Prejudice and Competitive Signaling

Sue H. Mialon*

Abstract

In a signaling game between an employer and applicants, there is prejudice if some applicants are pre-judged and dismissed without receiving a chance to reveal their qualifying signals. We find that low quality of signal encourages employer's prejudice to ignore the informative signals on the basis of irrelevant characteristics such as gender. Prejudice works as a means of pre-screening the applicants in order to improve the quality of signals from pre-selected applicants at the expense of excluded applicants. Prejudice differs from discrimination in that prejudice reduces total amount of available information in the economy, whereas discrimination only redistributes signaling incentives among the applicants. Unlike discrimination, prejudice occurs mainly due to the inefficiency in signaling. Hence, in order to reduce prejudice, we advocate policies that enhance the quality of signals.

JEL Classification Numbers: J71, J78, D82, K31

Keywords: Competitive Strategic Signaling, Prejudice, Discrimination.

*Sue H. Mialon, Department of Economics, Emory University, Atlanta, GA 30322 (e-mail: smialon@emory.edu, phone: 404-712-8169). I am grateful to Maria Arbatskaya, Daniel S. Hamermesh, R. Preston McAfee, Hugo Mialon, Kaz Miyagiwa, Malathi Velamuri, and Seung Han Yoo for their helpful comments.

1 Introduction

Prejudice is "a preconceived judgment or opinion, an adverse opinion or learning formed without just grounds or before sufficient knowledge" (Merriam Webster dictionary). Hence, prejudice implies decision-making without having all relevant information. Typically, in a situation when a person (an applicant) has valuable but unverifiable private information, the applicant has an incentive to reveal the hidden information through signaling. What has not received much attention, however, is that successful signaling also requires that the receiver of the information (an employer) should have an incentive to utilize the signaled information. While signaling models assume that the employer always fully utilizes the signaled information, the existence of prejudice implies that this may not be the case in reality.

Consider an example in which two applicants, a male and a female, are considered for a job. In order to determine if they are qualified, the employer must acquire and review their signals. However, instead of acquiring both applicants' signals, the employer may look into only the male candidate's information and not the information from the female candidate. In this case, there is *prejudice*, not just because gender was a factor in the decision, but because the useful information about the female candidate's qualification was ignored as a result of gender-based pre-screening. Whether the female candidate was qualified, or even more qualified than the male candidate, would never be known.

This paper develops a model of prejudice as under-utilized signaling due to the lack of interest in the signaled information from the employer. We investigate what makes the employer unwilling to fully utilize the information (henceforth, full information). Why would the employer want to forgo potentially useful information in favor of an uninformative index (following Spence [1973]) such as gender in decision-making? To the employer, pre-screening based on an index is preferable only if signals are not very helpful in identifying high-quality candidates. Hence, prejudice is often the evidence of inefficiency in the signaling system that makes full information not worthwhile.

Signaling may not work efficiently when the applicants need to signal strategically in competition for scarce positions. We show how competition affects the quality of signals in Section 3.2. In this model, the return to signaling is scarce. In the example of two applicants, there is a shortage of the return when only one position is available. In this case, an applicant's chance of being hired depends on whether or not the other applicant also has a qualifying signal. This makes the two applicants' signaling decisions interdependent. Because not all qualified applicants can be hired, the applicants expect that their signaling effort can be wasted. The expected outcome under competition can be as discouraging as the one under prejudice. Under prejudice, the effort is wasted if the applicant does not possess the favorable characteristic, whereas under competition for scarce positions, the effort is wasted if the applicant does not win the competition. As receiving a return to their signaling effort becomes uncertain, the applicants' incentive to signal decreases. As the applicants' signaling effort decreases, the employer's value of full information under competition diminishes. If the employer expects the quality of information under full information to deteriorate substantially, he chooses prejudice.

The employer uses prejudice as a mechanism to alleviate the negative impact of competition when competition destroys the applicants' signaling incentives. The main role of prejudice is to *pre-screen* the applicants and reduce the number of applicants to consider in order to weaken the value-destroying competition among the applicants.

The benefit of full information comes from allocating signaling incentives to *every* applicants. In contrast, prejudice always results in forfeiture of incentives for some applicants. Thus, prejudice is inefficient if society produces more valuable information by keeping opportunities open for every applicant. However, this paper points out that full information can be more inefficient than prejudice if competition scatters the incentives only to ensure that no one receives a sufficient incentive to make a meaningful effort. Especially, full information can be very inefficient if the signals are more correlated with the index, such as race or gender, than with the applicants' quality. For example, suppose that it is prohibitively costly for minorities to signal, but not for others. Then, regardless of whether or not there is prejudice, minorities with high ability face equally adverse signaling environment as they will never have a sufficient incentive to signal due to either economic hardship or prejudice. In this case, there is no essential difference between full information and prejudice.

In order to maintain the efficiency of full information under competition, the signals needs to be closely correlated with the hidden quality of applicants, but not with their current economic status, race, or gender. Unless the signals under full information are informative, any efforts to reduce prejudice would be futile. Thus, we emphasize the importance of policies to improve the quality of signals. For instance, Head Start programs or the Every Student Succeeds Act may help minorities overcome adverse signaling environment. If these policies can make signaling equally accessible to all applicants, independently of their economic status or race, then any difference in their signaling outcomes could be correctly attributed to a difference in the hidden quality. In that case, competitive signaling can work more efficiently.

The remainder of this paper is organized as follows. Section 2 briefly summarizes the contributions of this paper to the literature. Section 3 presents a benchmark case in which there is no competition among applicants and a main model of a strategic signaling with a shortage in the return to signaling. Section 4 considers extensions to a generalized model of prejudice and the long-term effects of prejudice via intergenerational wealth transfer. We discuss the implications of related policies. Section 5 presents our conclusions.

2 Contributions to the literature

This paper is the first to model prejudice, in distinction from discrimination, as a purposeful act of ignoring potentially valuable information about applicants in favor of a pre-judgment based on an uninformative index. This paper also analyzes how the expectation of discrimination interacts with prejudice.

On discrimination, there is an abundance of theoretical literature. While some theories, such as in Becker, (1957) describe discrimination as a matter of preference, many others spurred by the seminal papers of Phelps (1972) and Arrow (1973) explains group inequality as a result of rational choice based on various reasons that motivate the use of group index to infer some relevant quality of applicants in the presence of incomplete information about the unobservable quality.

In models of statistical discrimination, group inequality can arise in equilibrium even if there is no exogenous difference between groups. Coate and Loury (1993) explains that the inequality is a result of group-wide coordination onto different equilibria, i.e., coordination failure, whereas Moro and Norman (2004) explains that it is a result of group-wide specialization over different types of jobs. Moro and Norman (2004) are the first to explicitly model discrimination as a "relative" advantage that induces a greater payoff for one group at the expense of another. In their general equilibrium model, the production technology requires two complementary inputs. If too many workers invest in the skills for one input, the marginal product of each input adjusts accordingly. Discrimination is an outcome that two groups specialize in different tasks. Mialon and Yoo (2017) introduce interdependent signaling decisions of applicants to explain how employers benefit from the beliefs of discrimination. They find that employers may prefer discrimination because it reduces the overall risks of hiring. They show how employers can actively influence the formation of beliefs on discrimination. Fang and Moro (2011) provide an excellent survey of the literature on statistical discrimination.[1]

The framework of fixed wage in this paper plays a similar role of "posted wage" in Lang et al.(2008). This commonly-observed labor market practice makes employers unable to adjust their posted wage ex post upon observing many more candidates with a qualifying signal than the available number of positions. Thus, under this framework of posted wages, the competition among workers are expected to be persistent. Lang et al.(2008) focuses on how posted wages amplifies slight racial preferences to produce significant racial discrimination and segregation in labor markets.

In order to explain prejudice, the present paper develops a framework of strategic and interdependent signaling decisions by the applicants. In this sense, this paper is closely related to Mialon and Yoo (2017) which analyzes discrimination in the framework of interdependent signaling. However, this paper shows that *prejudice* fundamentally differs from *discrimination*. Prejudice is essentially a choice to reduce competition in signaling, whereas discrimination only redistributes signaling incentives among the applicants for any given degree of competition. The incentive for prejudice arises when competitive signaling is inefficient, whereas the incentive for discrimination is not necessarily due to the inefficiency. The incentive for discrimination exists for any level of competition. With or without discrimination, prejudice occurs if full information under competitive signaling is inefficient. In contrast, as in Mialon and Yoo (2017), discrimination is mainly a result of "self-fulfilling beliefs." Section 4 highlights the different implications of non-prejudice policies and non-discrimination policies.

[1] Given the myriad of papers on discrimination, many interesting studies had to be skipped in this literature review simply because they are not closely related to this paper. For example, Mailath et al (2000) model discrimination as an outcome of search frictions. Chaudhuri and Sethi (2008) discuss the relationship between integration and discrimination. Blume (2005) shows that discrimination is a result of firms' learning regarding worker investment decisions. Fryer (2007) finds that a discriminated worker who overcomes initial discrimination can actually benefit from the discrimination. For the literature on explaining what makes the group inequality persistent, see Dulauf (1996), Benabou (1996), Mookerjee et al (2010), and Bowles et al. (2014), for examples.

3 Model

Consider a signaling game between an employer R (the receiver of signals) and n applicants (the senders of signals), $n \geq 2$. An applicant earns $W > 0$ if hired, and nothing, otherwise.[2] Each applicant is characterized by an unobservable ability $A \in \{H, L\}$ and an observable characteristic $\theta \in \Theta = \{1, 2, ...n\}$. Assume that there is one applicant for each θ.[3] An applicant θ has a high (H) ability with a probability of $q > 0$. Applicants can make an effort to signal their ability. The effort $e \in [0, 1]$ generates an observable signal of high ability σ_H with a probability e and a signal of no qualification σ_L with a probability $1 - e$. Let $C_{\theta A}(e) \in [0, \infty)$ and $c_{\theta A}(e)$ be type θA applicant's cost and marginal cost of investing in e, respectively. $C_{\theta A}(e)$ is strictly convex, twice differentiable, $C_{\theta A}(0) = 0$, $C_{\theta L}(e) \geq C_{\theta H}(e)$, and $c_{\theta L}(e) > c_{\theta H}(e)$. We assume that $\frac{c_{1L}(e)}{c_{1H}(e)} \geq \frac{c_{2L}(e)}{c_{2H}(e)} \geq ... \geq \frac{c_{nL}(e)}{c_{nH}(e)}$ for a given $e > 0$.

The payoff for the employer R depends only on the applicant's ability A. R receives $V > 0$ from hiring a $H-$type applicant, and 0 for selecting a $L-$type applicant. The qualifying signal σ_H is observable from θ only if R reviews the application at the cost of $s \geq 0$, whereas θ is costlessly observable. In some cases, R may pre-select only a subset Θ_P of applicants to review based on θ, $\Theta_P \subset \Theta$, $\Theta_P \neq \Theta$. In that case, R observes the signals only from the pre-selected type $\theta \in \Theta_P$.

Throughout the paper, the timing of the game between the employer and the applicants is as follows. At Stage 0, nature chooses the ability A for each applicant θ and determines k available positions for the employer. In the benchmark case, $k \geq n$ and, thus, there is no competition among the applicants. In contrast, the main model in Section 3.2 introduces competition due to insufficient positions, $k < n$. In this case, applicants foresee a non-negligible chance of failing to get a position even if they are qualified and strategically determine their signaling effort. At Stage 1, the employer determines whether to consider the entire pool of applicants (T) or to review only a subset (P) based on the θ. In the case of P, n^* of θ applicants are pre-selected with probability p_θ, $n^* < n$. Without observing the choice between T and P,[4] applicants

[2] We take $W > 0$ as exogenously given, because prejudice occurs in many non-market situations as well as in markets. For example, in the case of college admissions, the payoff W from being selected for scarce positions is not a market wage.

[3] n is the maximum number of *differentiable* characteristics. The number of θ that are actually used in differentiating the applicants can be much less than n, as shown later in this section.

[4] This information on T or P is publicly unavailable because employment discrimination and prejudice based on indices such as race and gender is illegal. Even in the case of non-hiring deci-

determine their effort e to signal their ability and receive a signal for a given effort e. At Stage 2, the signals (σ_H, σ_L) become available to R. If P had been chosen at Stage 1, R can observe the signals only from the pre-selected $\theta \in \Theta_P$. If T had been chosen, however, all applicants' signals are observed. At Stage 3, based on the available signals σ_H or σ_L from the applicants, R determines hiring.

For simplicity, our analysis focuses on the cases where, in equilibrium, R never hires an applicant with the signal σ_L. This allows us to focus on the value of information in a qualifying signal σ_H in explaining R's incentive for prejudice.[5] In addition, we assume that the cost of processing information is negligible, $s \approx 0$, for simplicity.

3.1 Benchmark

Suppose that $k \geq n$. Since there are more positions than the number of applicants, the applicants expect no competition. At Stage 3, observing that type θ applicant has a signal σ_H, R's expected payoff from hiring the applicant is $\mu_\theta(H|\sigma_H)V - W$, where $\mu_\theta(H|\sigma_H) := \frac{qe_{\theta H}}{qe_{\theta H}+(1-q)e_{\theta L}}$, and $e_{\theta A} > 0$ is the probability of observing σ_H from the type θA applicant. Since $k \geq n$, the employer R is inclined to hire *any* applicant θ so long as $\mu_\theta(H|\sigma_H)V - W > 0$. Then, at Stage 1, R reviews θ with a probability $p_\theta > 0$ only if $\mu_\theta(H|\sigma_H)V - W > 0$ is expected, for $\theta \in \Theta_P$, whereas $p_{\theta'} = 0$, for $\theta' \notin \Theta_P$. If R expects $\mu_\theta(H|\sigma_H)V - W > 0$ for all types, then he will choose T to review entire applications.[6]

If the employer plans not to use some θ's information of signal σ_H in hiring, he will choose not to acquire the signal in the first place by excluding those θ in the review process with a choice of P, instead of wasting the information of σ_H after acquiring it. Thus, in the present framework, there is no wasteful acquisition of σ_H. R prioritizes the decision between T and P in determining the value of seeking for the signal σ_H from more applicants.

sions, social norms against prejudice and discrimination often make the information about T or P unavailable.

[5] The main qualitative results about prejudice in this paper are not affected by this behavioral assumption. If an applicant with σ_L can be hired as well, it reduces applicants' incentive to signal given that signaling is costly. Thus, allowing the possibility of hiring an applicant with σ_L makes signaling less informative and less valuable to the employer. In this paper, prejudice is caused by ineffective signaling and linked to insufficient signaling efforts by the applicants. Thus, when an applicant with σ_L can be hired, in equilibrium, prejudice is more likely to be prevalent.

[6] For a negligible but positive s, the employer does not review an applicant θ if $\mu_\theta(H|\sigma_H)V - W = 0$ because the cost of reviewing type θ can be saved from dropping the θ at Stage 1.

Let α be the probability that the employer chooses T to review entire applications, i.e., $\alpha = 1$ if T, and $\alpha = 0$ if P. Then, for a type θA applicant, the expected payoff from investing in e is $E(e|\theta, A) = e[\alpha + (1-\alpha)p_\theta]W - C_{\theta A}(e)$. The optimal effort level $e_{\theta A} > 0$ satisfies

$$x_\theta \equiv [\alpha + (1-\alpha)p_\theta]W = c_{\theta A}(e_{\theta A}). \tag{1}$$

x_θ determines the signaling incentive for θ. For any given $x_\theta > 0$, $e_{\theta H} > e_{\theta L}$, since $c_{\theta L}(e) > c_{\theta H}(e)$, and $c'_{\theta A} > 0$ for $e > 0$.

The employer R's expected payoff from using the type θ applicant's information is $U_\theta := qe_{\theta H}(V - W) - (1-q)e_{\theta L}W$. The employer uses full information (T) if he expects that $\mu_\theta(H|\sigma_H)V - W > 0$ from *all* θ, and the expected payoff is

$$E(T) = \sum_{\theta \in \Theta} \underbrace{qe_{\theta H}(V - W) - (1-q)e_{\theta L}W}_{U_\theta}, \tag{2}$$

whereas the expected payoff from using only a subset Θ_P is $E(P) = \sum_{\theta \in \Theta_P} p_\theta U_\theta$. As long as $\mu_\theta(H|\sigma_H)V - W > 0$, $U_\theta > 0$. (1) and (2) show that for $k \geq n$, U_θ is independent of the efforts chosen by the other type θ', $\theta \neq \theta'$. The equilibrium depends on the applicants' expectation of the chance to be reviewed, α. From (1), when $\alpha = 1$, $x_\theta = W$. Define $g_\theta(x_\theta) = c_{\theta L}^{-1}(x_\theta)/c_{\theta H}^{-1}(x_\theta)$. To make applicants' signaling effort intrinsically valuable to the employer, we need the following assumptions.

Assumption 1 $\frac{q(V-W)}{(1-q)W} > g_\theta(W)$ *for all* θ.

Assumption 2 $g_\theta(x_{\theta 1}) > g_\theta(x_{\theta 2})$ *for* $x_{\theta 2} > x_{\theta 1}$.

From (1), as long as applicants expect a sufficiently high chance of being reviewed ($\alpha + (1-\alpha)p_\theta$), they make a sufficient effort to induce $U_\theta > 0$. Then, *all* of the signals are worthwhile to look into for the employer and, thus, full information occurs in equilibrium.[7] In this case, if prejudice occurs in equilibrium, it can be only due to self-fulfilling expectations of a low p_θ, just as predicted in typical statistical discrimination models. Otherwise, the employer has no reason to choose P ($\alpha = 0$) and refuse to review the information from some applicants θ'. In particular, prejudice against a type θ' applicant requires that all applicants, both θ and θ', $\theta \neq \theta'$, correctly expect that only $p_{\theta'}$ will be low while p_θ is not.

[7] A detailed proof of the equilibrium is provided in the Appendix.

Clearly, this benchmark model does not explain prejudice properly. First of all, this model does not explain *why* they should believe that only $p_{\theta'}$ will be lower. Second, any equilibrium involving an expectation of a low $p_{\theta'}$ is not stable. Given that $k \geq n$, the employer is always willing to hire any applicant as long as the applicant makes a sufficient effort. Knowing this, there is no reason why the applicants should persistently believe that type θ' will not be given a chance even if they make a sufficient effort to signal their quality.

Prejudice (and discrimination) is always a relative concept: offering a type θ applicant an opportunity to prove his qualification becomes prejudice, only if the same opportunity is not available to other type θ' applicant for no proper reason. This requires *comparison* of the applicants. The framework in the benchmark is inadequate to explain prejudice, because the employer's decision for each applicant is independent. Therefore, in the following, we consider a new model in which the employer needs to consider all types *in comparison*.

3.2 Competitive signaling

Suppose now that $k < n$. When R chooses T to review all, at Stage 3, if more than k applicants turn out to have a qualifying signal σ_H, there is a competition among the qualified applicants. In selecting the best k candidates, for the employer R, the value of an applicant's signal is determined relatively in comparison with the other applicants' signals. In order to be hired, an applicant not only needs a qualifying signal, but also her signal must be of higher quality than that of her competitors. This makes all the applicants' signaling decisions interdependent.

Let M be the set of the m applicants with σ_H at Stage 3. R selects an applicant $\theta \in M$ with a probability $\pi_{\theta m} \in [0, 1]$. There is *discrimination* when $\pi_{\theta m} \neq \pi_{\theta' m}$, for any *ex ante* identical $\theta, \theta' \in M$. In contrast, *prejudice* occurs when R is only interested in the signal of the type $\theta \in \Theta_P$ and does not wish to learn about an applicant's qualifying signal if the type is θ', $\theta' \notin \Theta_P$. Therefore, prejudice differs from discrimination in the level of information that R has.

When P is chosen, R chooses n^* of applicants θ to include in the subset Θ_P for a review, $k \leq n^* < n$. If $n^* = k$, $p_\theta = 1$ for $\theta \in \Theta_P$ and any of those θ with σ_H will be hired. On the other hand, if $n^* > k$, then p_θ depends on how prejudice is implemented. This is because, when $n^* > k$, if there are more qualified applicants than k, competition may exist even within the n^*pre-selected applicants under P. In this section, we consider a simple case that P entails no expectation of competition, implying that if $n^* > k$, $p_\theta < 1$ for some $\theta \in \Theta_P$. By doing so, the main analysis in this section focuses on identifying the impact of competition under full information T and how it promotes prejudice. In Section 4, however, we incorporate a more general framework of prejudice that allows competition among the n^* pre-selected applicants and show that the same qualitative results hold. Hence, in general, P can be understood as an environment with less competition compared to T.

3.2.1 Applicants' signaling effort

Consider the signaling decision for type $1H$ applicant in an example of $n = 3$ and $k = 1$. If R chooses P, $1H$ expects to be selected with a $p_1$probability. If R chooses T to review all, $1H$ with a qualifying signal faces competition when $m > 1$ and expects to be hired with a π_{1m} probability, whereas she expects to be hired for sure if she is the only one with σ_H. Let F_2 and F_3 denote the probabilities that the type 2 and 3 applicants have a qualifying signal σ_H, respectively, where $F_\theta = qe_{\theta H} + (1-q)e_{\theta L}$, $\theta = 2, 3$. For given α, p_1,and π_{1m}, the expected payoff for type $1H$ applicant is

$$E(e_{1H}; \mathbf{e}_{-1}) = e_{1H} W \left[\begin{array}{c} \alpha\{F_2 F_3 \pi_{13} + (F_2 + F_3 - 2F_2F_3)\pi_{12} \\ +(1 - (F_2 + F_3 - F_2F_3))\} + (1-\alpha)p_1 \end{array} \right] - C_{1H}(e_{1H}), \quad (3)$$

where $\mathbf{e}_{-1}$ denotes the efforts by the competitors of type 1 applicant.

Generalizing this, under T, we can see that for any given m applicants with a qualifying signal σ_H, (i) if $m > k$, an applicant faces competition even after obtaining σ_H, but (ii) if $m \leq k$, an applicant with σ_H is guaranteed to be hired. When $m > k$, let $F_{-\theta Cm}$ denote the probability that type θ applicant with σ_H faces competition with $m-1$ other applicants with σ_H. $F_{-\theta Cm}$ is a function of $\mathbf{e}_{-\theta}$. The expected payoff for type θA applicant is

$$E(e_{\theta A}; \mathbf{e}_{-\theta}) = e_{\theta A} \left[\alpha w_\theta + (1-\alpha)p_\theta\right] W - C_{\theta A}(e_{\theta A}), \quad (4)$$

$$\text{where } w_\theta \equiv \sum_{m=k+1}^{n} (F_{-\theta Cm} \cdot \pi_{\theta m}) + \left(1 - \sum_{m=k+1}^{n} F_{-\theta Cm}\right). \quad (5)$$

Other things being equal, the probability of competition $\sum_{m=k+1}^{n} F_{-\theta Cm}$ increases as k decreases. If $k = 3$, for example, $\sum_{m=k+1}^{n} F_{-\theta Cm}$ is a probability that there are at least three more applicants with σ_H other than θ. If $k = 2$, it is a probability that there are at least two more applicants with σ_H other than θ.

The optimal effort level $e^*_{\theta A} > 0$ is chosen to satisfy

$$x_{\theta C} \equiv \left[\alpha w_\theta(\mathbf{e}_{-\theta}, \pi_{\theta m}) + (1-\alpha)p_\theta\right] W = c_{\theta A}(e^*_{\theta A}), \tag{6}$$

for $x_{\theta C} > 0$, and w_θ is defined in (5). If $x_{\theta C} = 0$, $e^*_{\theta H} = e^*_{\theta L} = 0$. Since $F_{-\theta Cm}$ depends on $\mathbf{e}_{-\theta}$, the signaling efforts by competing applicants $-\theta$, (6) determines type θA's best response for a given $\mathbf{e}_{-\theta}$,

$$e^*_{\theta A} = c^{-1}_{\theta A}(\left[\alpha w_\theta(\mathbf{e}_{-\theta}, \pi_{\theta m}) + (1-\alpha)p_\theta\right] W). \tag{7}$$

Note that if $k \geq n$, $\pi_{\theta m} = 1$ for all m, and $w_\theta = 1$ for all θ. Thus, $x_{\theta C}$ becomes x_θ in (1). Compared to (1), under competition, the signaling incentive $x_{\theta C}$ in (7) is weighted by $w_\theta \leq 1$, because hiring is not guaranteed ($\pi_{\theta m} \leq 1$) even after obtaining the qualifying signal σ_H. Since $w_\theta \leq 1$, other things being equal, $x_{\theta C} \leq x_\theta$, with a strict inequality for at least one θ with $\pi_{\theta m} < 1$. Thus, applicants' incentive to signal is lower under competitive signaling. Proposition 1 summarizes this result.

Proposition 1 *Other things being equal, competition reduces applicants' incentive to signal.*

When $k < n$, an applicant's signaling effort imposes a negative externality on the other applicants. An applicant θ with σ_H is guaranteed for a position only if there are at most $k-1$ other applicants with σ_H. However, if more than $k-1$ other applicants have σ_H, the applicant θ can get the position only with a $\pi_{\theta m} \leq 1$ probability. Thus, the applicant's incentive to signal decreases.

On the other hand, P relieves such a competitive pressure by not acquiring all the signals. Type θ applicant expects $x_{\theta C} = p_\theta W$ with a p_θ chance of pre-selection. Hiring is guaranteed as long as the applicant with σ_H has been pre-selected as one of the n^* applicants. As a result, an applicant's best response becomes independent of the efforts of the other applicants, eliminating the competition effect. That is, an important function of pre-selection P is to reduce the impact of competition. For this reason, if competition destroys the applicants' incentive to signal and the value of signaling, R may choose P in order to eliminate the negative impact of competition.

3.2.2 Employer's optimal hiring strategy

Discrimination at Stage 3 First, consider R's decision concerning the optimal $\pi^o_{\theta m}$ if more than k applicants turn out to have σ_H after choosing T to acquire all applicants' signal. For example, suppose that $n = 3, k = 2$, and R observes that all three applicants $\theta = 1, 2, 3$ have σ_H. R's payoff from selecting type 1 is $E(1|\sigma_H,\sigma_H,\sigma_H) = \mu(1 = H|\sigma_H, \sigma_H, \sigma_H)V - W$, where $\mu(1 = H|\sigma_H, \sigma_H, \sigma_H) = \frac{qe_{1H}F_2F_3}{F_1F_2F_3} = \frac{qe_{1H}}{F_1} = \mu_1(H|\sigma_H)$, and $F_\theta = qe_{\theta H} + (1-q)e_{\theta L}$. Thus, $E(1|\sigma_H,\sigma_H,\sigma_H) > 0$, as long as $\mu_1(H|\sigma_H)V - W > 0$. However, an important difference in competitive signaling is that R also cares about which candidate is more likely to be qualified. $E(1|\sigma_H,\sigma_H,\sigma_H) > E(2|\sigma_H,\sigma_H,\sigma_H)$ if and only if $\mu_1(H|\sigma_H) > \mu_2(H|\sigma_H)$, or

$$\frac{e_{1L}}{e_{1H}} < \frac{e_{2L}}{e_{2H}}. \tag{8}$$

(8) implies that type 1's signal σ_H is more likely from an $H-$type than is type 2's signal σ_H. Thus, type 1 is preferable to type 2 if R believes that (8) holds. When many applicants have equally qualifying signals, R hires the applicants with a lower probability of signaling errors.

In general, when m applicants have σ_H, a type θ applicant is preferable to type θ' if $e_{\theta L}/e_{\theta H} < e_{\theta' L}/e_{\theta' H}$. If $c_{1L}/c_{1H} > c_{2L}/c_{2H} > ... > c_{nL}/c_{nH}$, other things being equal, the employer's rational belief about $e_{\theta L}/e_{\theta H}$ is $e_{1L}/e_{1H} < e_{2L}/e_{2H} < ... < e_{nL}/e_{nH}$. Based on this belief, if $m > k$, the employer hires the first k of θ who belong to M with $\pi^o_{\theta m} = 1$, whereas the rest of the types $\theta' \in M$ are not chosen, $\pi^o_{\theta' m} = 0$. Instead, if $c_{\theta L}/c_{\theta H} = c_{\theta' L}/c_{\theta' H}$, it is expected that $e_{\theta L}/e_{\theta H} = e_{\theta' L}/e_{\theta' H}$ for all $\theta, \theta' \in M$. In this case, $\pi^o_{\theta m} \in [0, 1]$. If $\pi^o_{\theta m} \neq k/m$ for some $\theta \in M$, there is *discrimination* against the applicants with equally qualifying signals.

Pre-screening at Stage 1 With the optimal strategy $\pi^o_{\theta m}$, at Stage 1, R's expected payoff from acquiring all of the signaled information is

$$E(T) = \sum_{\theta=1}^{n} w_\theta U_\theta, \tag{9}$$

where w_θ is defined in (5) and $U_\theta = qe_{\theta H}(V - W) - (1-q)e_{\theta L}W$. Compared to (2), as a result of competition, the value of signal U_θ is now weighted by w_θ. For some θ, $w_\theta < 1$. Thus, competition reduces the value of signals for the employer.

Now consider R's incentive to review only a subset Θ_P of applicants at Stage 1. The employer's expected payoff from P is

$$E(P) = \sum_{\theta \in \Theta_P} p_\theta U_\theta, \tag{10}$$

since $p_{\theta'} = 0$ for all $\theta' \notin \Theta_P$. Compared to (9), (10) shows that in the case of prejudiced selection, p_θ replaces w_θ for the weight of U_θ for the pre-selected $\theta \in \Theta_P$. The question is whether p_θ encourages the selected applicants' signaling efforts and increases U_θ more than w_θ does.

$E(P)$ increases as the U_θ of the n^* pre-selected applicants in Θ_P increases. Thus, R sets n^* to include as many applicants with a high U_θ as possible and chooses p_θ to maximize the allocation of the weight on higher U_θ. Since $U_\theta = qe_{\theta H}(V - W) - (1 - q)e_{\theta L}W$ is decreasing in $e_{\theta L}/e_{\theta H}$, other things being equal, the employer expects that $U_1 \geq U_2 \geq ... \geq U_n$ for a given belief of $e_{1L}/e_{1H} \leq e_{2L}/e_{2H} \leq ... \leq e_{nL}/e_{nH}$. Then, the employer's optimal strategy of pre-screening is

$$\begin{array}{ll} p^o_\theta = 1 & \text{if } \frac{e_{\theta L}}{e_{\theta H}} < \frac{e_{kL}}{e_{kH}} \\ p^o_\theta \in [0,1] & \text{if } \frac{e_{\theta L}}{e_{\theta H}} = \frac{e_{kL}}{e_{kH}} \\ p^o_\theta = 0 & \text{if } \frac{e_{\theta L}}{e_{\theta H}} > \frac{e_{kL}}{e_{kH}} \end{array} \tag{11}$$

Example 1. Consider a case of $k = 2$ positions and $n = 4$ applicants, $\theta \in \{1, 2, 3, 4\}$ with $c_{1L}/c_{1H} > c_{2L}/c_{2H} = c_{3L}/c_{3H} = c_{4L}/c_{4H}$. The employer R expects $e_{1L}/e_{1H} < e_{2L}/e_{2H} = e_{3L}/e_{3H} = e_{4L}/e_{4H}$. Suppose that $\theta = 1, 3, 4$ have σ_H (i.e., $m = 3$). If R acquires all signals under T, at Stage 3, R may hire 1 and 3, i.e., $\pi^o_{13} = \pi^o_{33} = 1$, and $\pi^o_{43} = 0$. On the other hand, if R chooses P at Stage 1, R may pre-select type 1 with $p^o_1 = 1$, and types 2 and 3 with $p^o_2 + p^o_3 = 1$, whereas type 4 is excluded in the review, $p^o_4 = 0$. Then, R will never be able to discover that type 4 also had σ_H. Type 4 experiences discrimination under T and prejudice under P. The difference between $\pi^o_{43} = 0$ and $p^o_4 = 0$ is that under T, she expects no chance of getting the position $\pi^o_{43} = 0$ only if there is competition (i.e., $m > 2$) and, thus, she has some incentive to signal for the cases of no competition (i.e., $m \leq 2$), whereas under P, she has no incentive to signal as she expects no chance of showing her qualifying signal, $p^o_4 = 0$, at all.

3.2.3 Equilibrium prejudice against *ex ante* identical applicants

In this section, we characterize the equilibrium under competitive signaling, especially when the applicants are *ex ante* identical in that no one has an advantage in signaling with a lower cost, i.e., $c_{\theta L}/c_{\theta H}(e) = c_{\theta' L}/c_{\theta' H}(e) = c_L/c_H(e)$ for all $\theta \neq \theta'$. The case of heterogenous applicants is analyzed in Section 4. The equilibrium depends on the applicants' self-fulfilling beliefs of $\pi_{\theta m}$ and p_θ.

Equilibrium belief of non-discriminatory $\pi_{\theta m}$ Suppose that applicants expect equal treatment at Stage 3 if R reviews all of the applications (T). Under the belief of $\pi_{\theta m} = k/m$, all applicants have an equal incentive to signal and expect the same probability of encountering competition with other applicants, $F_{-\theta Cm} = F_{-Cm}$, for all θ. Let x_E^* be the incentive under T for any θ. Then, from (6), for all θ,

$$x_E^* = w_E W, \text{ and} \tag{12}$$

$$w_E = 1 - \sum_{m=k+1}^{n} F_{-Cm}(x_E^*) \left(1 - \frac{k}{m}\right) < 1. \tag{13}$$

Let $e_E^* = (e_H(x_E^*),\, e_L(x_E^*))$ denote the efforts by the applicants at x_E^*. At e_E^*, R's expected payoff from full information T is

$$E(T; \pi^o_{\theta m} = k/m,\, e_E^*) = n w_E U_E, \text{ and} \tag{14}$$

$$U_E = q e_H(x_E^*)(V - W) - (1-q) e_L(x_E^*) W. \tag{15}$$

Instead, if P, for a pre-selected type $\theta \in \Theta_P$, the incentive to signal is $x^p_{\theta C} = p^o_\theta W$. All other types $\theta' \notin \Theta_P$ do not make an effort since $x^p_{\theta' C} = 0$. At $e_P = (e_{\theta H}(x^p_{\theta C}),$ $e_{\theta L}(x^p_{\theta C}),\, e_{\theta' H}(0) = e_{\theta' L}(0) = 0)$, R's expected payoff from P is,

$$E(P; p^o_\theta,\, e_P) = k U_p, \tag{16}$$

where $U_P = q e_{\theta H}(x^p_{\theta C})(V - W) - (1-q) e_{\theta L}(x^p_{\theta C}) W$.

Prior to characterizing the equilibrium, we need to define when competitive signaling is considered efficient. For $k < n$, competition reduces applicants' signaling incentives (Proposition 1). If signaling is efficient, signal σ_H must be valuable to the employer even after this, provided that applicants optimize their effort for a given signaling technology $c_{\theta A}(e)$. When applicants are *ex ante* identical, U_E represents the value of each applicant's competitive signal σ_H that is unaffected by any other distortions in incentives. Each applicant's signal σ_H is valuable under competition if $U_E > 0$. This is summarized by the following remark.

Remark 1 *Competitive signaling is efficient if* $U_E > 0$.

Proposition 2 *Suppose that all applicants are ex ante identical.*

1. *The employer utilizes full information* T *in equilibrium only if* $U_E > 0$.

2. If $U_E < 0$, *the unique equilibrium is prejudice.*

Prejudiced review optimizes the signaling outcomes from the pre-selected applicants. In equilibrium prejudice, $n^* = k$, and $\theta \in \Theta_P = \{1, 2, .k\}$ is pre-selected with $p^o_\theta = 1$ (see the proof of Proposition 2 in Appendix). The favored θ receives a higher signaling incentive under prejudice, i.e., $x^p_{\theta C} = W > x^*_E$. However, the employer has to bear a risk of having unqualified applicants among $\theta \in \Theta_P$ and missing a chance to find $H-$types among the unselected. Under full information T, the employer can avoid the risk as he determines hiring after observing σ_H. The only problem of is that the information is obtained at the expense of its quality because competition lowers applicants' signaling incentive (Proposition 1). Thus, the employer benefits from having more information only if the reduction in the quality of σ_H is not too severe in that $nw_E U_E > 0$ or $U_E > 0$. Proposition 2 indicates that prejudice is an outcome of low expected quality of signal σ_H under full information, the inefficiency of competitive signaling. Thus, what reduces U_E facilitates prejudice.

Corollary 1 *Prejudice is an outcome of inefficient competitive signaling.*

Proposition 3 *Consider ex ante identical applicants. Other things being equal, prejudice is more likely to occur in equilibrium if*

1. *the competition increases (*k *decreases),*

2. *the cost of signaling* $c_{\theta A}$ *increases,*

3. *the value of* $H-$*type's skill,* V*, decreases, or*

4. *the probability that an applicant is of* $H-$*type,* q*, decreases.*

Intense competition due to a low k facilitates prejudice by lowering U_E. When only few positions are available, equal opportunity for everyone means a very slim chance of getting a position for each applicant. This lowers signaling incentive x_E^* (or w_E) and may result in an insufficient effort, making σ_H not so informative to the employer, $U_E < 0$. Therefore, a decrease in k is likely to induce prejudice. Alternatively, if inefficient signaling technology makes the signaling itself highly expensive, there will be an insufficient level of effort under competition, which induces $U_E < 0$. Similarly, a low V or q promotes prejudice. A low V implies that the skill of the $H-$type is not a lot more valuable than the skill of the $L-$type, which lowers the value of signals. Hence, prejudice is more likely to be prevalent in an occupation that does not require distinctive skills.

The results in Propositions 2 and 3 imply that equal opportunity policies need to be implemented with a caution. In the current framework, equal opportunity is enforced when the employer utilizes full information T. According to our analysis, a policy that promotes equal opportunities (T) is efficient only if competitive signaling is efficient. Then, spreading out the signaling incentives to everyone does not destroy the value of signaling, i.e., $U_E > 0$. However, if signaling itself is inefficient, $U_E < 0$ and signals are not so informative that simply pushing for equal opportunities aggravates the inefficiency under T. Instead, an efficient policy would be the one that improves the value of U_E, for instance, by lowering the cost of signaling.

Belief of discriminatory $\pi^o_{\theta m}$ This section shows how the expected discrimination of *ex ante* identical applicants at Stage 3 promotes prejudice. A full description of the equilibrium in this case, however, is deferred to Section 4 after we introduce the generalized framework of prejudice.

Because there are only k positions, applicants may expect unequal treatment at Stage 3, even if they have the same qualifying signals. To see the impact of discrimination, consider a belief that R hires the k applicants by the order of θ when m applicants have σ_H. Let us divide the applicants into two groups I and J such that I is a group of favored, first i applicants, $i \leq k$, and J is a group of the rest, discriminated j applicants, $j > k$. This belief of discrimination against j type implies that $\pi^o_{im} = 1 \geq \pi^o_{jm}$. Any i type with σ_H expects to be hired for sure, whereas hiring of j type with σ_H is not guaranteed. If $m \leq k$, there is no competition among the m applicants and, thus, applicant j can be hired with $\pi^o_{jm} = 1$. However, if $m > k$,

$\pi^o_{jm} = 1$ only when not all of the qualifying signals are from more preferred types than j. In summary, for $i \leq k < j$,

$$x_{iC} = W > x_{jC} = w_j W,$$
$$\text{where } w_j = \sum_{m=k+1}^{n} F_{-jCm} I_{jmk} + \left(1 - \sum_{m=k+1}^{n} F_{-jCm}\right) < 1, \tag{17}$$

$I_{jmk} = 1$ is an indicator for the cases in which j is one of the k preferred types among the $m \in M$ applicants, and $I_{jmk} = 0$, otherwise, for $m > k$. Let $e^*_{UE} = (e_{iH},\ e_{iL},\ e^d_{jH},\ e^d_{jL})$ be the effort levels by i and j applicants when they expect full information T under this belief of discriminatory $\pi^o_{\theta m}$. Due to the chance of no competition $(1 - \sum_{m=k+1}^{n} F_{-jCm} > 0)$, $x_{jC} > 0$ for all j, implying that $e^d_{jH} > 0$ and $e^d_{jL} > 0$. The employer's expected payoff from full information is

$$E(T; \pi^o_{im}, \pi^o_{jm},\ e^*_{UE}) = kU^*_p + \sum_{j=k+1}^{n} w_j U_j(e^d_{jH}, e^d_{jL}). \tag{18}$$

U^*_p, the value of signal from favored type i, is the same as U_p under prejudice P in (16) when $x_{iC} = W$. Type i's signaling incentive x_{iC} is the same as that of the pre-selected θ under P, $x_{iC} = W = x^p_{\theta C}$. Particularly, $x_{iC} = W$ is free from competition. Thus, discrimination at Stage 3 makes the favored type i applicants immune to competition by treating them like the pre-selected ones under prejudice. Yet, by keeping the opportunities open to type j, full information with discrimination generates extra signaling values, $\sum_{j=k+1}^{n} w_j U_j(e^d_{jH}, e^d_{jL})$, as shown in (18).

Discrimination lowers the discriminated type j's signaling incentive disproportionately. Since j has a greater chance to be hired than $j+1$ in competition, it is expected that $w_j > w_{j+1} > ... > w_n$ and $U_j > U_{j+1} > ... > U_n$. Consider the least favored applicant n's problem. For n, the chance of surviving competition is zero $(I_{nmk} = 0)$ because whenever $m > k$, there are always k other, more preferred applicants. This lowers n's signaling effort. If that leads to $U_n(e^d_{nH}, e^d_{nL}) < 0$, full information cannot occur in equilibrium. This is because, whenever $U_j(e^d_{jH}, e^d_{jL}) < 0$ for some j, the employer can always do better by excluding those j applicants in the review. Thus, the employer has an incentive to deviate from T. However, if $U_n(e^d_{nH}, e^d_{nL}) > 0$, then $U_j(e^d_{jH}, e^d_{jL}) > 0$ for *all* $j = k+1, ..n$. Thus, $\sum_{j=k+1}^{n} w_j U_j(e^d_{jH}, e^d_{jL}) > 0$, and the employer is always better off using full information.

Proposition 4 *Under the belief of discrimination against ex ante identical applicants, full information is obtainable in equilibrium only if the weakest signal from the least preferred applicant n is valuable to the employer, i.e.,* $U_n(e^d_{nH}, e^d_{nL}) > 0$.

Proposition 5 *Discrimination facilitates prejudice.*

An interesting question is whether discrimination (i.e., unequal $\pi^o_{\theta m}$) promotes prejudice. While non-discrimination makes every applicant's signal equally weak, discrimination weakens the signals from j only. Without discrimination, all applicants (i or j) equally share the risk of no-selection, $1 - k/m > 0$. In contrast, with discrimination, only the discriminated j bears the risk of no-selection, $1 - \pi^o_{jm} > 0$, after obtaining σ_H. In Proposition 5, we prove that with discrimination, prejudice becomes more tempting to the employer because the weakest signal is weaker $x_{nC} < x^*_E$. The most discriminated applicant n's signal would have higher quality if there is no discrimination. Since it is the weakest signal that motivates prejudice, discrimination facilitates prejudice. The employer is more inclined to disregard the information from n compared to the case without discrimination.

Proposition 5 also implies that, in some cases, prejudice occurs simply because of discrimination. Suppose that $U_E > 0 > U_n(e^d_{nH}, e^d_{nL})$. Without discrimination, signaling is efficient to promote full information. However, just because of discriminatory belief, the employer will choose prejudice. Given that the applicants are ex ante identical, discrimination is simply a distortion in their incentives. Hence, prejudice is not due to the inefficiency in competitive signaling, but due to a distortion in the applicants' incentives from discrimination. For such a case, non-discriminatory policy can effectively reduce prejudice. Corollary 2 summarizes this.

Corollary 2 *If* $U_E > 0 > U_n(e^d_{nH}, e^d_{nL})$*, non-discrimination policy reduces prejudice.*

While $U_n(e^d_{nH}, e^d_{nL}) < 0$ is all it takes to trigger the incentive for prejudice, we need to ask whether the employer would want to get rid of competition under P in this case. This is because, from (18), for some discriminated j (who are near i), the signal still generates $U_j(e^d_{jH}, e^d_{jL}) > 0$. Since those positive signal values are attainable under competition and discrimination, the employer may wish to keep the competition within the pre-selected applicants $\theta \in \Theta_P$ after choosing P. If so, the employer will not exclude all of the discriminated applicants, implying that $n^* > k$, and will decide hiring after observing σ_H from the pre-selected n^* applicants. In the next section, we introduce this generalized framework of prejudice that allows limited competition within the pre-selected applicants under P in order to derive the equilibrium prejudice under discrimination.

4 Extension and Discussion

4.1 Generalized Prejudice

Suppose now that under P, the employer R chooses k out of the pre-selected n^* applicants after observing σ_H at Stage 3. If $m \leq k$, the m applicants with σ_H face no competition and are hired with a probability 1 at Stage 3. If $m > k$, however, the n^* pre-selected applicants still face competition under P and R selects type $\theta \in \Theta_P$ with a $\delta_{\theta m}$ probability. This structure of P is the same as the one of T shown in the Section 3.2, except that P *narrows down* the competition only within the pre-selected n^* applicants, for $n^* < n$. The rest of $n - n^*$ applicants who are excluded from the outset experience prejudice. Here we show that prejudice is expected to be more prevalent under this general type of pre-selection that allows limited competition under P.

Under P, type θA applicant's optimal effort $e^*_{\theta AP} > 0$ satisfies

$$x^*_{\theta CP} \equiv [\alpha w_\theta + (1-\alpha) p_\theta w_{\theta P}]\, W = c_{\theta A}(e^*_{\theta AP}), \tag{19}$$

$$\text{where } w_{\theta P} = \sum_{m=k+1}^{n^*} (F_{-\theta Pm} \cdot \delta_{\theta m}) + \left(1 - \sum_{m=k+1}^{n^*} F_{-\theta Pm}\right), \tag{20}$$

and $F_{-\theta Pm}$ is the probability that θ is one of the m applicants with σ_H who face competition under P. The other applicants who were excluded in the Stage 1 make no efforts. The difference between (7) and (19) is that now the incentive under P is weighted by $w_{\theta P} \leq 1$. $w_{\theta P}$ depends on the probability of surviving competition, $\delta_{\theta m}$, among n^* applicants. The determination of $\delta_{\theta m}$ is the same as $\pi^o_{\theta m}$ in Section 3.2.2.

To describe the equilibrium, suppose that applicants expect discriminatory $\pi^o_{\theta m}$ under T as in the previous section and discriminatory $\delta_{\theta m}$ under P.[8] First, the employer pre-selects n^* by eliminating any unfavored applicant l who has $U_l < 0$ under T. Given the order of discrimination, these excluded applicants will be close

[8] Equilibrium under the belief of non-discriminatory $\pi^o_{\theta m}$ is provided in the Appendix.

to n. Among the pre-selected Θ_P, the favored group I expects no competition as before, whereas the discriminated group J among Θ_P now expects *weaker* competition as a result of reduced competition among n^* applicants. In competition with $m-1$ other applicants, for any given F_{-jPm}, j expects the same chance of being selected I_{jmk}, for $m = k+1, k+2, ..., n^*$, because the employer's order of discrimination against $j \in J \subset \Theta_P$ remains the same as before, regardless of whether or not competition is among a smaller set of applicants. That is, $\delta_{im} = 1 \geq \delta_{jm} = I_{jmk}$, $i \leq k < j$. However, reduced competition to n^* increases the chance of *no competition* $\left(1 - \sum_{m=k+1}^{n^*} F_{-jPm}\right)$. Thus, pre-selected j has a higher signaling incentive under P, i.e., $w_{jP} > w_j$.

Proposition 6 *Prejudice is more likely to occur in equilibrium when competition is available for the pre-selected* n^* *applicants.*

Compared to T, under P with limited competition, the discriminated applicants among the pre-selected ones face less competition, which enhances their signaling incentive. Those discriminated applicants now receive an incentive to signal, which would not have been possible without competition under P ($n^* = k$). As the quality of signals from the discriminated enhances, R's incentive to choose P increases. In the meantime, the favored I group applicants are always guaranteed to receive the highest incentive for signaling through discrimination in selection under competition, which secures top quality from their signals. That is, this structure ensures the best quality of signal from the favored i applicants, while enhancing the positive incentive effects of competitive signaling for the discriminated applicants. Thus, prejudiced review P is more attractive to the employer with the possibility of limited competition. Put it differently, the employer can use P to optimally reduce the level of competition for a given k if n is too large. The degree of prejudice (optimal n^*) in equilibrium varies depending on the level of self-fulfilling expectation of exclusion, and $k \leq n^* < n$.

4.2 Heterogeneous Applicants

4.2.1 Intergenerational Effects of Prejudice

In this section, we introduce heterogeneity among the applicants and explain the long-term effects of prejudice. To motivate the origin of heterogeneity in a dynamic context of prejudice, consider an intergenerational wealth transfer that affects the *ex ante* identical applicants' cost of signaling. Suppose that the game in our main model is repeated for two generations, parents and children. The equilibrium derived in the previous section can be understood as the outcome of the game in the parent generation. Each type in the parent generation accumulates their wealth obtained from signaling and their children inherit the type and the wealth. Even if there is no prejudice or discrimination, shortage of available positions k results in an asymmetric wealth distribution at the end of the game. With the inherited wealth, the children of the k applicants who succeeded in signaling have access to a lower cost of signaling, unlike the other $n-k$ applicants. Without prejudice or discrimination, the selection of k is random, whereas in the presence of prejudice or discrimination, the selection is determined by θ. Prejudice and/or discrimination in the parent generation expedite and aggravate the heterogeneity in the children, creating advantages for the favored i type children. Unlike type j child, with the inherited wealth, type i child now has access to a lower cost technology in signaling. Suppose that the resulting heterogeneity is $c_{1L}(e)/c_{1H}(e) \geq c_{2L}(e)/c_{2H}(e) \geq ... \geq c_{nL}(e)/c_{nH}(e)$ for a given $e > 0$ in the beginning of the game in children's generation.

The most significant impact of the heterogeneity is to set the expectation of *discriminatory* $\pi_{\theta m}$ and $\delta_{\theta m}$. According to Proposition 5, this promotes prejudice. Given the cost disadvantage, type $j \in J$ expects an unfavorable selection at Stage 3, and this reduces an incentive to signal for j. The disadvantage in signaling technology decreases $j's$ signaling effort further. While $j's$ disadvantage in signaling cost has nothing to do with $j's$ ability H, it creates the expectation of lower quality in $j's$ signal σ_H than the signal of type i who is equal with j in every other sense, by forming the expectation of $e_{iL}/e_{iH} < e_{jL}/e_{jH}$. The expectation encourages discrimination against j, which further lowers the expected payoff from utilizing type j's signal and thus enhances the employer's incentive for prejudice against j. This shows how economic inequality derived from prejudice and discrimination reproduces and aggravates prejudice and discrimination.

To show the effect of heterogeneity in comparison with the case of homogeneous cost, suppose now that the heterogeneity in the cost takes the form of a mean-preserving spread of the homogeneous cost. That is, for a given e, the average signaling cost of heterogenous applicants is the same as the cost of homogeneous applicants. On average, type i (type j) applicant has more (less) efficient signaling technology than the applicants with homogeneous cost. For the same level of effort, type j has to incur a higher cost than what he would have had with homogeneous cost. Thus, type j's effort is lower than what would have been under homogeneous cost. As a result, the value of signal from the discriminated type j is lower. The negative impact on the disadvantaged $j \in J$ is increasingly higher as j gets closer to n. Thus, it is more likely that $U_n(e^h_{nH}, e^h_{nL}) < 0$ when the costs are heterogeneous. This triggers R's incentive for prejudice (Proposition 4). Hence, prejudice is more likely to occur when the costs of signaling differ across the applicants.

In summary, in a more realistic situation of heterogeneous applicants, prejudice is expected to be prevalent for two reasons. First, as shown in Proposition 5, expectation of discrimination at Stage 3 facilitates prejudice. Second, on average, the negative impact on the most discriminated applicant n is greater as signaling becomes more costly for the discriminated. In the context of intergenerational wealth transfer, this implies that prejudice occurred in one generation has a long-term effect on the other generations by generating long-lasting economic inequality and permanent reduction in the signaling incentives of disadvantaged types.

4.2.2 Policy Implications

When the applicants are heterogeneous, an effective policy to reduce prejudice would need to be discriminatory in the sense that a greater level of support is necessary for the group who faces more disadvantages. We use the following example to explain this policy implication.

Example 3. Suppose that there are $n = 3$ applicants, $\theta = 1, 2, 3$, $k = 2$, and $c_{1L}(e)/c_{1H}(e) > c_{2L}(e)/c_{2H}(e) > c_{3L}(e)/c_{3H}(e)$. Under T, $R's$ optimal strategy is to hire types 1 and 2 whenever they have σ_H and $U_1(W) > U_2(W) > 0$ at $x_1 = x_2 = W$. Type 3 applicant with σ_H is hired only if at least one of the two does not have σ_H. Thus, type 3 expects that $x_3 = (1 - F_1F_2)W$. Then, prejudice occurs if $U_3(x_3) < 0$, since $E(P, e^h) = U_1(W) + U_2(W) > E(T, e^h) = U_1(W) + U_2(W) + U_3(x_3)$. Consider a non-discriminatory policy that equally improves the signaling cost of all applicants. Such a policy may not be effective in preventing prejudice. As a result of increased effort, U_θ increases uniformly. However, this policy does not eliminate prejudice unless it induces $U_3 > 0$. In fact, this shows that a more efficient policy would be simply improving on type 3 applicant's signaling environment. Consider a policy that particularly improves the cost ratio for type 3 applicant so that $c_{1L}(e)/c_{1H}(e) >$

$c_{2L}(e)/c_{2H}(e) = c_{3L}(e)/c_{3H}(e)+\varepsilon$. As a result, $U_1(W)$ and $U_2(W)$ are unaffected, but with a lower cost and higher effort, U_3 has been improved. If the increase is sufficient to induce $U_3(x_3\ ) > 0$, then $E(T,\, e^h) > E(P,\, e^h)$, and thus, prejudice can be avoided.

Example 3 shows that non-prejudice policy often promotes discrimination, instead of discouraging it, highlighting the difference between non-prejudice policy and non-discrimination policy. Similarly, a discriminatory anti-discrimination policy can be more effective than a neutral anti-discrimination policy in fighting prejudice. Consider an example in the context of affirmative action. Suppose that signaling is a lot more costly for the H-type minorities than it is for the $H-$type majority. In this case, color-blind affirmative action that imposes a neutral treatment of all types does not improve efficiency, if minorities were unsuccessful in getting admitted mainly because of the adverse signaling environment.

As a result of affirmative action bans for higher education, in California, Michigan, and Texas, any race-based decision is now prohibited in the college admission process. Antonovics and Backes (2013) report that, under a system that emphasized grades and test scores, the rate of under-representation would have been twice as great, if the admission standard of the University of California system had not been adjusted to incorporate race-relevant factors. This shows that test scores (i.e., competitive signals) are not race-neutral. When test scores are more correlated with the race of the applicants than their hidden ability, the difference in test scores may merely represent the racial difference in testing environment, e.g., how easy it is for them prepare the test. The adverse signaling environment alone can sufficiently lower their incentive for signaling and lead to prejudice. If so, blind color-blind affirmative action would be inefficient. As Aristotle argued, treating "like cases alike" is required for equality and "identical treatment is not equal treatment,[..] if individuals are not similarly situated."[9]

4.3 Effectiveness of Competitive Signaling

Our analysis shows that one of the key factors that facilitate prejudice is the inefficiency of competitive signaling. When signaling itself is not so efficient, competition among the applicants can easily destroy their signaling incentive. In this case, simply insisting on equal opportunities under competitive signaling may not be optimal.

[9] "Minority Report," *The Economist* (Oct. 23, 2013)

A competitive system is effective and desirable only if it can offer meaningful opportunities to signal their ability for the applicants who would not have had the same opportunities otherwise. To get the most of the competitive system, the signals need to clearly represent the hidden quality of the applicants and nothing else, being uninfluenced by other irrelevant factors like economic hardship. However, in reality, we often observe that economic hardship significantly limits the possibility of high-quality underprivileged applicants (j types) to signal. Moreover, the economic hardship status of an applicant is often highly correlated with the index type θ of the applicant. Consequently, prejudice occurs simply because the underprivileged applicants cannot use signaling properly.

Hence, the main message of this paper is to advocate policies that enhance the quality of information under competitive signaling. The policies will have to ensure that the signals are free from the influence of all other irrelevant factors. With those policies, reducing prejudice would improve efficiency. Head Start programs or the Every Student Succeeds Act would be good examples. As these policies make the signaling environment more homogeneous for all applicants, more underprivileged applicants can participate in signaling, and any observed signaling outcomes will correctly represent the hidden quality of all applicants.

We have assumed that the cost of processing information is negligible, $s \approx 0$, for simplicity. In reality, however, the cost is not negligible.[10] When $s > 0$, reviewing only a subset of the applicants generates extra cost savings. Thus, the set of parameters for which prejudice is attractive to R will become larger. Then, in order to prevent prejudice, competitive signaling system needs to be much more efficient.

5 Conclusion

This paper models prejudice in a signaling game between an employer and n applicants who differ in an irrelevant but observable characteristic. Prejudice occurs when the employer chooses to ignore the informative signals from some applicants based on the irrelevant characteristic. In equilibrium, prejudice arises when the value of signal is low because either signaling is inefficient or discrimination distorts the sig-

[10] According to Society for Human Resource Management (SHRM), in 2011, the estimated average cost-per-hire is $4,285 for organizations with 1000 or more employees and $3,079 for organizations with fewer than 1000 employees. "Executive Brief: What Factors Influence Cost-Per-Hire?" (www.shrm.org/research/benchmarks/)

naling incentives. Signaling itself is inefficient if it is too costly for applicants to make a meaningful effort. When signaling is inefficient, competition for limited positions easily becomes the reason for insufficient incentive to signal. Even without discrimination, prejudice against identical applicants can occur if equally spreading signaling incentives across the applicants does not make any applicant's signal valuable. In this context, we find that equal opportunity policies to ensure competitive signaling may not be always efficient and can actually facilitate prejudice.

Even if signaling is efficient, however, discrimination facilitates prejudice by distorting applicants' signaling incentives. When discrimination is expected, prejudice against some of *ex ante* identical applicants occurs just because the discriminated applicants lose incentives to make sufficient effort. Considering a mechanism of intergenerational wealth transfer that determines the cost of signaling for applicants, we argue that the impact of discrimination is long-lasting and that prejudice against one type in one generation is likely to be inherited to the next generation.

Overall, an effective policy to reduce prejudice is to improve the quality of information from signaling so that the employer can be more interested in using full information. To improve the quality of information in competitive signaling, we argue the importance of policies such as Head Start programs and the Every Student Succeeds Act.

Prejudice is inefficient in this paper because, in selection, prejudice uses criteria that are unrelated to ability. Hence, even though we find some cases in which prejudice may improve the quality of information by concentrating the signaling incentives onto some favored types, the main point of this paper is that such an improvement is possible only when the competitive system is severely inefficient.

In this paper, applicants' investment on effort is purely for the purpose of signaling the innate ability, H or L. The ability itself is unaffected by the effort, however. Thus, prejudice that fundamentally depletes some applicants' signaling incentive does not affect the productive potential of an economy, in the sense that the productivity can be easily restored through a proper incentive re-alignment and policy intervention. However, if the investment can enhance the ability, as in the case of human capital investment, the long-term impact of prejudice would be much more dire. Discouraged investment effort would destroy some of the productive potential of the economy. In this case, a policy intervention to reduce prejudice would be imperative in order to maintain and enhance the productivity of the economy.

6 Appendix

1. Equilibrium in the benchmark case when $k \geq n$

Define $\underline{p_\theta}$ as the level of pre-selection probability at which $g_\theta(\underline{p_\theta}W) = \frac{q(V-W)}{(1-q)W}$ and, thus, $U_\theta(g_\theta(\underline{p_\theta}W)) = 0$. Then, as long as each of θ applicant believes that $p_\theta > \underline{p_\theta}$, $x_\theta = p_\theta W$ generates a sufficient incentive for signaling, which induces a sufficient effort, $\frac{q(V-W)}{(1-q)W} = g_\theta(\underline{p_\theta}W) > g_\theta(p_\theta W)$ and makes reviewing the application of θ worthwhile for R, $U_\theta(g_\theta(p_\theta W)) > 0$. Thus, the employer acquires all of the signaled information (T). Then, in equilibrium, $x_\theta = W$, $e_{\theta A} = c_{\theta A}^{-1}(W)$. However, if some θ expect $p_\theta \leq \underline{p_\theta}$, they do not make enough efforts and the value of signal from θ decreases, $U_\theta < 0$. Hence, the self-fulfilling expectation of a low p_θ can lead to an equilibrium of prejudice (P) in which $p_\theta = 0$ for some θ.

2. Proof of Proposition 2

Full information occurs in equilibrium if $E(T; \pi^o_{\theta m} = k/m,\, e^*_E) \geq E(P; p_\theta,\, e^*_E)$. At e^*_E, $R's$ payoff of using partial information from the pre-selected, P, and randomizing among the n^* applicants with p^o_θ, $k \leq n^* < n$, and $p^o_{\theta'} = 0$ for other θ' is

$$E(P; p^o_\theta, e^*_E) = kU_E. \tag{21}$$

From (14) and (21), if $U_E < 0$, T can never be optimal for a given e^*_E. Thus, the equilibrium necessarily involves prejudice. When $U_E > 0$, however, T is chosen in equilibrium if $nw_E > k$.

The condition $nw_E > k$ can be rewritten as $1 - k/n > \sum_{m=k+1}^{n} F_{-Cm}(1 - k/m)$. This holds because $1 - k/n \geq (1 - k/m)$ and $1 > \sum_{m=k+1}^{n} F_{-Cm}$, which induces $1 - k/n > \sum_{m=k+1}^{n} F_{-Cm}(1 - k/n) > \sum_{m=k+1}^{n} F_{-Cm}(1 - k/m)$. Thus, $E(T, \pi^o_{\theta m} = k/m, e^*_E) > E(P, p^o_\theta = k/n^*, e^*_E)$ as long as $U_E > 0$. Intuitively, this means that for the same expected effort e^*_E, $R's$ expected payoff is always greater from T than P if the signals are informative, $U_E > 0$. Thus, if $U_E > 0$, in equilibrium, R reviews all applications T expecting e^*_E and the applicants make efforts e^*_E under the expectation of T and equal treatment $\pi^o_{\theta m} = k/m$ for $k < m$, and $\pi^o_{\theta m} = 1$, otherwise.

Prejudice occurs in equilibrium if $E(P; p^o_\theta,\, e_P) \geq E(T; \pi_\theta,\, e_P)$. When the applicants expect P, for given $x^p_{\theta C} = p^o_\theta W$, $p^o_\theta = k/n^*$, their effort levels are $e_P = (e_{\theta H}(x^p_{\theta C})$, $e_{\theta L}(x^p_{\theta C})$, $e_{\theta' H}(0) = e_{\theta' L}(0) = 0)$. At e_P, R's expected payoff from P is $E(P; p^o_\theta$, $e_P) = kU_p$ as in (16). If T is chosen instead, R receives

$$E(T;\pi^o_{\theta m}=k/m,e_P)=n^*w_\theta(e_P)U_p, \tag{22}$$

$$\text{where } w_\theta(e_P)=\begin{array}{ll} 1-\sum_{m=k+1}^{n^*}F_{-Cm}(1-k/m)<1 & \text{if } n^*>k, \\ 1 & \text{if } n^*=k. \end{array} \tag{23}$$

The choice of n^* affects the level of e_P and U_p. With Assumption 1, $U_p(W)>0$ when $n^*=k$ and $p^o_\theta=1$ for all $\theta\in\Theta_P$. If $n^*>k$, it must be that $p^o_\theta<1$, which lowers U_p. Consider $\underline{p}=k/n^*<1$ at which $U_p(\underline{p}W)=0$ for a given n^* and such a $\underline{p}$ exists. Then, for a lower n^*, $p^o_\theta>\underline{p}$, and $U_p(p^o_\theta W)>0$, since U_p is an increasing function of $x^p_{\theta C}$. The employer never wants to choose a higher n^* as it will lower U_p. Then, for any such $n^*>k$, $U_p(p^o_\theta W)>0$ P is never optimal for the employer since $n^*w_\theta(e_P)>k$. On the other hand, if $n^*=k$, $w_\theta(e_p)=1$, $e_P=(e_{\theta H}(W)$, $e_{\theta L}(W)$, $e_{\theta' H}(0)=e_{\theta' L}(0)=0)$, and $E(T;\pi^o_{\theta m}=1,e_P)=E(P;p^o_\theta=1,\ e_P)$ for given $s=0$. (For any slightly positive information cost $s=\varepsilon>0$, however, P is always preferred.) Thus, prejudice occurs in equilibrium only if $n^*=k$. In this case, $x^p_{\theta C}=W>x^*_E>x^p_{\theta' C}=0$.

3. Proof of Proposition 3

Let $x_0>0$ be the level of signaling incentive at which $U_E(x_0)=0$. That is, $U_E(x_0)=qmc_H^{-1}(x_0)(V-W)-(1-q)mc_L^{-1}(x_0)W=0$. Since $U_\theta(W)>0$, this implies that $x_0<W$. Similarly, we can define an effort ratio $e_0\equiv e_{0L}/e_{0H}=mc_L^{-1}(x_0)/mc_H^{-1}(x_0)>0$ that induces $U_E(x_0)=0$. Then, $e_0=q(V-W)/(1-q)W$. The employer would find the signaling information from applicant θ valuable only if $e_\theta<e_0$, or $x_\theta>x_0$.

(1) $U_E(x^*_E)$ is increasing in $x^*_E=w_EW$. From (13), $w_E=[1-\sum_{m=k+1}^{n}F_{-Cm}(1-k/m)]$. To show the effect of k, suppose that w_E is differentiable with respect to k. Then, $dw_E/dk=\sum_{m=k+1}^{n}F_{-Cm}(1/m)-\sum_{m=k+1}^{n}(1-k/m)\left[dF_{-Cm}/dk+(F'_{-Cm})dw_E/dk\right]$. The first term $\sum_{m=k+1}^{n}F_{-Cm}(1/m)>0$ is the direct effect of increasing k. This is positive since an increase in k increases the signaling incentive by increasing the chance of selection. The second term $dF_{-Cm}/dk<0$ shows that increasing k directly lowers the probability of facing competition. The third term $(dF_{-Cm}/dw_E)dw_E/dk>0$ is the indirect effect. As the incentive for signaling increases, a higher level of effort increases the chance of having a qualifying signals from competing appli-

cants, and thus, there is a higher chance of facing competition. Then, $dw_E/dk = \frac{\sum_{m=k+1}^{n} F_{-Cm}(1/m) - \sum_{m=k+1}^{n}(1-k/m)[dF_{-Cm}/dk]}{1+\sum_{m=k+1}^{n}(1-k/m)(F'_{-Cm})} > 0$. Thus, an increase in k increases w_E and x_E^*. The same argument can be applied to the case when w_E is not differentiable.

(2) If signaling is very costly, in order to attain e_0, a high incentive is required, i.e., $x_0 \approx W$. Then, distributing the incentives across all applicants can easily lower x_E^* to the level below x_0, resulting in $e_L(x_E^*)/e_H(x_E^*) > e_0$. Thus, R does not find it worthwhile to examine all information. In this case, reducing the cost is an effective remedy as it will lower x_0, which is more likely to induce $x_E^* > x_0$, and $U_E > 0$.

(3) A lower V lowers $q(V-W)/(1-q)W$. Thus, as V decreases, it is more likely that $e_0 < e_{\theta L}(x_E^*)/e_{\theta H}(x_E^*)$ and $U_E(x_E^*) < 0$, other things being equal. Therefore, prejudice is more likely to occur for a low V.

(4) A lower q affects the signaling incentive $x_E^* = [1 - \sum_{m=k+1}^{n} F_{-Cm}(1-k/m)]W$ as well as $q(V-W)/(1-q)W$. The probability of competition $\sum_{m=k+1}^{n} F_{-Cm}$ depends on how likely other applicant θ' will have σ_H. The probability of observing σ_H from applicant θ', $F_{\theta'} = qe_{\theta' H} + (1-q)e_{\theta' L}$, is an increasing function of q. Thus, a decrease in q increases the incentive to signal x_E^* by decreasing the probability of competition and lowers $q(V-W)/(1-q)W$. Let e_{L0} be the level at which $e_{L0} = q_L(V-W)/(1-q_L)W$ for a lower $q_L < q$. Then, $e_{L0} = mc_L^{-1}(x_{L0})/mc_H^{-1}(x_{L0}) = q_L(V-W)/(1-q_L)W < e_0$. Since $\frac{e_{\theta L}}{e_{\theta H}}$ decreases as x increases, it must be that $x_{L0} > x_0$ for $q_L < q$. That is, for a low q, a higher level of incentive x_{L0} is required to get $U_E(x_{L0}) = 0$. Therefore, other things being equal, prejudice is more likely to occur for a low q.

4. Proof of Proposition 5

Without discrimination, the incentive to deviate from full information depends on the level of U_E, whereas under discrimination, it is the level of $U_n < 0$ that leads to a deviation from full information. Thus, it suffices to show that $w_E > w_n$ to show that discrimination facilitates prejudice. By definition, $w_E > w_n$ if and only if $\sum_{m=k+1}^{n} F_{-nCm} > \sum_{m=k+1}^{n} F_{-Cm}(1-\frac{k}{m})$. If other applicants' efforts are the same, for given $\mathbf{e}_{-n}$, the probability that applicant n faces competition is $F_{-nCm} = F_{-Cm}$. Thus, $F_{-nCm} > F_{-Cm}(1-\frac{k}{m})$ for any $m > k$. That is, for any given $\mathbf{e}_{-n}$, other things being equal, n expects a lower chance of being selected under discrimination. This lowers n's incentive to make efforts and increase the competitor's efforts under discrimination. That is, the competitors' efforts $\mathbf{e}'_{-n}$ under discrimination are higher than $\mathbf{e}_{-n}$ under non-discrimination. Since F is an increasing function of the efforts, the probability that n faces competition under discrimination $F_{-nCm}(\mathbf{e}'_{-n})$ is higher than that under non-discrimination $F_{-Cm}(\mathbf{e}_{-n})$. Then, the probability that the applicant n is *not* selected is higher under discrimination, $\sum_{m=k+1}^{n} F_{-nCm}(\mathbf{e}'_{-n}) > \sum_{m=k+1}^{n} F_{-Cm}(\mathbf{e}_{-n}) > \sum_{m=k+1}^{n} F_{-Cm}(\mathbf{e}_{-n})(1-\frac{k}{m})$.

5. Proof of Corollary 2

Suppose that $U_n(e^d_{nH}, e^d_{nL}) < 0$. Then, under the expectation of discrimination, full information cannot hold in equilibrium. Non-discriminatory policy establishes the expectation of equal $\pi^o_{\theta m}$. Since $U_E > 0$, full information is supported in equilibrium as a result of non-discriminatory policy. Hence, non-discrimination reduces prejudice.

When $U_n(e^d_{nH}, e^d_{nL}) > 0$, with discrimination, even though full information Pareto dominates prejudice, prejudice can arise in equilibrium as a result of self-fulfilling expectation. Since it is the expectation of prejudice that leads to the equilibrium, non-discriminatory policy does not affect the expectation. Nonetheless, since $w_E > w_n$ from Proposition 5, $U_E > U_n(e^d_{nH}, e^d_{nL}) > 0$. Non-discriminatory policy ensures that applicants expect non-discriminatory $\pi^o_{\theta m}$ at Stage 3 to induce the payoff-dominating U_E, if full information were to be expected.

6. Equilibrium: generalized prejudice under the belief of nondiscriminatory $\pi_{\theta m}$

Let x^*_E and x^*_{EP} be the levels of the incentive under T and P for all θ. $E(T; \pi^o_{\theta m} = k/m, e^*_E)$ is described in (14). In the case of P, under the expectation of $\delta_{\theta m} = k/m$, $F_{-\theta Pm} = F_{-Pm}$ for all $\theta \in \Theta_P$, $m > k$. Then, for those pre-selected applicants

$$x^*_{EP} = w_{EP} W, \text{and} \tag{24}$$

$$w_{EP} = 1 - \sum_{m=k+1}^{n^*} F_{-Pm}\left(1 - \frac{k}{m}\right) < 1. \tag{25}$$

Compared to w_E, the only difference in w_{EP} is the probability of facing competition $\sum_{m=k+1}^{n^*} F_{-Pm}$ since only n^* applicants are in consideration. For any given effort level e, other things being equal, the probability of facing competition under n^* is lower. This makes the applicants expect that $w_{EP}(e) > w_E(e)$, which increases the efforts of the n^* applicants. Let $e^*_{EP} = (e_H(x^*_{EP}),\ e_L(x^*_{EP}), 0, 0)$ denote the vector of effort levels by the applicants at x^*_{EP}. At e^*_{EP}, R's expected payoff from P is

$$E(P;\delta_{\theta m}=k/m,\, e^*_{EP}) = n^* w_{EP} U_{EP}, \tag{26}$$

$$\text{where } U_{EP} = qe_H(x^*_{EP})(V-W) - (1-q)e_L(x^*_{EP})W. \tag{27}$$

Since U is an increasing function of x^*_{EP}, $U_{EP}(x^*_{EP}) > U_E(x^*_E)$.

There are three types of equilibria. (1) P is chosen with $n^* = k$. In equilibrium, applicants correctly anticipate $n^* = k$ and make efforts at e^*_P. At e^*_P, $E(P;p_\theta, e^*_P) = kU_P = E(P;\delta_{\theta m} = k/m,\, e^*_P) = E(T;\pi_{\theta m} = k/m,\, e^*_P)$. For a slightly positive information cost s, P with $n^* = k$ is preferred by R.

(2) P is chosen with $n^* > k$. For any given n^*, in equilibrium, applicants correctly anticipate n^* and make efforts at e^*_{EP}. At e^*_{EP}, the expected payoff from P is given by (26). In equilibrium, if P is chosen, it must be that $U_{EP} > 0$. Otherwise, the employer is better off by excluding some of the n^*applicants. At e^*_{EP}, increasing n^* does not increase $E(P)$ because other applicants $\theta' \notin \Theta_P$ are not making any efforts. Given that $U_{EP} > 0$, it is an equilibrium only if lowering n^* does not improve the payoff, i.e., $n^* w_{EP} U_{EP} > (n^*-1)\widetilde{w_{EP}} U_{EP}$, where $\widetilde{w_{EP}} := 1 - \sum_{m=k+1}^{n^*-1} F_{-Pm}\left(1 - \frac{k}{m}\right)$. This holds only if $\sum_{m=k+1}^{n^*-1} F_{-Pm}\left(1-\frac{k}{m}\right) > n^* F_{-Pn^*}\left(\frac{n^*-k}{n^*}\right)$ at e^*_{EP}.

(3) T is chosen. Applicants correctly anticipate T and make efforts at e^*_E. At e^*_E, T is an equilibrium only if $U_E > 0$. When $U_E > 0$, excluding the applicant n (moving to P with competition among $n-1$ applicants) results in $(n-1)w_{EP}(e^*_E)U_E$. Thus, T is an equilibrium only if $(n-1)w_{EP} < nw_E$.

7. Proof of Proposition 6

Let $e^*_{UP} = (e^*_{iH},\, e^*_{iL},\, e^*_{jHP},\, e^*_{jLP}, 0, 0)$ be the vector of efforts by groups I, J, and L of type $A = H, L$ for $i \in I$, $j \in J$, $l \in L$, and $i \le k < j < l \le n$ that satisfy $x^*_{iCP} \equiv W = c_{iA}(e^*_{iA})$, $x^*_{jCP} \equiv w_{jP}W = c_{jA}(e^{*P}_{jA})$, and $e^*_{lHP} = e^*_{lLP} = 0$,where $w_{jP} = \sum_{m=k+1}^{n^*}(F_{-jPm}\cdot I_{jmk}) + (1 - \sum_{m=k+1}^{n^*} F_{-jPm})$. The payoff from P with limited competition for the employer is

$$
\begin{aligned}
&E(P; p^o_{im} = p^o_{jm} = 1, p^o_{lm} = 0,\, \delta^o_{im} = 1, \delta^o_{jm} = I_{jmk}, e^*_{UP}) \\
&\quad = kU^*_p + \sum_{j=k+1}^{n^*} w_j U_j(e^*_{jHP}, e^*_{jLP}) \\
&\quad = E(T; \pi^o_{im} = 1, \pi^o_{jm} = I_{jmk}, e^*_{UP}).
\end{aligned}
$$

$E(P) > E(T)$ for a slightly positive s as P incurs a less cost in obtaining signals.

Suppose that $n^* = l^*$ and for the last applicant $l^* \in \Theta_P$, $U_{l^*} > 0$ at e^*_{UP}. Then, the employer cannot do better by deviating to a lower $n^* = l^* - 1$ since $\sum_{j=k+1}^{l^*} w_j U_j(e^*_{jHP}, e^*_{jLP}) > \sum_{j=k+1}^{l^*-1} w_j U_j(e^*_{jHP}, e^*_{jLP})$ given that $w_{(l^*-1)P} > w_{l^*P}$ and $U_{l^*-1} > U_{l^*} > 0$. Deviating to a higher $n^* = l^* + 1$ does not affect the payoff since $e^*_{lHP} = e^*_{lLP} = 0$. Thus, pre-selecting $n^* = l^*$ applicants and excluding the rest ($l^* + 1$ through n) in the review is an equilibrium. In turn, e^*_{UP} is optimal for given expectation of P with $n^* = l^*$.

Let n^0 be the largest integer that allows $U_{n^0} > 0$ under P when the rest of $l > n^0$ applicants in the order are eliminated in the review. By construction, $w_{(n^0-1)P} > w_{n^0P}$ and $U_{n^0-1} > U_{n^0} > 0$. Thus, the employer has no incentive to use less information than n^0 for given $e^*_{UP} = (e^*_{iH}, e^*_{iL}, e^*_{jHP}, e^*_{jLP}, 0, 0)$, for $i \leq k < j \leq n^0 < n$. Including more than n^0 applicants does not improve $E(P)$ either. Thus, the n^0 is the largest n^* that makes $E(P) > E(T)$. Equilibrium P exists for any $k \leq n^* \leq n^0$ as a result of self-fulfilling expectation.

When $U_n(e^d_{nH}, e^d_{nL}) > 0$, equilibrium may exhibit prejudice as a result of self-fulfilling expectation of $n^* < n$. When $U_n(e^d_{nH}, e^d_{nL}) \leq 0$, full information cannot arise in equilibrium as the employer can do better by deviating to P with $n^* = n - 1$. For the equilibrium prejudice, the optimal n^* varies in the range $k \leq n^* \leq n^0$, depending on the level of self-fulfilling expectation of exclusion. Without the possibility of competition under P, equilibrium prejudice occurs at $n^* = k$. As a result of allowing competition under P, the set of parameters that induce prejudice in equilibrium becomes larger, encompassing the cases when $n^* > k$. Hence, prejudice is more prevalent with the possibility of competition under P.

References

[1] Antonovics, K. and B. Backes (2014), "The Effect of Banning Affirmative Action on College Admissions Policies and Student Quality," *Journal of Human Resources*, 49(2), pp.295-322.

[2] Arrow, K.J. (1973), *The Theory of Discrimination*, Discrimination in Labor Markets, edited by Ashenfelter, O. and A. Rees, Princeton University Press, pp.3–33.

[3] Becker, G.S. (1957), *The Economics of Discrimination*, Chicago: The University of Chicago Press.

[4] Benabou, R. (1996), "Equity and Efficiency in Human Capital Investment: The Local Connection," *Review of Economic Studies*, 63, pp.237-264.

[5] Bowles, S., G. Loury, and R. Sethi (2014), "Group Inequality," *Journal of European Economic Association*, 12(1), pp.129–152.

[6] Blume, L. (2005), "Learning and statistical discrimination," *American Economic Review,* 95(2), pp.118-121.

[7] Chaudhuri, S. and R. Sethi (2008), "Statistical Discrimination with Peer Effects: Can Integration Eliminate Negative Stereotypes?" *Review of Economic Studies*, 75, pp.579-596.

[8] Coate, S. and G.C. Loury (1993), "Will affirmative-action policies eliminate negative stereo-types?" *American Economic Review,* 83, pp.1220-1240.

[9] Durlauf, S.N. (1996), "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1, pp.75-93.

[10] Eeckhoudt, L. and C. Gollier (1995), "Demand for risky assets and the monotone probability ratio order", *Journal of Risk and Uncertainty*, 11, pp.113-122.

[11] Fang, H. and A. Moro (2011), "Theories of Statistical Discrimination and Affirmative Action: A Survey," *Handbook of Social Economics*, Vol. 1A, edited by Jess Benhabib, Matthew O. Jackson and Alberto Bisin, North-Holland, pp.133-200.

[12] Fryer, R. (2007), "Belief-Flipping in a Dynamic Model of Statistical Discrimination," *Journal of Public Economics*, 91(5), pp.1151-1166.

[13] Lang, K., M. Manove, and W.T. Dickens, (2005), "Racial Discrimination in labor markets with posted wage offers," *American Economic Review*, 95, pp.1327-1340.

[14] Mailath, G., Samuelson, L., and A. Shaked (2000), "Endogenous Inequality in Integrated Labor Markets with Two-sided Search," *American Economic Review*, 90(1), pp.46-72.

[15] Mialon, S. and S.H. Yoo (2017), "Incentives for Discrimination," *Journal of Economic Behavior and Organization*, 136, pp.141-160.

[16] Mookherjee, D., S. Napel, and D. Ray, (2010), "Aspirations, Segregation, and Occupational Choice," *Journal of the European Economic Associations*, 8, pp.139-168.

[17] Moro, A., and P. Norman (2003), "Affirmative Action in a Competitive Economy," *Journal of Public Economics, 87,* pp.567–594.

[18] Moro, A., and P. Norman (2004), "A General Equilibrium Model of Statistical Discrimination," *Journal of Economic Theory*, 114(1), pp.1–30.

[19] Phelps, E. (1972), "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62, pp.659-661.

[20] Spence, M. (1974), "Job Market Signaling," *Quarterly Journal of Economics*, 87, pp.355–374.

[21] The Society for Human Resource Management (SHRM) (2011), "Executive Brief: What Factors Influence Cost-Per-Hire?" www.shrm.org/research/benchmarks/.