Prejudice, Bias and Identity Neutral Policy*

Parimal K. Bag† Bibhas Saha‡ Shiva Sikdar§

July 20, 2020

Abstract

How does identity blind hiring, as opposed to sighted hiring, influence matching between high ability candidates and high value jobs? Job seekers might face constraints in signaling their abilities for lack of wealth and being denied education. Adding to this problem are the prejudices and biases of employers as some believe White (advantaged) applicants, on average with more wealth, contain a greater proportion of high abilities compared to Black (disadvantaged) applicants. Given such distorted beliefs, jobs/skills matching may improve or worsen under blind hiring, and there may co-exist unfilled vacancies and unemployment. The positive outcome, from job seekers' point of view, is when employers collectively view the blind pool of the uneducated applicants to have enough high ability candidates worth risking filling up their high value slots rather than leaving the positions vacant. The negative outcome occurs when the blind pool is perceived to contain a low percentage of high ability candidates, so the advantaged high ability candidates with no wealth remain unemployed while they would have been assigned to high value jobs under sighted hiring.

Finally, if education is more than a mere signal and enhances productivity in the low value jobs, unemployment will be less. But there is also a possibility that the low ability workers of the employers-favored advantaged group choose not to educate themselves even when they have the necessary wealth. Identity neutral hiring may or may not eliminate such perverse incentives.

***JEL* Classification Number:** J78, I24. **Key Words:** Prejudice, bias, ability, social identity, wealth inequality, education, job market signaling, identity neutral policy, job assignment.

*__Acknowledgements:__ This is a significant rewrite of an earlier version of the manuscript with the same title. We thank the editor-in-charge and two anonymous reviewers for helpful comments. The work has received partial funding from the Singapore Ministry of Education in the form of a Tier 1 grant. Mistakes, and especially the views on prejudice and bias, are entirely ours.

†Department of Economics, National University of Singapore, Faculty of Arts and Social Sciences, AS2 Level 6, 1 Arts Link, Singapore 117570; E-mail: ecsbpk@nus.edu.sg

‡Durham University Business School, Durham, U.K., DH1 3LB; E-mail: b.c.saha@durham.ac.uk

§Keele Business School, Keele University, Keele ST5 5BG, U.K.; E-mail: shivasikdar@gmail.com

"Should you change your name to get a job? The extreme measures some people take to fight hiring bias might surprise you. But some say it's the only way."

– Careers Job Search Psychology, BBC, September 15, 2016

(http://www.bbc.com/capital/story/20160915-should-you-change-your-name-to-get-a-job)

1 Introduction

Starting with the Civil Rights Act of 1964 prohibiting any form of discrimination in employment, legal positions of various states in USA regarding affirmative action, to expand job opportunities and university admissions for minorities, have fluctuated a great deal;[1] see the history of affirmative action policies from the 1960s (https://www.aaaed.org/aaaed/history_of_affirmative_action.asp).[2] For discrimination in jobs, not only is one's ethnicity a concern, but gender is also an issue. Discrimination against females in employment and enterprise continues to persist even now, as highlighted on May 25, 2016 by the American Economic Association.[3] Latest, in the Harvard admissions lawsuit for alleged discrimination against Asian-American applicants, a federal judge in Massachusetts ruled in favor of Harvard saying, to achieve diversity, the university authorities can use race, among other characteristics, to rank applicants.[4]

In the UK and continental Europe, affirmative action has never been favored. The UK Equality Act 2010 prohibits any form of direct or indirect discrimination, including 'positive discrimination'. Thus, UK and EU policies are color-blind, what we call *identity neutral*, for the Act's explicit requirement that employers demonstrate neutrality. In contrast, India and Malaysia have constitutional provisions for preferential treatment in education and government jobs for historically disadvantaged communities.[5]

In this paper, we focus on an important aspect of affirmative action, the effectiveness of *identity neutral* hiring vis-a-vis *identity dependent* (or sighted) hiring. Specifically, under the two alternative modes of hiring, how employers treat workers of unknown ability, while trying to assign them to jobs that strictly requires high ability, is our central concern. If some employers

[1] *Grutter v. Bollinger, 539 U.S. 306 (2003)* (https://supreme.justia.com/cases/federal/us/539/306/) had ruled in favor of affirmative action in university admissions of underrepresented minority groups, while nine states banned race-based affirmative action at all public universities (California (1996), Texas (1996), Washington (1998), Florida (1999), Michigan (2006), Nebraska (2008), Arizona (2010), New Hampshire (2012), and Oklahoma (2012)) through voter referenda; see "What Can We Learn from States That Ban Affirmative Action?" available at https://tcf.org/content/commentary/what-can-we-learn-from-states-that-ban-affirmative-action/.

[2] Sowell (2008) offers an international perspective.

[3] See https://www.aeaweb.org/research/can-strong-peer-networks-gender-gap-entrepreneurship, and in particular, the report "Gender Equality In Silicon Valley Will Take A Generation, Say Women Founders And Funders" (https://www.forbes.com/sites/clareoconnor/2016/05/13/gender-equality-in-silicon-valley-will-take-a-generation-say-women-founders-and-funders/#7cfbbb52535b).

[4] See Oct 1, 2019 New York Times article, https://www.nytimes.com/2019/10/01/us/harvard-admissions-lawsuit.html.

[5] See Deshpande (2006) and Abdullah (1997). Brazil, Canada, China, Japan and Nepal also have similar policies. Japan has policies to help the Burakumin; in China, tribals and certain ethnic communities have to meet lesser requirements than the majority of the Chinese population in education and public sector jobs. For USA, the Federal government and contractors are obligated to hire minorities, though they do not have explicit job reservations.

are prejudiced, the sense of which we make precise shortly, can the effect of their prejudice be mitigated through identity neutral/blind policy, and in addition can *positive* discrimination also be a ground for imposing blind hiring? What are the costs of ensuring equal treatment via identity neutral hiring? These are some of the questions we analyze in a model of job matching within the ability signaling framework of Spence (1975) and Hendel et al. (2005).

We consider a static model of job assignment with the number of jobs exactly equal to the number of workers. Each firm has two jobs to allocate; one job is high value, requiring a high ability worker; any mismatch will result in a significant loss to the firm. The other job is low value that requires no specific ability. The high value job is also costly to leave unfilled. A section of the job candidates, who are not wealth/credit constrained, are able to signal their *high* ability by acquiring costly education; such workers receive priority in being assigned to high value jobs. However, the number of candidates who signal their high ability are fewer than the number of high value jobs. So the employers have to decide whether they will leave their high value jobs vacant, or fill them from the pool of remaining workers, for whom they do not observe any 'education' signal and hence, cannot tell whether an uneducated worker is of high ability but lacked wealth or is of low ability. They assess the workers' likelihood of being high ability using their observable traits, such as their social (i.e., ethnic, racial or religious) identity, which we classify into two groups – advantaged and disadvantaged. The first group is more wealthy, but the two groups are identical in terms of the distribution of ability.

Employers know the proportion of the wealthy people, but may hold different beliefs about the ability distribution, in each group. This is where the notions of prejudice and bias come in. Some employers believe that a member of the disadvantaged group is *less likely* to be of high ability than she truly is; this is *prejudice* by our definition. Likewise, the same or other employers may believe that a member of the advantaged group is *more likely* to be of high ability than she truly is; this is what we call *bias*. Bias leads to favoritism toward the advantaged group and prejudice leads to discrimination against the disadvantaged group. There are also employers whose beliefs are neutral to one or both groups, i.e., same as the true probabilities. Prejudice and bias matter only when it comes to a decision of assigning an uneducated worker to an unfilled high value job. In the assignment of a known 'high ability' worker to a high value job, or an unknown ability worker to a low value job, employers' beliefs are immaterial. Hence, ours is a 'mild' form of prejudice and bias.

We depart from the existing literature on discrimination in two ways. First, while we look at the important issue of identity neutral (or color blind) hiring, we do so without any explicit redistribution goal imposed on the part of the government.[6] Second, the prejudice and bias in our model are exogenous, which we assume to originate from society and may perpetuate through employers' own experience and sharing other employers' experience. We do not try to formalize this process, as it would require a dynamic framework and fall outside the scope of the

[6]See, for instance, Fryer and Loury (2013) for an analysis of identity-blind affirmative action policies with redistribution objectives.

present paper. One may speculate that such an endeavor may take the route of the statistical discrimination model of Coate and Loury (1993), or the 'echo chamber' formulation of Levy and Razin (2017, 2019).

Prejudice and bias can be cultural (as in India's caste system), preference related (Becker, 1957), asymmetric information driven (Arrow, 1973), or perception based (Bertrand and Mullainathan, 2004; Banerjee et al., 2009). Ours can be seen as a combination of asymmetric information and misperception. In Section 2, a plausible scenario is discussed where the behavioral traits such as *confirmation bias* can perpetuate an initial misperception toward a group. In the concluding section, we also present an informal sketch of how our model can incorporate in a minimalist way the formation of beliefs. Clearly, exogenous prejudice and bias can be seen as a limitation of our model, but empirical evidence is sufficient to justify basing our policy evaluation on them.[7]

In short, we assume that employers can be of four categories – being prejudiced to the advantaged, biased to the advantaged, and neutral to one or both groups. We rule out prejudice against the advantaged and bias toward the disadvantaged, on empirical and behavioral grounds.[8]

In our model, if abilities were costlessly observable there would be no unemployment, and no jobs/vacancies would remain unfilled, rendering the employers' beliefs irrelevant. But unobservability of ability implies a strictly positive probability of unemployment for those who do not or cannot acquire education. After filling some of the high value vacancies with the available educated (and thus, revealed to be of high ability) workers, the remaining high value vacancies can only be filled by a random draw from the uneducated pool. Here, the employers' beliefs are going play to a crucial role. Those employers who are prejudiced against the disadvantaged do not fill their high value vacancies if they are matched with an uneducated disadvantaged group worker. Likewise, employers who are neutral to the advantaged group will also leave their high value vacancies unfilled if matched with an uneducated advantaged worker. This latter practice can be called *positive* discrimination. Employers who are neutral to both groups will hire only from the disadvantaged group, because their conditional belief of an uneducated worker being high type will be greater for the disadvantaged group due to their low wealth level.

Under the EU and British law neither type of discrimination – positive or negative – is allowed at the hiring stage, although minorities and disadvantaged groups can be selectively encouraged to apply. Our model shows why concerns for either type of discrimination may trigger public demands for identity neutral hiring. In reality, by hiding the candidates' ethniicity/gender information at the short-listing of interview stage (which is mandatory in UK) firms comply with the identity neutrality requirement. In the subsequent selection stage, although the employers learn the group

[7] In our defense, we embrace the not-so-uncommon perception that is discussed in the media, as noted in the introductory headline. Several important contributions have already emphasized the presence of prejudice and bias; see Bertrand and Mullainathan (2004), Banerjee et al. (2009), and some selected papers on social identity and discrimination edited by Chen and Mengel (2016).

[8] In a dynamic setting, whether these beliefs can persist or not is an important question. Reasons for persistence of wrong beliefs have been analyzed by various authors, e.g., Black (1995), Lang et al. (2005), and Fryer and Jackson (2008).

identity of the candidate, procedures are supposed to be followed to prevent discrimination. In our theoretical model, this is compressed as a single shot blind hiring. The employer draws an uneducated worker without knowing the latter's group identity and must fill the vacancy. This occurs if the former's expected belief of the latter being of high type exceeds a critical level; otherwise the employer does not draw any candidate and leaves his vacancy unfilled.

We show that, under plausible conditions, the identity blind policy reduces unemployment, but fails to eliminate it altogether. Employers who are neutral to both groups and employers who are both prejudiced and biased are most likely to respond positively to the policy. But those who are prejudiced against the disadvantaged and neutral to the advantaged will not budge at all; they will leave their vacancies unfilled. Thus, there are limitations of the identity blind policy. Furthermore, if the government wants to promote employment of the disadvantaged ahead of the advantaged out of social and economic inequality concerns, as seems to be the case in the public sector hirings of several countries (e.g., India, Malaysia, South Africa, Canada), identity blind policy will entail a loss of positive discrimination. In some circumstances, identity blind hiring can also *increase* unemployment if the extent of prejudice is deep; highly prejudiced employers may prefer to leave their vacancies unfilled and unemployment would spread among the advantaged as well.

One may argue that our results are partially affected by the assumption that education is purely a signaling device. So we extend our model to a scenario where education enhances productivity in the low value jobs.[9] Low ability workers then will have an incentive to acquire education for their own benefit, rather than to imitate the high ability types. Therefore, all the wealthy workers will acquire education, high or low, depending on their abilities. Employers, in turn, will be able to tell that lack of education is entirely due to wealth constraint. This will help avoid discrimination of the advantaged group workers (in the hands of the neutral belief employers). But the discrimination of the disadvantaged workers in the hands of the prejudiced employers would continue, although the overall rate of unemployment will be lower in this case. Identity blind policy can also eliminate unemployment fully, unlike in the case of pure signalling.

An interesting possibility arises here if the outside opportunities of the two groups are different and economic returns to low education is smaller than the reservation income. Low ability advantaged group workers may not get educated at all, while their disadvantaged group counterparts would still acquire education, because otherwise they will face a greater threat of unemployment. This asymmetric education profile may also survive under identity blind hiring as part of a perfect Bayesian equilibrium. Empirically, it may not be an unrealistic scenario to see that the curse of discrimination may motivate some ethnic groups to invest more in human capital than groups who are favored by the society at large.

Finally, some comments on the government's policy choice are in order. Our aim here is to compare the costs and benefits of two alternative hiring policies, primarily in terms of unemploy-

[9]The qualitative results do not change if we allow education to increase productivity in both jobs. We discuss this point in the relevant section.

ment and education outcomes. The distributional implications have been studied along the way as well. However, we chose not to explicitly introduce a utilitarian social welfare function to determine the policy choice. The additional complexity of welfare based analysis would be justified, if we were to recommend an *optimal* policy from a host of policies including mandated preferential hiring of the disadvantaged (as in the USA for the Federal government), and job reservations (as in India for backward castes) etc. Nonetheless, implicit in our analysis is a notion of social welfare, in which the government attaches equal weight to both groups' welfare, which are directly dependent on their expected incomes. In the relevant sections of the paper, we do comment on how the government's policy choice might be affected, if it weights two groups unequally. For example, we argue that when the identity blind policy backfires and increases unemployment, a government may still adopt an identity blind policy if it values equal treatment more than anything else; to eliminate unemployment in this circumstance, the government should offer subsidies to the firms while recommending the blind policy.

Literature review. Our paper is close in spirit to Hendel et al. (2005) who first introduced credit constraints in a Spencian education signaling model. Using a static model and then extending it to a dynamic setup, they show that an educational subsidy can hurt low income individuals, for reasons of losing out on the implicit positive discrimination. We adapt their static model by allowing for two (differentially) credit constrained groups and two different types of employers. Our focus is on the employment policy rather than the education policy.[10]

Hopkins (2012) studies ability signaling in a matching and labor market tournament. Workers signal their types to vertically differentiated firms, the latter types being observable. Under separation, assortative matching occurs and an agent's job market outcome depends on her relative ability ranking within the population. As the labor market competition increases, the higher (lower) ability types increase (decrease) investment if wages are sticky; however, with flexible wages, investment and wages fall for all abilities. In contrast to our paper, social identities of the workers do not differ in Hopkins (2012), nor are employers prejudiced/biased.

In the empirical literature, the most potent argument against an identity neutral policy is that of not creating a 'level playing field'. Even if we ignore that, there is a subtle argument of losing out on the positive discrimination that is possible only under identity-based decision making. We should also emphasize that the scope for voluntary (and rational) positive discrimination arises only in an ability signaling model, but not in a statistical discrimination model like that of Coate and Loury (1993). In Coate and Loury's model, ex post, there is nothing to infer about a worker's productivity, if she had not invested in education. She could only be given a low wage commensurate to her low productivity. But in our signaling model, she still can be offered a higher wage after being assigned to a high value job.

[10]For university admissions in the USA, Fryer et al. (2008), and Ray and Sethi (2010) highlighted some negative consequences of switching to color-blind policies from race-based affirmative action. Kojima (2012), on the other hand, shows that affirmative action can hurt minority students, the converse being that a color-blind policy can benefit them.

Before we proceed to the formal analysis, we would like to emphasize how identity neutrality can be viewed very differently depending on the context. Fryer and Loury (2013) examined if identity blind policies can achieve the same goal of improving diversity in the allocation of productive opportunities that affirmative action explicitly pursues. In our case, identity neutrality is motivated not so much to hide away from sighted policy, but to redress the ills of prejudice and bias that manifest in a sighted environment. So, while in Fryer and Loury, identity blindness is definitely more constraining (relative to sighted policy) for their redistribution objective, for us identity neutrality is expected to yield some positive benefits.

After introducing the model next, the analysis comparing the two hiring policies are contained in sections 3 and 4. Section 5 concludes. A separate Supplementary file contains the analysis of an ideal world free of all prejudice and bias.

2 Job matching model

Workers. Consider an economy with two identifiable ethnic groups, 'advantaged' and 'disadvantaged', denoted by $i = A, D$, with N_A and N_D working age members respectively, whom we call candidates or workers interchangeably. For convenience, assume $N = N_A + N_D$ to be an even number. Each candidate is born with an innate ability a_j, high or low, $j \in \{h, \ell\}$, which is going to matter in the jobs they will be assigned to by an employer. The probability of being of high ability ($j = h$) is p, *independent* of ethnic and wealth backgrounds. We assume $p > 1/2$ to suggest that the economy is fairly productive. All candidates have an outside option, e.g., self employment, that gives a basic income of y_0. But they prefer to be hired by a firm and earn a higher wage.

We assume a particular type of wealth inequality for simplification: people either have zero wealth or a substantially large wealth, $\omega \in \{0, W\}$.[11] Suppose β_i fraction of group i do not have any wealth, with $\beta_D > \beta_A$. Furthermore, $\beta_D N_D + \beta_A N_A > N/2$, i.e., there are more poor people than the rich.

Neither ability nor wealth is observable. Ability is signaled only through education, which, in turn, requires wealth, in the absence of a well-functioning credit market or government subsidy on tuition. There are two components to education cost – tuition fee, F, and effort cost, which is lower for the high ability type. Specifically, education cost is given by:

$$c(e, j) = \begin{cases} F + \frac{1}{a_j} e & \text{if} \quad e > 0, \quad j = \ell, h, \\ 0 & \text{if} \quad e = 0, \end{cases}$$

where $a_h > a_\ell > 0$. A wealthy and ability j candidate's payoff is $u_j = W - c(e, j) + w$, where w is the lifetime discounted wages he earns.[12] A wealthless candidate's utility is $u = w$. All

[11] If wealth continuously varies across individuals, multiple equilibria may arise (Hendel et al., 2005).

[12] We do not carry out a dynamic analysis of education and employment. Instead, returns to education are compressed into a single period.

candidates are risk neutral.

Firms. There are $(N_A + N_D)/2$ firms, each having a portfolio of two jobs (or tasks), τ, $\tau = H, L$ – one requiring high ability and for the other, ability does not matter. There is no strict complementarity or substitutability between the two jobs, as each job independently produces output $y(\tau, j)$, where

$$y(H, h) > y(L, \ell) = y(L, h) > y(H, \ell).$$

Matching of the H-type job with a high ability worker is most rewarding. Equally, a mismatch of a high value job with low ability yields the lowest output $y(H, \ell)$ which may necessarily be positive. When $y(H, \ell) < 0$ and no candidate of known ability h is found, a firm may leave such jobs unfilled. We make such 'no hiring' option costly for H-type jobs by imposing a 'shut-down' or 'leaving vacant' cost of $\kappa > 0$. The costs can arise due to losing corresponding customer base to rival firms. For L-type jobs, there is no cost of leaving them vacant.

That $y(L, h) = y(L, \ell) > 0$ is assumed to reinforce the intrinsically low productivity nature of the L-type job. Ability does not make any difference.

The wage policy is education based. When a worker is able to perfectly signal her ability h via education, she will be assigned to an H-type job and paid a wage of $\alpha y(H, h)$, where $\alpha < 1$ reflects rent extraction by the employer due to the 'hold up' rendered by the matching mechanism rather than a fully competitive labor market. But for an uneducated worker, wage is the minimum wage w_0. For simplicity, we set $w_0 = y(L, \ell) \geq y_0$, where y_0 is the reservation wage of workers. We assume α to be such that $\alpha y(H, h) > w_0$, so despite the hold-up, an employer cannot push the wage of a high ability candidate to a low ability candidate's wage.

■ **Job matching and hiring protocol.** As is clear, the number of total jobs equals the total number of workers $(N_A + N_D)$. Further, the total number of high ability candidates is greater than the total number of high value jobs. Yet, due to wealth inequality, the number of workers who are able to signal their high ability through education is strictly less than the number of high value jobs. So, for some high value jobs there will not be enough educated workers, resulting in matching inaccuracy. The following bound on the pool of rich, high ability workers will be assumed.

Assumption 1. $E = p[(1 - \beta_A)N_A + (1 - \beta_D)N_D] < N/2 < pN$.

If all the rich and high ability candidates signal their ability through education, then E is also the equilibrium pool of educated workers. Randomly selected E number of firms are matched with one educated worker from the above pool. Then,

$$\nu := N/2 - E = N/2 - p[(1 - \beta_A)N_A + (1 - \beta_D)N_D]$$

number of firms will have to fill their high value jobs with uneducated workers or leave these positions vacant. They will fill them provided the expected profit is no less than $-\kappa$.

We assume, however, that the *ex post* cost of a mis-match is significant:

$$y(H, \ell) - w_0 = y(H, \ell) - y(L, \ell) < -\kappa. \tag{1}$$

That is, mismatch for a high-type job drives the *ex post* loss above the cost of holding the job unfilled. Of course, *ex ante*, assigning an uneducated worker to an H-type job can still be optimal. This is where the employers' beliefs and the employment policy – sighted or blind hiring – will be important.

After the hiring decisions for H-type jobs are completed, L-type jobs are filled simply by drawing from the remaining pool. As ability does not matter for L-type jobs, random assignment or otherwise does not make a difference for such jobs.

■ **Employers' beliefs: Prejudice and bias.** Employers form their beliefs about a worker's productivity potentially through two signals. The first signal is received about the two groups – advantaged and disadvantaged, and the second signal is about an individual worker. If the individual worker is appropriately educated, the worker's ability is judged to be high, regardless of what group-specific prior the employer had.[13] If the worker is uneducated, the employer relies on group-specific belief about the worker's productivity.

The employers' group-specific beliefs are formed by an exogenously given perception, which we call *social stereotype*, and independent group signals observed by them. Social stereotypes are formed through complex social interactions. These interactions help formation of certain beliefs about the likelihood of a worker being high or low productivity depending on group identity. We assume the following stereotype.

Social stereotype. Unless evidenced otherwise (through education signals), a candidate from the disadvantaged group is believed to be less likely to be of high type than one from the advantaged group. This will be our starting point.[14]

Each employer also draws an independent signal about each group – 'negative' or 'positive'. This could be due to an employer's own labor market survey and/or his own past experience. By negative (positive) signal of a group we mean that the employer thinks the (actual) likelihood of a high ability worker is lower (higher) than the ex ante probability. The signals and the stereotype, together, determine the employers' individual beliefs, including bias and prejudice.

[13] We will mainly focus on wealth-constrained revealing equilibria. Note that we denote this a revealing equilibrium, rather than semi-separating, as those high-ability workers choosing no education are forced to do so given their wealth constraint and not as part of a mixed strategy.

[14] As already discussed in footnote 7, there is strong evidence of stereotypes. Alternatively, we could have assumed unprejudiced, unbiased employers and allow them to become prejudiced or biased through individual experience and social interactions. This formulation, as shown in the Supplementary file, yields the result that rational employers, before they become prejudiced or biased, would then discriminate positively in high value jobs favoring uneducated (or less educated) disadvantaged (say, Black) candidates over similarly educated advantaged (say, White) candidates. Clearly, such a position appears to be an unreasonable starting premise from point of view of the empirical evidence of how the job market works in the real world. Bertrand and Mullainathan (2004), Banerjee et al. (2009), and Chen and Mengel (2016), among others lend support to our hypothesis that employers are often prejudiced and biased.

Since bias and prejudice are in reference to the *neutral beliefs*, let us denote the conditional probability of an uneducated worker of group i being high ability as:

$$\rho_i = \text{Prob}(h|e = 0, i) = \frac{\beta_i p}{\beta_i p + (1 - p)} \ (< p_i), \quad i = A, D.$$

The conditional probability of an uneducated worker being low ability is $1 - \rho_i$.

If there were no social stereotype and no signals about each group, the employer's hiring decision would be guided by the above neutral beliefs. If there were signals but no stereotype, employers can deviate from the neutral beliefs in either direction. Finally, if there was only social stereotype but no signals, all employers will be prejudiced against the disadvantaged group and biased in favor of the advantaged.

In the presence of both the stereotype and signals, an employer's belief can move in any direction or remain neutral. The following rules specify how the signals interact with the stereotype.

1. If a negative signal is received about group D, the social stereotype about group D is reinforced due to the familiar notion of 'confirmation bias'.[15] The employer then believes that, in group D, the share of high ability workers is λp, where $\lambda < 1$, and the low ability workers' share is $(1-\lambda p)$. The workers' ability distribution in group D is $(\lambda p, 1-\lambda p)$. This is the case of 'prejudice'.

   The conditional belief of such prejudiced employers that an uneducated disadvantaged group worker is of high ability type is

   $$q_D = \frac{\beta_D \lambda p}{\beta_D \lambda p + 1 - \lambda p} \ (< \rho_D).$$

2. If a positive signal is received about group D, it helps to overturn the stereotype. The workers ability distribution in group D is then reset to the one followed by nature $(p, 1-p)$. The belief becomes neutral. The employers will use their neutral belief ρ_D.

3. If a negative signal is drawn about group A, the stereotype is overturned. The employer revives the neutral beliefs as $(p, 1-p)$, which is given by ρ_A.

4. If a positive signal is received about group A, the 'confirmation bias' of the stereotype kicks in. The employer then believes that the likelihood of a high ability worker in group A is δp, where $\delta > 1$. The workers' ability distribution in group A is believed to be $(\delta p, 1-\delta p)$. We call this case one of 'bias'.

   The conditional belief of a biased employer that an uneducated advantaged group worker is of high ability type is

   $$q_A = \frac{\beta_A \delta p}{\beta_A \delta p + 1 - \delta p} \ (> \rho_A).$$

[15] The experimental literature provides ample evidence on the existence of confirmation bias – see, for instance, Jones and Sugden (2001). Further, Charness and Dave (2017) show that, even in the presence of financial interests, confirmation bias exists.

Since the employers draw two signals, one for each group, there are four types of employers in society, distributed as follows.

1. Category 1: Suppose r_1 proportion of employers are prejudiced against D and neutral towards A.
2. Category 2: r_2 proportion of employers are neutral towards both groups.
3. Category 3: r_3 proportion of employers are neutral towards D and biased in favor of A.
4. Category 4: Finally, r_4 ($\equiv 1 - r_1 - r_2 - r_3$) proportion of employers are prejudiced against D and biased in favor of A.

With stereotype as the default system of beliefs, the composition (r_1, r_2, r_3, r_4) will depend on the actual draw of signals of employers. Employers steadfastly stick to their revised beliefs after observing their private signals and act on these beliefs. In fact, they can and do calculate the proportions (r_1, r_2, r_3, r_4), just like the job candidates, from the primitives, the underlying distribution of private signals. We take (r_1, r_2, r_3, r_4) as given and common knowledge in the rest of the paper, which can be justified for a large economy. The heterogenous beliefs can arise through "echo chambers" and characterize labor markets, as argued in Levy and Razin (2017).[16] Note that our assumption of a *common* and unspecified distribution of private signals is an abstraction but it gives rise to a similar heterogeneity of beliefs as in the echo chamber story of Levy and Razin (2017). We leave the distribution unspecified to allow our analysis to connect the extent and composition of prejudice and bias to policies.

3 Prejudice and bias

■ **Benchmark case - neutral beliefs.** Let us first consider the case where all employers believe that the probability of a worker being high ability is p, regardless of their group identity. We confine our discussion to a revealing equilibrium where only the (wealthy and) high ability candidates perfectly signal their talent through education.[17] All others remain uneducated.

Under the policy of sighted hiring, employers will apply their conditional probability, in accordance with the equilibrium, ρ_A or ρ_D to assign an uneducated worker to an H-type vacancy. Assign a group i uneducated worker to a vacant H-type job, if

$$E\pi = \rho_i y(H,h) + (1-\rho_i) y(H,\ell) - w_0 = \rho_i y(H,h) + (1-\rho_i) y(H,\ell) - y(L,\ell) \geq -\kappa$$

$$\text{or,} \qquad \rho_i \geq \frac{y(L,\ell) - \kappa - y(H,\ell)}{y(H,h) - y(H,\ell)} \equiv \hat{\rho}. \tag{2}$$

[16] See also Levy and Razin (2019). Because our starting employer beliefs involve both prejudices and biases, we allow the stereotypes to either persist or get partly or fully corrected. Levy and Razin (2017), on the other hand, generate polarised beliefs endogenously in a dynamic model.

[17] One can also call it a semi-separating equilibrium, because the wealth constrained individuals cannot be separated from the low ability individuals.

Since $\beta_D > \beta_A$ implies $\rho_D > \rho_A$, employers will have a preference for disadvantaged group candidates, because the wealth constraint is more severe for them and the likelihood of failing to signal high ability through education is greater for them than the advantaged group. This is a case of *positive discrimination* by employers in favor of the disadvantaged. A formal result demonstrating this is developed in the Supplementary file.

■ **Sighted hiring under prejudice and bias.** When some employers are prejudiced against the disadvantaged group and some are biased in favor of the advantaged, a number of possibilities arise as to filling the H-type vacancies. The most interesting case arises when the prejudiced employers (i.e., when categories 1 and 4) prefer to hold the H vacancy unfilled than hire a group D uneducated worker (under sighted hiring). This will happen if

$$q_D y(H,h) + (1-q_D) y(H,\ell) - y(L,\ell) < -\kappa.$$

This is where the prejudice bites. The above inequality reduces to $q_D \leq \hat{\rho}$, or equivalently,

$$\lambda < \min\{1,\ \frac{\hat{\rho}}{p[\beta_D(1-\hat{\rho})+\hat{\rho}]}\}. \tag{3}$$

That means, in these employers' *perception*, the likelihood of a high ability worker in group D (λp) has to be sufficiently small.

As for the group A workers, there are two likely scenarios: (i) $\rho_A \geq \hat{\rho}$ so that, in the neutral belief environment, group A uneducated workers were employable for the H-type jobs. The neutral belief employers (categories 1 and 2) will continue to hire them, and the biased ones (categories 3 and 4) will also do the same, but with more enthusiasm. (ii) $\rho_A < \hat{\rho}$, so that in the neutral belief environment the group A workers were not hired. Now their prospect of being hired may improve if they are matched with a biased employer. This is an interesting case and we will focus on this.

Suppose $\rho_A < \hat{\rho}$ so that employers who are neutral to group A will not assign uneducated workers from that group to H jobs. But those who are biased in favor of A, will do so if:

$$q_A y(H,h) + (1-q_A) y(H,\ell) - y(L,\ell) \geq -\kappa.$$

The above inequality reduces to $q_A \geq \hat{\rho}$, or equivalently,

$$\delta \geq \max\{1,\ \frac{\hat{\rho}}{p[\beta_A(1-\hat{\rho})+\hat{\rho}]}\}. \tag{4}$$

That is, in the biased employers' *perception* the likelihood of a high ability worker in group A (δp) has to be sufficiently higher than nature's draw (p). It is noteworthy that category 1 employers, who are neutral towards group A and prejudiced against group D, will assign neither groups' uneducated workers to H jobs. From either groups of workers they see their expected profit to be negative.

Vacancies in H slots. Note that, given the distribution of wealth and high value jobs in the

economy, $\nu = N/2 - E$ such jobs remain vacant after all educated workers are hired. The number of uneducated individuals from group i is: $U_i = N_i - p(1-\beta_i)N_i$, $i = A, D$, with $U \equiv U_A + U_D$. Hence, the probability that any employer is matched with an uneducated candidate from group i is U_i/U. Given the distribution of category of employers, without further loss of generality, we have employer of category j with $r_j\nu$ ($j = 1, 2, 3, 4$) number of H-type vacancies. The prejudice/bias of each of these employers implies that category 1 employers hold all these positions, $r_1\nu$, vacant rather than employing an uneducated candidate from either background. Category 2 employers will leave $r_2\nu U_A/U$ high value slots vacant, while category 3 employers will hire uneducated worker from either background to fill their vacant H-type positions. Finally, category 4 employers will leave $r_4\nu U_D/U$ such positions vacant. Hence, the total number of vacant high value jobs is $\nu(r_1 + r_2 U_A/U + r_4 U_D/U)$.

Proposition 1 (Prospects for the uneducated)**.** *Assume sighted hiring, and suppose* $\rho_A < \hat{\rho} < \rho_D$. *Further,* λ *and* δ *satisfy conditions* (3) *and* (4)*, respectively. Then the following results obtain in trying to fill in the high value slots with uneducated workers:*

(i) *Employers who are neutral towards group* A *and prejudiced against group* D *(category 1 employers) will not hire at all; they will leave their vacancies unfilled.*

(ii) *Employers who are neutral to group* D *and biased to group* A *will hire from either group and have their vacancies always filled.*

(iii *) Employers who are prejudiced against group* D *(categories 1 and 4 employers) will not hire from group* D*, and the employers who are biased in favor of group* A *(categories 3 and 4 employers) will hire from group* A*.*

(iv) *Total unemployment rate will be* $\nu(r_1 + r_2\frac{u_A}{U} + r_4\frac{u_D}{U})$.

This is a full manifestation of prejudice and bias under sighted hiring, hurting the disadvantaged group. What is interesting is that prejudiced/biased but otherwise rational employers disregard the more stringent wealth constraints faced by the disadvantaged when recruiting uneducated workers for high value positions. The positive discrimination that should favor the disadvantaged disappears from their hiring and gives rise to vacancies. Given this, a blind hiring protocol is likely to mitigate the adverse impacts of prejudice and bias, which we will soon examine.

Education Choice

Now we see that, compared to the benchmark case of all employers having neutral beliefs, group D workers will face some probability of unemployment (in the hands of the prejudiced employers), and group A workers will have an improved prospect of employment, because the biased employers will assign them to their vacant H jobs. Overall, both groups will face some

probability of unemployment. The expected wages of the two groups' uneducated workers are:

$$\begin{aligned} Ew_D &= \left[\frac{N/2}{U} + \frac{(N/2)-E}{U}(r_2+r_3)\right] y(L,\ell) + \left[1 - \frac{N/2}{U} - \frac{(N/2)-E}{U}(r_2+r_3)\right] y_0, \\ Ew_A &= \left[\frac{N/2}{U} + \frac{(N/2)-E}{U}(r_3+r_4)\right] y(L,\ell) + \left[1 - \frac{N/2}{U} - \frac{(N/2)-E}{U}(r_3+r_4)\right] y_0. \end{aligned}$$

It is easy to see that if $r_2 < (\geq) r_4$, $Ew_D < (\geq) Ew_A$. This implies, from the incentive compatibility conditions:

$$\alpha y(H,h) - F - \frac{e}{a_h} \geq Ew_i, \quad \text{and} \quad \alpha y(H,h) - F - \frac{e}{a_\ell} \leq Ew_i,\ i = A, D,$$

that the acquired education level of group D will be higher than the education level of group A if $r_4 > r_2$. That is, if the proportion of the employers who are both prejudiced against the disadvantaged and biased in favor of the advantaged is greater than the proportion of the employers who are neutral to both groups, the disadvantaged will have to spend more effort and acquire greater education to avoid the curse of prejudice.

The education levels of the two groups are determined by the following two sets of equations. For group D

$$\begin{aligned} e_2^D \quad | \quad \alpha y(H,h) &= F + \frac{e_2}{a_h} + Ew_D \\ e_1^D \quad | \quad \alpha y(H,h) &= F + \frac{e_1}{a_\ell} + Ew_D, \end{aligned}$$

and for group A

$$\begin{aligned} e_2^A \quad | \quad \alpha y(H,h) &= F + \frac{e_2}{a_h} + Ew_A \\ e_1^A \quad | \quad \alpha y(H,h) &= F + \frac{e_1}{a_\ell} + Ew_A. \end{aligned}$$

Since $a_h > a_\ell$ we have $e_2^i > e_1^i$ $(i = A, D)$. The equilibrium education level must be between e_1^i and e_2^i for each group i. With appropriate out-of-equilibrium belief restrictions, we can set e_2^i as the separating equilibrium education level for the high type. Furthermore, as already noted, if $r_4 > r_2$, then $Ew_D < Ew_A$, which implies that $e_2^D > e_2^A$ – a disadvantaged group high ability worker will have to acquire greater education than her counterpart in the advantaged group. The cost of signaling is higher for the disadvantaged group.

Proposition 2 (More to prove, less to gain)**.** *If $r_4 > r_2$, the uneducated from the advantaged group receives a higher expected wage than the disadvantaged group, i.e., $Ew_A > Ew_D$. Also, the disadvantaged group workers will have to acquire more education than their advantaged group counterparts.*

This proposition thus reinforces Proposition 1 demonstrating how prejudice and bias demands

more from the disadvantaged, in terms of education, and yet gives them less in return.

■ **Identity blind hiring.** How can identity blind hiring impact on the hiring practice? First, note that there will be no impact on group 3, who are biased in favor of group A and neutral towards group D. Their beliefs are q_A and q_D and both are greater than $\hat{\rho}$. Therefore, their expected profit is positive, regardless of the group they hire from. So anonymity does not affect them.

Second, group 1 employers, who are neutral towards group A workers and prejudiced against group D workers will not change their behavior either. Previously, they were not assigning any workers to H-type jobs, because $\rho_A < \hat{\rho}$ and also $q_D < \hat{\rho}$. Under identity blind hiring their expected profit is

$$E\pi = \pi_A \underbrace{[\rho_A y(H,h) + (1-\rho_A) y(H,\ell)]}_{< -\kappa + y(L,\ell)} + \pi_D \underbrace{[q_D y(H,h) + (1-q_D) y(H,\ell)]}_{< -\kappa + y(L,\ell)} - y(L,\ell)$$

Therefore, for this group, $E\pi < -\kappa$ and this group will continue to leave their vacancies unfilled.

Third, identity blind policy may have a positive effect on category 2 employers who are neutral towards both groups. We have seen in the benchmark case of no prejudice or bias that this group was *positively discriminating* in favor of the disadvantaged group. But, now, they hire from both groups if $\pi_A \rho_A + \pi_D \rho_D \geq \hat{\rho}$. In other words, positive discrimination will disappear.

Fourth, likewise for the category 4 employers, who are biased in favor of group A and prejudiced against group D, identity blind policy may have significant implications. They will stop discriminating against group D workers if $\pi_A q_A + \pi_D q_D \geq \hat{\rho}$. This condition follows from the requirement $E\pi \geq -\kappa$.

This indicates that, as compared to the sighted environment, categories 2 and 4 employers change their behavior and now hire uneducated workers to fill all their high value jobs. Hence, identity blind policy reduces unemployment and the number of vacant positions by $\nu(r_2 U_A/U + r_4 U_D/U)$ relative to the sighted hiring protocol.

While the conditions for the positive impact of the identity blind hiring policy may be met, one thing becomes clear. The behavior of group 1 employers will remain unaffected – they will leave their vacancies unfilled. This is a limitation of the identity blind policy.

Proposition 3 (Positive effect: identity blind hiring)**.** *Suppose* $\pi_A \rho_A + \pi_D \rho_D \geq \hat{\rho}$ *and* $\pi_A q_A + \pi_D q_D \geq \hat{\rho}$*, or equivalently*

$$\pi_D \geq \max\left\{\frac{\hat{\rho} - \pi_A \rho_A}{\rho_D}, \quad \frac{\hat{\rho} - \pi_A q_A}{q_D}\right\}. \tag{5}$$

Then, identity blind hiring policy leads to more vacancies being filled in, but the probability of unemployment is not driven to zero; nor can the policy prevent discrimination in all cases.

1. *Employers who are neutral towards group* A *and prejudiced against group* D *(i.e., category 1 employers) will continue to leave their* H*-type jobs* ($r_1 v$) *unfilled if they do not find educated workers.*

2. *Employers neutral towards group* D *but biased in favor of group* A *(i.e., category 3 employers) continue to hire candidates from either group.*

3. *Employers belonging to categories 2 and 4 will now hire workers from either groups. Previously, in the sighted environment category 2 employers were not hiring from group* A *and category 4 employers were not hiring from group* D*.*

4. *Total unemployment is* $r_1 v$*; hence, unemployment falls by* $\frac{v}{u}(r_2 u_A + r_4 u_D)$ *relative to the sighted hiring protocol.*

Under blind hiring, prejudiced and biased employers can no longer "punish" discriminatingly the disadvantaged, wealthless and hence uneducated workers by refusing them high value jobs. As one would expect, despite their skepticism some would now agree to risk the average ability of the uneducated and fill in their high value positions (rather than incur the cost of leaving those vacant) because on average the recruited uneducated are of sufficiently high expected abilities given their optimistic beliefs about the advantaged uneducated candidates' skills. Surprisingly, though, it is those employers who are wrong only about the disadvantaged but not about the advantaged (i.e., not biased but prejudiced) who may refuse to fill in their high value slots. This is because these employers believe the loss caused by the relatively high proportion of low ability candidates among the disadvantaged (significantly low λ), together with a high percentage of them not having wealth (i.e., β_D significantly larger than β_A), makes the pool of uneducated job applicants a high risk proposition. This is so especially because the same employers do not entertain any overoptimism about the advantaged group's high abilities (relative to neutral beliefs), denying them the buffer of a sufficient number of their uneducated applicants being high ability types. On balance, these employers will leave their high value slots unfilled. If this group of employers is significant in size, identity blind hiring may not be able to lift up the economy much from the clutches of inefficiency of sighted hiring. In contrast, under sighted hiring, the same employers would have recruited uneducated applicants from the advantaged group for their high value slots.

Here, we may add that although our main aim is to contrast the two hiring policies in terms of efficiency, we should be able to comment on the distributional aspect as well. As the identity blind policy fails to alter the 'no hiring' behavior of category 1 employers, even under the best circumstances, we can clearly see that both groups will face a positive probability of unemployment. However, as group D will have more uneducated workers (because it is less wealthy on average) the expected number of unemployed will be larger for group D. Although we did not assume any explicit social welfare function, implicit in our analysis is that the government weights the welfare of the two groups equally. If one allows a greater weight on the welfare of group D, one might argue that identity blind policy does not go very far, and may suggest other policies like direct

affirmative actions such as mandated hiring of the disadvantaged workers. Clearly, that will be outside the scope of our paper.

Proposition 3 paints a positive role of identity blind policies. However, if $\pi_A \rho_A + \pi_D \rho_D < \hat{\rho}$, category 2 employers, who are neutral towards both groups, expect the chance of an uneducated h-type candidate drawn from the entire uneducated pool to be relatively low. Thus, these employers will now leave their high value slots vacant if matched with an uneducated worker, and unemployment will increase by $r_2 \nu U_D / U$ relative to the sighted case. Similarly, category 4 employers will leave their high value slots vacant when matched with an uneducated worker if $\pi_A q_A + \pi_D q_D < \hat{\rho}$; this results in increased unemployment under identity blind hiring by $r_4 \nu U_A / U$ as compared to sighted hiring. Hence, we have:

Proposition 4 (Negative effect: identity blind hiring)**.** *Identity blind hiring can increase unemployment under certain conditions.*

1. *If* $\pi_A \rho_A + \pi_D \rho_D < \hat{\rho}$*, category* 2 *employers leave their high value slots vacant if matched with an uneducated worker and unemployment increases by* $r_2 \nu U_D / U$ *relative to the sighted case.*

2. *If* $\pi_A q_A + \pi_D q_D < \hat{\rho}$*, category* 4 *employers do not hire any uneducated worker for high value positions, thereby increasing unemployment by* $r_4 \nu U_A / U$ *under identity blind hiring policy.*

3. *If both the above conditions are satisfied, total unemployment is* $(r_1 + r_2 + r_4)\nu$*. Hence, overall unemployment increases under identity blind hiring by* $\frac{\nu}{U}(r_2 U_D + r_4 U_A)$ *relative to sighted hiring.*

In a way, as Proposition 4 shows, identity blind policy backfires, but also restores some parity in the treatment of the two groups, because group A workers will now be the victims, as much as the group D workers. From a social welfare point of view, if the government values 'equal treatment in bad times' more than the 'equal treatment in good times', then it would press on with the identity blind policy even under the circumstances specified in Proposition 4. In this case, to eliminate unemployment, the government must also offer some subsidy to induce all firms to hire. On the other hand, a government, which is weary of overall unemployment or unwilling to offer subsidy, should refrain from the identity blind policy in such circumstances.

Education choice

Suppose condition (5) holds so that category 2 and 4 employers will both respond positively and fill their vacancies under identity blind hiring. Previously under sighted hiring, category 4 employers were not hiring uneducated disadvantaged group workers and category 2 employers were not hiring uneducated advantaged workers. Only category 1 employers will leave their vacancies unfilled as before. Thus, not only the probability of employment increases for each group, but

both groups also face the same probability of employment as well as the same probability of unemployment. Therefore, their common expected wage is

$$E\tilde{w} = \left[\frac{N/2}{U} + \frac{(N/2)-E}{U}(1-r_1)\right] y(L,\ell) + \left[1 - \frac{N/2}{U} - \frac{(N/2)-E}{U}(1-r_1)\right] y_0.$$

The equilibrium education level must satisfy the following incentive compatibility conditions:

$$\begin{aligned} \alpha y(H,h) - F - \frac{e}{a_h} &\geq E\tilde{w}, \\ \alpha y(H,h) - F - \frac{e}{a_\ell} &\leq E\tilde{w}. \end{aligned}$$

Clearly, the equilibrium choice of education must be between the two bounds given by the two inequalities. Let us say the equilibrium education is given by $\tilde{e}_2$ where e_2 solves the following equation:

$$\alpha y(H,h) = F + \frac{\tilde{e}_2}{a_h} + E\tilde{w}.$$

It is easy to check that $\tilde{e}_2 < \min\{e_2^A, e_2^D\}$, as $E\tilde{w} > \max\{Ew_A, Ew_D\}$. That means when the identity blind policy is beneficial, workers from either group will have to acquire less and the same education.

Proposition 5 (Education choice under beneficial blind hiring)**.** *Blind hiring will always lead to equal education choice for either groups. Further, suppose condition (5) holds so that blind hiring reduces discrimination and unemployment. Then, the chosen level education will be unambiguously lower than the sighted hiring level. Hence, the signaling cost will be lower for the economy.*

The above Proposition spells out the most beneficial case of blind hiring. Although blind hiring does not eliminate discrimination and unemployment fully, it does reduce unemployment and signaling cost. The reason it cannot fully eliminate discrimination is that the policy cannot prevent employers from holding vacancies unfilled.

On the other hand, if $q_A n_A + q_D n_D < \hat{\rho}$ and $\rho_A n_A + \rho_D n_D < \hat{\rho}$, or equivalently,

$$n_D \leq \min\left\{\frac{\hat{\rho} - n_A \rho_A}{\rho_D}, \quad \frac{\hat{\rho} - n_A q_A}{q_D}\right\}, \tag{6}$$

only category 3 employers fill all their vacancies through identity blind hiring. All other employers leave their H-type vacancies unfilled if matched with an uneducated candidate. The latter's common expected wage is

$$E\hat{w} = \left[\frac{N/2}{U} + \frac{(N/2)-E}{U} r_3\right] y(L,\ell) + \left[1 - \frac{N/2}{U} - \frac{(N/2)-E}{U} r_3\right] y_0.$$

Note that the expected wage is strictly lower than both Ew_D and Ew_A of the sighted hiring case because category 2 and 3 employers are not hiring now. Therefore, by the analogous reasoning as before, it can be shown that there will be an adverse implication for education. Both groups of

workers will choose a significantly higher (and common) level of education than the sighted hiring level. Denote $\hat{e}_2$ as the equilibrium level of education, given by

$$\alpha y(H, h) = F + \frac{\hat{e}_2}{a_h} + E\hat{w}.$$

Since $E\hat{w} < \min\{Ew_D, Ew_A\}$, it must be that $\hat{e}_2 > \max\{e_2^D, e_2^A\}$. Hence, the signaling cost will rise under blind hiring, when the latter raises unemployment.

Proposition 6 (Education choice under detrimental blind hiring)**.** *When condition (6) holds, not only does blind hiring raise unemployment compared to the sighted hiring case, but it also raises the level of education necessary for signaling high ability.*

4 Productivity enhancing education

Now let us assume that education contributes to worker productivity. To retain partially the pure signalling aspect of education, we make the simplifying assumption $y(H, h)$, the productivity of high ability workers in high value jobs, is unaffected by education. As before, assignment of a low ability worker to an H-type job is costly, regardless of what education the worker might have acquired, i.e., $y(H, \ell) < 0$. That is, assignment of high ability workers to H-type jobs remains crucial.[18]

Education, we assume, is going to enhance the productivity of the workers in L-type jobs depending on the level of education of the worker (but not his ability).[19] That is, $y(L, e) > y(L, \ell)$ if $e > 0$. As we have assumed when $e = 0$, $y(L, \ell) = y(L, h) = w_0$, we can set $y(L, e)$ as

$$y(L, \ell, e) = y(L, \ell)\phi(e) \quad \text{where } \phi(0) = 1, \text{ and } \phi(e) > 1 \text{ for } e > 0$$
$$\text{and} \quad y(L, h, e) = y(L, h)\phi(e) = y(L, \ell)\phi(e) \quad \text{where } \phi(0) = 1, \text{ and } \phi(e) > 1 \text{ for } e > 0.$$

The assumption that education enhances productivity only in the L jobs is mainly for simplification. We could allow productivity in the H jobs to increase with education (for the h-types). However, it is easy to see that the h-type (wealthy) workers will then acquire their income maximizing level of education, or the separating level of education, whichever is higher. The fact that the h-type cannot ever escape the burden of separation (from ℓ) will always force them to choose a minimum level of education, although we will see below that separation will be less burdensome now due to the ℓ group's incentive to choose their own income maximizing education, rather than imitate the h type. Therefore, qualitatively we do not lose anything by assuming that education does not enhance productivity in the H-type jobs.

[18] A simple justification would be that for high-end jobs the role of intrinsic ability of the worker is of paramount importance and years of education, as in Spence's signalling model, makes little difference on the margin.

[19] For example, for operators in call centers in India who respond to complains and enquiries from overseas customers in the USA or UK, better fluency in speaking and writing English equips them to dispense with the calls faster, improving customer satisfaction.

Of course, we need to also ensure that an h-type will not deviate to an ℓ-type's education level and get hired for an L job. We ensure it by assuming that a worker's productivity in an L-type job will not exceed $y(H, h)$. In other words, $y(L, e) \to \bar{y}$ as $e \to \infty$ and $\bar{y} < y(H, h)$. When an educated ℓ-ability worker is assigned to an L-type job, his wage is $w_\ell = y(L, \ell) + \alpha[y(L, \ell)\{\phi(e) - 1\}] = y(L, \ell)[1 + \alpha(\phi(e) - 1)] = y(L, \ell)[1 + \tilde{\alpha}(e)]$.

■ **Key implications.** Let us first take note of the key implications of education contributing to productivity. Consider an equilibrium where all wealthy candidates acquire education. The high ability candidates acquire e_2 and low ability candidates acquire e_1, where $e_1 < e_2$. There are two noteworthy implications. First, now that the (wealthy) candidates reveal their types, no e_1-level educated candidate will be hired to H-vacancies; instead they will be employed to L-vacancies only. The H-vacancies will be filled in as before by e_2-level educated workers first, and then by drawing from the uneducated pool, based on the employer's belief. Therefore, as a possibility if there are more e_1-level educated workers than L-type jobs, then some of them will have to remain unemployed. However, we will not consider such a possibility, because then it is implausible to think of an economy where more than half of its population is 'rich'.[20]

Second, for a neutral employer the conditional probability of an uneducated worker being of h ability is same as the unconditional probability p, because the pool of uneducated workers does not contain wealthy individuals, and likewise no wealthy candidate is uneducated. Since the conditional probability is group invariant (for the neutral employers), the hiring practice, whether sighted or blind, does not matter; they yield the same outcome. We should recall that in Section 3, where education did not affect productivity, neutral employers positively discriminated under sighted hiring in favor of the disadvantaged group. The positive discrimination disappears now.

For the same reason as above, a prejudiced employer's conditional belief will be λp for group D and δp for group A.

Third, an implication of the neutral employers' conditional belief (of h type) being just p is that if all firms were neutral there will be full employment and all vacancies will be filled. To see this, denote the group e_2-level educated workers as E_2 and the group of e_1-level educated workers as E_1. The size of E_2 is $p[(1 - \beta_A)N_A + (1 - \beta_D)N_D]$. Similarly, the size of E_1 will be $(1-p)[(1-\beta_A)N_A + (1-\beta_D)N_D]$. Assuming that less than half of the population is wealthy (for each group), we must have $E_1 + E_2 < N/2$. After E_2 workers are assigned to H-type jobs and E_1 workers to L-type jobs, the remaining uneducated, $(N - E_1 - E_2)$, workers will be distributed into the vacant H-type jobs and the remaining L-type jobs. The expected profit from an H-type task assignment will be strictly positive (since $p > \rho_D > \hat{\rho}$), and therefore all H-type vacancies will be filled. Thus, neutral belief induces full employment. This is starkly different from the previous case of *positive* discrimination, where neutral employers did not hire group A uneducated workers.

[20] In this exceptional scenario, some low educated workers will be unemployed, but all uneducated workers may be employed (if the extent of prejudice is insignificant). Also, some H-vacancies will remain unfilled, for not having enough 'uneducated' workers.

Lemma 1 (Efficiency of neutral beliefs)**.** *Suppose all employers hold neutral beliefs and education enhances productivity. Then no employer will discriminate positively or otherwise. No vacancies will remain unfilled and no worker will be unemployed. Therefore, the mode of hiring – sighted or blind – does not make any difference.*

This result is starkly different from the previous case when education did not enhance productivity.

Fourth, what if some employers have **prejudice** or **bias**? Given that their conditional beliefs of an uneducated worker being h-type are λp or δp (depending on the group they come from), only the prejudiced employers (i.e., category 1 and 4 firms) will leave their H-vacancies unfilled, if they are matched with a group D worker and if $\lambda p < \hat{\rho}$. In other words, their discriminatory behavior is less likely now than in the previous case, because $\lambda p > q_D$. The biased employer's hiring of group A workers will not be affected. So, we can say that the chance of discrimination and thereby the risk of unemployment are both smaller when education enhances productivity – a sign of efficiency improvement.

Assuming $\lambda p < \hat{\rho}$, we calculate the probability of unemployment for group D workers as follows.[21] The probability of being picked up for a vacant H-slot is $\frac{N/2-E_2}{N-E_1-E_2}$ and also being matched with a category 1 or category 4 employer is $(r_1+r_4)\frac{N/2-E_2}{N-E_1-E_2}$. Thus, their expected wage is

$$Ew_D = (r_1+r_4)\frac{N/2-E_2}{N-E_1-E_2}y_0 + (1-r_1-r_4)y(L,\ell). \tag{7}$$

Now we will go into the question of education choice.

Education Choice

We will first consider the neutral belief case. Let us define e_2 as the education level of the high ability. Since there is no discrimination against any group, workers' group identity is irrelevant in the incentive or participation constraints. So we write the incentive compatibility conditions for the high ability workers as

$$\alpha y(H,h) - \frac{e_2}{a_h} - F \geq y(L,\ell)(1+\tilde{\alpha}(e_1)) - \frac{e_1}{a_h} - F \tag{8}$$

$$\alpha y(H,h) - \frac{e_2}{a_h} - F \geq y(L,\ell) \tag{9}$$

and for the low ability workers as

$$y(L,\ell)(1+\tilde{\alpha}(e_1)) - \frac{e_1}{a_\ell} - F \geq \alpha y(H,h) - \frac{e_2}{a_\ell} - F \tag{10}$$

$$y(L,\ell)(1+\tilde{\alpha}(e_1)) - \frac{e_1}{a_\ell} - F \geq y(L,\ell). \tag{11}$$

[21] All group A workers – uneducated or educated – will be employed.

The individual rationality conditions are:

$$\begin{aligned} \alpha y(H,h) - \frac{e_2}{a_h} - F &\geq y_0 \\ y(L,\ell)(1+\tilde{\alpha}(e_1)) - \frac{e_1}{a_\ell} - F &\geq y_0. \end{aligned}$$

Since $y(L,\ell) \geq y_0$, if the incentive constraints are satisfied the individual rationality constraints will also be satisfied. For the low ability type, the optimal level of education is e_1^* that maximizes $y(L,\ell)(1+\tilde{\alpha}(e_1)) - \frac{e_1}{a_\ell} - F$. Therefore, e_1^* would be given by $y(L,\ell)\tilde{\alpha}'(e_1) - \frac{1}{a_\ell} = 0$. Once we obtain e_1^*, we can set e_2^* from the equation $\frac{e_2}{a_h} = \alpha y(H,h) - y(L,\ell)(1+\tilde{\alpha}(e_1^*)) + \frac{e_1^*}{a_h}$. As $y(H,h) > y(L,\ell)(1+\tilde{\alpha}(e_1^*))$, the incentive compatible level of e_2^* will exceed e_1^*:

$$e_2^* = e_1^* + a_h\big[\underbrace{\alpha y(H,h) - y(L,\ell)(1+\tilde{\alpha}(e_1^*))}_{>0}\big]. \tag{12}$$

The incentive compatible education choices e_1^* and e_2^* are demonstrated in Fig. 1. The equilibrium e_2^* will always be to the right of e_1^*. The low type's optimal e_1^* corresponds to the peak of the $y(L,\ell)(1+\tilde{\alpha}(e_1)) - \frac{e_1}{a_\ell}$ curve which we denoted as the $y(e_1)$ curve. That $e_2^* > e_1^*$ ensures that h type will not choose e_1^*. Likewise, ℓ type will also not deviate to e_2^* because that would give less utility (or net income) than the highest utility achieved by choosing e_1^*. This is clear from Fig. 1.

An important aspect of this equilibrium is that e_2^* will be smaller than what was chosen under the pure signalling case of education (with neutral beliefs). The reason is that previously the sole purpose of choosing education for the high type was to separate from the low type, who had a strong incentive to imitate the high type due to high wage differential. But now the low type has an optimal level of education, say e_1^*, and its net income is strictly greater than the previous income w_0 (or $y(L,\ell)$). The incentive constraint of the previous case is given by the dashed line and the previous education choice (of the high type) was e_2 as noted in Fig. 1. Now the cost of signalling has fallen for the high type, because the low type has its own incentive to choose e_1^* rather than imitate the high type.

Choosing e_1^* and e_2^* must also be individually rational for the low and high types respectively. That means, by not acquiring education the workers can earn at least y_0 from self-employment or an outside sector. We show in Fig. 1 that the incentive compatibility condition for the low type is satisfied, because its highest utility strictly exceeds $y(L,\ell)$ (denoted as w_0). Since $w_0 > y_0$, the individual rationality conditions are easily satisfied.

Prejudice. While the above discussion of education is presented for the benchmark case of all firms being neutral, it can be easily seen that in the presence of prejudice and bias the education result will not change, although now for group D workers there is a positive probability of unemployment and their reservation wage (i.e., alternative to acquiring education) falls from $y(L,\ell)$ to Ew_D. This will be reflected in the modification of the second incentive constraints of

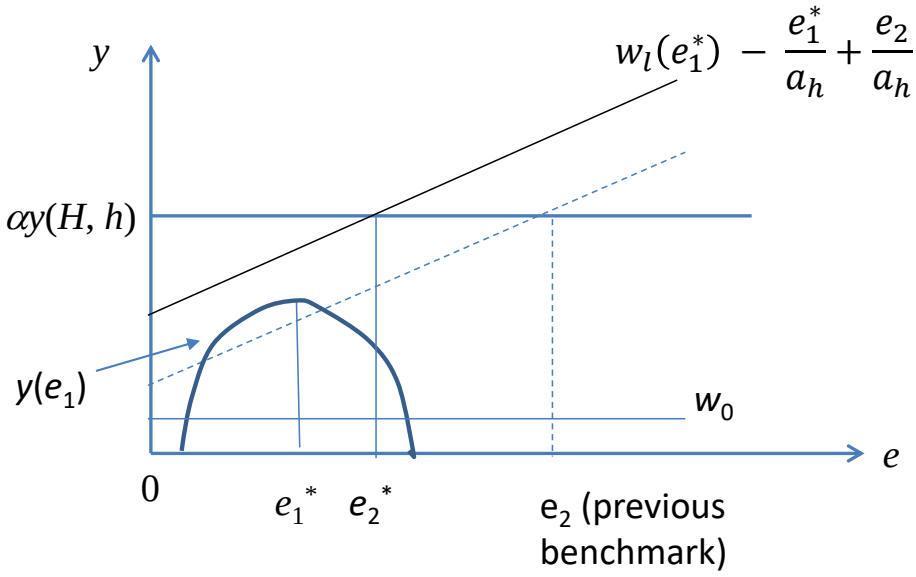

Figure 1: The case of productivity of enhancing education

group D workers, namely (9) and (11) as follows:

$$\alpha y(H, h) - \frac{e_2}{a_h} - F \geq Ew_D, \tag{13}$$

$$y(L, \ell)(1 + \tilde{\alpha}(e_1)) - \frac{e_1}{a_\ell} - F \geq Ew_D. \tag{14}$$

For group A workers the incentive constraints will not change. Since e_1^* gives a much higher payoff than $y(L, \ell)$ to a low type, and e_2^* gives even a much higher payoff to a high type worker, the incentive constraint for choosing a positive education is easily met (for either types). Therefore, the equilibrium levels of education e_1^* and e_2^* of the neutral belief case will still be the outcome despite prejudice and bias. We skip the graphical presentation of this case, but the point made is that regardless of the new obstacles of prejudice and bias, faced by 'not acquiring' education does not alter the incentives to be educated, when education is not merely a signal, but productivity enhancing at least for some jobs.

Lemma 2. *When education enhances productivity in the low tasks, the low ability wealthy candidates will acquire e_1^* and the high ability wealthy candidates will acquire e_2^* levels of education, with $e_1^* < e_2^*$. Furthermore, e_2^* is strictly less than the education level the high type candidates chose in the earlier case where education did not enhance productivity. Moreover, the education levels of both the high and low ability candidates will not vary across groups, regardless of blind or sighted hiring, and employers' beliefs.*

■ **Identity blind.** As is clear from the discussion above, identity blind hiring will not affect the education choices. But it may alter the hiring behavior of some employers, in particular the ones

prejudiced against the disadvantaged group. In other words, category 2 and category 3 employers' behavior will not change. Category 1 and category 4 employers who are prejudiced against group D may now change their hiring.

Category 1 employers will fill up the H slots, rather than leave them vacant if:[22]

$$E\pi_{G1} = \pi_D[\lambda p y(H,h)+(1-\lambda p)y(H,\ell)]+\pi_A[py(H,h)+(1-p)y(H,\ell)]-y(L,\ell) \geq -\kappa \Leftrightarrow \lambda p\pi_D+p\pi_A \geq \hat{\rho},$$

where $\pi_D = \frac{\beta_D N_D}{\beta_D N_D+\beta_A N_A}\pi_A = \frac{\beta_A N_A}{\beta_D N_D+\beta_A N_A}$. Likewise, category 4 employers will also fill up the vacancies if:

$$E\pi_{G4} = \pi_D[\lambda p y(H,h)+(1-\lambda p)y(H,\ell)]+\pi_A[\delta p y(H,h)+(1-\delta p)y(H,\ell)]-y(L,\ell)] \geq -\kappa \Leftrightarrow \lambda p\pi_D+\delta p\pi_A \geq \hat{\rho}.$$

From the two inequalities it is clear, if the above inequality is satisfied for category 1 employers, then the inequality for category 4 employers is automatically satisfied (because $\delta > 1$). Then the identity blind policy is beneficial. Alternatively, if the above inequality is violated for category 4, then the identity blind policy is ineffective for both category 4 and 1 employers.

Proposition 7. *Suppose education enhances productivity and some firms are not neutral.*

1. *The probability of unemployment is strictly positive under sighted hiring for the disadvantaged group workers if* $\lambda p < \hat{\rho}$. *But the education choice remains unaffected.*
2. *Under identity-blind hiring, there will be no unemployment if* $(\lambda\pi_D + \pi_A)p \geq \hat{\rho}$.
3. *Identity-blind hiring will not reduce the probability of unemployment, if* $(\lambda\pi_D + \delta\pi_A)p < \hat{\rho}$.

The general message is that when education enhances productivity, the signaling aspect of education becomes less important. At least some workers – low ability workers in our model – will choose individually optimal education level. Therefore, the signaling cost of the high ability type is reduced. This is generally true regardless of employers' beliefs being neutral or non-neutral. The problems of the labor market do not much affect education decisions.

As for employment, the disadvantaged group will be discriminated against by those employers who are prejudiced against them – category 1 and 4 employers. Of these two groups, category 1 employers are neutral toward group A, but prejudiced against D. If their expected probability of an uneducated worker being H type exceeds $\hat{\rho}$, then they will fill their vacancies. Category 4 employers will also fill their vacancies, because their expected belief will even be higher, as they are biased toward group A. The condition for full employment reduces to the share of the poor people of group D in the total poor population exceeding a critical level. So, in short if the poor people come mostly from group D, then the identity-blind policy is likely to be more effective.

[22] This follows since $P(h|e = 0, D) = \lambda D$ and $P(h|e = 0, A) = p$ given the beliefs of category 1 employers. The number of h and ℓ-types in group D who are uneducated are, respectively, $\beta_D N_D \lambda p$ and $\beta_D N_D(1-\lambda p)$. Hence, these employers will hire an uneducated worker under identity-blind hiring protocol if $\frac{\beta_D N_D}{\beta_A N_A+\beta_D N_D}[\lambda p y(H,h) + (1-\lambda p)y(H,\ell)] + \frac{\beta_A N_A}{\beta_A N_A+\beta_D N_D}[py(H,h) + (1-p)y(H,\ell)] - y(L,\ell) \geq -\kappa$.

■ Incentive constraints and asymmetric education

Sighted hiring. A key requirement for the low ability worker to acquire education is $y(e_1^*) \geq y(L, \ell)$. If not, the low ability worker will not find education and enhancing productivity worthwhile. We will go back to the case where only the high type gets educated. Now we consider an interesting possibility, where $Ew_D < y(e_1^*) < y(L, \ell)$. This possibility can arise if the cost of education (F) is substantial and the productivity-induced wage gain is marginal.

In this case, the equilibrium stated in Proposition 7 will not work. The disadvantaged group low ability workers will acquire education, but the advantaged group low ability workers will not. So when an advantaged group uneducated worker is matched with a category 1 or category 2 employer (both are neutral to group A) for a high value slot, he will be judged unemployable, because their conditional belief about group A workers will be given by ρ_A, and by assumption $\rho_A < \hat{\rho}$. So, for group A uneducated workers now there is a positive probability of unemployment due to *positive* discrimination. For group D uneducated workers the positive probability of unemployment is due to prejudice-induced discrimination, as before, which they can escape by acquiring education high or low depending on their ability.

So what would be the equilibrium education profile and the assignment outcome in this case? To simplify our presentation, we make an additional assumption that the advantaged group's reservation income is $y(L, \ell)$ (or equivalently the minimum wage w_0). That is, if they are unemployed they can always get $y(L, \ell)$ from their self-employment option, whereas for the disadvantaged group's reservation income is $y_0 (< y(L, \ell))$ as before. With this assumption in place, we propose that the following education profile will be part of an equilibrium: From group A the h type acquires e_2^A and ℓ type acquires zero education. From group D the h type acquires e_2^*, as given by eq. (12), and the ℓ type acquires e_1^* as defined earlier.

The above equilibrium education profile is given by the following incentive compatibility conditions. For group A the incentive constraints for h and ℓ types respectively are:

$$\begin{aligned} \alpha y(H, h) - \frac{e_2^A}{a_h} - F &= y(L, \ell), \\ \alpha y(L, \ell) &< \alpha y(H, h) - \frac{e_2^A}{a_\ell} - F. \end{aligned} \tag{15}$$

For group D the incentive compatibility conditions for the h type are

$$\begin{aligned} \alpha y(H, h) - \frac{e_2}{a_h} - F &= y(L, \ell)(1 + \tilde{\alpha}(^*e_1)) - \frac{e_1^*}{a_h} - F, \\ \alpha y(H, h) - \frac{e_2}{a_h} - F &> Ew_D, \end{aligned}$$

and for the ℓ type are

$$\begin{aligned} y(L,\ell)(1+\tilde{\alpha}(e_1^*)) - \frac{e_1^*}{a_\ell} - F &\geq \alpha y(H,h) - \frac{e_2}{a_\ell} - F, \\ y(L,\ell)(1+\tilde{\alpha}(e_1^*)) - \frac{e_1}{a_\ell} - F &\geq Ew_D. \end{aligned}$$

Thus, we have an interesting education disparity between the two groups. In the advantaged group, low ability individuals do not acquire education, because they are favored by the bias of some employers and their outside option is also attractive. On the other hand, the disadvantaged group, which faces discrimination and low outside opportunity, will acquire education to escape discrimination. Furthermore, within each group the education disparity will also be different. Within group A, e_2^A will have to be significantly as ℓ types will try to imitate the h type. Within group D, the difference between e_1^* and w_2^* would be far less. We have already noted in Lemma 2, that e_2^A will be strictly greater than e_2^*. Thus, we will have the combined education profile of the wealthy population of this economy as $\{0, e_1^*, e_2^*, e_2^A\}$.

Proposition 8 (Possibilities under sighted hiring)**.** *Assume education enhances productivity and some employers have non-neutral beliefs; further,* $Ew_D < y(e_1^*) < y(L,\ell)$*. Then in equilibrium, the low ability wealthy individuals of the advantaged group will not acquire education and face 'positive' discrimination in the hands of category 1 and 2 employers (who are neutral to group* A*) along with poor group* A *workers. On the other hand, the low ability wealthy individuals of the disadvantaged group will acquire education, and only the poor members of this group will face discrimination in the hands of the prejudiced employers (i.e., category 1 and category 4 employers). There is a positive probability of unemployment for either group.*

Identity blind hiring. Can the identity blind policy eliminate the above unemployment? The answer is yes, if we meet the right condition. But the problem is, as soon as the identity blind policy ensures full employment, the above mentioned education disparity of Proposition 8 will disappear. For the education disparity to survive, we must have unfilled vacancies by some employers strongly influenced by prejudice against the disadvantaged group. Note that the employers cannot see the identity of the candidates. But category 1 and category 4 employers hold prejudice against group D. In equilibrium, of course their updated belief of an uneducated group D to be of high type is λp. Toward group A uneducated workers, category 1 and category 4 employers' beliefs are ρ_A and q_A respectively. So for these employers to leave their vacancies unfilled, we must have

$$\lambda p n_D + \rho_A n_A < \hat{\rho} \qquad \text{(for Category 1)} \tag{16}$$

$$\lambda p n_D + q_A n_A < \hat{\rho} \qquad \text{(for Category 4).} \tag{17}$$

On the other hand, for the identity blind policy to work for category 2 and category 3 em-

ployers, who will fill their vacancies, we must have the following inequalities to hold for them:

$$p n_D + \rho_A n_A \geq \hat{\rho} \qquad \text{(for Category 2)} \tag{18}$$

$$p n_D + q_A n_A \geq \hat{\rho} \qquad \text{(for Category 3).} \tag{19}$$

It is clear that if condition (17) holds, then condition (16) will also hold, because $q_A > \rho_A$. Similarly, if condition (18) holds, then condition (19) also holds. Therefore, our asymmetric education acquisition for the low ability workers to survive under identity blind policy we must have:

$$p n_D + \rho_A n_A \geq \hat{\rho} > \lambda p n_D + q_A n_A. \tag{20}$$

The above inequality is consistent, if

$$\lambda < 1 - \frac{(q_A - \rho_A) n_A}{p n_D} \equiv 1 - \frac{\beta_A n_A (\delta - 1)}{(\beta_A \delta p + 1 - \delta p)(\beta_A p + 1 - p)}. \tag{21}$$

The second term on the right-hand side expression of the above is positive, since $\delta > 1$. But it is not *a priori* clear that it will always be a fraction. However, one can suitably choose n_D large enough (or equivalently n_A) small enough to find a range of λ. Clearly, the extent of prejudice has to be deep enough for the prejudiced employers to hold their vacancies unfilled for this equilibrium to work. While the workers of both groups face unemployment, the consequence of unemployment is asymmetric. The disadvantaged group faces a much lower expected income for not acquiring education, and this prompts them to acquire some education and secure assignment in low jobs, where their education enhances productivity and income. For the advantaged group the income from low education is not high enough to exceed what they are assured of from self-employment or blind (and probabilistic) assignment in high value jobs. Hence, they do not care to acquire education.

As for the high ability types of either group, they will acquire a high level of education. But will that be identical? It is not obvious. Suppose the policy is 'blind hiring of all employees including the educated workers'. That means, the employer can see education, but cannot ascertain the group identity of the candidate (although by our assumption no prejudice or bias will apply in that case). We can go through the incentive constraints of each group. For group A, the low ability workers will choose $e_1 = 0$ and face an assured income of $y(L, \ell)$. By deviating from $e_2 > 0$ to $e_1 = 0$ the high type will face just $y(L, \ell)$. This will require the high type of group A to choose e_2^A as given by equation (15). On the other hand, for the disadvantaged group the incentive compatible education profile is (e_1^*, e_2^*) as argued in Lemma 2. We know that $e_2^* < e_2^A$. But if the educated workers pool is also blind, then there will be a problem of attributing beliefs to three different levels of education, namely (e_1^*, e_2^*, e_2^A). Therefore, we can impose an additional restriction on the employers' out-of-equilibrium belief that if $e \in [e_1^*, e_2^A)$ the employers will believe that they are facing a low ability worker. This will prompt the high ability workers from the disadvantaged group to acquire e_2^A, the same level of education as the advantaged group workers.

Proposition 9 (Muted influence of blind hiring)**.** *Suppose condition* (20) *holds. Then the asymmetric education profile for the low ability wealthy individuals between the two groups will persist even under identity blind hiring. The advantaged group low ability type will not acquire education, and the disadvantage group low ability type will acquire education* e_1^**. The high ability individuals of either group will acquire* e_2^A *level of education.*

The above proposition highlights a 'negative' result. Prejudiced-induced unemployment and education disparity may not go away even with identity blind hiring, even if education enhances productivity. But if the blind hiring does eliminate unemployment, then of course education acquisition will be symmetric. One can set the appropriate conditions for that efficient outcome. However, the sole purpose of this section was to demonstrate the possibility of an asymmetric education outcome.

5 Conclusion

We presented a model of job matching in the presence of employer prejudice and bias, and have shown that there will always be some unemployment as a section of employers would leave their vacancies unfilled. Identity blind hiring can mitigate this problem under plausible conditions, but there are situations when identity blind hiring can also make things worse. Policy makers need to be mindful of these situations.

There are some limitations to our work. First, in countries such as India, Malaysia, South Africa, and Canada there are explicit reservations in government jobs mandating hiring a certain percentage of positions from disadvantaged group members. Our model is not appropriate for such situations. Second, we studied prejudice only at the stage of hiring. Prejudice and unequal treatments are of significant concerns for promotion and career progression as well, as is reflected in the complaints of 'glass ceiling'. Prejudice may affect firms' decisions on employee-specific investments, which can place an advantaged group employee in a favorable position in the race for promotion. Third, as discussed earlier we have taken the belief distribution of the employers as exogenous. One may argue that over time an employer who is biased toward the advantaged group will earn less profits on average than a neutral employer. This will create an inconsistency of the employer's belief with his own profit experience. Clearly, the issue needs to be resolved by making the beliefs endogenous. Below we briefly discuss one possible extension that would allow the beliefs to be formed endogenously over time.

Suppose firms are long lived and the workers live for one period. Education is a pure signal. Every period, N workers are born and their wealth and abilities are i.i.d., i.e., independently and identically distributed. The group size also remains stable over time, as does the wealth distribution. The same number of educated workers receive the high value jobs, and the remaining vacancies are filled up by drawing from the uneducated workers' pool. After the draw, the firm decides whether to hire the worker or leave the post unfilled.

Suppose initially all employers have the *neutral* belief. In the very first period, after all educated workers are hired, from the uneducated pool only group D workers will be hired to fill the remaining vacancies (positive discrimination) and all group A uneducated workers will be left unemployed. At the end of the first period, all firms including the ones who have hired only the educated workers post their experiences in a common repository. Feedback in the form of such posting will be clearly positive about group A, because only the proven high ability individuals of group A have been hired. But the feedback about group D will be mixed, ironically for the very act of the positive discrimination. Thus, from the repository, an aggregate information is generated for each group. Clearly, group A starts with only positive report and group D with mixed reports.

From the second period onward, employers consider their own experience and the aggregate information to form their beliefs about each group. Here, employers may vary in terms of how they assign relative weights to their own experience and the aggregate information. The employers who attach greater weight to their own experiences, are likely to have their beliefs eventually converge to the true probability, i.e. neutral belief, assuming they try from the pool of uneducated workers from both groups.[23] It is also important to stress that people who largely rely on their own experience will also earn higher profits on average, which in turn will reinforce reliance on their own belief. Thus, in the absence of any preference for discrimination (in the sense of Becker) such employers will uphold the neutral beliefs in the long run equilibrium of our model.

Other employers who value the aggregate information more than their own can form non-neutral beliefs and even sustain them for many periods. The reason is that the aggregate information may suffer from the initial bias we mentioned, i.e., the mixed report about group D and positive report about group A. In the extreme case, where a large number of employers engage in herding, early mistakes may be sustained for long time, if not perpetuated.

Of course, we need to be mindful that the firms whose beliefs are non-neutral will suffer in terms of profits, either by forgoing good opportunities (not hiring from group D uneducated pool) or by over-expecting from group A. Their negative experiences should eventually be reflected in the aggregate information and one can hope that such non-neutral beliefs will eventually disappear in the very long run. Until that happens, our analysis can be seen as a snapshot of the hiring in play.

By now a large literature has emerged on social learning. The rational inattention literature (see for instance, Caplin and Dean, 2015) shows that the process of learning is very complex as agents tend to rely on both own and aggregate or social information, depending on the relative costs of learning from alternative sources. There are additional cognitive issues as well as the coarseness or reliability of social information. Even the aggregation of information is a complex problem. While we wish to see the non-neutral beliefs wither away in the long run (and they probably do, as can be seen from the progresses made on gender and race issues over the last few

[23] Along with belief adjustment such employers are also likely to modify the distribution of weights with increasing reliance on individual experience.

decades), their persistence can be more than temporary.

References

Abdullah, F.H. (1997). Affirmative Action in Malaysia. *Ethnic Studies*, 15, 189-221.

Arrow, K. (1973). The Theory of Discrimination. In: *Discrimination in Labor Markets*, edited by Orley Ashenfelter and Albert Rees, Princeton University Press.

Banerjee, A., Bertrand, M., Datta, S. and Mullainathan, S. (2009). Labor Market Discrimination in Delhi: Evidence from a Field Experiment. *Journal of Comparative Economics*, 37, 14-27.

BBC (2016). "Should you change your name to get a job?..." Source: http://www.bbc.com/-capital/story/20160915-should-you-change-your-name-to-get-a-job

BBC (2017). "Is it easier to get a job if you're Adam or Mohamed?" Source: http://www.bbc.co.uk/news/uk-england-london-38751307

Becker, G. (1957). *The Economics of Discrimination*. The University of Chicago Press.

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94, 991-1013.

Black, D. (1995). Discrimination in an Equilibrium Search Model. *Journal of Labor Economics*, 13, 309-334.

Caplin, A. and Dean, M. (2015). Revealed Preference, Relative Inattention, and Costly Information Acquisition, *American Economic Review*, 105, 2183-2203.

Charness, G. and Dave, C. (2017). Confirmation Bias with Motivated Beliefs. *Games and Economic Behavior*, 104, 1-23.

Chen, Y. and Mengel, F. (2016). *Social Identity and Discrimination: Introduction to the Special Issue. European Economic Review*, 90, 1-3.

Coate, S. and Loury, G.C. (1993). Will Affirmative Action Policies Eliminate Negative Stereotypes? *American Economic Review*, 83, 1220-1240.

Deshpande, A. (2006). "Affirmative Action in India and the United States." *World Development Report* on *Equity and Development*, Background Papers.

Fryer, R.G. and Jackson, M. (2008). A Categorical Model of Cognition and Biased Decision Making. *Contributions in Theoretical Economics*, 8, Issue 1.

Fryer, R.G. and Loury, G.C. (2013). Valuing Diversity. *Journal of Political Economy*, 121, 747-774.

Fryer, R.G., Loury, G. and Yuret, T. (2008). An Economic Analysis of Color-Blind Affirmative Action. *Journal of Law, Economics and Organization*, 24, 319-355.

Goldberg, M.S. (1982). Discrimination, Nepotism and Long-Run Wage Differentials. *Quarterly Journal of Economics*, 97, 307-319.

Hendel, I., Shapiro, J. and Willen, P. (2005). Educational Opportunity and Income Inequality. *Journal of Public Economics*, 89, 841-870.

Hopkins, E. (2012). Job Market Signalling of Relative Position, or Becker Married to Spence. *Journal of the European Economic Association*, 10, 290-322.

Jackson, M.O. (2009). Social Structure, Segregation, and Economic Behavior. 2007 Nancy Schwartz Memorial Lecture.

Jones, M. and Sugden, R. (2001). Positive Confirmation Bias in the Acquisition of Information. *Theory and Decision*, 50, 59-99.

Kojima, F. (2012). School Choice: Impossibilities for Affirmative Action. *Games Economic Behavior*, 75, 685-693.

Krishna, K. and Robles, V.F. (2016). Affirmative Action in Higher Education in India: Targeting, Catch Up and Mismatch. Working paper, Department of Economics, Penn State University (http://grizzly.la.psu.edu/~kkrishna/Papers/).

Lang, K., Manove, M. and Dickens, W. (2005). Racial Discrimination in Labor Markets with Posted Wage Offers. *American Economic Review*, 95, 1327-1340.

Levy, G. and Razin, R. (2017). The Co-evolution of Segregation, Polarised Beliefs and Discrimination: The Case of Private vs. State Education. *American Economic Journal: Microeconomics*, 9, 141-170.

Levy, G. and Razin, R. (2019). Echo Chambers and their Effects on Economic and Political Outcomes. *Annual Review of Economics*, 11, 303-328.

The New York Times (2019). *Harvard does not discriminate against Asian-Americans in admissions, Judge rules.* Retrieved 8 October 2019, from

https://www.nytimes.com/2019/10/01/us/harvard-admissions-lawsuit.html.

Ray, D. and Sethi, R. (2010). A Remark on Color-Blind Affirmative Action. *Journal of Public Economic Theory*, 12, 399-406.

Sowell, T. (2004). *Affirmative Action Around the World: An Empirical Study*, Yale University Press, New Haven and London.

Spence, A.M. (1973). Job Market Signaling. *Quarterly Journal of Economics*, 87, 355-374.