



HAL
open science

Holdup and hiring discrimination with search friction

Sheng Bi, Yuanyuan Li

► **To cite this version:**

Sheng Bi, Yuanyuan Li. Holdup and hiring discrimination with search friction. 2016. halshs-01277548

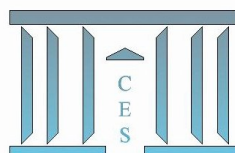
HAL Id: halshs-01277548

<https://shs.hal.science/halshs-01277548>

Submitted on 23 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Holdup and hiring discrimination with search friction

Sheng Bi, Yuanyuan Li

2016.02



Holdup and hiring discrimination with search friction*

BI Sheng[†], LI Yuanyuan[‡]

30th, December

Abstract

A holdup problem on workers' skill investment can arise when employers adopt discriminatory hiring norm to extract higher than socially optimal profit. When hiring priority is determined by both productivity-dependent (skill level) and -independent characteristics (discrimination), skill investment decision becomes strategic between the discriminated and favored group. We consider frictional markets with either posted or bargained wage (fixed sharing rule). With posted wage, depending on market tightness there may be equilibrium or multiple equilibria on skill investment. With discriminatory hiring, if in equilibrium both groups stay high skilled, both are worse off and firms better off; In any equilibrium where one group underinvest, the other group remain high skilled and are better off, while firms are worse off with discrimination. With bargained wage, similar equilibrium where the favored group underinvest exists, and firms incur cost for an intermediate range of bargaining power when they discriminate.

JEL Classifications: J7, J42

Key words: Discrimination; Directed Search; Pre-matching Investment

*We are very grateful for the discussions and encouragement from Professor Bernhard Eckwert, Professor François Langot, Professor Anna Zaharieva, Professor Bertrand Wignolle, Nikolai Brandt, Dr. Andreas Szcztkowski, as well as participants in BIGSEM Colloquium, BIGSEM Workshop, EDEEM Annual meeting, and GAINS seminar. Yuanyuan LI would like to thank the EDEEM program for the financial support, and BI would like to thank for the financial sponsorship from University of Paris 1 and University of Bielefeld, and Deutsch-Französische Hochschule.

[†]University of Paris 1, and University of Bielefeld; E-mail: brbsheng@gmail.com

[‡]University of Bielefeld, and University of Paris 1

1 Introduction

A holdup problem arises when some investment is sunk ex ante by one party, and the payoff is shared with that one party's trading partner. Since cost has no other use once sunk, that trading partner will have every incentive to squeeze the profit at the ex post stage. In an important study on such a problem in a labor market with search friction, Acemoglu and Shimer (1999b) show that with firms' sinking capital and ex post wage bargaining, the equilibrium is always inefficient, since wages paid ex post can be so high such that firms' ex ante incentive of investment is harmed; while if firms are able to post wages to direct workers' search, then the holdup problem to firms' investment no longer appears; the efficiency can be achieved, because wage posting allows workers to observe offers and choose where to apply, and it induces workers to optimize their expected payoff from application by making trade-off between every wage they observe and the probability of obtaining it. Within conventional wage posting framework, we spot another source of inefficiency in a holdup problem where workers sink skill investment cost: when the market is crowded for the firms, by adopting a discriminatory hiring norm firms are able to expropriate higher than socially optimal level of profit, and this has the consequence of discouraging the investment incentives for both the favored and discriminated groups. We analyze the impact of such rent seeking behavior of firms on the structure of market segmentation, and on the workers' skill investment incentives.

When discrimination is absent, the wage posting economy with workers' ex ante skill investment attains efficiency in the equilibria, and we show which equilibrium emerges depends on the rivalry between the log return to skill and the market tightness (workers/firms ratio) which measures the degree of market competition. The fundamental reason behind this efficiency result is that skill achievement is a quality which can be legally written into the wage contracts. It is a different story when other (binary) characteristics which are not closely related to productivity, such as gender, race, height, origin etc. enter also into firms' preference. Under equal pay legislation, posted wages can not be conditioned explicitly on these characteristics; however, if firms still select workers according to their preference on these characteristics, a separating equilibrium can result where separate firms post different levels of wages, and workers of different groups sort themselves and apply to different wages: the market is then endogenously segregated. On the side of firms, they have incentive to adopt such discriminatory hiring norm, when workers' return to skill investment is sufficiently high; in that case discrimination allows them to grasp higher than the socially optimal level of operating profit. On the side of the workers, it proves that both the discriminated group and favored group are worse off: for the former, it is because discrimination discretely reduces the labor market opportunity of these workers, who anticipate discrimination, then demand lower wages, which makes them cheaper to hire; for the latter, it is so because when firms are able to hire the discriminated workers cheaply, it is as if firms enjoy larger "market power", which allows them to suppress further the undiscriminated workers' expected payoff. Naturally, anticipating discrimination, all groups expect lower payoff from search, jeopardizing their skill investment incentives.

A key feature of our study is the multidimensionality of characteristics based on which workers are ranked. On one hand, there is ranking by productivity-dependent type identity: workers are either high skilled (type H) or low skilled (type L); high skilled have priority to low skilled simply because such ranking gives firms higher profit. On the other hand, there is ranking by productivity-independent group identity: workers belong either to the favored (group a) or the discriminated group (group b). The resulting ranking schedule has the following order: $aH \succ bH \succ aL \succ bL$. It reads: given any skill level, group a are preferred to group b ; the high skilled are always preferred to low skilled. Under such an "intertwined" ranking order, the skill investment decision for different groups becomes strategically interdependent. Focusing on Nash

pure strategy equilibrium on skill investment, in the wage posting economy, we find that depending on the value of market tightness there can be equilibrium or multiple equilibria on skill investment due to that interdependence. Compared to the case without discrimination, when the market is very crowded (market tightness is small) for the firms, discrimination is profitable for firms and all the workers are worse off; as the tightness further increases, both group can choose low skill and in equilibrium whenever one group underinvest, the other group remain high skilled and are better off, while the firms are worse off with discrimination. In particular, the equilibrium where the favored group underinvest, while the discriminated group choose to remain high skilled exists; And in this case firms' profits drop since workers' underinvestment in skill leads to lower average productivity in the economy compared to the case where discrimination is absent.

In the economy where wages are bargained (determined according to a fixed sharing rule) after matching hence do not direct search, we find similar equilibrium where the favored group underinvest, hence earn lower expected payoff compared to the case without discrimination within a certain region of bargaining power; in such an equilibrium, surplus is transferred from firms and favored group to discriminated group. Firms' profits are piecewise monotone, because increase of workers' bargaining power can increase workers' incentive of skill investment, hence discretely improves the market skill composition and average productivity. We also find that there is an intermediate range of workers' bargaining power for values of which firms are worse off by discriminating, due to discouraged skill investment from discriminated group. All in all, the key difference between wage posting and wage bargaining is that the actual wage now exogenously pegs on the productivity, and firms can no longer manipulate their market power by translating their discriminatory preference into constantly lower wages.

1.1 Relation to the literature

Job search process is an important channel through which discrimination keeps functioning in the labor market. Several papers have highlighted the impact of discrimination through job search channel to the wages gaps. To name a few, Pendakur and Woodcock (2010) show that the existent glass ceilings for the immigrant and minority workers may be attributed by large measure to their poor access to the jobs in high-wage firms; As well, in an important article from Ritter and Taylor (2011), they show that most of the disparity in unemployment rate could not be explained by cognitive skills that emerge at an early stage, although for wage gap it could be the case. This result concerning the unemployment disparity is confirmed by the finding that this disparity is still significant even for workers of similar skill levels.

Our work is most closely related to the directed search literature¹. In this literature, search frictions are derived endogenously through agents' sequential strategic interactions. Taking into account strategic interaction allows search externality to be internalized. The resulting economy remains competitive, albeit with a non-Walrasian market structure, and prices play an allocative role to achieve efficiency. To the best of our knowledge, among the discrimination literature with search friction, only two of them are built upon wage posting context. Lang, Manove, and Dickens (2005, hereafter LMD) show that a discriminatory hiring rule could lead to labor market segmentation and significant wage gap with even a negligible difference in productivity; however, the discriminated group turn out to have lower unemployment rate, which is in sharp contrast with evidence. Merlino (2012) aims at improving the result of LMD (2005). He considers further the

¹This literature is sometimes also termed as wage posting game with coordination friction

pre-matching investment from the firms' side, and obtain technology dispersion and realistic unemployment gap. His results rely on the strong assumption that there is more discrimination in the high technology sector, and he is silent on the workers' skill levels. Our paper differs from theirs, in that our focus is to analyze how hiring discrimination could distort workers' skill investment incentives and the structure of market segmentation.

While the setup of wage bargaining (no information of level of wage before matching) is more prevalent, it neglects an important trade-off that the workers make to some extent in their search for jobs: the wage and the probability of obtaining it. This endogenous link between wage and employment probability is especially important, since wages convey information on whether the employers discriminate. Having information of wages available before matching, workers are able to adjust accordingly their search strategy to avoid being discriminated. Workers apply to certain wage only when their expected payoff (wage times the employment probability) from this application attains certain level, and a high wage which attracts also the favored group discretely lowers the probability of employment for the discriminated group to such an extent that the expected payoff for the latter at these high wage firms does not meet the expected market payoff. This setup is supported by Lang and Lehmann (2012) and Heckman (1998), who mention that workers do not apply randomly and they actually avoid prejudiced employers to some extent, which implies between-group search externality is taken into account by the discriminated workers. Moreover, it is well known that within-group search externality may be prevalent when wages are bargained; while in wage posting context, we are able to abstract from search externality and focus on discrimination. Hall and Krueger (2010) use U.S. data to show that fraction of posted and bargained wages are both around one third. They also document a negative relationship between the education level and precise information concerning the expected pay. Brenzel, Gartner and Shnabel (2013) focus on the employer's side of the study in Germany, and show that around two thirds of the wages are posted, and the bargained wages are more likely set for those with higher education and qualification. The message is that not only is wage posting a prevalent wage determination process in the labor market, more importantly, it is also dominant in the relatively low skilled sector.² Within our context, employers can not post wages contingent on workers' group identity which is irrelevant to productivity, which could be understood as due to the functioning of the equal opportunity legislation.

Literature addressing discrimination problem in random search context is vaster. However, to have tractable such model convenient for linking to evidence, the introduced discrimination is usually taste-based, hence to obtain realistic outcome may often require making compromise on assuming ex ante differences in parameters governing relevant characteristics. Rosen (1997) is an exception and shows that discrimination can result even if there are no differences across groups. Job opportunities arrive stochastically, minority workers choose reservation productivities above which they accept the job; To avoid majority workers who are always preferred, they choose to accept jobs even with low reservation wages. Although private information is the key element in Rosen (1997)'s model, search externality remains the main channel for the functioning of the discrimination mechanism. Our focus is on how the ranking order of firms contributes to strategic interdependence in workers' skill investment decisions, and search externality is internalized when search is directed.

There is also the important statistical discrimination literature³ which emphasizes the role of asymmetric information on qualities related to the productivity. One strand of this literature derives group inequalities

²It is consistent with our knowledge that the more skilled workers, whose number is comparatively small, usually receive more attention and protections.

³We refer readers to the survey from Fang & Moro (2010)

endogenously even in the absence of ex ante group difference on relevant characteristics. Their mechanism is that decision makers' asymmetric beliefs on relevant characteristics of members for different groups could subsequently dim unfavored agents' incentive on investment on payoff-relevant technology, which in turn justifies the firms initial beliefs. Our context is different from this literature mainly in the point that, instead of relying on the information friction which plays central role in generating the pessimistic outcome, we work through a sequential game where agents could correctly anticipate the pessimistic outcomes, hence choose to react accordingly in a rational way.

The paper is organized as follows. Section 2 analyzes the case without discrimination. We then move to the economy with discrimination in Section 3. In section 4, we consider the case of wage bargaining. Section 5 discusses, and then we conclude.

2 The model without discrimination

We start with a context without hiring discrimination. Consider an economy populated by two kinds of agents, the workers and the firms. The number of workers is N ,⁴ with the index $i \in \{1, 2, \dots, N\}$, and the number of firms is M , with the index $j \in \{1, 2, \dots, M\}$. Define the market tightness as $\beta \equiv \frac{N}{M}$.

We introduce a pre-matching investment stage in a standard wage posting game. Each job seeker makes a skill investment decision before entering into the labor market. This skill choice is assumed to be binary, such that if the worker decides to become highly skilled, an investment cost E_H is paid, and otherwise E_L , with $E_H > E_L$. A highly skilled job seeker who pays E_H is capable of producing y_H ; while a low skilled could only deliver y_L . It would be useful to understand the formulation in the following way: workers who enter labor market after a longer period of training at school expect to receive higher expected income compared to those who spend a shorter period in schooling and enter the market at an earlier stage; the opportunity cost for the former is E_H , and for the latter is E_L . We assume that workers' skill level is public information. The costs $\{E_L, E_H\}$ and productivities $\{y_L, y_H\}$ are exogenous, but should satisfy some conditions which will be specified later.

Firms are ex ante identical. Having observed the distribution of skill attainment of job seekers, they post wages conditional on skills. If firms choose to attract a high skilled worker, they post w_H , and the surplus after matching is $y_H - w_H$; in case a low-skilled worker is searched for, w_L is announced and the surplus is $y_L - w_L$. We emphasize that skill is a characteristic of workers which the wage contracts can be conditioned on; this is in sharp contrast to other qualities such as gender, race, height etc. which, under equal pay legislation, should not be conditioned on; so when firms distinguish workers according to these latter qualities, the wage contract becomes "incomplete";⁵ by this, we will say that firms discriminate.

The timing follows that of a standard wage posting game, augmented by a pre-matching skill investment stage (Stage 0):

Stage 0: *Workers choose skill level, and pay either E_H or E_L .*

Stage 1: *Firms observe skills of job seekers, and announce the wage (w_L, w_H) .*

⁴As noted by Lang, Manove, and Dickens (2005), the number N could be regarded as the expected number of entrants (job seekers) from the firms' perspective.

⁵Incompleteness of contract is the source of inefficiency for the holdup problem. See Acemoglu and Shimer (1999) for related literature.

Stage 2: Workers observe the wages offers, and choose which wage to apply to.

Stage 3: Firms select workers from the received applications, and they select workers with same skill levels with equal probability. Then the production is carried on, and payoffs are realized.

We will focus on subgame perfect equilibria. Firms choose wages to maximize profits, and workers choose firstly the skill level and then which wage to apply to, in order to maximize the expected payoff from search.

2.1 Specification of the Strategies, matching probabilities, and payoff functions

To write agents' payoffs, it is a routine procedure in the directed search literature to first derive the matching functions. This section provides a quick summary for the general understanding of the context.

Define a type- t job seeker i 's strategy as a vector of probabilities $\Theta_t^i = (\theta_t^{i1}, \dots, \theta_t^{iM})$, where θ_t^{ij} is the probability with which the type- t worker i applies to firm j , and $t \in \{L, H\}$. It holds that $\sum_j \theta_t^{ij} = 1$ for any i and t . As in the literature, it is convenient⁶ to proceed with a transformation of variable. We define q_t as *expected number of applications received per firm*; it is also called the expected queue length.

Denote q^j as the queue length of firm j , and q_t^j as the queue length of the type- t workers in firm j . If a firm attracts both high and low skilled, we have $q^j = q_L^j + q_H^j$, where q_L^j and q_H^j are the queue length of the corresponding workers in firm j . Since we only consider symmetric equilibria, for a given firm j , θ_t^{ij} has the same value for any type t job seeker, so we denote $\theta_L^{ij} = \theta_L^j$ and $\theta_H^{ij} = \theta_H^j$ for any j . By definition, q_t^j is the number of workers of type t in firm j times their application probability: $q_H^j = N_H \times \theta_H^j$, and $q_L^j = N_L \times \theta_L^j$ for any j , where N_L and N_H are the total number of low skilled and high skilled workers respectively.

Firms. A particular firm j matches with a worker if after the search stage at least one worker appears, which happens with probability $1 - (1 - \theta_t^j)^{N_t}$. The probability that no job seeker sends application to this firm is $(1 - \theta_t^j)^{N_t}$, and $1 - (1 - \theta_t^j)^{N_t}$ is the probability of receiving at least one application from type t workers. According to the aboved defined relationship $q_t^j = N_t \theta_t^j$, the probability $(1 - (1 - \theta_t^j)^{N_t})$, goes to $(1 - e^{-q_t^j})$ when $N_t \rightarrow \infty$. This probability is increasing in q_t^j , which means that the more the expected number of applicants, the higher the probability that the firm could fill the vacancy. The firm chooses wage to maximize their expected profit, which is the product of the probability of meeting a worker of this type and the net surplus, $(1 - e^{-q_t^j}) \times (y_t - w_t^j)$, where $y_t \in \{H, L\}$. In the following we may occasionally suppress the superscript j whenever it does not raise confusion.

As shown by Shi (2006), in case workers of both skill levels appear in the market, it is optimal for firms to post both w_L and w_H to attract both skill types. Furthermore, firms rank the high skilled in priority to the low skilled, that is, firms will only consider hiring the low skilled workers when they did not receive any application from high skilled workers, an event which happens with probability e^{-q_H} . Then the expected profit (*from attracting both types of workers*) is

⁶When the number of firms and workers are large, it is no longer convenient to operate with the workers' application strategy θ_t^i , because it will tend to zero in the symmetric mixed strategy equilibrium.

$$(1 - e^{-q_H}) \times (y_H - w_H) + e^{-q_H} (1 - e^{-q_L}) \times (y_L - w_L)$$

Job seekers. Job seekers observe all the wages \mathbf{w} announced by firms, and choose which wage to apply to. Consider a particular job seeker. Conditional on visiting a particular firm j , his probability of employment in that firm is $\frac{1 - (1 - \theta_H^j)^{N_H}}{N_H \theta_H^j}$ if he is high skilled, and $(1 - \theta_H^j)^{N_H} \times \frac{1 - (1 - \theta_L^j)^{N_L}}{N_L \theta_L^j}$ if low skilled (see Appendix for more details). And these probabilities become $\frac{1 - e^{-q_H^j}}{q_H^j}$ and $e^{-q_H^j} \frac{1 - e^{-q_L^j}}{q_L^j}$ when $N \rightarrow \infty$ and $M \rightarrow \infty$. Notice that $\frac{1 - e^{-q_t}}{q_t}$ is decreasing in q_t : the higher the expected number of applicants in this firm competing this job with him, the lower the probability with which this job seeker will be employed. Also notice that the employment probability of the low skilled workers is a product of e^{-q_H} and $\frac{1 - e^{-q_L}}{q_L}$, where the former governs the between-group competition effect, and $\frac{1 - e^{-q_L}}{q_L}$ governs the within-group competition effect.

We remark that since q is a function of job seekers' application strategy, it depends on w . We now look more closely into their causal relationship. We should distinguish two terms: (1) each job seeker's *expected payoff from application*, and (2) her *expected "market" payoff*. The expected payoff from application is the payoff that a worker receives when applying to a certain firm, namely, a product of the wage and the probability of obtaining it at that firm, namely, $\frac{1 - e^{-q_H}}{q_H} \times w_H$ for the high skilled, and $e^{-q_H} \frac{1 - e^{-q_L}}{q_L} \times w_L$ for the low skilled. The expected market payoff, denoted by U_t , is the maximum level of the expected payoff from application in the equilibrium. U_t is regarded as an aggregate variable, which is assumed to be invariant with respect to any variation of an individual agent's strategy. Consider a particular type H job seeker. He is willing to send application to a particular firm j , if and only if his expected payoff $\frac{1 - e^{-q_H^j}}{q_H^j} \times w_H^j$ from doing so is equal or greater than the expected market payoff U_H . By the definition that U_H is the maximum level attainable, we have

$$q_H^j \begin{cases} > 0 & \text{if } \frac{1 - e^{-q_H^j}}{q_H^j} \times w_H^j = U_H \\ = 0 & \text{if } \frac{1 - e^{-q_H^j}}{q_H^j} \times w_H^j < U_H \end{cases}$$

Job seekers make trade-off between the wage and the probability obtaining it. To highlight the dependence of q_t on w_t , we could rewrite the above expressions as

$$q_H^j \begin{cases} > 0 & \text{if } w_H^j > U_H \\ = 0 & \text{if } w_H^j \leq U_H \end{cases}$$

Indeed, the employment probability $\frac{1 - e^{-q_H}}{q_H}$ is a number which belongs to the interval $(0, 1)$; if the wage is too low such that $w_H^j \leq U_H$, it will attract no workers, q_H is zero; if $w_H^j > U_H$, there is always a positive value q_H which satisfies $\frac{1 - e^{-q_H^j}}{q_H^j} \times w_H^j = U_H$.

We now formalize the notion of equilibrium and will proceed in two steps. In the first, we state the notion of equilibrium for the wage posting subgame given a skill distribution of workers. In the second, we distinguish two notions on how workers optimally choose their skill level. For the wage posting subgame, the solution concept will be standard subgame-perfect competitive equilibrium (SPCE) similar to LMD (2005). As emphasized by LMD (2005), it is '*a simplification of standard subgame-perfection in which the aggregate*

variables are assumed constant with respect to the changes in the strategy of an individual agent.’ Given the fraction of $\alpha \in [0, 1]$ of high skilled workers, an equilibrium, symmetric among workers, consists of the expected market payoffs (U_H, U_L) , each firms’ strategy (w_H^*, w_L^*) , and workers’ strategies (q_H^*, q_L^*) , that satisfy the following requirements: (i) each firm’s posted wage (w_H^*, w_L^*) is a best response to the other firms’ strategy and to the workers’ strategies (q_H^*, q_L^*) , on the assumption that the market expected payoff $(U_H(w), U_L(w))$ ⁷ remains fixed at $U_H(w^*)$ and $U_L(w^*)$, and is invariant to the firm’s own wage; and (ii) $q_t^*(w)$ with $t \in \{L, H\}$ is a best response of each worker to any wage vector w , and to the choice of $q_t^*(w)$ by all other workers.

Now we turn to workers’ skill decision which determines the value of α . Workers choose to be high skilled whenever $U_H - E_H > U_L - E_L$, then the fraction of high skilled workers is such that $\alpha = 1$. Workers choose to be low skilled whenever $U_H - E_H < U_L - E_L$, then the fraction of high skilled workers is such that $\alpha = 0$. Workers are indifferent between high and low skill when $U_H - E_H = U_L - E_L$. In the case with indifference, we should make distinctions on two notions specified as follows:

- Notion 1 (perfect correlation): Whenever indifferent, all workers (within a group) choose either high or low skilled.
- Notion 2 (no perfect correlation): Whenever indifferent, α represents individual worker’s probability of choosing to be high skilled.

Under Notion 1, all workers in one group, whenever indifferent between two alternatives (L or H), will randomize towards the same direction: that is to say, we consider the group of workers as a whole, or there is perfect correlation on their skill choices; as a result, α does not represent an individual’s probability of choosing high skilled and takes either the value 0 or 1, $\alpha \in \{0, 1\}$. Under Notion 2, each worker can randomize between high and low skill whenever they are indifferent; since each worker can end up either high or low skilled, after this randomization, high and low skilled workers can be present at the same time, in contrast to Notion 1; The equilibrium value of α , denoted by α^* , will be endogenously determined by the indifference equation, in which case $\alpha^* \in (0, 1)$.

In the following analysis in this section without discrimination, we will derive equilibrium under both notions because technically Notion 1 could be regarded as a subproblem of Notion 2. However, we will focus exclusively on the study under Notion 1 when we turn to the section with discrimination. By doing so, we are able to abstract from some equilibrium which only arises under theoretical rigor but at the same time yields insights to a limited extent and induces unnecessary complexity in analysis.

Solution of the wage posting subgame. According to Burdett, Shi and Wright (2001), the definition of equilibrium is equivalent to a problem where firms choose wages to maximize their expected profit, taking into account the best responses of other firms as well as of the job seekers. As we consider a large economy, in which a single firm’s deviation does not alter the expected market payoff U_t , the market payoff can be taken as given in the stage where a firm maximizes its profit, and will later be determined endogenously. Thus, we consider the optimization problem in which a (deviating) firm j chooses w_t with $t \in \{L, H\}$ to maximize its expected profit, taking expected market payoff U_t (other firms’ responses) and the functional relationship between w_t and q_t (job seekers’ responses) as given. When the firm attracts a single skill type,

⁷“ w ” denotes the wage vector of all the firms.

the equilibrium can be solved by⁸

$$\begin{aligned} \max_{w_t} \quad & (1 - e^{-q_t}) \times (y_t - w_t) \\ \text{s.to} \quad & \frac{1 - e^{-q_t}}{q_t} \times w_t = U_t \end{aligned}$$

for $t \in \{L, H\}$. For a given U_t , solving w_t from the constraint, substituting it into the objective function, and maximizing with respect to q_t , we can obtain an optimal functional relationship between q_t^* and U_t . Using this obtained relationship, with the help of the constraint, we then achieve an optimal functional relationship between w_t^* and q_t^* . And since in symmetric equilibrium all firms will post the same wages, so that all workers will apply to each firm with equal probability, by definition of q we have $q_H^j = q_H^* = \frac{N}{M} = \beta$ and $q_L^j = q_L^* = \frac{N}{M} = \beta$, when either all workers are high skilled or all workers are low skilled. This program is applicable under both Notion 1 and relevant situations of Notion 2. Under Notion 2, in addition, when workers are indifferent, both skills can be present at the same time, then firms attract both skill types as in Shi (2006), the problem is

$$\begin{aligned} \max_{w_H, w_L} \quad & (1 - e^{-q_H}) \times (y_H - w_H) + e^{-q_L} (1 - e^{-q_H}) \times (y_L - w_L) \\ \text{s.to} \quad & \frac{1 - e^{-q_H}}{q_H} \times w_H = U_H \\ & e^{-q_H} \frac{1 - e^{-q_L}}{q_L} \times w_L = U_L \end{aligned}$$

Solving the program, we could obtain $U_H = e^{-q_H^*} (y_H - w_H^*) + e^{-q_H^* - q_L^*} y_L$, while $U_L = e^{-q_H^* - q_L^*} y_L$; at the equilibrium, we have $q_H^* = \frac{N_H}{M}$, $q_L^* = \frac{N_L}{M}$ and $q_H^* + q_L^* = \beta$. At last, it is important to remark that q_t depends on w_t continuously, as remarked by Shi (2002). In this way, a marginal change of wage w_t can only lead to a marginal modification on the expected number of applicants q_t . By the definition of q , under Notion 2, we have $q_H^*(\alpha) = \frac{N_H}{M} = \alpha\beta$, and $q_L^*(\alpha) = \frac{N_L}{M} = (1 - \alpha)\beta$. Once the skill investment choice pins down the value of α^* , we obtain $q_H^* = \alpha^*\beta$, and $q_L^* = (1 - \alpha^*)\beta$. In the next section, we establish the decentralized market equilibrium and examine its properties. It is convenient to start with Notion 2.

2.2 Decentralized Market Equilibrium without discrimination under Notion 2

Firms' wage offers are conditioned on job seekers' skill levels, so we first consider the skill investment decision of job seekers at first stage. Denote α as the fraction of the job seekers who choose to invest in high skill, so the remaining fraction $(1 - \alpha)$ is low skilled. Under Notion 2, α is also the probability with which a job seeker chooses to invest in high skill, by virtue of the Law of Large Number. Let α^* denote the equilibrium fraction of high skilled job seekers on the total population. There are three cases:

Case (1). $\alpha^* = 1$. *All job seekers invest in high skill.*

Case (2). $\alpha^* \in (0, 1)$. *Some invest in high skill, while the remaining in low skill.*

Case (3). $\alpha^* = 0$. *All invest in low skill.*

With Case (1) and Case (3), there exists only one skill level in the market, and since skills can be conditioned on wages, there is only one wage posted in equilibrium. However, the market with Case (2) features two skill levels. Shi (2006) establishes that in this case it is optimal for firms to attract both skill types, while ranking the high skilled in priority to the low skilled. We now show that the rivalry between the

⁸The program could also be understood as a deviating firm's profit maximization, taking all the other firms' best response as given. The other firms' best response is w_t^* , which satisfies $\frac{1 - e^{-q_t^*}}{q_t^*} w_t^* = U_t$.

market competition (captured by market tightness β) and the magnitude of the return to skill ratio $\frac{y_H - y_L}{E_H - E_L}$ are crucial in the determination of which of the above three cases may prevail.

Proposition 1. (return to skills) Given the return to skill ratio $\frac{y_H - y_L}{E_H - E_L}$, define $\hat{\beta}$ as $\frac{y_H - y_L}{E_H - E_L} = e^{\hat{\beta}}$.

- (i) when $0 < \beta \leq \hat{\beta}$, the unique equilibrium is such that all job seekers choose high skill, i.e. $\alpha^* = 1$.
- (ii) when $\beta > \hat{\beta}$, the equilibrium consists of a unique value $\alpha^* \in (0, 1)$ which satisfies $\frac{y_H - y_L}{E_H - E_L} = e^{\alpha^* \beta}$.
- (iii) when $\frac{y_H - y_L}{E_H - E_L} \leq 1$ such that $\frac{y_H - y_L}{E_H - E_L} = e^{\hat{\beta}}$ has no positive solution on $\hat{\beta}$, the unique equilibrium is $\alpha^* = 0$.

Proof. In the appendix.

When the value of return to skill $\frac{y_H - y_L}{E_H - E_L}$ is sufficiently large compared to e^{β} , which measures the intensity of competition of the market, job seekers find it a dominant strategy to invest in high skills; There is no incentive for them to deviate, and the output is highest among all the equilibria. When the value of $\frac{y_H - y_L}{E_H - E_L}$ is moderate, there exists an equilibrium where job seekers are indifferent from being high skilled or low skills; all firms find it optimal to attract both skill types; the output is lower compared to the previous equilibrium. At last, when the value of return to skill is sufficiently low, it does not provide them incentive to sink this fixed cost against the risky job search game they are going to play; the equilibrium level of output turns out to be the lowest.

2.2.1 Decentralized Market Equilibrium without discrimination under Notion 1

Under Notion 1, there is only one skill type present in the market. When firms attract a single skill type, they solve

$$\begin{aligned} \max_{w_t} \quad & (1 - e^{-q_t}) \times (y_t - w_t) \\ \text{s.to} \quad & \frac{1 - e^{-q_t}}{q_t} \times w_t = U_t \end{aligned}$$

In equilibrium, for the high skilled we have: $U_H = e^{-q_H^*} y_H$, $w_H^* = \frac{e^{-q_H^*} q_H^*}{1 - e^{-q_H^*}} y_H$, where $q_H^* = \beta$; and for the low skilled: $U_L = e^{-q_L^*} y_L$, $w_L^* = \frac{e^{-q_L^*} q_L^*}{1 - e^{-q_L^*}} y_L$, where $q_L^* = \beta$. So that we have

$$\alpha = \begin{cases} 1 & \text{when } e^{-\beta} y_H - E_H > e^{-\beta} y_L - E_L \\ 0 \text{ or } 1 & \text{when } e^{-\beta} y_H - E_H = e^{-\beta} y_L - E_L \\ 0 & \text{when } e^{-\beta} y_H - E_H < e^{-\beta} y_L - E_L \end{cases}$$

It turns out that the threshold which makes workers indifferent is the same as $\hat{\beta}$ established under Notion 2.

2.3 Constrained efficient allocations

The objective of this section is to find the efficient allocations in the centralized market, and evaluate whether the decentralized market attains its efficiency. The social planner maximizes the aggregate output,

subject to the same matching friction as in the decentralized equilibrium. More precisely, the social planner chooses the fraction of workers to be high-skilled, divides firms into different groups to attract distinct compositions of workers, and assigns workers to match with a certain group of firms. With the same matching friction as before, the social planner is restricted to treat workers of the same skill level in the same way, and assures that workers of the same skill type must match with firms from the same group with the same probability.

We use the superscript “ p ” to label the equilibrium allocation chosen by the planner. Let α be the fraction of high skilled workers the social planner chooses, $\alpha \in [0, 1]$. If the optimal arrangement is $\alpha^p = 1$, all job seekers are high skilled, and only one type of firms exists - those which attract high skilled workers. It is similar for $\alpha^p = 0$. If $\alpha^p \in (0, 1)$, there are both high and low skilled job seekers and it is optimal for the planner to assign all firms to post wages for both the high and low skilled (shown in Shi (2006)). Furthermore, in the last case the planner can also manage the priority of firms’ hiring workers with different skills - whether to prefer high skilled to low skilled or otherwise. Let R be the probability with which the firms rank high skilled workers in priority to the low skilled. And q_t is the expected number of applicants in a firm, $t \in \{L, H\}$, which governs how the planner assigns workers’ applications. Thus, the social planner’s problem is to maximize the following aggregate output

$$M \times \left[\begin{array}{l} (1 - e^{-q_H}) (R \times (1 - e^{-q_L}) + e^{-q_L}) y_H \\ + (1 - e^{-q_L}) ((1 - R) (1 - e^{-q_H}) + e^{-q_H}) y_L \\ - N \times (\alpha E_H + (1 - \alpha) E_L) \end{array} \right]$$

If at least one high skilled visits a certain firm, with probability $(1 - e^{-q_H})$, the firm hires this high skilled, either with probability 1 when no low skilled worker shows up, which happens with probability e^{-q_L} , or with probability R if there is at least one low skilled who shows up at the same firm, which occurs with probability $(1 - e^{-q_L})$; it is similar for the case with low productivity. Since the firms and workers of the same skill are all identical from the planner’s perspective, we have $q_H = \alpha\beta$ and $q_L = (1 - \alpha)\beta$. The above objective includes all cases with different values of α . Solving the problem, we can see that the optimal ranking is that firms always prefer high skilled workers, i.e., $R^p = 1$, and we have the following proposition

Proposition 2. The equilibrium labor allocation and skill investment choice are socially optimal.

Proof. In the Appendix.

In the proof, we could also verify that the threshold $\hat{\beta}^p$ for skill investment coincides with $\hat{\beta}$ in the decentralized economy. That is, when $0 < \beta \leq \hat{\beta} = \log \frac{y_H - y_L}{E_H - E_L}$, it is socially optimal that workers all invest in high skill; when $\beta > \hat{\beta}$, it is socially optimal that a fraction α^* of workers invest in high, while the rest invest in low skill; and $\frac{y_H - y_L}{E_H - E_L} \leq 1$ such that $\hat{\beta}$ has no positive real solution, all invest in low skill. In the rest of the paper, we mainly focus on the first case, so that whenever workers are discouraged to underinvest, it is due to the effect of discrimination.

3 The model with hiring discrimination

We now introduce discrimination. Consider an economy where workers can be partitioned into two groups, group a and group b , according to certain trait which is irrelevant to productivity. Gender, for example, is such one possible binary partition of labor force. Denote the fraction of group a as γ , and the fraction of

group b as $1 - \gamma$. The two group of workers are ex ante identical in all other aspects. Discrimination modifies the matching functions of agents. Specifically, in order to formulate discrimination, we introduce a term x called hiring (ranking) rule specified by firms. To be precise, x could be understood as the probability with which the group a workers are selected when workers from both groups are present. The probability that a group a worker is employed by this firm is

$$F_a(q_a, q_b, x) = \frac{1 - e^{-q_a}}{q_a} \times [x(1 - e^{-q_b}) + e^{-q_b}]$$

Analogously, the probability that a group b worker is employed by this firm is

$$F_b(q_a, q_b, x) = \frac{1 - e^{-q_b}}{q_b} \times [(1 - x)(1 - e^{-q_a}) + e^{-q_a}]$$

To understand these expressions, we have to notice that now when job seekers are considering their probability of being hired, they have to take into account of the impact from the competition with the other group. The parts $\frac{1 - e^{-q_a}}{q_a}$ and $\frac{1 - e^{-q_b}}{q_b}$ capture the within group competition, while the remaining parts with x capture the between group competition.

When $x = 1$, firms hire group b workers only when none of the group a is present. Firms' preference is such that group a are preferred to group b , although both groups have identical productivity. This is what we call hiring discrimination. The employment probability for group a and group b workers become respectively,

$$F_a(q_a, q_b, 1) = \frac{1 - e^{-q_a}}{q_a}$$

$$F_b(q_a, q_b, 1) = \frac{1 - e^{-q_b}}{q_b} \times [e^{-q_a}]$$

Another interesting example is $x = \frac{1}{2}$. The employment probability for group a and group b workers becomes respectively

$$F_a\left(q_a, q_b, \frac{1}{2}\right) = \frac{1 - e^{-q_a}}{q_a} \times \left[\frac{1}{2} \times (1 - e^{-q_b}) + e^{-q_b}\right]$$

$$F_b\left(q_a, q_b, \frac{1}{2}\right) = \frac{1 - e^{-q_b}}{q_b} \times \left[\frac{1}{2} \times (1 - e^{-q_a}) + e^{-q_a}\right]$$

When $q_a = q_b$, we have $F_a(q_a, q_b, \frac{1}{2}) = F_b(q_a, q_b, \frac{1}{2})$, both workers have identical employment probability, and it is as if there is no discrimination among workers. Furthermore, we have $F_a(q_a, q_b, 1) \geq F_a(q_a, q_b, \frac{1}{2})$ for any positive value of (q_a, q_b) . Then the employment probability of workers from the preferred group (group a) is higher under discrimination than that in the case without discrimination. Similarly, when $F_b(q_a, q_b, 1) \leq F_b(q_a, q_b, \frac{1}{2})$ for any positive values of (q_a, q_b) , i.e. the employment probability of the discriminated group (group b) is lower when there is hiring discrimination. In fact, the employment probability of group a (group b) is increasing (decreasing) in x . Thus, for a given positive pair of (q_a, q_b) , there exists $\hat{x} \in (0, 1)$ such that $F_a(q_a, q_b, \hat{x}) = F_b(q_a, q_b, \hat{x})$, and \hat{x} could be considered as the hiring rule without discrimination. Indeed x measures the intensity of firms' discriminatory preference. Given q_a and q_b , for $x \in [0, \hat{x})$, firms discriminate group a , and for $x \in (\hat{x}, 1]$ firms discriminate against group b . The closer x approaches to the extremes of the interval $[0, 1]$, the more intensive the hiring discrimination is. In the rest of paper, we focus on the case $x = 1$ such that group a achieve absolute priority to group b .

3.1 The case of strong discrimination: $x = 1$

Formally, we introduce two assumptions as in Merlino (2012). These assumptions help introduce some heterogeneity which is not related to productivity among the labor pool.

Assumption 1: Firms are not allowed to post wages which are dependent on the group identity.

Assumption 2: Firms prefer group a in the sense that firms only hire workers from group b when no group a workers are present, i.e. $x = 1$.

Same as the case without discrimination, workers visit a firm only when they can obtain the expected market payoff from applying to that firm. We denote in this section the expected market payoff of high skilled job seekers from group a and group b as U_{aH} and U_{bH} respectively. The above assumptions imply

$$U_{aH} = F_{aH}(q_{aH}, q_{bH}) \times w_H = \frac{1 - e^{-q_{aH}}}{q_{aH}} \times w_H$$

$$U_{bH} = F_{bH}(q_{aH}, q_{bH}) \times w_H = \frac{1 - e^{-q_{bH}}}{q_{bH}} e^{-q_{aH}} \times w_H$$

For $\beta \leq \hat{\beta}$, all workers choose to be high-skilled if there were no discrimination. To analyze how does discrimination have impact on workers' expected payoff from search hence the skill investment incentives, it is important to first study the wage posting subgame with discrimination given a skill distribution. Without loss of generality, we start with the case where both groups choose to be high skilled. In the next section, we review the results from LMD (2005), where they study the case with discrimination but no difference in workers' skill levels (or productivity).

3.2 Existing results revisited and reinterpreted

In a context where there are two groups of workers with identical productivity (skill level) and firms strongly prefer group a to group b . LMD (2005) show that any subgame-perfect competitive equilibrium (SPCE) is separating.

Separating equilibrium. LMD (2005) show that there is no wage to which both groups of job seekers apply. More precisely, no wage can maximize firms' profit while attracting both groups of workers simultaneously (with the expected payoff from application attaining the corresponding market payoff). The equilibrium is separating. That is, there are some firms posting a higher level of wage attracting only the preferred group a , whereas the rest of firms offering a lower wage which is applied only by the discriminated group b (see Proposition 2 in LMD (2005)). Notice that the discriminated group have always the choice of applying to the high wage firms, however, they choose not to do so at all, because they anticipate discrimination in these firms. The most essential results of LMD (2005) are summarized as follows:

(i) *At the firms attracting group a workers, expected profit for the firms and the expected payoff for the workers in the equilibrium are*

$$\pi_{aH}^S = \left(1 - e^{-q_{aH}^S} - q_{aH}^S e^{-q_{aH}^S}\right) \times y_H$$

$$U_{aH}^S = e^{-q_{aH}^S} \times y_H$$

$$w_{aH}^S = \frac{q_{aH}^S e^{-q_{aH}^S}}{1 - e^{-q_{aH}^S}} \times y_H$$

ii) At the firms attracting group b workers, expected profit for the firms and the expected payoff for the workers in the equilibrium are

$$\pi_{bH}^S = \left(1 - e^{-q_{bH}^S}\right) \left(1 - e^{-q_{aH}^S}\right) \times y_H,$$

$$w_{bH}^S = U_{aH}^S$$

$$U_{bH}^S = \frac{1 - e^{-q_{bH}^S}}{q_{bH}^S} w_{bH}^S$$

Separating equilibrium requires that firms be indifferent between attracting group a and group b workers, that is

$$\pi_{aH}^S = \pi_{bH}^S$$

which helps to determine q_{bH}^S and q_{aH}^S jointly.

(iii) Furthermore, we have $w_{aH}^S > w_{bH}^S$. $q_{aH}^S > \beta > q_{bH}^S$, both q_{aH}^S and q_{bH}^S are increasing in β and γ , and both q_{aH}^S and q_{bH}^S are independent of y_H .

We make some important remarks on the features of the separating equilibrium. Firstly, the resulted equilibrium allocations are incentive compatible. For any particular bH job seeker, by deviating to applying for w_{aH}^S , the best they can get is $e^{-q_{aH}^S} \times w_{aH}^S$ (when none of the group a shows up in the firm which this bH worker deviates to apply to). However, this deviating payoff is strictly lower than sticking to applying to w_{bH}^S owing to the following relationship:

$$\begin{aligned} e^{-q_{aH}^S} \times w_{aH}^S &= e^{-q_{aH}^S} \times \frac{q_a^S \times e^{-q_a^S}}{1 - e^{-q_a^S}} \times y_H = \frac{q_a^S \times e^{-q_a^S}}{1 - e^{-q_a^S}} \times U_{aH}^S \\ &= e^{-q_{bH}^S} U_{aH}^S = e^{-q_{bH}^S} w_{bH}^S < \frac{1 - e^{-q_{bH}^S}}{q_{bH}^S} w_{bH}^S = U_{bH}^S \end{aligned}$$

. As for any particular aH job seeker, by deviating to w_{bH}^S , the best they can get is $w_{bH}^S = U_{aH}^S$, which is as good as what he could get if he does not deviate. Secondly, we do not have the reservation wage structure. A reservation wage structure requires that workers apply to any wage which gives them an expected payoff higher than certain reservation value, in our case it would imply group bH should apply to both the low and high wages, however, it is not the case. This is because the expected payoff from applying to the high wage is a strictly dominated strategy for group b : the expected payoff from applying to high wages is too low to match their expected market payoff U_{bH}^S . Following are several noteworthy properties of such an equilibrium.

Results from LMD (2005): Compared to the context without discrimination, (1) Both groups have lower expected payoff. (2) All firms earn higher profits. (3) The expected payoff of group a and group b are such that $U_{aH}^S > U_{bH}^S$.

Group b workers are worse off, because of firms' discriminatory hiring norm. Anticipating discrimination, group b demand lower expected payoff, which makes them cheaper to be employed. This in turn increases firms' market power in hiring group a . Group a understand that if they demand high wages, firms will threat to switch to hiring group b workers instead. Hence group a workers demand also low wages, and are worse off too. Now more about firms. Apart from the mechanisms just described, firms are able to earn high profits because in the regime with discrimination market is segmented, which allows the firms to face less competition in each segment. As a general remark, discrimination enables firms to extract higher profit by holding up job seekers' skill investment and providing all the job seekers lower expected payoff.

Furthermore, since $U_{aH}^S > U_{bH}^S$, it suggests that the group b job seekers, being discriminated, are hurt to a larger extent. So that group b ' incentive of skill investment is distorted further downwards. We enter more detailed discussions in the following section.

3.3 Analysis under our context

In the last section, we interpreted the equilibrium of the wage posting subgame given that all workers choose to be high skilled. In this section, we study how discrimination leads to different incentives of skill investment for these two groups respectively, and attempt to find the corresponding equilibrium.

An important observation is that the skill decision for group a and group b is strategic, and this is a direct consequence of the coexistence of ranking through the productivity-dependent (skill) and -independent traits (discrimination). Ranking by skills requires that the high skilled workers have the priority; While ranking by productivity-independent traits means that group a have the priority. Although multidimensional characteristics are involved, these two ranking schedules yield a unique market hierarchy:

$$aH \succ bH \succ aL \succ bL$$

. It reads as follows: high skilled group a (aH) are preferred to high skilled group b (bH), who are preferred to low skilled group a (aL), who are then preferred to low skilled group b (bL). How does the strategic interdependence arise? Take group a as an example for explanation. Although they are always ranked prior to group b due to discrimination, whenever they contemplate to lower skill investment, they understand that they will be ranked behind the high skilled group b ; then the term $e^{-q_{bH}}$ which captures the competition from bH will appear in their payoffs.

Given the strategic interdependence in payoffs from skill investment, we adopt Nash equilibrium as the solution concept for the skill investment game. We proceed under Notion 1, and regard a whole group as making decision collectively and simultaneously, then all the workers in one group either end up high or low skilled. And we will focus on pure strategy of each group. Let α_g be the probability of group g 's choosing to be high skilled, for $g = a$ or b . We have either $\alpha_g = 1$ or $\alpha_g = 0$. We have the following four possibilities as equilibrium:

- (P1) $\alpha_a = 1$ and $\alpha_b = 1$: group a - high, group b - high
- (P2) $\alpha_a = 1$ and $\alpha_b = 0$: group a - high, group b - low
- (P3) $\alpha_a = 0$ and $\alpha_b = 1$: group a - low, group b - high
- (P4) $\alpha_a = 0$ and $\alpha_b = 0$: group a - low, group b - low

To decide the skill investment, workers take into account firms' best response in the wage posting stage to infer the expected payoff from application, and compare the payoffs net of the cost of skill investment. In the wage posting subgame, when facing all workers with identical skill level (as in case (P1) and (P4)), firms' optimal strategy is the same as stated in LMD (2005); that is, some firms post a higher wage which is only applied by group a , whereas the rest post a lower wage which is only applied by group b . When there are both low and high skilled workers (as in case (P2) and (P3)), firms post wages conditional on skill level, and it is optimal for firms to attract both skill levels and rank the high skilled in priority to low skilled, as in Shi (2006). We then proceed to find workers' best response in the skill investment stage, and in turn

the equilibrium in this discriminatory context with skill investment. We will use P1, P2, P3, P4 as the superscript for corresponding equilibrium allocations. When $\alpha_a = 1$ and $\alpha_b = 1$, workers are composed of type aH and bH . Firms post wages separately. We have the following payoffs for aH and bH respectively,

$$U_{aH}^{P1} = e^{-q_{aH}^{P1}} y_H - E_H$$

and

$$U_{bH}^{P1} = \frac{1 - e^{-q_{bH}^{P1}}}{q_{bH}^{P1}} e^{-q_{aH}^{P1}} y_H - E_H$$

where we recall that $q_{aH}^S = q_{bH}^{P1} < \beta < q_{aH}^{P1} = q_{bH}^S$.

When $\alpha_a = 1$ and $\alpha_b = 0$, workers are composed of type aH and bL . Firms post wages conditional on skills, attract both types at the same time, and rank the high skilled in priority to the low skilled. As for the payoffs, we have

$$U_{aH}^{P2} = e^{-q_{aH}^{P2}} (y_H - y_L) + e^{-q_{aH}^{P2} - q_{bL}^{P2}} y_L - E_H$$

and

$$U_{bL}^{P2} = e^{-q_{aH}^{P2} - q_{bL}^{P2}} y_L - E_L$$

where $q_{aH}^{P2} = \gamma\beta$, $q_{bL}^{P2} = (1 - \gamma)\beta$, and $q_{aH}^{P2} + q_{bL}^{P2} = \beta$.

When $\alpha_a = 0$ and $\alpha_b = 1$, workers are composed of different skill levels, aL and bH , we have similarly

$$U_{aL}^{P3} = e^{-q_{aL}^{P3} - q_{bH}^{P3}} y_L - E_L$$

and

$$U_{bH}^{P3} = e^{-q_{bH}^{P3}} (y_H - y_L) + e^{-q_{aL}^{P3} - q_{bH}^{P3}} y_L - E_H$$

where $q_{aL}^{P3} = \gamma\beta$, $q_{bH}^{P3} = (1 - \gamma)\beta$, and $q_{aL}^{P3} + q_{bH}^{P3} = \beta$.

When $\alpha_a = 0$ and $\alpha_b = 0$, workers are composed of type aL and bL . Both are of same skill level, firms will discriminate and post wages as in LMD (2005), and workers' payoffs are

$$U_{aL}^{P4} = e^{-q_{aL}^{P4}} y_L - E_H$$

and

$$U_{bL}^{P4} = \frac{1 - e^{-q_{bL}^{P4}}}{q_{bL}^{P4}} e^{-q_{aL}^{P4}} y_L - E_L$$

Since the solution of q in LMD (2005) is independent of y_H and y_L , we have $q_{bH}^{P1} = q_{bL}^{P4} < \beta < q_{aL}^{P4} = q_{bH}^{P1}$. The payoff matrix is as follows

	bH	bL
aH	$\text{bH: } \frac{1 - e^{-q_{bH}^{P1}}}{q_{bH}^{P1}} e^{-q_{aH}^{P1}} y_H - E_H$ $\text{aH: } e^{-q_{aH}^{P1}} y_H - E_H$	$\text{bL: } e^{-q_{aH}^{P2} - q_{bL}^{P2}} y_L - E_L$ $\text{aH: } e^{-q_{aH}^{P2}} (y_H - y_L) + e^{-q_{aH}^{P2} - q_{bL}^{P2}} y_L - E_H$
aL	$\text{bH: } e^{-q_{bH}^{P3}} (y_H - y_L) + e^{-q_{aL}^{P3} - q_{bH}^{P3}} y_L - E_H$ $\text{aL: } e^{-q_{aL}^{P3} - q_{bH}^{P3}} y_L - E_L$	$\text{bL: } \frac{1 - e^{-q_{bL}^{P4}}}{q_{bL}^{P4}} e^{-q_{aL}^{P4}} y_L - E_L$ $\text{aL: } e^{-q_{aL}^{P4}} y_L - E_H$

A pure-strategy Nash equilibrium on skill investment consists of a profile of actions with the property that no single group as a whole can achieve a higher payoff by unilateral deviation. The equilibrium depends on the value of β . For example, holding group a high skilled, when we decide whether group b choose to be high or low skill, we must compare U_{bH}^{P1} and U_{bL}^{P2} . And we find that there exists at least one threshold $\hat{\beta}_2$ which is determined by $U_{bH}^{P1} = U_{bL}^{P2}$ and is such that $\hat{\beta}_2 < \hat{\beta}$; this implies that depending on different values of β , the group b may choose high or low skill, leading to potentially different equilibrium. For tractability, we should introduce the following two conditions which guarantee that any such threshold as $\hat{\beta}_2$ is unique:⁹

Assumption 3.1 The equality $e^{-q_{aH}^{P1}(\hat{\beta}_1)}y_H - E_H = e^{-\hat{\beta}_1}y_L - E_L$ admits a unique solution $\hat{\beta}_1$.

Assumption 3.2 The equality $\frac{1-e^{-q_{bH}^{P1}(\hat{\beta}_2)}}{q_{bH}^{P1}(\hat{\beta}_2)}e^{-q_{aH}^{P1}(\hat{\beta}_2)}y_H - E_H = e^{-\hat{\beta}_2}y_L - E_L$ admits a unique solution $\hat{\beta}_2$.

The results on the equilibrium are summarized as follows:

Proposition 3: Under the above assumptions, there exist two thresholds $\hat{\beta}_2$ and $\hat{\beta}_1$ with $0 < \hat{\beta}_2 < \hat{\beta}_1 < \hat{\beta}$, such that

(1) When $0 < \beta < \hat{\beta}_2$, there exists a unique equilibrium in which both group a and group b invest in high skill, (aH, bH) . At $\beta = \hat{\beta}_2$, both (aH, bH) and (aH, bL) can be equilibrium.

(2) When $\hat{\beta}_2 < \beta < \hat{\beta}_1$, there exists a unique equilibrium in which group a invest in high skill while group b in low skill, (aH, bL) . At $\beta = \hat{\beta}_1$, both (aL, bH) and (aH, bL) can be equilibrium.

(3) When $\hat{\beta}_1 < \beta \leq \hat{\beta}$, there exist multiple equilibria. Either group a invest in high skill and group b in low skill, or group a invest in low skill and group b invest in high skill, (aL, bH) or (aH, bL) .

Proof. In the Appendix.

In fact, if one group choose to be low skilled, the best response of the other group is always to be high skilled, while the best response to the other's high skill choice depends on the two thresholds. Furthermore, the rise of market tightness β makes workers have stronger incentive to deviate from high skill, and group b is more prone to deviate compared to group a , in the sense that the threshold of β at which group b begins to contemplate to invest in low skill is lower compared to group a . Interestingly, for values of β close to $\hat{\beta}$, there exists an equilibrium where the preferred group a choose low skill, while the discriminated group b choose high skill. We have the following results on the comparison of workers' expected payoff and firms' profits compared to the case without discrimination.

Corollary. Compared to the case without discrimination,

(1) In (aH, bH) equilibrium, firms always earn higher expected profits; In (aH, bL) equilibrium and (aL, bH) , firms earn lower expected profits.

(2) In (aH, bH) equilibrium, both aH and bH workers earn lower expected payoff; In (aH, bL) equilibrium, group aH (group bL) earn higher (lower) expected payoff; in (aL, bH) equilibrium, group bH (group aL) earn higher (lower) expected payoff.

Proof. In the Appendix.

This corollary tells that firms can be worse off with discrimination. Indeed, when workers anticipate discrimination, their investment incentive may be downwards distorted, and some group may end up choosing to underinvest. In equilibrium, whenever one group underinvest and the other group remain high skilled, the

⁹When y_H is sufficiently large compared to y_L , these two conditions are satisfied. Take $\hat{\beta}_1$ for example. The related assumption is satisfied when $\frac{e^{-q_{aH}^{P1}(\hat{\beta})} \frac{dq_{aH}^{P1}}{d\beta}}{e^{-\hat{\beta}}} < \frac{y_H}{y_L}$. Since q_{aH}^{P1} does not depend on y_H and y_L , there exists always a pair of y_L and y_H such that this condition is satisfied.

firms turn out to earn lower expected profits compared to the case without discrimination. This is simply due to the fact that underinvestment in skills discretely drags down the average productivity of the economy.

4 Comparison with fixed sharing rule (Wage Bargaining)

In this section, we shut down the channel through which firms use wages to influence workers' choices on applications, and examine whether the inefficiency can be alleviated. Notice that in this section, the workers only choose the amount of skills to obtain, not where to search.

Consider an economy with the same discriminatory ranking as previous, but the wage is determined by ex post bargaining after a job seeker meets an employer. The timing of the economy now is as follows: firstly, workers decide skill levels simultaneously; secondly, workers and firms get matched according to the matching technology; thirdly, the matched worker-firm pair bargain à la Nash to determine how to share the output y . The simplest form of Nash bargaining widely used in literature is equivalent to a fixed sharing rule of output. If we denote the bargaining power for all workers as ψ , then from the output y_t , workers receive ψy_t , and firms receive $(1 - \psi) y_t$. We focus on the case where ψ is the same for both skill levels, otherwise there is too much degree of freedom.

For the ease of comparison, we require that the matching technology here is the same as in previous section. The hiring norm is as previous Group $aH \succ$ Group $bH \succ$ Group $aL \succ$ Group bL . The corresponding employment probability for different types of workers is inherited, so that the employment probability is respectively $\frac{1-e^{-q_{aH}}}{q_{aH}}$ for aH , $e^{-q_{aH}} \frac{1-e^{-q_{bH}}}{q_{bH}}$ for bH , $e^{-q_{aH}-q_{bH}} \frac{1-e^{-q_{aL}}}{q_{aL}}$ for aL , and $e^{-q_{aH}-q_{bH}-q_{aL}} \frac{1-e^{-q_{bL}}}{q_{bL}}$ for bH . The expected payoff (after skill investment) is just the employment probability times ψy_t ; for example, we have for aL people $e^{-q_{aH}-q_{bH}} \frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L$ as the expected payoff from search.

Now, we specify the expected queue lengths q_{aH} , q_{bH} , q_{aL} and q_{bL} parametrically. Recall the definition of queue length is $\frac{\text{nb. of workers}}{\text{nb. of vacancies}}$, then we have $q_{aH} = \gamma \eta_a \beta$ for aH workers, $q_{bH} = (1 - \gamma) \eta_b \beta$ for bH workers, $q_{aL} = \gamma (1 - \eta_a) \beta$ for aL workers, and $q_{bL} = (1 - \gamma) (1 - \eta_b) \beta$ for bL workers, where η_a represents the fraction of high skilled group a , and η_b the fraction of high skilled group b . The values of η_a and η_b depend on the comparison between the expected payoff from investing in high or low skill:

$$\eta_a \begin{cases} = 1 & \text{if } \frac{1-e^{-q_{aH}}}{q_{aH}} \times \psi y_H - E_H > e^{-q_{aH}-q_{bH}} \frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L - E_L \\ = 0 & \text{if } \frac{1-e^{-q_{aH}}}{q_{aH}} \times \psi y_H - E_H < e^{-q_{aH}-q_{bH}} \frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L - E_L \\ \in \{0, 1\} & \text{if } \frac{1-e^{-q_{aH}}}{q_{aH}} \times \psi y_H - E_H = e^{-q_{aH}-q_{bH}} \frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L - E_L \end{cases}$$

and

$$\eta_b \begin{cases} = 1 & \text{if } e^{-q_{aH}} \frac{1-e^{-q_{bH}}}{q_{bH}} \times \psi y_H - E_H > e^{-q_{aH}-q_{bH}-q_{aL}} \frac{1-e^{-q_{bL}}}{q_{bL}} \times \psi y_L - E_L \\ = 0 & \text{if } e^{-q_{aH}} \frac{1-e^{-q_{bH}}}{q_{bH}} \times \psi y_H - E_H < e^{-q_{aH}-q_{bH}-q_{aL}} \frac{1-e^{-q_{bL}}}{q_{bL}} \times \psi y_L - E_L \\ \in \{0, 1\} & \text{if } e^{-q_{aH}} \frac{1-e^{-q_{bH}}}{q_{bH}} \times \psi y_H - E_H = e^{-q_{aH}-q_{bH}-q_{aL}} \frac{1-e^{-q_{bL}}}{q_{bL}} \times \psi y_L - E_L \end{cases}$$

To keep consistency with the previous section (under Notion 1), we require that whenever indifferent, the whole group will choose either high or low skill, so that η is either 0 or 1 in that case. We will also assume that the group a are the majority: $\gamma \geq \frac{1}{2}$.¹⁰

¹⁰We think this case with $\gamma \geq \frac{1}{2}$ is more empirically relevant, when we are talking about gender and racial discrimination for example. The case $\gamma < \frac{1}{2}$ could be also analogously derived according to the proof of the following proposition.

We focus on Nash equilibrium as the solution concept. In the current context, only workers make skill investment decisions, firms do not post wages since ψ is exogenous. We consider each group, i.e. group a or group b , as a whole when they are making decisions. Each group of workers invest in skills simultaneously. Due to the discriminatory rule, the payoffs from skill investment for different groups of workers are interdependent. This renders the skill investment strategic. The following proposition helps explain how workers' expected payoffs vary with respect to ψ :

Proposition 4. Let $\gamma \geq \frac{1}{2}$ so that group a is the majority. There are four thresholds $\hat{\psi}_{aL,b} \leq \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$, such that

(1) For $\psi \in [0, \hat{\psi}_{aL,b})$, the unique pure strategy nash equilibrium is (aL, bL) ; for $\psi = \hat{\psi}_{aL,b}$, the equilibrium can be (aL, bL) or (aL, bH) . (2) For $\psi \in (\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a})$, the unique pure strategy nash equilibrium is (aL, bH) ; (3) For $\psi \in (\hat{\psi}_{bH,a}, \hat{\psi}_{bL,a})$, there is no pure strategy nash equilibrium; for $\psi = \hat{\psi}_{bL,a}$, the equilibrium is (aH, bL) ; (4) For $\psi \in (\hat{\psi}_{bL,a}, \hat{\psi}_{aH,b})$, the unique pure strategy nash equilibrium is (aH, bL) ; at the point $\psi = \hat{\psi}_{aH,b}$, the equilibrium can be (aH, bL) or (aH, bH) . (5) For $\psi \in (\hat{\psi}_{aH,b}, 1)$, the unique pure strategy nash equilibrium is (aH, bH) . (6) Define the threshold $\hat{\psi}$ of skill investment without discrimination as $\hat{\psi}y_H \frac{1-e^{-\beta}}{\beta} - E_H = \hat{\psi}y_L \frac{1-e^{-\beta}}{\beta} - E_L$; then $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$.

Proof. In the Appendix.

It is not always true that the group a workers are always better off. Notably, there is an equilibrium similar as before where group a underinvest: It could be observed from the payoff matrix (provided in the proof of Proposition 4) that in the region $\psi \in (\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a})$ group a (group b) workers obtain lower (higher) expected payoff compared to the case without discrimination. This is also closely related to the fact that when $\gamma \geq \frac{1}{2}$ the within-group competition in group a is fiercer. In general, firms' payoff is written as follows:

$$\pi = \begin{cases} (1 - e^{-q_{aH} - q_{bH}}) (1 - \psi) y_H & \text{if all high skilled} \\ (1 - e^{-q_{aH} - q_{bH}}) (1 - \psi) y_H & \text{if both high and low skilled} \\ + e^{-q_{aH} - q_{bH}} (1 - e^{-q_{aL} - q_{bL}}) (1 - \psi) y_L & \\ (1 - e^{-q_{aL} - q_{bL}}) (1 - \psi) y_L & \text{if all low skilled} \end{cases}$$

According to Proposition 4, we can determine the exact values of the queue lengths in the expression. Firms' profit will be piecewise monotone because although ψ increases continuously, the skill composition hence the average productivity of the market improves discretely with respect to this bargaining power. The fact that $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$ suggests that although firms can gather higher profits for $\psi < \hat{\psi}$, they encounter loss for $\psi \geq \hat{\psi}$ compared to the case without discrimination. The reason is that strategic competition between the group a and group b deters the discriminated group's skill investment decision (in the sense that group b may still choose to be low skilled when ψ is sufficiently high), which pulls down the market's average productivity and makes firms' expected profit dim.

It is interesting to notice that our simple result that discrimination is costly for firms at high skilled sector (when wages are bargained) questions the plausibility of key assumption of Merlino (2012) that "there is more discrimination in the high technology sector". Although Merlino (2012) mentioned bunches

of empirical evidence in support of this assumption¹¹, our simple result suggest that firms are simply better off not discriminating when wages are principally bargained, since the loss in profit from discriminating in the high skilled sector may surpass the gain from discriminating in the low skilled sector. All in all, the key difference between wage posting and wage bargaining is that the ex post wage now exogenously pegs on the productivity, and firms can no longer manipulate their market power by translating their discriminatory preference into constantly lower wages.

5 Discussion

Free entry. LMD (2005) have shown that their economy under discrimination with workers' identical in productivity can be generalized to take into account firms' free entry. Specifically, we consider a stage where firms sink capital after observing workers' skills. Each firm has different capital cost with $C_1 < C_2 < \dots < C_M < y_L$. Then firms which earn expected non-positive profits after the reduction of capital cost would simply not enter into the market. In the paper, we observe that the equilibrium is unique with respect to β , which has a one-one relationship with M - the number of firms in the market, so that the results in the paper could carry through with free entry. All firms in the market expect positive net profits. When there are different skill groups, this result could also carry through, because the equilibrium profit of firms is still an increasing function of β .

Heterogeneity in skill investment cost. Some preliminary attempts from us suggest that our context could be generalized to a situation where workers are heterogenous in their skill investment cost (although more complicated): let the low skill investment cost be zero ($E_L = 0$) for all workers, and the high skill investment cost be, for simplicity, of two values $E_{H,1} < E_{H,2}$; there are still two levels of productivity: y_L and y_H . Focus on the corresponding $\hat{\beta}$ and define it as $\hat{\beta} = \log \frac{y_H - y_L}{E_{H,2} - E_L}$. If the contracts can be contingent on $E_{H,1}$ and $E_{H,2}$, the submarkets for type $E_{H,1}$ workers and type $E_{H,2}$ workers are separated, and all the results in the paper carry through for the workers of cost $E_{H,2}$; as for the workers of cost $E_{H,1}$, their skill investment cost is lower, hence they have stronger incentive to remain high skilled; then for values of β close to $\hat{\beta} = \log \frac{y_H - y_L}{E_{H,2} - E_L}$, some equilibrium which exists in the $E_{H,2}$ submarket may not exist in the $E_{H,1}$ submarket. If the contracts can not be contingent on $E_{H,1}$ and $E_{H,2}$, both type $E_{H,1}$ and type $E_{H,2}$ are in the same market and will compete; as a result, there may exist a region of β where both high skilled group a and group b , as well as both low skilled group a and group b , exist at the same time. The extent of the skill investment game is in turn larger, because, for example, a particular group " $a, E_{H,1}$ "'s skill investment decision should be a best response of other groups: " $a, E_{H,2}$ ", " $b, E_{H,1}$ ", and " $b, E_{H,2}$ ". If heterogeneity in skill investment cost is managed, it is possible to extend the model to multiple skill levels. Shi (2006) shows that in such a model with multiple skill levels free of discrimination, the result that firms always rank the high skilled workers in priority to the workers with lower skills can be generalized. The difficulty under the context with discrimination, as just stated, is on the extent of the game.

6 Conclusion

¹¹See Merlino (2012) page 4 for more relevant reference.

In this paper, we study a holdup problem where firms can use discriminatory hiring norms to extract higher than socially optimal profits. We find that when firms rank workers according to both productivity-dependent and productivity-independent characteristics, skill investment becomes strategic between the discriminated and the favored group. In case wages are posted, we suggest that depending on the market tightness there may be equilibrium or multiple equilibria on skill investment; in some equilibrium the discriminated group can obtain higher expected payoff compared to the case without discrimination¹² and firms can be worse off. We also consider fixed sharing rule (bargained wage) and make a comparison. Similar equilibrium, where favored group underinvest while the discriminated group remain high skilled, exists; however, the discriminated group are in general worse off compared to the case without discrimination in the sense that they may still choose to underinvest when ψ is sufficiently high. Firms' profits are piecewise monotone because the skill composition hence the average productivity of the market improves discretely with respect to the bargaining power, and profit loss may be incurred with discrimination within an intermediate range of bargaining power.

References

- [1] Acemoglu, Daron & Shimer, Robert (1999a): "Efficient Unemployment Insurance," *Journal of Political Economy*, University of Chicago Press, vol. 107(5), pages 893-928, October.
- [2] Acemoglu, Daron & Shimer, Robert (1999b): "Holdups and Efficiency with Search Frictions," *International Economic Review*, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 40(4), pages 827-49, November.
- [3] Albrecht, James & Tan, Serene & Gautier, Pieter & Vroman, Susan (2004): "Matching with multiple applications revisited," *Economics Letters*, Elsevier, vol. 84(3), pages 311-314, September.
- [4] Brenzel, Hanna & Gartner, Hermann & Schnabel, Claus (2014): "Wage bargaining or wage posting? Evidence from the employers' side," In: *Labour Economics*, Vol. 29, August, S. 41-48.
- [5] Burdett, Kenneth & Shi, Shouyong & Wright, Randall (2001): "Pricing and Matching with Frictions," *Journal of Political Economy*, University of Chicago Press, vol. 109(5), pages 1060-1085, October.
- [6] Cao, Melanie & Shi, Shouyong (2000): "Coordination, matching, and wages," *Canadian Journal of Economics*, Canadian Economics Association, vol. 33(4), pages 1009-1033, November.
- [7] Clark, AE & Oswald, AJ (1996): "Satisfaction and comparison income," *Journal of public economics* 61 (3), 359-381.
- [8] Clark, AE (1997): "Job satisfaction and gender: why are women so happy at work?" *Labour economics* 4 (4), 341-372
- [9] Delacroix, Alain & Shi, Shouyong (2013): "Pricing and signaling with frictions," *Journal of Economic Theory*, Elsevier, vol. 148(4), pages 1301-1332.
- [10] Neal, Derek A & Johnson, William R. (1996): "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, University of Chicago Press, vol. 104(5), pages 869-95, October.

¹²Recall that without discrimination, it is socially optimal.

- [11] Fang, Hanming & Moro, Andrea (2010): "Theories of Statistical Discrimination and Affirmative Action: A Survey," NBER Working Papers 15860, National Bureau of Economic Research, Inc.
- [12] Galenianos, Manolis & Kircher, Philipp & Virág, Gábor (2011): "Market Power And Efficiency In A Search Model," *International Economic Review*, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 52(1), pages 85-103, 02.
- [13] Hall, Robert E. & Krueger, Alan B. (2012): "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search," *American Economic Journal: Macroeconomics*, American Economic Association, vol. 4(4), pages 56-67, October.
- [14] Heckman, James J. (1998): "Detecting Discrimination," *Journal of Economic Perspectives*, 12(2): 101-16.
- [15] Jacquet, Nicolas L. & Tan, Serene (2012): "Wage-vacancy contracts and coordination frictions," *Journal of Economic Theory*, Elsevier, vol. 147(3), pages 1064-1104.
- [16] O'Neil, June (1990): "The Role of Human Capital in Earnings Differences between Black and White Men," *Journal of Economic Perspectives*, American Economic Association, vol. 4(4), pages 25-45, Fall.
- [17] Philipp, Kircher, & Manolis, Galenianos (2009): "Directed Search with Multiple Job Applications," *Journal of Economic Theory*, 114(2), pp. 445-471.
- [18] Lang, Kevin & Michael, Manove (2003): "Wage announcements with a continuum of worker type," *Annales d'Economie et de Statistique*, ENSAE, issue 71-72, pages 223-244.
- [19] Lang, Kevin & William, T. Dickens (2005): "Racial Discrimination in Labor Markets with Posted Wage Offers," *American Economic Review*, 95 (4), 1327-1340.
- [20] Lang, Kevin & Michael, Manove (2011): "Education and Labor Market Discrimination," *American Economic Review*, 4 (101), 1467-96.
- [21] Lang, Kevin & Jee-Yeon K. Lehmann (2012): "Racial Discrimination in the Labor Market: Theory and Empirics," *Journal of Economic Literature*, American Economic Association, vol. 50(4), pages 959-1006, December.
- [22] LI, Fei & Can, Tian (2013): "Directed Search and Job Rotation," *Journal of Economic Theory*, Volume 148 (3), 1268-1281.
- [23] Merlino, Luca Paolo (2012): "Discrimination, Technology and Unemployment," *Labour Economics*, 19(4), pp. 557-567
- [24] Pendakur, Krishna & Woodcock, Simon (2010): "Glass Ceilings or Glass Doors? Wage Disparity Within and Between Firms," *Journal of Business & Economic Statistics*, American Statistical Association, vol. 28(1), pages 181-189.
- [25] Peters, Michael (2010): "Noncontractible Heterogeneity in Directed Search," *Econometrica*, Econometric Society, vol. 78(4), pages 1173-1200, 07.
- [26] Ritter, Joseph A. & Lowell, J. Taylor (2011): "Racial Disparity in Unemployment," *Review of Economics and Statistics*, 93 (1), 30-42

- [27] Rosen, Asa (1997): "An equilibrium search-matching model of discrimination," *European Economic Review*, Elsevier, vol. 41(8), pages 1589-1613, August.
- [28] Shimer, Robert (2005): "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy*, University of Chicago Press, vol. 113(5), pages 996-1025, October.
- [29] Shi, Shouyong (2001): "Frictional Assignment. I. Efficiency," *Journal of Economic Theory*, Elsevier, vol. 98(2), pages 232-260, June.
- [30] Shi, Shouyong (2002): "Product Market and the Size-Wage Differential," *International Economic Review*, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 43(1), pages 21-54, February.
- [31] Shi, Shouyong (2002): "A Directed Search Model of Inequality with Heterogeneous Skills and Skill-Biased Technology," *Review of Economic Studies* 69, 467-491.
- [32] Shi, S. (2006). Wage differentials, discrimination and efficiency. *Contributions to Economic Analysis*, 275, 83-115.
- [33] Tan, Serene (2012): "Directed Search and Firm Size," *International Economic Review*, Vol. 53(1), 95-113.
- [34] Virág, Gábor (2011): "High profit equilibria in directed search models," *Games and Economic Behavior*, Elsevier, vol. 71(1), pages 224-234, January.

Appendix

A1. Derivation of matching probabilities.

We now derive a job seeker's matching probability and expected payoff.

Job seekers. Having observed all the wage $\mathbf{w} = \{w^1, w^2, \dots, w^M\}$ announced by the firms, job seekers choose which firm (or wage) to visit (or to apply for). Consider a particular job seeker i 's problem, where $i \in \{1, 2, \dots, N\}$. This job seeker thinks in the following way: Suppose I visit firm j , then conditional on the fact that my application is sent to j , what is the probability that I could be employed? It depends upon the number of *the other job seekers* who also send their job application to the same firm competing with me on this job in firm j . This number (of *the other job seekers*) is a random variable which has a realisation from the set $\{0, 1, \dots, N-1\}$ and has a Binomial distribution. To see why it is the case, we use k to represent the realized number of competitors. If $k = 0$, which happens with probability $(1 - \theta^j)^{N-1}$, then the job seeker i will be chosen by the firm with probability 1, because this job seeker is the only candidate. If $k = 1$, which happens with probability $(N-1) \times (\theta^j)^1 (1 - \theta^j)^{(N-1)-1}$, this job seeker i will be chosen by the firm with probability $\frac{1}{2}$, because now the firm receives two applications, hence has two candidates, among whom i is one. Generalising, if $k = \hat{k}$, which happens with probability $C_{N-1}^{\hat{k}} \times (\theta^j)^{\hat{k}} (1 - \theta^j)^{N-1-\hat{k}}$, then this job seeker i will be chosen with probability $\frac{1}{\hat{k}+1}$, because the firm j has $\hat{k} + 1$ candidates at disposal.

The employment probability for the workers is $\sum_{k=0}^{N-1} C_{N-1}^k (\theta^j)^k (1 - \theta^j)^{N-1-k} \frac{1}{k+1}$. This expression could be simplified to $\frac{1 - (1 - \theta^j)^N}{N\theta^j}$ ¹³. Hence the job seeker's expected pay off is $\frac{1 - (1 - \theta^j)^N}{N\theta^j} \times w^j$.

A2. Proofs of propositions

Proposition 1. (return to skills) Given the return to skill ratio $\frac{y_H - y_L}{E_H - E_L}$, define $\hat{\beta}$ as $\frac{y_H - y_L}{E_H - E_L} = e^{\hat{\beta}}$.

(i) when $0 < \beta \leq \hat{\beta}$, the unique equilibrium is such that all job seekers choose high skill, i.e. $\alpha^* = 1$.

(ii) when $\beta > \hat{\beta}$, the equilibrium consists of a unique value $\alpha^* \in (0, 1)$ which satisfies $\frac{y_H - y_L}{E_H - E_L} = e^{\alpha^* \beta}$.

(iii) when $\frac{y_H - y_L}{E_H - E_L} \leq 1$ such that $\frac{y_H - y_L}{E_H - E_L} = e^{\hat{\beta}}$ has no positive solution on $\hat{\beta}$, the unique equilibrium is $\alpha^* = 0$.

Proof. We will prove only case (i) while the proof of case (ii) and (iii) are highly similar. Notice first that $\frac{y_H - y_L}{E_H - E_L} \geq e^\beta$ is equivalent to $e^{-\beta} y_H - E_H \geq e^{-\beta} y_L - E_L$.

We prove firstly that the deviation to low skill is not optimal. By this, we prove that a proportion ϵ of workers' deviating to low skill is suboptimal. And it suffices to show that after deviation, the deviator can not get higher expected payoff. Before deviation, the expected payoff is $e^{-q_H^*} y_H - E_H$, where $q_H^* = \beta$. After deviation, the expected payoff becomes $e^{-q_H^D - q_L^D} y_L - E_L$, where $q_H^D + q_L^D = \beta$. However, under the condition $e^{-\beta} y_H - E_H \geq e^{-\beta} y_L - E_L$, the expected payoff after deviation is weakly lower.

For the uniqueness. We should furthermore show that for the case of $\alpha = 0$ and $\alpha \in (0, 1)$, there will be profitable deviation. When $\alpha = 0$, the expected payoff from search is $e^{-\beta} y_L - E_L$. If there is a fraction ϵ deviating to high skill, then the expected income for the deviator becomes $e^{-\epsilon \beta} (y_H - y_L) + e^{-\beta} y_L - E_H$.

¹³One way of deriving it could be seen in Melanie Cao & Shouyong Shi, 2000. "Coordination, matching, and wages". It could also be checked by change of variable.

Then this expected payoff after deviation is greater than the expected payoff before deviation because. $e^{-\epsilon\beta}(y_H - y_L) > e^{-\beta}(y_H - y_L) \geq E_H - E_L$. So the deviation is profitable for the deviators. When $\alpha \in (0, 1)$, the expected income from search is $e^{-\beta}y_L - E_L$ for the low skilled, and $e^{-\hat{\alpha}\beta}(y_H - y_L) + e^{-\beta}y_L - E_H$ for the type H job seekers, where $\hat{\alpha}$ should be pinned down by workers' indifference condition $e^{-\hat{\alpha}\beta}(y_H - y_L) = E_H - E_L$. However, this condition is incompatible for any $\alpha < 1$ with our condition $e^{-\beta}y_H - E_H \geq e^{-\beta}y_L - E_L$. So that it is impossible that job seekers are indifferent from being high or low skilled.

All in all, we have proved that when the configuration of parameters is such that $e^{-\beta}y_H - E_H > e^{-\beta}y_L - E_L$, the only equilibrium is all the job seekers choose to obtain high skill, i.e. $\alpha^* = 1$. Q.E.D.

Proposition 2. The equilibrium labor allocation and skill investment choice are socially optimal.

Proof. We derive the proof for Notion 2, the proof for Notion 1 can be analogously derived. By definition of q , we have $q_H = \frac{N_H}{M}$ and $q_L = \frac{N_L}{M}$ - all firms will attract both skill types. Since $N_H = \alpha N$ and $N_L = (1 - \alpha)N$, we have $q_H = \alpha\beta$ and $q_L = (1 - \alpha)\beta$. When $\alpha = 1$ or $\alpha = 0$, there is only one skill level present in the market; when $\alpha \in (0, 1)$, there are both high and low skilled. Define a priority rule $R \in [0, 1]$, which is the probability of choosing high skilled job applicants when both high and low skilled are present in the same firm. The planner chooses α , which hence determines q_H and q_L , to maximize the aggregate output

$$M \times \left[\begin{array}{l} (1 - e^{-q_H})(R \times (1 - e^{-q_L}) + e^{-q_L})y_H \\ + (1 - e^{-q_L})((1 - R)(1 - e^{-q_H}) + e^{-q_H})y_L \\ - N \times (\alpha E_H + (1 - \alpha)E_L) \end{array} \right]$$

The objective program can be rearranged to the following way

$$M \times \left[\begin{array}{l} (1 - e^{-q_L})(1 - e^{-q_H}) \times [Ry_H + (1 - R)y_L] \\ + e^{-q_L}y_H + e^{-q_H}y_L - e^{-\beta}y_H - e^{-\beta}y_L \\ - N \times (\alpha E_H + (1 - \alpha)E_L) \end{array} \right]$$

Whenever we are with corner solutions on α , i.e. $q_H = 0$ or $q_L = 0$, the expression does not depend on R . When the solution on α is interior, q_H and q_L are both positive. And if we maximize with respect to R , we have $(1 - e^{-q_H})(1 - e^{-q_L})(y_H - y_L) > 0$, so that setting $R^p = 1$ is the optimal choice. By doing so, we could further reduce the objective to

$$M \times [(1 - e^{-q_H})y_H + e^{-q_H}(1 - e^{-q_L})y_L] \\ - N \times (\alpha E_H + (1 - \alpha)E_L)$$

Now the derivative with respect to α yields

$$\begin{array}{l} \overset{=N}{\widehat{M}\beta} \times [e^{-\alpha\beta}(y_H - y_L)] \\ - N \times (E_H - E_L) \end{array}$$

where the first line represents the marginal gain from assigning 1% more workers to the high skilled section, and the second line represents the corresponding marginal loss. For values of β such that the marginal gain surpasses the marginal loss, the planner will set $\alpha = 1$, in which case a threshold $\hat{\beta}^p$ is determined by $e^{\hat{\beta}^p}(y_H - y_L) = E_H - E_L$, such that for values of β not larger than this threshold pinned down by the log-return to skills, the planner finds socially optimal to assign all workers to the high skilled sector. When the skill to return is such that $(y_H - y_L) < E_H - E_L$, for all values of β the marginal gain will be lower than the marginal loss, the planner will choose $\alpha^p = 0$. At last if β satisfies $\beta > \hat{\beta}^p$, such that

the log-return to skill investment is not high enough to offset the market competition (captured by β), there is an $\alpha^p \in (0, 1)$ (an interior solution) which is determined by $e^{-\alpha^p \beta} (y_H - y_L) = E_H - E_L$ such that the planner will recommend ex ante identical workers to randomize on skill choice and a fraction α^p will end up high skilled. It is straightforward to notice that the correspond values of $q_H^p = \alpha^p \beta$ and $q_L^p = (1 - \alpha^p) \beta$ correspond to the equilibrium allocation.

As a summary, we have shown that the threshold for skill investment $\hat{\beta}^p$ coincides with $\hat{\beta}$, and α^p coincides with α^* . Q.E.D.

Proposition 3: Under the above assumptions, there exist two thresholds $\hat{\beta}_2$ and $\hat{\beta}_1$ with $0 < \hat{\beta}_2 < \hat{\beta}_1 < \hat{\beta}$, such that

(1) *When $0 < \beta < \hat{\beta}_2$, there exists a unique equilibrium in which both group a and group b invest in high skill, (aH, bH). At $\beta = \hat{\beta}_2$, both (aH, bH) and (aH, bL) can be equilibrium.*

(2) *When $\hat{\beta}_2 < \beta < \hat{\beta}_1$, there exists a unique equilibrium in which group a invest in high skill while group b in low skill, (aH, bL). At $\beta = \hat{\beta}_1$, both (aL, bH) and (aH, bL) can be equilibrium.*

(3) *When $\hat{\beta}_1 < \beta \leq \hat{\beta}$, there exist multiple equilibria. Either group a invest in high skill and group b in low skill, or group a invest in low skill and group b invest in high skill, (aL, bH) or (aH, bL).*

Proof. Holding group b high skilled, group a's best response depends on the comparison between $U_{aH}^{P1}(\beta)$ and $U_{aL}^{P3}(\beta)$. On one hand, since $q_{aH}^{P1}(\beta) > \beta$, we have

$$U_{aH}^{P1}(\hat{\beta}) = e^{-q_{aH}^{P1}(\hat{\beta})} y_H - E_H < e^{-\hat{\beta}} y_H - E_H = e^{-\hat{\beta}} y_L - E_L = U_{aL}^{P3}(\hat{\beta})$$

On the other hand, since q_{aH}^{P1} is increasing in β , when $\beta \rightarrow 0$, we have $U_{aH}^{P1}(\beta) \rightarrow y_H - E_H$ which is greater than $U_{aL}^{P3}(0) = y_L - E_L$. Due to the continuity of $U_{aH}^{P1}(\beta)$ and $U_{aL}^{P3}(\beta)$ on β , there exists at least one $\hat{\beta}_1 < \hat{\beta}$ such that $U_{aH}^{P1}(\hat{\beta}_1) = U_{aL}^{P3}(\hat{\beta}_1)$. According to the assumptions on the uniqueness of the intersection point, we have $U_{aH}^{P1}(\beta) > U_{aL}^{P3}(\beta)$ for $\beta < \hat{\beta}_1$, and $U_{aH}^{P1}(\beta) < U_{aL}^{P3}(\beta)$ for $\beta > \hat{\beta}_1$.

Holding group b low skilled, group a's best response depends on the comparison between U_{aH}^{P2} and U_{aL}^{P4} . It turns out that for all values of $\beta < \hat{\beta}$

$$U_{aH}^{P2} = e^{-\gamma \beta} (y_H - y_L) + e^{-\beta} y_L - E_H > e^{-\beta} y_H - E_H > e^{-q_{aH}^{P4}} y_L - E_L = U_{aL}^{P4}$$

where the last inequality is due to the fact that $e^{-\beta} y_H - E_H \geq e^{-\beta} y_L - E_L$ for all $\beta \leq \hat{\beta}$ and $q_{aH}^{P4} > \beta$, implying that choosing high skill is a dominant strategy for group a when group b choose low skill.

Holding group a high skilled, group b's best response depends on the comparison between U_{bH}^{P1} and U_{bL}^{P2} . On one hand, we have

$$U_{bH}^{P1}(\hat{\beta}) = \frac{1 - e^{-q_{bH}^{P1}(\hat{\beta})}}{q_{bH}^{P1}(\hat{\beta})} e^{-q_{aH}^{P1}(\hat{\beta})} y_H - E_H < e^{-\hat{\beta}} y_H - E_H = e^{-\hat{\beta}} y_L - E_L = U_{bL}^{P2}(\hat{\beta})$$

where the first inequality is due to the fact that $\frac{1 - e^{-q_{bH}^{P1}(\hat{\beta})}}{q_{bH}^{P1}(\hat{\beta})} e^{-q_{aH}^{P1}(\hat{\beta})} < e^{-q_{aH}^{P1}(\hat{\beta})}$ and $q_{aH}^{P1}(\beta) > \beta$. On the other hand, since $q_{aH}^{P1}(\beta)$ and $q_{bH}^{P1}(\beta)$ are increasing in β , when $\beta \rightarrow 0$, we have $U_{bH}^{P1}(\beta) \rightarrow y_H - E_H$ which is greater than $U_{bL}^{P2}(0) = y_L - E_L$. Due to the continuity of $U_{bH}^{P1}(\beta)$ and $U_{bL}^{P2}(\beta)$ on β , there exists at least one $\hat{\beta}_2 < \hat{\beta}$ such that $U_{bH}^{P1}(\hat{\beta}_2) = U_{bL}^{P2}(\hat{\beta}_2)$. According to the assumptions on the uniqueness of the intersection point, we have $U_{bH}^{P1}(\hat{\beta}_2) > U_{bL}^{P2}(\hat{\beta}_2)$ for $\beta < \hat{\beta}_2$, and $U_{bH}^{P1}(\hat{\beta}_2) < U_{bL}^{P2}(\hat{\beta}_2)$ for $\beta > \hat{\beta}_2$.

Holding group a low skilled, group b's best response depends on the comparison between U_{bH}^{P3} and U_{bL}^{P4} . It turns out that for all values of $\beta \leq \hat{\beta}$ we have

$$\begin{aligned}
U_{bH}^{P3} &= e^{-(1-\gamma)\beta} (y_H - y_L) + e^{-\beta} y_L - E_H \\
&> e^{-\beta} y_H - E_H \\
&\geq e^{-\beta} y_L - E_L \\
&> \frac{1 - e^{-q_{bL}^{P4}(\beta)}}{q_{bL}^{P4}(\beta)} e^{-q_{aL}^{P4}(\beta)} y_L - E_L
\end{aligned}$$

where the last inequality is due to the fact that $e^{-\beta} > e^{-q_{aL}^{P4}} > \frac{1 - e^{-q_{bL}^{P4}(\beta)}}{q_{bL}^{P4}(\beta)} e^{-q_{aL}^{P4}(\beta)}$, implying that choosing high skill is a dominant strategy for group b when group a choose low skill.

To summarize, for values of $\beta \in (0, \hat{\beta}_2)$, both groups choosing high skill, i.e. (aH, bH) , is the unique equilibrium; for $\beta \in (\hat{\beta}_2, \hat{\beta}_1)$, group a choosing high skill and group b choosing low skill, i.e. (aH, bL) is the unique equilibrium; When $\beta \in (\hat{\beta}_1, \hat{\beta})$, both (aH, bL) and (aL, bH) are possible to appear as equilibrium. Q.E.D.

Corollary. Compared to the case without discrimination,

(1) In (aH, bH) equilibrium, firms always earn higher expected profits; in (aH, bL) equilibrium and (aL, bH) , firms earn lower expected profits.

(2) In (aH, bH) equilibrium, both aH and bH workers earn lower expected payoff; in (aH, bL) equilibrium, group aH (group bL) earn higher (lower) expected payoff; in (aL, bH) equilibrium, group bH (group aL) earn higher (lower) expected payoff.

Proof. (1) In (aH, bH) equilibrium, the result is proved in LMD (2005). We prove the case for (aH, bL) equilibrium. In (aH, bL) equilibrium, firms' profits are

$$\pi_{aH, bL}^{P2} = \left(1 - e^{-q_H^{P2}} - q_H^{P2} e^{-q_H^{P2}}\right) y_H + \left[(q_H^{P2} + 1) e^{-q_H^{P2}} - e^{-\beta} (\beta + 1)\right] y_L$$

This term is smaller than the profit without discrimination, because

$$\begin{aligned}
\pi_{aH, bL}^{P2} &= \left(1 - e^{-q_H^{P2}} - q_H^{P2} e^{-q_H^{P2}}\right) y_H + \left[(q_H^{P2} + 1) e^{-q_H^{P2}} - e^{-\beta} (\beta + 1)\right] y_L \\
&< \left(1 - e^{-q_H^{P2}} - q_H^{P2} e^{-q_H^{P2}}\right) y_H + \left[(q_H^{P2} + 1) e^{-q_H^{P2}} - e^{-\beta} (\beta + 1)\right] y_H \\
&= (1 - e^{-\beta} - e^{-\beta} \beta) y_H
\end{aligned}$$

where the inequality uses the fact that $(x + 1) e^{-x}$ is a decreasing function and $q_H^{P2} < \beta$. The proof for the case of (aL, bH) equilibrium can be analogously reproduced.

(2) For the case (aH, bH) , it follows from LMD (2005). For the case of (aH, bL) . We have

$$\begin{aligned}
U_{aH}^{P2} &= e^{-\gamma\beta} (y_H - y_L) + e^{-\beta} y_L - E_H \\
&> e^{-\beta} y_H - E_H \\
&> e^{-q_{aH}^S} y_H - E_H
\end{aligned}$$

where the first inequality comes from $e^{-\gamma\beta} - e^{-\beta} > 0$.

$$\text{and } U_{bL}^{P2} = \begin{aligned} &e^{-\beta} y_L - E_L \\ &< e^{-\beta} y_H - E_H \end{aligned}$$

where the inequality comes from the fact that $\beta \leq \hat{\beta}$. The proof for the case of (aL, bH) equilibrium can be analogously reproduced. **Q.E.D.**

Proposition 4. Let $\gamma \geq \frac{1}{2}$ so that group a is the majority. There are four thresholds $\hat{\psi}_{aL,b} \leq \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$, such that

(1) For $\psi \in [0, \hat{\psi}_{aL,b})$, the unique pure strategy nash equilibrium is (aL, bL) ; for $\psi = \hat{\psi}_{aL,b}$, the equilibrium can be (aL, bL) or (aL, bH) . (2) For $\psi \in (\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a})$, the unique pure strategy nash equilibrium is (aL, bH) ; (3) For $\psi \in (\hat{\psi}_{bH,a}, \hat{\psi}_{bL,a})$, there is no pure strategy nash equilibrium; for $\psi = \hat{\psi}_{bL,a}$, the equilibrium is (aH, bL) ; (4) For $\psi \in (\hat{\psi}_{bL,a}, \hat{\psi}_{aH,b})$, the unique pure strategy nash equilibrium is (aH, bL) ; at the point $\psi = \hat{\psi}_{aH,b}$, the equilibrium can be (aH, bL) or (aH, bH) . (5) For $\psi \in (\hat{\psi}_{aH,b}, 1)$, the unique pure strategy nash equilibrium is (aH, bH) . (6) Define the threshold $\hat{\psi}$ of skill investment without discrimination as $\hat{\psi}y_H \frac{1-e^{-\beta}}{\beta} - E_H = \hat{\psi}y_L \frac{1-e^{-\beta}}{\beta} - E_L$; then $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$.

Proof. The payoff matrix is as follows.

	bH	bL
aH	$bH: e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H$ $aH: \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H$	$bL: e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$ $aH: \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H$
aL	$bH: \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H$ $aL: e^{-(1-\gamma)\beta} \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$	$bL: e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$ $aL: \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$

Define $\hat{\psi}_{aH,b}$ by $e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H = e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$.

Define $\hat{\psi}_{aL,b}$ by $\frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H = e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$.

Define $\hat{\psi}_{bH,a}$ by $\frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H = e^{-(1-\gamma)\beta} \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$.

Define $\hat{\psi}_{bL,a}$ by $\frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H = \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$.

When $\gamma \geq \frac{1}{2}$, it can be verified that $\hat{\psi}_{aL,b} \leq \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$. (1) We first prove for values of $\psi \in [0, \hat{\psi}_{aL,b})$, the unique pure strategy Nash equilibrium is (aL, bL) : group a choose low skill, group b choose low skill. Holding group b high skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{aL,b} < \hat{\psi}_{bH,a}$. Holding group b low skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{aL,b} < \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{aL,b} < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be low skilled, because $\psi < \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{aL,b}$, the equilibrium can be (aL, bH) or (aL, bL) .

(2) We prove for values of $\psi \in (\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a})$, the unique Nash pure strategy equilibrium is (aL, bH) . Holding group b high skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{bH,a}$. Holding group b low skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{bH,a} < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{bH,a}$, the unique equilibrium is (aL, bH) .

(3) We prove for values of $\psi \in (\hat{\psi}_{bH,a}, \hat{\psi}_{bL,a})$, there is no pure strategy Nash equilibrium is (aL, bH) . Holding group b high skilled, group a choose to be high skilled, because $\psi > \hat{\psi}_{bH,a}$. Holding group b low

skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{bL,a}$, the unique equilibrium is (aH, bL) .

(4) We prove for values of $\psi \in (\hat{\psi}_{bL,a}, \hat{\psi}_{aH,b})$, there is a unique pure strategy Nash equilibrium (aH, bL) . Holding group b high skilled, group a choose to be high skilled, because $\psi > \hat{\psi}_{bH,a}$. Holding group b low skilled, group a choose to be high skilled, because $\psi > \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{aH,b}$, the equilibrium can be (aH, bL) or (aH, bH) .

(5) We prove for values of $\psi \in (\hat{\psi}_{aH,b}, 1)$, there is a unique pure strategy Nash equilibrium (aH, bH) . Holding group b high skilled, group a choose to be high skilled. Holding group b low skilled, group a choose to be high skilled. Holding group a high skilled, group b choose to high skilled, because $\psi < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$.

(6) At last, notice that $\hat{\psi} = \frac{E_H - E_L}{y_H - y_L} \times \left(\frac{1 - e^{-\beta}}{\beta}\right)^{-1}$, $\hat{\psi}_{bL,a} = \frac{E_H - E_L}{y_H - y_L} \times \left(\frac{1 - e^{-\gamma\beta}}{\gamma\beta}\right)^{-1}$, and $\hat{\psi}_{aH,b} = \frac{E_H - E_L}{y_H - y_L} \times \left(e^{-\gamma\beta} \frac{1 - e^{-(1-\gamma)\beta}}{(1-\gamma)\beta}\right)^{-1}$. It is straightforward to verify that $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$ because $\left(\frac{1 - e^{-\gamma\beta}}{\gamma\beta}\right)^{-1} < \left(\frac{1 - e^{-\beta}}{\beta}\right)^{-1} < \left(e^{-\gamma\beta} \frac{1 - e^{-(1-\gamma)\beta}}{(1-\gamma)\beta}\right)^{-1}$. Q.E.D.