the precise structure of the prior information. Finally, we sweep over the possible thresholds, solving the corresponding linear program at each point. In this manner, we obtain both a rule to screen candidates and a specific threshold policy for distributing funds to candidates with sufficiently high post-screening value.

We further consider an extension of the above problem in which policymakers have explicit value for diversity. For example, instead of simply finding a max-utility policy, one might maximize utility subject to the constraint that a particular group—such as

those who traditionally have had limited access to credit markets—achieve at least a fixed minimum utility. In the United States, those with unmet demand for credit are disproportionately black and Hispanic [1], heightening the value of diversity considerations in allocation decisions. We show that this (and related) extensions can be incorporated into our general algorithmic approach in a straightforward manner.

To demonstrate the potential value of augmenting allocation decisions with a screening phase, we apply our methods to both synthetic datasets and one with real measures of creditworthiness. In particular, we examine the potential benefits of screening as a function of the cost and value of information. Especially when we impose a diversity constraint, we find that screening strategies can significantly outperform a naive strategy that simply attempts to satisfy the constraint without screening any applicants.

For concreteness, we frame our discussion in terms of lending decisions, though our approach applies to many allocation settings. For example, it is often challenging to accurately assess household wealth—particularly in countries where informal and irregular work is more common—and, in turn, to appropriately target the distribution of government subsidies [22]. The simple strategy of distributing funds to those families clearly in need can systematically overlook populations with harder-to-verify financial status. As with lending, one can judiciously allocate some of the budget to more extensively screen certain applicants, ensuring funds are ultimately distributed to those who can benefit the most.

## 2 RELATED WORK

Several papers address the problem of *active feature acquisition* [19, 20, 25], where one can selectively purchase missing data to improve the overall out-of-sample performance of a statistical model. We consider the related problem of acquiring features to identify specific, high-value individuals. While there is some shared intuition between the two settings—that one should seek information on individuals most likely to alter downstream decisions—the technical approach we take is different, in large part because our end goal is optimal allocation rather than statistical learning.

In a related, recent stream of research, Bakker et al. [2] and Noriega-Campero et al. [22] likewise consider a feature acquisition problem, but with fairness constraints. In their setting, the decision maker must acquire additional features for each individual to ensure classification decisions have similar error rates across groups—including parity in false negative and false positive rates—a common measure of fairness in the machine learning community [8, 15, 17]. Our approach to the problem differs in three important respects. First, we adopt the perspective of constrained utility maximization. Past work has shown that directly equalizing error rates can lead to outcomes that, counterintuitively, may harm the very groups they were designed to protect [9, 10, 18]. We avoid such deleterious outcomes by instead framing the problem explicitly in terms of group-specific utilities: fairness is encoded into our requirement that the decision maker must allocate some minimum amount of utility to each group. Second, we focus on one-shot screening decisions, in which decision makers simply choose whether or not to acquire information on each individual, rather than sequentially deciding how much information to acquire based on the results

of each past acquisition decision. Our one-shot formulation maps to the binary decision structure (i.e., to screen or not to screen) common in many institutions and leads to different optimization challenges. Third, we directly model the tradeoff between screening and allocation decisions by tying both to a common budget constraint (i.e., more screening means less funds are available to ultimately distribute to individuals).

Elzayn et al. [12] also consider equitable ways to allocate resources, but in a setting where learning occurs by repeatedly allocating resources instead of by purchasing information. In comparison to our work, they focus on solving the problem of censored feedback, where the learner may not understand, and thus never allocate resources to, groups who did not previously receive resources.

Finally, our work touches on research from the fair division and allocation literature [5, 21, 26], which considers how to share resources while satisfying fairness properties defined between individuals. The former devises mechanisms wherein strategic agents self-divide the resources fairly, and the latter studies the existence of allocations that jointly satisfy various fairness notions. In contrast to our work, that line of research is particularly concerned with individual incentives, strategic action, and equilibrium effects.

## 3 A MODEL OF SCREENING AND ALLOCATION

We model screening and allocation decisions as a sequential process in which a budget-constrained lender first selects a (possibly random) subset of candidates to further screen from a pool of  $n$  applicants, and then, based on the information revealed in that screening phase, selects a second (possibly random) subset of candidates to receive a loan.

We assume the value of lending to an applicant  $i$  is given by the random variable  $U_i$ . These utilities are intended to capture the full social value of providing loans, and we imagine the lender aims to optimize social welfare, as in the case of a government agency. In general, the lender has only partial information about  $U_i$ . More specifically, if the lender chooses not to screen an applicant, we assume the lender knows only the applicant’s conditional expectation  $\mu_i = \mathbb{E}[U_i \mid X_i = x_i]$  given their pre-screening covariates  $x_i$ , such as credit score for those applicants who have traditional credit histories. On the other hand, if the lender opts to screen an applicant, they learn  $\mathbb{E}[U_i \mid X_i = x_i, \tilde{X}_i = \tilde{x}_i]$ , where  $\tilde{x}_i$  denotes the additional information one gains through screening. For example,  $\tilde{x}_i$  might encode applicant  $i$ ’s history of paying their electricity or phone bills—information that is often feasible to acquire with some extra effort and which is a good indicator of creditworthiness [1].

When deciding whom to screen, we assume the lender knows the distribution of  $\mathcal{D} = (D_1, \dots, D_n)$ , where  $D_i = \mathbb{E}[U_i \mid X_i = x_i, \tilde{X}_i]$ . That is, the lender knows how their estimate of utility could change if they decide to screen each applicant, where these distributions may depend on the available pre-screening covariates. A lender may, for example, thus choose only to screen applicants whose estimate is likely to substantially change given additional information. We further assume the lender must pay a fixed cost  $c_S$  for screening an applicant and a cost  $c_A$  for underwriting a loan. For simplicity we assume these costs do not vary across applicants, though it is

straightforward to extend to the more general case; see the online appendix for details.

Based on knowledge of the above information and cost structure, the lender selects a randomized strategy to screen applicants. That is, the lender chooses a vector  $\mathcal{P} = (p_1, \dots, p_n)$ , meaning that each applicant  $i$  is selected to be screened independently with probability  $p_i$ . Let  $\mathcal{S} = (S_1, \dots, S_n)$  indicate which applicants are ultimately screened under this policy; therefore,  $S_i \in \{0, 1\}$  is a Bernoulli random variable with probability of success  $p_i$ .

Given this randomized screening policy, we can now write the information  $\hat{\mathcal{U}} = (\hat{U}_1, \dots, \hat{U}_n)$  the lender has at the end of the screening phase as follows:

$$\hat{U}_i = \begin{cases} \mathbb{E}[U_i | X_i = x_i] & \text{if } S_i = 0, \\ \mathbb{E}[U_i | X_i = x_i, \tilde{X}_i] & \text{if } S_i = 1. \end{cases} \quad (1)$$

In other words,  $\hat{U}_i$  is the lender's post-screening estimated utility of giving a loan to applicant  $i$ . In particular, if  $S_i = 1$  (i.e., the applicant is screened), the lender's estimate changes from  $\mathbb{E}[U_i | X_i = x_i]$ , the estimate based only on applicant  $i$ 's pre-screening covariates  $x_i$ , to  $\mathbb{E}[U_i | X_i = x_i, \tilde{X}_i = \tilde{x}_i]$ , which incorporates the post-screening information  $\tilde{x}_i$ .

Finally, the lender chooses an allocation policy, denoted  $\mathcal{A}(\hat{\mathcal{U}}) = (A_1, \dots, A_n)$ , where  $A_i \in [0, 1]$  specifies the probability a loan is (independently) offered to each applicant. Importantly,  $\mathcal{A}$  is a function of the lender's post-screening utility estimates  $\hat{\mathcal{U}}$ . For example, the lender might give loans to the individuals with the highest post-screening utility estimates, up to the budget constraint.

Combining all of the above, the lender's optimization problem is to choose screening and allocation policies  $(\mathcal{P}^*, \mathcal{A}^*)$  that maximize expected welfare,

$$(\mathcal{P}^*, \mathcal{A}^*) \in \arg \max_{\mathcal{P}, \mathcal{A}} \mathbb{E} \left[ \sum_{i=1}^n U_i A_i \right], \quad (2)$$

subject to being budget-balanced in expectation,<sup>1</sup>

$$\mathbb{E} \left[ \sum_{i=1}^n c_S S_i + c_A A_i \right] \leq B, \quad (3)$$

where  $B$  is a fixed, non-negative constant.

In some settings, decision makers may value diversity in their allocations. For example, they may wish to ensure a certain minimum number of loans are provided to groups that historically have been excluded from credit markets. One can encode this policy preference directly into the utilities, in which case the resulting optimization problem would incorporate one's value for diversity. That approach, however, requires decision makers to agree upon these utilities to interpret the results, which can be challenging.

Here we take a complementary approach that explicitly allows value for diversity to differ across decision makers. Suppose the applicant pool is partitioned into  $m$  groups  $\{G_1, \dots, G_m\}$ ; for example, if  $m = 2$ , we might partition candidates into those who traditionally have had access to credit markets and those who have not. Then we require the selected policy  $(\mathcal{P}^*, \mathcal{A}^*)$  to allocate at least  $\Lambda_j \geq 0$

utility to group  $G_j$ , where these utilities do not themselves include any value for diversity:

$$\mathbb{E} \left[ \sum_{i \in G_j} U_i A_i \right] \geq \Lambda_j. \quad (4)$$

In practice, as we discuss below, one would solve this optimization problem for a range of  $\Lambda$ , which traces out the Pareto frontier of possible policies across different group constraints, corresponding to different values for diversity. The diversity condition above is expressed in terms of utility, but we might, alternatively, simply lower bound the *number* of loans  $\sum_{i \in G_j} A_i$  given to members of each group  $G_j$ . This alternative constraint can be handled in a straightforward manner by our algorithm detailed below.

### 3.1 A stylized example

We illustrate the above ideas in the context of a simple, stylized example. Suppose a lender must decide how best to allocate loans among an applicant pool of 13 people, with an overall budget of \$2,000. Providing a loan costs  $c_A = \$400$ , and additional screening of an applicant costs  $c_S = \$50$ .

Further suppose that five of the applicants are able to provide rich credit histories, and the expected utility of giving each of them a loan is  $\mu_i = \$750$ . For this group, additional screening would not provide any more information. Imagine that the other eight applicants do not have formal credit histories, and the utility of giving them a loan is accordingly lower due to the risk of default, with  $\mu_i = \$500$ . However, the lender knows that these applicants come in two types that could be disambiguated through additional screening. More specifically, the lender knows that after screening, individuals in this group can be divided into those with expected utility \$1,000 (with 50% chance) and those with expected utility \$0 (with 50% chance). The high-utility group could, for example, correspond to those who demonstrate a history of consistently paying their bills on time.

The naive strategy that does not screen any individuals would allocate five loans to hit the budget constraint, since  $5 \times \$400 = \$2,000$ . All five loans would go to applicants with rich credit histories ( $\mu_i = \$750$ ) over those without ( $\mu_i = \$500$ ). In this case, the total expected utility of the no-screening allocation is  $5 \times \$750 = \$3,750$ .

But in this scenario, one can improve overall utility by screening all eight applicants without formal credit histories and then granting loans to the high-utility applicants that are identified. Under that strategy, we expect four of the eight screened applicants to be identified as high utility (\$1,000), and so the expected utility of the allocation is  $4 \times \$1,000 = \$4,000$ , greater than the expected utility of \$3,750 under the no-screening strategy. Finally, the expected cost of the strategy is  $8 \times \$50 = \$400$  for screening plus  $4 \times \$400 = \$1,600$  for distributing the loans, totaling \$2,000 and satisfying the budget constraint. In this example, one can thus improve overall utility—and even allocate more loans to the group that *a priori* appears worse—by incorporating additional screening into the decision-making process.

## 4 FINDING OPTIMAL POLICIES

We now derive an efficient algorithm to find optimal screening and allocation policies  $(\mathcal{P}^*, \mathcal{A}^*)$ , subject to the budget and diversity

<sup>1</sup>By requiring the budget constraint to hold only in expectation—rather than exactly—we are able to find a computational tractable solution to the problem. In practice, such a constraint means that agencies are allowed some flexibility as long as they don't overspend on average, which, we believe, is often a realistic requirement.

constraints. We start by showing that over the full space of policies, it is optimal to allocate resources according to a *threshold policy*, with loans dispersed to individuals having post-screening expected utility  $\hat{U}_i$  above a group-specific threshold  $t_j$  for  $j = 1, \dots, m$ . Then, for each threshold policy, we show the optimal screening policy can be obtained by solving a linear program (LP), a type of optimization problem with linear objective function and linear constraints, for which there exist fast solution methods [3]. As a result, we can find an optimal combined screening and allocation policy by sweeping over threshold policies and solving the corresponding LP for each such policy.

We begin by formally defining threshold policies.

**Definition 1** (Threshold Policy). A *threshold policy* is an allocation policy  $\mathcal{A}$  for some fixed  $t_j \in \mathbb{R} \cup \{-\infty, \infty\}$  and  $\alpha_j \in [0, 1]$ ,  $j = 1, \dots, m$ , such that

$$A_i = \begin{cases} 1 & \hat{U}_i > t_{g_i}, \\ \alpha_{g_i} & \hat{U}_i = t_{g_i}, \end{cases}$$

where  $g_i$  denotes the group membership of individual  $i$ .

Threshold policies deterministically allocate resources to those with post-screening expected utilities above a fixed, group-specific threshold  $t_j$ . Randomization (i.e., allocating resources with probability  $\alpha_j$ ) at the threshold  $t_j$  may be necessary to exactly satisfy the budget constraint, which is important because for an individual not screened, the distribution of  $\hat{U}_i$  is concentrated at a single point.

Theorem 2 below formally states that it is sufficient to restrict to the set of threshold policies when searching for a globally optimal screening and allocation policy.

**Theorem 2.** Suppose the constrained optimization problem defined by Eqs. (2), (3), and (4) has a solution  $(\mathcal{P}^*, \mathcal{A}^*)$ . Then there is a threshold policy  $\mathcal{T}^*$  such that  $(\mathcal{P}^*, \mathcal{T}^*)$  is also a solution.

To see this, suppose that  $\mathcal{P}^*$  deterministically selects a subset of applicants to screen. Then, for each group  $G_j$ , it is clear one should allocate loans to the approximately  $k_j$  individuals in each group with the highest post-screening estimated utility, where  $k_j$  is the number of loans granted to each group under  $\mathcal{A}^*$ . Such a rank-based allocation can equivalently be written as a threshold rule with the same expected utility. The more general case, in which screening decisions are randomized, introduces some technical complications, but the spirit of the argument is similar. Full details can be found in the online appendix.

Now, given a threshold policy  $\mathcal{T}$  with thresholds  $t_1, \dots, t_m \in \mathbb{R}$  and boundary randomization probabilities  $\alpha_1, \dots, \alpha_m \in [0, 1]$ , we turn to finding an optimal companion screening policy. We show that such an optimal screening policy  $\mathcal{P}$  can be found by solving an LP. In particular, the LP has  $n$  decision variables  $p_1, \dots, p_n$ , with  $p_i \in [0, 1]$  specifying the probability that applicant  $i$  is screened; the objective equals the utility of the combined screening and allocation policy; and the constraints encode our budget and diversity conditions.

To construct the LP, we first define the following quantities that depend on both the threshold rule  $\mathcal{T}$  and the lender's prior

knowledge on the value of screening:

$$\begin{aligned} q_i &= \Pr(D_i > t_{g_i}), \\ e_i &= \mathbb{E}[U_i \mid D_i > t_{g_i}], \\ o_i &= \mathbb{1}_{\mu_i > t_{g_i}} + \alpha_{g_i} \cdot \mathbb{1}_{\mu_i = t_{g_i}}, \end{aligned}$$

where  $g_i$  denotes applicant  $i$ 's group membership. Recall that  $\mu_i$  denotes pre-screening expected utility and  $D_i$  denotes post-screening expected utility if applicant  $i$  is screened. Both  $\mu_i$  and the distribution of  $D_i$  are known in advance.

For fixed screening probabilities  $p = (p_1, \dots, p_n)$ , the expected utility of the corresponding screening and allocation policy can now be expressed as a linear function of  $p$ :

$$\sum_{i=1}^n [q_i \cdot e_i \cdot p_i + o_i \cdot \mu_i \cdot (1 - p_i)]. \quad (5)$$

The first summand reflects the expected utility associated with applicant  $i$  if they were screened, and the second summand reflects the expected utility if they were not screened.

Our goal is to maximize the expression in Eq. (5) subject to the budget and diversity conditions, which we now show can also be expressed as linear constraints on  $p$ . In terms of the constants defined above, the budget constraint can be written as

$$\sum_{i=1}^n [c_S \cdot p_i + c_A \cdot q_i \cdot p_i + c_A \cdot o_i \cdot (1 - p_i)] \leq B. \quad (6)$$

Likewise, the  $j$  diversity constraints in Eq. (4) become

$$\sum_{i \in G_j} [q_i \cdot e_i \cdot p_i + o_i \cdot \mu_i \cdot (1 - p_i)] \geq \Delta_j, \quad (7)$$

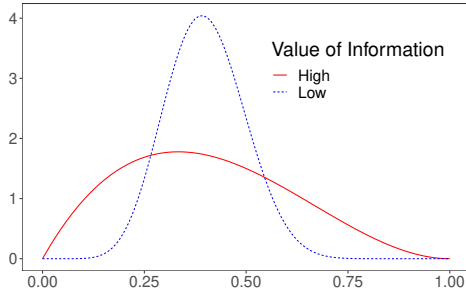
for  $j = 1, \dots, m$ .

Together, the objective given by Eq. (5), with constraints defined by Eqs. (6) and (7), define a linear program with decision variables  $\{p_i\}$ , the solution of which—if one exists—gives a screening policy  $\mathcal{P}^*$  that is optimal for the threshold policy  $\mathcal{T}$ . A jointly optimal screening and allocation policy  $(\mathcal{P}^*, \mathcal{T}^*)$  can accordingly be found through a grid search over all threshold policies  $\mathcal{T}$  in a (discretized) space  $\mathbb{R}^m \times [0, 1]^m$ . For each choice of policy  $\mathcal{T}$ , defined by threshold and randomization parameters  $t_1, \dots, t_m$  and  $\alpha_1, \dots, \alpha_m$ , we find the corresponding optimal screening policy  $\mathcal{P}$  by solving the LP described above. Then, among the resulting screening-allocation pairs, the one maximizing utility is guaranteed to be globally optimal.

## 5 EXPERIMENTS

We now investigate the value of our sequential screening and allocation approach through two simulation exercises, one based on synthetic data and another based on real loan data. First, with the synthetic data, we examine the optimal policies as we vary both the cost of screening and the value of the resulting information. Then, with the real-world data, we illustrate how a decision maker could, in practice, operationalize our screening and allocation approach to pursue equity when distributing limited resources.

In each experiment, we allocate loans to members of two groups. One of the groups—which we call the “targeted” group—can be screened for more information at some cost, and we imagine there is social value to distributing more resources to this group. For



**Figure 1: For an applicant with pre-screening likelihood of being creditworthy  $x_i = 0.4$ , the distribution of post-screening creditworthiness,  $\Pr[Y_i = 1 \mid X_i = 0.4, \tilde{X}_i]$ . Under a regime with high value of information (red line), the post-screening distribution has higher variance than in a setting with low value of information (blue line).**

example, the targeted group may be comprised of those who traditionally have not had ready access to the banking system and accordingly have limited formal credit history. For simplicity, we further assume that those in the non-targeted group cannot be screened, perhaps because they already have complete credit histories.

We define the utility of lending to applicant  $i$  to be

$$U_i = \begin{cases} a & Y_i = 1, \\ b & Y_i = 0, \end{cases} \quad (8)$$

where  $Y_i \in \{0, 1\}$  indicates whether they would pay back the loan, and  $a$  and  $b$  are fixed, known constants. In both of our experiments, we set  $a = \$1,000$  and  $b = -\$200$ , meaning there is \$1,000 of social utility when an applicant receives and pays back a loan and  $-\$200$  utility when an applicant defaults on a loan, for example because defaulting could trigger further financial distress. As above, we imagine the lender is a government institution or other agent attempting to maximize total social utility. We further assume that the pre-screening covariate  $x_i \in [0, 1]$  specifies the lender's pre-screening estimate of applicant  $i$ 's likelihood to repay a loan; in other words,  $x_i = \Pr(Y_i = 1 \mid X_i = x_i)$ . The covariate  $x_i$  can thus be translated into (pre-screening) expected utility  $\mu_i$  by Eq. (8):

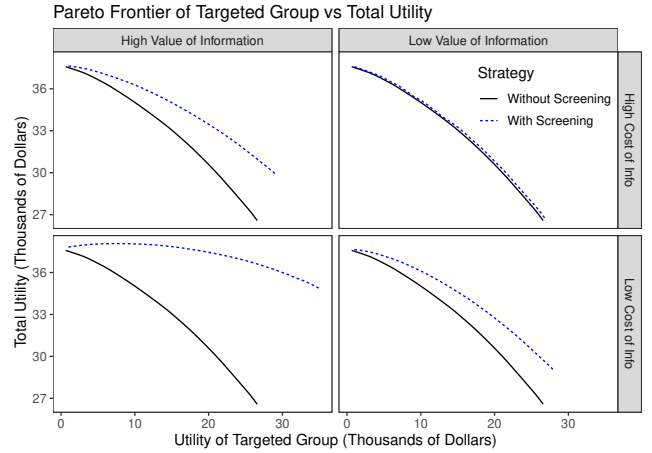
$$\mathbb{E}[U_i \mid X_i = x_i] = ax_i + b(1 - x_i).$$

## 5.1 Synthetic data

We illustrate the value of screening in four regimes of low vs. high cost of information paired with low vs. high value of information. To do so, we created four synthetic datasets, each comprised of  $n = 500$  individuals.

For all four datasets, we first evenly split the population into targeted and non-targeted groups. For each applicant in the targeted group, we generated their pre-screening probability of repayment  $x_i$  (or, equivalently, their observed pre-screening covariates) by independently drawing from a beta distribution<sup>2</sup> with mean 0.5

<sup>2</sup>We use the beta as it is a familiar, easily-parameterized function with support on  $[0, 1]$ .



**Figure 2: Comparison of our optimal screening strategy (blue line) with one without screening (black line) in four regimes with different costs and values of information.**

and count parameter 50.<sup>3</sup> Similarly, for each applicant in the non-targeted group, we generated their pre-screening probability of repayment by independently drawing from a beta distribution with mean 0.70 and count parameter 50. The higher mean repayment probability for members of the non-targeted group corresponds to them being, on average, more creditworthy.

Now, for each applicant in the targeted group, the lender may elect to screen them. As a result of screening, the lender receives an improved estimate  $\tilde{x}_i$  of the applicant's repayment probability, so that:

$$\mathbb{E}[U_i \mid X_i = x_i, \tilde{X}_i = \tilde{x}_i] = a\tilde{x}_i + b(1 - \tilde{x}_i).$$

We assume  $\tilde{x}_i$  is drawn from a beta distribution with mean  $x_i$  (i.e., the lender's pre-screening estimate of the applicant's repayment probability). In the high-information scenario, we set the count parameter for this beta distribution to be 5; and in the low-information scenario, we set it equal to 25.

Figure 1 shows these two post-screening information distributions for an applicant with  $x_i = 0.4$ . As illustrated in the plot, the high-information distribution (red line) has higher variance than the low-information distribution (blue line), and so screening is more likely to reveal very high risk and very low risk applicants in the high-information setting. Finally, we set the cost of screening  $c_S$  to be \$25 in the low-cost scenario and \$100 in the high-cost scenario, the cost of a loan to be  $c_A = \$1,000$ , and the total budget to be  $B = \$50,000$ .

Given these four datasets, we computed the optimal screening and allocation strategies using the algorithm described above.<sup>4</sup> In our original problem formulation, the diversity constraint in Eq. (4) specified only that we lower bound the utility of each group. To

<sup>3</sup>In terms of the  $\alpha$  and  $\beta$  shape parameters often used to parameterize beta distributions, the mean is  $\alpha/(\alpha + \beta)$  and the count parameter is  $\alpha + \beta$ . A higher count corresponds to a lower variance.

<sup>4</sup>Because only one group can be screened in our experiments, we use a faster variant of our optimization algorithm, described in the online appendix. For any fixed set of parameters, this approach returns the optimal screening and allocation policy within a few seconds with the open-source LP solver SCS [23].

better understand the impact of diversity on utility, we modify this constraint to be a strict equality for the utility of the targeted group and set the lower bound on utility to be \$0 for the non-targeted group. Thus, across a range of exactly satisfied utilities for the targeted group, we find the strategy that maximizes overall utility.

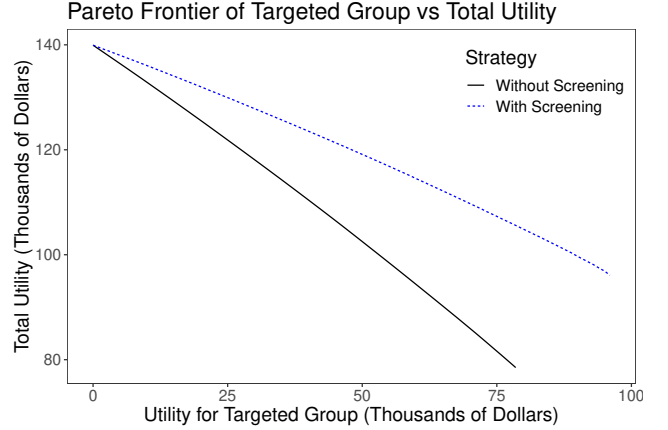
Figure 2 shows the results of this analysis, with the blue lines tracing out the Pareto frontiers for each of the four scenarios we consider. For comparison, the black lines show the corresponding result under a strategy that does not screen any applicants. Specifically, for any fixed utility constraint on the targeted group, the optimal no-screening policy first allocates loans to the  $k$  individuals in the targeted group most likely to repay based on their pre-screening estimates  $x_i$ , where  $k$  is chosen to satisfy the utility constraint; and then any remaining budget is used to allocate loans to those in the non-targeted group most likely to repay.

When either the value of information is high (left column) or the cost of screening is low (bottom row), we find that screening can be a valuable tool to improve utility. Notably, screening plays a more important role in these examples as we demand greater utility be allocated to the targeted group, since the no-screening strategy ends up dispersing loans to relatively high-risk applicants in the targeted group even though more creditworthy applicants in that group could be identified for little marginal cost. As one might expect, the gap between the screening and no-screening strategies is particularly large when both information is valuable and cheap. Indeed, in the high-value, low-cost setting (lower-left panel), one can achieve substantial diversity with little drop in overall utility.

## 5.2 Empirical credit data

We next apply our approach to the German Credit Dataset [16], which includes a variety of individual-level socioeconomic and financial characteristics (e.g., age, employment status, and credit history) on a sample of  $n = 1,000$  people, of whom 700 are deemed creditworthy. We define the targeted group to be those who currently do not own their residence, a subpopulation that comprises 28% of the dataset. In this case, 60% of individuals in the targeted group are creditworthy compared to 74% in the non-targeted group. For members of the targeted group, we assume the lender, prior to screening, only knows the targeted group’s overall base rate of creditworthiness. If, however, the lender chooses to screen an applicant in the targeted group, they learn  $\hat{x}_i$ , the applicant’s likelihood of being creditworthy conditional on all the available features in the dataset. For members of the non-targeted group, we assume this full estimate of creditworthiness is available prior to screening, and that there is no opportunity to obtain additional information.<sup>5</sup> Finally, we translate estimates of creditworthiness to estimates of utility via Eq. (8), in line with our simulations above.

Figure 3 shows the result of applying our screening and allocation algorithm to this dataset, where we assume the cost of screening  $c_S$  is \$100 (equal to our high cost regime in the synthetic datasets), the cost of a loan  $c_A$  is \$1,000, and the total budget is  $B = \$150,000$ . Like before, we compare the Pareto frontier of our approach (blue



**Figure 3: For the German Credit Dataset, comparison of our optimal screening strategy (blue line) with one without screening (black line).**

line) to that of a naive policy in which the lender does not screen applicants (black line). As with the synthetic datasets above, we find that the optimal policies with screening substantially outperform those without screening, particularly when we enforce a diversity constraint. For example, when we require \$50,000 of utility to come from allocating loans to the targeted group, the maximum total utility under the no-screening policy is \$102,000, compared to \$119,000 under the screening policy, an increase of 17%.

## 6 DISCUSSION

Many creditworthy individuals often have difficulty gaining access to traditional credit markets due to lack of formal financial histories, an issue that can exacerbate existing socioeconomic disparities. To address this gap, we developed a simple and efficient algorithm for a budget-constrained decision maker to screen applicants and then allocate a limited resource, an approach that we find offers substantial benefits on both real and synthetic lending datasets. This joint screening-plus-allocation approach is especially useful in settings where a targeted subset of the population—those most in need of an intervention—are also those for whom the least information is available *a priori*, a common situation in many social welfare programs.

Past research has shown that a dearth of high-quality data for various subgroups of the population can lead to poor models in a variety of domains [14, 24], including text analysis [4, 7, 13], facial recognition [6], and automated hiring [11]. Looking forward, our combined data acquisition and decision-making approach provides one framework to address this challenge by jointly modeling the cost of data collection and the value for subsequent improvements in downstream decisions.

## ACKNOWLEDGEMENTS

We thank Sam Corbett-Davies, Bernardo García Bulle Bueno, and Zhiyuan Jerry Lin for helpful comments and discussions. The online appendix is available at: <https://arxiv.org/abs/1911.02715>.

<sup>5</sup>More specifically, at the start of this exercise, we train a logistic regression model on the full dataset predicting creditworthiness as a function of the available covariates. Then, for members of the targeted group,  $\hat{x}_i$  is the model-estimated probability of creditworthiness for applicant  $i$ ; and for the members of the non-targeted group, that same model estimate is available prior to screening.

## REFERENCES

- [1] Apaam et al. 2018. 2017 FDIC National survey of unbanked and underbanked households. *US Federal Deposit Insurance Corporation* (2018).
- [2] Michiel A Bakker, Alejandro Noriega-Campero, Duy Patrick Tu, Prasanna Sattigeri, and Kush R Varshney. 2019. On Fairness in Budget-Constrained Decision Making. In *KDD '19: Explainable AI (XAI)*.
- [3] Dimitris Bertsimas and John N Tsitsiklis. 1997. *Introduction to linear optimization*. Vol. 6. Athena Scientific Belmont, MA.
- [4] Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In *Advances in Neural Information Processing Systems*. 4349–4357.
- [5] Steven J Brams and Alan D Taylor. 1996. *Fair Division: From cake-cutting to dispute resolution*. Cambridge University Press.
- [6] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*. 77–91.
- [7] Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science* 356, 6334 (2017), 183–186.
- [8] Alexandra Chouldechova. 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data* 5, 2 (2017), 153–163.
- [9] Sam Corbett-Davies and Sharad Goel. 2018. The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning. *arXiv preprint arXiv:1808.00023* (2018).
- [10] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. 2017. Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 797–806.
- [11] Jeffrey Dastin. 2018. Amazon scraps secret AI recruiting tool that showed bias against women. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>
- [12] Hadi Elzayn, Shahin Jabbari, Christopher Jung, Michael Kearns, Seth Neel, Aaron Roth, and Zachary Schutzman. 2019. Fair algorithms for learning in allocation problems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, 170–179.
- [13] Nikhil Garg, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences* 115, 16 (2018), E3635–E3644.
- [14] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumeé III, and Kate Crawford. 2018. Datasheets for datasets. *arXiv preprint arXiv:1803.09010* (2018).
- [15] Moritz Hardt, Eric Price, and Nati Srebro. 2016. Equality of opportunity in supervised learning. In *Advances In Neural Information Processing Systems*. 3315–3323.
- [16] Hans Hofmann. 1994. Statlog (German Credit Data) Data Set. *UCI Repository of Machine Learning Databases* (1994).
- [17] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. 2017. Inherent trade-offs in the fair determination of risk scores. In *Proceedings of Innovations in Theoretical Computer Science (ITCS)*.
- [18] Lydia T Liu, Sarah Dean, Esther Rolf, Max Simchowitz, and Moritz Hardt. 2018. Delayed impact of fair machine learning. *arXiv preprint arXiv:1803.04383* (2018).
- [19] Prem Melville, Maytal Saar-Tsechansky, Foster Provost, and Raymond Mooney. 2004. Active feature-value acquisition for classifier induction. In *Fourth IEEE International Conference on Data Mining (ICDM'04)*. IEEE, 483–486.
- [20] Prem Melville, Maytal Saar-Tsechansky, Foster Provost, and Raymond Mooney. 2005. An expected utility approach to active feature-value acquisition. In *Fifth IEEE International Conference on Data Mining (ICDM'05)*. IEEE, 4–pp.
- [21] Hervé Moulin. 2004. *Fair division and collective welfare*. MIT press.
- [22] Alejandro Noriega-Campero, Michiel A Bakker, Bernardo Garcia-Bulle, and Alex'Sandy' Pentland. 2019. Active Fairness in Algorithmic Decision Making. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. ACM, 77–83.
- [23] Brendan O'Donoghue, Eric Chu, Neal Parikh, and Stephen Boyd. 2016. Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications* 169, 3 (2016), 1042–1068.
- [24] Alexandra Olteanu, Carlos Castillo, Fernando Diaz, and Emre Kiciman. 2019. Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in Big Data* 2 (2019), 13.
- [25] Maytal Saar-Tsechansky, Prem Melville, and Foster Provost. 2009. Active feature-value acquisition. *Management Science* 55, 4 (2009), 664–684.
- [26] William Thomson. 2011. Fair allocation rules. In *Handbook of social choice and welfare*. Vol. 2. Elsevier, 393–506.