Wage Differentials, Discrimination and Efficiency*

Shouyong Shi

Department of Economics

University of Toronto

150 St. George Street, Toronto,

Ontario, Canada, M5S 3G7

(email: shouyong@chass.utoronto.ca).

2004

Abstract

In this paper I construct a search model of a large labor market in which workers are heterogeneous in productivity and (homogeneous) firms post wages and a ranking of workers to direct workers' search. I establish the following results. First, the wage differential is *negatively* related to productivity when the productivity differential is small, while a positive relationship emerges when the productivity differential is large. Second, as the productivity differential decreases to zero, the reverse wage differential increases and so it remains strictly positive in the limit. Third, high-productivity workers are not discriminated against even when they have a lower wage, because they always have a higher priority in employment and higher expected wage than low-productivity workers. Fourth, the equilibrium is socially efficient, and so the wage differential and the ranking are part of the efficient mechanism. Finally, I provide numerical examples to illustrate the wage distribution.

JEL classification: J3, J6, J7
Keywords: Search; Wage Differential; Discrimination.

1. Introduction

Standard economic theories view wage differentials as a compensation for workers' human capital or productivity. These theories have encountered great difficulties in explaining the large wage differentials in the US data. For example, Juhn et al. (1993) have found that all observable characteristics of workers' productivity, such as education, experience and age, can explain only one-third of the differential between the ninetieth and the tenth percentile of the wage distribution between 1963 and 1989. On the other hand, wage differentials seem to depend statistically on seemingly irrelevant background characteristics, such as workers' race, gender, and height. To allow for this anomalous dependence, the standard theory attributes it to discrimination.[1]

Given these difficulties, it is useful to explore alternative theories of the labor market. In a seminal paper, Mortensen (1982) provided one such alternative that emphasizes search frictions in the labor market. He characterized an efficient compensation scheme in a class of frictional markets and showed that prices (or wages) may depart significantly from Walrasian prices in order to achieve efficiency. Now, there is a large and still growing literature that explores the importance of search frictions in the labor market (see Mortensen, 2002, for the references).

The current paper follows this line of research. The main purpose is to illustrate that wage differentials are sometimes a bad indicator of productivity differentials and discrimination. I will establish the following results. First, a wage differential can be *negatively* related to the productivity differential and the size of this reverse wage differential *increases* when workers become more and more similar in productivity. Thus, even when the productivity differential approaches zero, a wage differential still exists in equilibrium. Second, in contrast to the standard interpretation, the reverse wage differential is not discrimination. On the contrary, higher productivity is always rewarded with a higher expected wage, which takes workers' employment probability

---

[1]See Altonji and Blank (1999) for a survey of the facts and the literature on discrimination in the labor market. A popular spinoff is the theory of statistical discrimination. It argues that when firms are uncertain about workers' fundamental characteristics, discrimination can be an equilibrium outcome, either because the characteristics on which discrimination is based are correlated with workers' fundamental characteristics, or because discrimination leads to self-fulfilling separation of worker types.

into account. Third, the equilibrium is socially efficient, and so the differentials in wages and employment probabilities are part of the efficient mechanism.

The model is one with directed search. It can be best described for the case where there are only two types of workers. A worker's type is determined by one observable skill, which I call productivity. The difference in productivity among workers can be very small. All firms are identical. They simultaneously post wages and ranking schemes for the workers. Each firm can post different wages for different types of workers, but is restricted to give identical workers the same ranking and the same wage. After observing firms' announcements, workers decide which firm to apply and they cannot coordinate their applications. After receiving the applicants, each firm selects one worker according to the ranking scheme and pays the posted wage. Search is directed because, when choosing the wage and the ranking scheme, each firm takes into account the effect of its announcements on the matching probability.

I show that there is a unique symmetric equilibrium where identical workers use the same application strategies. Every firm attracts both types of workers with positive probability, and so separation of the two types is not an equilibrium. Moreover, every firm gives high-productivity workers the priority in employment whenever the firm receives both types of workers. However, when the productivity differential is small, this employment advantage comes with a lower wage.

The reverse wage differential arises from the trade-off between the employment probability and wage. Workers care about the expected wage, which is the employment probability times the wage. By giving high-productivity workers a higher ranking, a firm can lower the actual wage by a discrete amount for these workers and yet still be able to attract them. The combination of a high ranking and low wage is optimal for a firm, because it enables the firm to increase the utilization of high-productivity workers. The combination is also attractive to high-productivity workers, provided that the combination yields higher expected wage. Indeed, the expected wage is always higher for high-productivity workers than for low-productivity workers.

Despite the non-standard features of actual wages, I show that the equilibrium is socially

efficient in the following sense: If a social planner tries to maximize expected aggregate output under the constraint of the same matching function as the one generated in the equilibrium, then the planner will choose the same allocation between workers and firms as in the equilibrium. The efficiency result is in accordance with Mortensen's (1982) general result on efficiency, in the sense that each worker's expected wage in the current model takes into account the expected crowding-out that the particular worker creates on other workers (see also Hosios, 1990). Notice that efficiency entails both the ranking of workers and the wage differential.

I extend the model to a market where there are many types of workers, characterize the equilibrium and use numerical examples to illustrate the equilibrium wage distribution.

This paper belongs to the search literature (see Mortensen, 2002). Most of the papers in this literature assume that search is not directed. For models of directed search, see Peters (1991), Acemoglu and Shimer (1999), Burdett et al. (2001), Julien et al. (2000), and Shi (2001, 2002).[2] These models have either homogeneous agents on both sides of the market, or heterogeneous agents on both sides of the market who are complementary with each other in production. The model in the current paper lies somewhere in between – it has heterogeneous workers and identical firms. Shimer (1997) constructs a model similar to mine with two types of workers. Although our results overlap to some extent, his focus is to contrast the effects of different mechanics of wage determination on the division of the match surplus. In particular, he does not emphasize the reverse wage differential that can arise in the directed search environment.

Search models are used to examine discrimination by Black (1995) and Bowlus and Eckstein (2002). They show that if some firms have prejudice against a subset of workers, then the search cost will support a wage differential in the steady state. In contrast to my model, search in these models is not directed. More importantly, my model does not rely on the exogenous prejudice to generate a large wage differential among similar workers. On the issue of discrimination, the model most closely related to mine is Lang et al. (2002). An important difference is that they restrict each firm to post only one wage for all workers whom the firm tries to attract; i.e., the

---

[2]The model of competitive search by Moen (1997) also falls into this category.

firm is not allowed to post different wages for different workers. As a result, the two types of workers are completely separated in their model. Such separation is not an equilibrium and is not efficient when each firm can post a different wage for each type of workers.

I will organize the paper as follows. In Section 2, I will describe the simple model with two types of workers and propose a candidate equilibrium. Section 3 will show that the candidate equilibrium is indeed the unique equilibrium. In Section 4, I will examine the properties of the equilibrium and show that the equilibrium is socially efficient. Section 5 will extend the model to incorporate many types of workers. I will then conclude in Section 6 and supply the necessary proofs in the Appendix.

2. The Model

2.1. Workers and Firms

Consider a labor market with a large number of workers, N. There are two types of workers, type T and type S. Type T workers are a fraction γ of all workers and type S a fraction $(1-\gamma)$, where $\gamma \in (0,1)$. Sometimes, I use the notation $\gamma_T = \gamma$ and $\gamma_S = 1 - \gamma$. A worker's type is observed immediately upon applying to a job. A type S worker produces y units of output and a type T worker produces $(1+\delta)y$, where $\delta > 0$. In a large part of the analysis in this paper, I will restrict δ to be sufficiently small. The purpose is to examine whether a small productivity difference can generate a large wage differential.

There are also a large number of firms, M, all of which are identical. For the moment, this number is fixed. Competitive entry of firms can be introduced easily and will be briefly discussed at the end of section 5.1. Each firm wants to hire only one worker. Denote the tightness of the market as $\theta = N/M$.

The recruiting game is as follows. First, all firms make their announcements simultaneously. Each firm i announces two wages, w_{iT} for type T workers and w_{iS} for types S workers, together with a rule that ranks the two types of applicants. Let $R_i \in \{1, 0, \Phi\}$ denote this ranking or priority rule, where $\Phi = [0,1]$. The firm selects type T workers first if it sets $R_i = 1$ and type

S workers first if it sets $R_i = 0$; If $R_i = \Phi$, the firm is indifferent between the two types of workers.[3] Once a firm posts the wages and the selection rule, it is committed to them. All workers observe all announcements and then decide which firm to apply to. This application decision can possibly be mixed strategies over the jobs. Let α_{ij} denote the probability with which a type j worker applies to firm i. After receiving the applicants, a firm selects a worker according to the announced ranking and pays the corresponding wage. The worker produces immediately, obtains the wage and the game ends.

Notice that a worker's strategy does not depend on the worker's identity. Thus, all workers of the same type must use the same strategy. This symmetry requirement on the workers' side reflects the realistic feature of the labor market that workers cannot coordinate their application decisions.[4] However, the coordination failure could be eliminated if firms could identify each worker and make an offer specifically to that worker. To preserve the coordination failure, I assume that each firm's offer and, in particular, the ranking rule should not depend on the identities of the workers whom the firm receives. This does not mean that all firms must use the same strategy. To the contrary, a firm's strategy is allowed to depend on the firm's identity. This allowance is necessary for examining the possibility of a separating equilibrium where two groups of firms each attract a distinct type of workers.

I will focus on the limit of the economy where N and M approach infinity while their ratio, θ, lies in the interior of $(0, \infty)$. The equilibrium in this limit is significantly easier to characterize than the finite economy, because a single firm's deviation does not affect workers' payoff from applying to other firms. More precisely, let p_{ij} be the probability with which a type j worker

---

[3]The ranking scheme is included in a firm's announcement to ease the description. However, firms do not have to post the ranking literally. For the model to work, all that is needed is that workers expect the firms to use the ranking scheme to select workers after workers apply. This expectation will be fulfilled since the ranking scheme in the equilibrium is compatible with the firms' ex post incentive.

[4]The set of asymmetric equilibria is large. In a model with two identical agents on each side of the market, Burdett et al. (2001) have shown that there are a continuum of asymmetric equilibria while there is a unique symmetric equilibrium.

who applies to a firm i gets the job. Define the "market wage" of a type j worker as follows:

$$E_j = \max_{i' \neq i} \left(p_{i'j} w_{i'j} \right).$$

(2.1)

In the limit described above, the effect of firm i's strategy on $p_{i'j}$ approaches zero (see Burdett et al., 2001). Thus, each firm takes E_j as given.

A worker maximizes the expected wage that he can obtain from applying to a job. Given E_j, a type j worker's strategy is to choose $\alpha_{ij} = 1$ if $p_{ij} w_{ij} > E_j$, $\alpha_{ij} = 0$ if $p_{ij} w_{ij} < E_j$, and $\alpha_{ij} \in (0,1)$ if $p_{ij} w_{ij} = E_j$. In the limit economy, it is convenient to express this strategy with a new variable $q_{ij} = \gamma_j N \alpha_{ij}$. Then,

$$q_{ij} \begin{cases} = \infty, & \text{if } p_{ij} w_{ij} > E_j \\ = 0, & \text{if } p_{ij} w_{ij} < E_j \\ \in (0, \infty), & \text{if } p_{ij} w_{ij} = E_j. \end{cases}$$

(2.2)

Since the sum of α_{ij} over i is one, then q_{ij} must satisfy:

$$\frac{1}{M} \sum_{i=1}^{M} q_{ij} = \gamma_j \theta, \quad j = T, S.$$

(2.3)

When all type j workers use the strategy q_{ij}, the expected number of type j applicants received by firm i is $\gamma_j N \alpha_{ij} = q_{ij}$. For this reason, I call q_{ij} the *queue length* of type j workers for firm i. Despite the coincidence between a worker's strategy and the queue length, one should not construed an individual choice of q as the worker's ability to influence other workers' or firms' decisions. When $q_{ij} > 0$, I say that the firm i *attracts* type j workers.

In the limit where the economy becomes infinitely large, the probability with which firm i attracts one or more type j worker is $1 - (1 - \gamma_j \alpha_{ij})^N \rightarrow 1 - e^{-q_{ij}}$. If firm i gives type j workers the selection priority, then the employment probability of a type j worker at the firm is:

$$\frac{1 - (1 - \gamma \alpha_{ij})^N}{N \gamma \alpha_{ij}} \rightarrow \frac{1 - e^{-q_{ij}}}{q_{ij}} \equiv G(q_{ij}).$$

On the other hand, if firm i give the other type $j' \neq j$ the priority, then the firm will consider a type j worker only if the firm receives no type j' applicants. This event occurs with probability

$e^{-q_{ij'}}$, in which case each type j worker applying to the firm is chosen with probability $G(q_{ij})$.

Thus, for a general priority rule R_i, a worker's employment probability at firm i is:

$$
\begin{aligned}
p_{iT} &= \left[R_i + (1 - R_i)e^{-q_{iS}}\right] G(q_{iT}), \\
p_{iS} &= \left[1 - R_i + R_i e^{-q_{iT}}\right] G(q_{iS}).
\end{aligned}
\tag{2.4}
$$

Notice that the function G is continuous and decreasing. Thus, when there are slightly more workers applying to a firm, each applicant is chosen by the firm with a slightly lower probability.

Now, consider a firm i's choices of (w_{iT}, w_{iS}, R_i). The firm's expected profit is:

$$
\begin{aligned}
\pi_i =\ & (1 - e^{-q_{iT}}) \left[R_i + (1 - R_i)e^{-q_{iS}}\right] \left[(1 + \delta)y - w_{iT}\right] \\
& + (1 - e^{-q_{iS}}) \left[1 - R_i + R_i e^{-q_{iT}}\right] (y - w_{iS}).
\end{aligned}
\tag{2.5}
$$

The firm's optimal choices solve the following problem:

$$
\max\ \pi_i \quad \text{subject to (2.2) for } j = T, S.
$$

The constraint (2.2) reflects the fact that the firm takes into account the effect of its choices on workers' decisions. Finally, the firm's ranking rule must be compatible with the firm's ex post incentive. That is, the following condition must hold:

$$
R_i = \begin{cases}
1, & \text{if } w_{iT} - w_{iS} < \delta y \\
0, & \text{if } w_{iT} - w_{iS} > \delta y \\
\Phi, & \text{if } w_{iT} - w_{iS} = \delta y.
\end{cases}
\tag{2.6}
$$

A (symmetric) equilibrium is defined as firms' strategies $(w_{iT}, w_{iS}, R_i)_{i=1}^{M}$, workers' strategies $(q_{iT}, q_{iS})_{i=1}^{M}$, and the numbers (E_T, E_S) such that the following requirements are met: (i) Given the firms' strategies and the numbers (E_T, E_S), each worker's strategy is given by (2.2); (ii) Given the numbers (E_T, E_S) and anticipating workers' responses, each firm's strategy is optimal and the ranking is compatible with the firm's incentive; and (iii) the numbers (E_T, E_S) obey (2.1).

2.2. A Candidate Equilibrium

Let me construct a candidate equilibrium. In this equilibrium, every firm ranks type T workers first and every firm attracts both types of workers. That is, $R_i = 1$, $q_{iT} > 0$ and $q_{iS} > 0$ for all i. Restricting the strategy to this particular type, a firm's maximization problem simplifies to:

$$
(P)\ \max_{(w_{iT}, w_{iS})} \pi_i = (1 - e^{-q_{iT}}) \left[(1 + \delta)y - w_{iT}\right] + e^{-q_{iT}} (1 - e^{-q_{iS}}) (y - w_{iS}),
\tag{2.7}
$$

7

subject to $G(q_{iT})w_{iT} \geq E_{iT}$ and $e^{-q_{iT}}G(q_{iS})w_{iS} \geq E_{iS}$. These constraints are necessary for $q_{iT} > 0$ and $q_{iS} > 0$ (see (2.2)) which are stipulated for the candidate equilibrium. Notice that both constraints must hold with equality. If one holds with strict inequality " $>$ ", then $q_i \to \infty$ and $G(q_i) \to 0$ which violate the corresponding constraint.

Using the constraints to substitute (w_{iT}, w_{iS}), I can write the firm's expected profit as follows:

$$\pi_i = \left(1 - e^{-q_{iT}}\right)(1 + \delta)y + e^{-q_{iT}}\left(1 - e^{-q_{iS}}\right)y - \left(q_{iT}E_{iT} + q_{iS}E_{iS}\right). \tag{2.8}$$

The first-order conditions for (q_{iT}, q_{iS}) yield:

$$ye^{-(q_{iT}+q_{iS})} = E_S, \tag{2.9}$$

$$\delta ye^{-q_{iT}} + E_S = E_T. \tag{2.10}$$

Since the solution to these equations does not depend on the firm index i, all firms use the same strategy in the candidate equilibrium. In this case, $\alpha_{iT} = \alpha_{iS} = 1/M$, which implies the following queue lengths:

$$q_T = \gamma\theta, \quad q_S = (1 - \gamma)\theta. \tag{2.11}$$

Finally, the first-order conditions and the constraints in problem (P) become:

$$E_S = ye^{-\theta}, \tag{2.12}$$

$$E_T = E_s + \delta ye^{-\gamma\theta} = ye^{-\theta}\left[1 + \delta e^{(1-\gamma)\theta}\right], \tag{2.13}$$

$$w_T = y\frac{\gamma\theta\left[\delta + e^{-(1-\gamma)\theta}\right]}{e^{\gamma\theta} - 1}, \quad w_S = y\frac{(1-\gamma)\theta}{e^{(1-\gamma)\theta} - 1}. \tag{2.14}$$

Proposition 2.1. *The market has a unique (symmetric) equilibrium, which is characterized by the following properties: (E1) All firms have the same strategy and rank type T workers above type S workers (i.e., $R = 1$); (E2) Each firm posts w_T for type T workers and w_S for type S workers, as given by (2.14); (E3) Each worker applies to every firm with the same probability, which yields the queue lengths (q_T, q_S) in (2.11) for every firm; and (E4) workers' expected wages satisfy (2.12) and (2.13).*

Let me discuss (2.8), (2.12) and (2.13), which will be useful later. Because all firms use the same strategy, I will suppress the firm's index i in this discussion. First, the condition (2.8) says that a firm's expected profit is equal to the difference between expected output and expected wage cost. Notice that, although a firm hires only one worker, the expected wage cost on type j workers is $q_j E_j$, as if the firm hires a number q_j of such workers at a wage rate E_j.

Second, the conditions (2.9) and (2.10) state that a worker's expected contribution to output, after subtracting the amount of other workers' expected output crowded out by this worker, is equal to the worker's expected market wage. Since adding a type S worker contributes to a firm's output only when the firm did not receive any other applicant, which occurs with probability $e^{-(q_T+q_S)}$, a type S worker's contribution to expected output is $ye^{-(q_T+q_S)}$. Similarly, a type T worker contributes to a firm's output by an amount $(1+\delta)y$ if the firm did not receive any other applicant, by δy if the firm received some type S applicants but no type T applicant, and by nothing if the firm received other type T applicants. Since the first case occurs with probability $e^{-(q_T+q_S)}$ and the second case occurs with probability $e^{-q_T}(1-e^{-q_S})$, then the expected contribution of a type T worker to output is $e^{-q_T}(\delta y + ye^{-q_S})$. Under (2.9), this equals $(E_S + \delta ye^{-q_T})$, as (2.10) states.

The equilibrium in Proposition 2.1 features complete mixing of the two types of workers, in the sense that every firm attracts both types of applicants. However, after workers' mixed strategy is played out, a firm may or may not received both types of applicants. If the firm does receive both types of workers, it chooses a high-productivity worker. If the firm receives only low-productivity workers, it selects one of them. Notice that the wages in (2.14) satisfy $w_T - w_S < \delta y$. That is, the ranking $R = 1$ is compatible with firms' ex post incentive.

To verify that the strategies described in Proposition 2.1 indeed constitute an equilibrium, it suffices to show that the following deviations by a single firm are not profitable:

D1 A deviation that intends to attract only type T workers.

D2 A deviation that intends to attract only type S workers.

D3 A deviation that attracts both types of workers but ranks type S workers first.

D4 A deviation that has no selection priority.

In the next section, I will accomplish this task, and more, by proving that the described equilibrium is the unique (symmetric) equilibrium.

3. The Candidate Is the Unique Equilibrium

Proposition 2.1 states that there is no equilibrium other than the one described in the proposition. To establish this result, I need to show that no other possible configuration of strategies forms an equilibrium. Let me partition the firms into two arbitrary groups, group A and group B. In addition to the equilibrium in Proposition 2.1, other possibilities are as follows.

N1. Complete separation of the firms into two groups, each attracting only one type of workers.

N2. Partial separation of type T workers: Group A firms attract only type T workers, while group B firms attract both types of workers and have one of the following ranking schemes: (N2a) ranking type T workers first; (N2b) ranking type S workers first.

N3. Partial separation of type S workers: Group A firms attract only type S workers, while group B firms attract both types of workers and have one of the following ranking schemes: (N3a) ranking type T workers first; (N3b) ranking type S workers first.

N4. No priority in a subset or all of the firms: A group of firms, say group A, give no priority, while group B firms use one of the following strategies: (N4a) attracting both types and giving no priority; (N4b) attracting both types and ranking type T workers first; (N4c) attracting both types and ranking type S workers first; (N4d) attracting only type T workers; (N4e) attracting only type S workers.

N5. No separation (i.e., all firms attract both types of workers) and all firms have strict ranking, but the ranking is different from the equilibrium one in one of the following ways: (N5a)

10

group A firms give priority to type S workers and group B to type T workers; (N5b) all firms rank type S workers first.

For each of these possibilities, I will use the same procedure to show that it is not an equilibrium. First, supposing that one of the above possibilities is an equilibrium, I will compute the wages posted by each firm, the expected number of applicants of each type whom a firm attracts, and expected wages in the market. Second, I will construct a single firm's deviation from this supposed equilibrium and toward the one described in Proposition 2.1. I will show that this deviation is profitable, and so the possibility is not an equilibrium.

By accomplishing this task, I also succeed in showing that the deviations D1 through D4 described earlier are not profitable against the equilibrium strategies. To see this, notice that I can let the size of group A firms approach zero. Then some of the above possibilities become a single firm's deviations from the equilibrium described in Proposition 2.1. In particular, case N2a becomes the deviation D1, case N3a becomes the deviation D2, case N5a becomes the deviation D3, and case N4b becomes the deviation D4. Because in these cases a firm can profit from making the strategies closer to the equilibrium strategies, the deviations D1 through D4 can be improved upon by further deviations toward the equilibrium strategies. Therefore, the strategies described in Proposition 2.1 form an equilibrium.

3.1. Separation Is Not an Equilibrium

Consider first the case of complete separation, i.e., Case N1. In this case, group A firms attract only type T workers and group B firms attract only type B workers. This possibility is an equilibrium in a similar model by Lang et al. (2002) who assume that each firm can post only one wage. When each firm can condition the wage on the type of the hired worker, complete separation is no longer an equilibrium.

The intuition is as follows. Suppose that complete separation is an equilibrium, as in Case N1. A firm in group A can maintain the same wage for type T workers and rank such workers first as in the supposed equilibrium, but chooses a wage to attract type S workers as well. Type

T workers will not change their strategy of applying to this firm, and so the expected profit from hiring a type T worker does not change. In the case where no type T worker shows up at the firm, the deviating firm can hire a type S worker and obtain additional profit. Thus, the deviation is profitable, provided that it is feasible and that it attracts type S workers.

To verify this intuition, let w_A be the wage posted by a group A firm (for type T workers) and w_B be the wage posted by a group B firm (for type S workers). Let a be the fraction of firms that are in group A. Then, the expected number of applicants is $q_A = \gamma\theta/a$ for a group A firm and $q_B = (1-\gamma)\theta/(1-a)$ for a group B firm. Let $\widehat{E}_T$ and $\widehat{E}_S$ be expected wages of these two types of workers, respectively, in the supposed equilibrium of separation. Then,

$$G\left(q_A\right)w_A = \widehat{E}_T, \quad G\left(q_B\right)w_B = \widehat{E}_S.$$

Because the two types of workers are completely separated, there is no crowding-out between them. Thus, the expected wage of each type of workers is equal to the worker's expected marginal contribution to output. That is,[5]

$$\widehat{E}_T = (1+\delta)ye^{-q_A}, \quad \widehat{E}_S = ye^{-q_B}. \tag{3.1}$$

Thus, the expected profit of a firm in the two groups is, respectively, as follows:

$$\pi_A = \left(1 - e^{-q_A}\right)\left[(1+\delta)y - w_A\right] = (1+\delta)y\left[1 - (1+q_A)e^{-q_A}\right],$$

$$\pi_B = \left[1 - e^{-q_B}\right]\left(y - w_B\right) = y\left[1 - (1+q_B)e^{-q_B}\right].$$

For a to be in $(0,1)$, a firm must be indifferent between being in the two groups. Thus, $\pi_A = \pi_B$. This requirement yields $q_A < q_B$, i.e., $a > \gamma$, for all $\delta > 0$. If $\delta = 0$, then $a = \gamma$.

Now consider the following deviation by a single firm in group A. The firm maintains the wage w_A for type T workers and still ranks these workers first. In contrast to the supposed equilibrium,

---

[5]To verify this, consider a type A firm's deviation to a wage w_A^d for type T workers and suppose that this deviating firm still attracts only type T workers (e.g., the firm sets zero wage for type S workers). Then, the first-order condition for this deviation yields $(1+\delta)ye^{-q_A^d} = \widehat{E}_T$, where q_A^d is the expected number of type T workers that the deviating firm will attract. For this deviation to be not profitable against the supposed equilibrium, we must have $q_A^d = q_A = \gamma\theta/a$. Thus, $\widehat{E}_T = (1+\delta)ye^{-q_A}$. The expression for $\widehat{E}_S$ can be obtained similarly by considering a group B firm's deviation.

the deviating firm posts w_{AS}^d for type S workers and ranks them below type T workers. It is clear that a type T worker will apply to the deviating firm with the same probability as in the supposed equilibrium, and so the expected number of type T workers whom the firm will receive is $q_A = \gamma\theta/a$. Let q_{AS}^d be the expected number of type S workers whom the deviating firm will receive. Then, the probability with which an individual type S worker who applies to the deviating firm will be selected is $e^{-q_A}G(q_{AS}^d)$. Let w_{AS}^d and the associated queue length q_{AS}^d satisfy the following conditions:

$$e^{-q_A}G(q_{AS}^d)w_{AS}^d = \widehat{E}_S, \quad ye^{-[q_A+q_{AS}^d]} = \widehat{E}_S. \tag{3.2}$$

The first condition requires the deviation to give a type S applicant the same expected wage as in the market, and the second condition requires the deviation to be the best of its kind so that a type S worker's expected output is equal to the expected wage in the market.

The deviation has the following features. First, the deviation indeed attracts type S workers. To see this, substituting $\widehat{E}_S$ from (3.1) into the second equation in (3.2) yields $q_{AS}^d = q_B - q_A$, which is positive as shown earlier. Second, the deviation is feasible in the sense that $w_{AS}^d < y$. To verify this, combine the two equations in (3.2) to obtain

$$w_{AS}^d = y\frac{q_{AS}^d}{e^{q_{AS}^d} - 1} < y.$$

Notice that the strict inequality implies that the deviating firm obtains a positive profit from hiring a type S worker when no type T worker shows up at the firm. Finally, the deviator's ranking of the two types of workers is compatible with the firm's incentive; i.e., the deviation satisfies $w_A - w_{AS}^d < \delta y$. To verify this, temporarily denote $\Delta = q_B - q_A \ (> 0)$. Substituting (w_{AS}^d, w_A) from the above, I can rewrite the compatibility condition as

$$\frac{1 - (1 + \Delta + q_A)e^{-(\Delta+q_A)}}{1 - e^{-q_A}} > \frac{1 - (1 + \Delta)e^{-\Delta}}{1 - e^{-\Delta}}.$$

Because $1 + \Delta < e^{\Delta}$, then $(1 + \Delta + q_A)e^{-(\Delta+q_A)} < e^{-q_A} + q_Ae^{-(\Delta+q_A)}$. Also, because $\frac{q_A}{e^{q_A}-1} < 1 < \frac{\Delta}{1-e^{-\Delta}}$, then

$$\frac{1 - (1 + \Delta + q_A)e^{-(\Delta+q_A)}}{1 - e^{-q_A}} > 1 - \frac{q_Ae^{-\Delta}}{e^{q_A} - 1} > 1 - \frac{\Delta e^{-\Delta}}{1 - e^{-\Delta}} = \frac{1 - (1 + \Delta)e^{-\Delta}}{1 - e^{-\Delta}}.$$

13

That is, the required condition holds.

Because the deviating firm's expected profit from type T workers is the same as in the supposed equilibrium and the additional expected profit from type S workers is strictly positive, the deviation is profitable. Thus, complete separation is not an equilibrium.

Similarly, partial separation as in Case N2 and Case N3 cannot be an equilibrium. A firm that attracts only one type of workers can increase its expected profit by deviating to a strategy that attracts both types of workers.

3.2. The Two Types of Workers Are Strictly Ranked by Firms

In all the cases examined so far, firms rank the two types of workers strictly. Now, I examine the cases in which some or all firms use no priority, i.e., the sub-cases of case N4. By showing that these cases are not an equilibrium, I establish the result that it is optimal for all firms to rank the two types of workers.

Suppose, to the contrary, that one of the cases N4a through N4e is an equilibrium. Because complete separation is not an equilibrium, one group of firms or both groups must attract both types of workers. Without loss of generality, assume that group A firms attract both types of workers. Let a be the fraction of firms in group A. For each firm in group A, let π_A be the firm's expected profit, w_{Aj} the wage for type j workers, and q_{Aj} the expected number of type j workers for the firm, where $j = T, S$. Let $\widehat{E}_j$ be the expected market wage of a type j worker. Because a type A firm does not rank the workers, the queue length of workers for the firm is $q_{AT} + q_{AS}$, which is temporarily denoted k. For both types of workers to apply to the firm, the following conditions must hold:

$$G\left(k\right) w_{Aj} = \widehat{E}_j, \quad j = T, S.$$

A group A firm must be indifferent between the two types of workers after both show up at the firm. That is, $w_{AT} - w_{AS} = \delta y$. The two conditions together solve:

$$w_{AT} = \frac{\widehat{E}_T}{G(k)}, \ w_{AS} = \frac{\widehat{E}_S}{G(k)}, \ G(k) = \frac{\widehat{E}_T - \widehat{E}_S}{\delta y}.$$

14

Moreover, $(\widehat{E}_T, \widehat{E}_S, w_{AT}, w_{AS})$ are all positive and the relationship $\widehat{E}_S \geq ye^{-k}$ holds.[6]

A group A firm can profit by deviating to a strategy that ranks type T workers first. Let (w_{AT}^d, w_{AS}^d) be the wages for the two types of workers posted by the deviating firm and (q_{AT}^d, q_{AS}^d) be the corresponding queue lengths. Let the deviation satisfy:

$$G(q_{AT}^d)w_{AT}^d = \widehat{E}_T, \quad e^{-q_{AT}^d}G(q_{AS}^d)w_{AS}^d = \widehat{E}_S,$$

$$ye^{-[q_{AT}^d + q_{AS}^d]} = \widehat{E}_S, \quad \delta ye^{-q_{AT}^d} = \widehat{E}_T - \widehat{E}_S.$$

Denote $k^d = q_{AT}^d + q_{AS}^d$. Because $\widehat{E}_S \geq ye^{-k}$, then $k^d \leq k$. Because $\delta yG(k) = \widehat{E}_T - \widehat{E}_S$, then $e^{-q_{AT}^d} = G(k)$. Since $e^{-z} < G(z)$ for all $z > 0$, then $G(k) < G(q_{AT}^d)$, which implies $k > q_{AT}^d$.

The deviator's expected profit is

$$\begin{aligned}\pi_A^d &= \left(1 - e^{-q_{AT}^d}\right)(1 + \delta)y - q_{AT}^d\widehat{E}_T + \left(e^{-q_{AT}^d} - e^{-k^d}\right)y - q_{AS}^d\widehat{E}_S \\ &= (1 + \delta)y - (1 + q_{AT}^d)\widehat{E}_T - q_{AS}^d\widehat{E}_S.\end{aligned}$$

The second equality is obtained from substituting $e^{-q_{AT}^d} = (\widehat{E}_T - \widehat{E}_S)/(\delta y)$ and $e^{-k^d} = \widehat{E}_S/y$. The expected profit from not deviating is $\pi_A = (1 - e^{-k})(1 + \delta)y - k\widehat{E}_T$. The gain from the deviation is

$$\pi_A^d - \pi_A = ye^{-k}\left[1 + \delta - \delta\left(1 + q_{AT}^d - k\right)e^{(k - q_{AT}^d)} - (1 + k^d - k)e^{(k - k^d)}\right].$$

The function $(1 - z)e^z$ is decreasing for all $z > 0$. Since $k > q_{AT}^d$ and $k \geq k^d$, then $\pi_A^d - \pi_A > 0$. That is, the deviation is profitable.

3.3. High-Productivity Workers Have the Priority

Since Cases N1 through N4 are not an equilibrium, all firms in an equilibrium must attract both types of workers and have a strict ranking of the two types. Now I show that the ranking must

---

[6]To show $\widehat{E}_S > 0$, suppose to the contrary that $\widehat{E}_S = 0$. Then $w_{AS} = 0$ and $\pi_A = (1 - e^{-k})y$. In this case, a single firm in group A can deviate to $w_{AS}^d = \varepsilon$ and $w_{AT}^d = w_{AT} + \varepsilon$, where $\varepsilon > 0$ is a sufficiently small number. It can be shown that this deviation is profitable. Because $k > 0$, then $\widehat{E}_S > 0$ implies that $(\widehat{E}_T, w_{AT}, w_{AS})$ are all positive. To show $\widehat{E}_S \geq ye^{-k}$, suppose $\widehat{E}_S < ye^{-k}$, instead. An individual firm in group A firm can deviate to $w_{AT}^d = 0$ and $w_{AS}^d > 0$. This deviation attracts only type S workers. Let w_{AS}^d and the associated queue length q_{AS}^d satisfy $G(q_{AS}^d)w_{AS}^d = \widehat{E}_S = ye^{-q_{AS}^d}$. Then, this deviation can be shown to be profitable.

be $R = 1$; that is, it is optimal for firms to rank the workers according to productivity. This is done by excluding the possibilities N5a and N5b as equilibrium.

Consider Case N5a first, where group A firms rank type S workers first and group B firms rank type T workers first. In this case, a firm's expected profit is higher in group B than in group A, and so the case is not an equilibrium. To see this, let (w_{iT}, w_{iS}) be the wages posted by a firm in group i, where $i = A, B$, and (q_{iT}, q_{iS}) be the corresponding queue lengths. Denote the total expected number of workers attracted by a group i firm as $k_i = q_{iT} + q_{iS}$, and the firm's expected profit as π_i. Denote a type j worker's expected market wages as $\widehat{E}_j$. Then, the maximization problem of a firm in each group yields:

$$(1 + \delta)ye^{-k_A} = \widehat{E}_T, \quad \delta ye^{-q_{AS}} = \widehat{E}_T - \widehat{E}_S,$$

$$ye^{-k_B} = \widehat{E}_S, \quad \delta ye^{-q_{BT}} = \widehat{E}_T - \widehat{E}_S.$$

These equations imply $q_{BT} = q_{AS}$. Also, a group B firm attracts more workers than a group A firm. More precisely, with the temporary notation $\Delta = k_B - q_{BT}$ (> 0), I can derive the following result from the above equations:

$$k_B - k_A = ln\left(\frac{1 + \delta e^{\Delta}}{1 + \delta}\right) > 0.$$

After computing a firm's expected firm in each group, I can obtain

$$\frac{\pi_B - \pi_A}{y} = \delta\left[1 - (1 + q_{BT})e^{-q_{BT}}\right] + e^{-k_B}\left[(k_A - q_{BT})\delta e^{(k_B - q_{BT})} - (k_B - k_A)\right].$$

The first term of this difference is positive. So is the second term.[7] Thus, $\pi_B > \pi_A$.

Finally, suppose that case N5b is an equilibrium, where each firm attracts both types of workers and ranks type S workers first. Let the wages posted by each firm be (w_T, w_S) and the

---

[7]To verify this, substitute k_A to obtain:

$$(k_A - q_{BT})\delta e^{(k_B - q_{BT})} - (k_B - k_A) = \delta\Delta e^{\Delta} - \left(1 + \delta e^{\Delta}\right)ln\frac{1 + \delta e^{\Delta}}{1 + \delta}.$$

This expression is an increasing function of Δ for all $\Delta > 0$ and, at $\Delta = 0$, it is equal to 0. Thus, the expression is positive for all $\Delta > 0$.

corresponding queue lengths be (q_T, q_S). Denote $k = q_T + q_S$ and denote a firm's expected profit as π. Then,

$$(1+\delta)ye^{-k} = \widehat{E}_T, \quad \delta y e^{-q_S} = \widehat{E}_T - \widehat{E}_S,$$

$$\pi = y + \delta y \left[1 + q_S\right] e^{-q_S} - (1+k)e^{-k}.$$

Moreover, since all firms attract the same expected number and composition of workers, $q_T = \gamma\theta$ and $q_S = (1-\gamma)\theta$. Consider a deviation by an individual firm to wages (w^d_{AT}, w^d_{AS}) that are intended to attract both types of workers and that give type T workers priority. Let (q^d_{AT}, q^d_{AS}) be the corresponding queue lengths and let the deviation be the best of its kind (i.e., let it satisfy the corresponding first-order conditions). Let (q^d_{AT}, q^d_{AS}) serve the roles of (q_{AT}, q_{AS}) in the above proof for case N5a, and let (q_T, q_S) serve the roles of (q_{BT}, q_{BS}). Then, the same proof shows that the deviation increases the firm's expected profit. In addition, it can be shown that the condition, $(1+\delta)y - w^d_{AT} > y - w^d_{AS}$, holds so that the ranking in the deviation is compatible with the firm's ex post incentive.[8] Therefore, it is profitable to deviate from the strategies in N5b to one that ranks type T workers above type S workers.

4. Properties of the Equilibrium and the Social Optimum

In this section I examine the properties of the equilibrium and show that the equilibrium is socially efficient. The following proposition can be readily confirmed from (2.11) through (2.14).

Proposition 4.1. *The equilibrium described in Proposition 2.1 has the following properties.*

(i) A firm's ex post profit from a type T worker is higher than from a type S worker.

---

[8]To verify this condition, use the first-order conditions to solve the queue lengths induced by the deviation as $q^d_{AT} = (1-\gamma)\theta$ and $q^d_{AS} = \gamma\theta - ln\left(1 + \delta - \delta e^{\gamma\theta}\right)$. The wages are $w^d_{AT} = \widehat{E}_T / G(q^d_{AT})$ and $w^d_{AS} = \widehat{E}_S e^{q^d_{AT}} / G\left(q^d_{AS}\right)$. Then the required condition holds if and only if the following condition holds:

$$0 < \delta - (1+\delta)\frac{(1-\gamma)\theta e^{-\gamma\theta}}{e^{(1-\gamma)\theta} - 1} + \frac{q^d_{AS}}{(1+\delta)\left(e^{\gamma\theta} - 1\right)}.$$

Using the solution for q^d_{AS} obtained earlier, I can show that the right-hand side of the above inequality is increasing in δ, and so it is greater than its value at $\delta = 0$ which is proportional to $\left[\gamma e^\theta - e^{\gamma\theta} + 1 - \gamma\right]$. The last expression is a concave function of γ, and it is equal to 0 at both $\gamma = 0$ and $\gamma = 1$. Thus, it is positive for all $\gamma \in (0,1)$.

(ii) A type T worker has a higher employment probability than a type S worker.

(iii) $E_T > E_S$, and $(E_T - E_S)$ is of the same order of magnitude as δ.

(iv) $\exists \delta_1 > \delta_0 > 0$ such that $w_T < w_S$ for all $\delta \in [0, \delta_0)$ and $w_T > w_S$ for all $\delta > \delta_1$.

(v) Define $\Delta = \frac{w_S}{w_T} - 1$. $\exists \delta_2 > 0$ such that, if $0 < \delta < min\{\delta_0, \delta_2\}$, then $\frac{d\Delta}{d\delta} < 0$, $\frac{d\Delta}{d\theta} > 0$, $\frac{d\Delta}{d\gamma} > 0$.

The properties (i), (ii) and (iii) are intuitive. Property (i) repeats an earlier result that it is optimal for all firms to rank the workers according to productivity. Such ranking gives each high-productivity worker a higher employment probability and higher expected wage than a low-productivity worker. Moreover, high-productivity workers get a higher expected wage, and the differential in the expected wage is of the same order of magnitude as the productivity differential.

However, the actual wage is not always higher for high-productivity workers. As stated in property (iv), only when the productivity differential is sufficiently large do high-productivity workers get a higher actual wage than low-productivity workers. When the productivity differential is small, workers with higher productivity get lower wages. This reverse wage differential arises from the ranking of workers. An increase in the ranking increases a worker's employment probability, and hence the expected wage, by a discrete amount. So, when a firm awards slightly more productive workers with a higher ranking, it can cut the wage for these workers by a discrete amount and yet still be able to attract them. By doing so, the firm can increase expected profit.

Moreover, the size of the reverse wage differential *increases* as the productivity differential (δ) decreases. Although unconventional, this result is quite intuitive. When the productivity advantage of one set of workers shrinks relative to other workers, maintaining the employment priority for these workers is optimal for a firm only if their wage is reduced. An implication of this result is that the reverse wage differential remains strictly positive as the productivity differential approaches zero. However, if the economy literally has $\delta = 0$, there is another equilibrium in which all workers are paid the same wage. This equilibrium with a uniform age is not selected as the limit outcome of a sequence of economies in which the productivity differential is positive.

The above results suggest that actual wages can sometimes be a bad indicator of workers' productivity. Thus, the standard practice in labor economics that attributes wage differentials to productivity differentials should be taken with caution. First, a large residual wage differential might be attributed to a statistically insignificant differential in productivity, as it is the case here when δ is small. In this case, it is futile trying to explain the wage differential by ever expanding the list of workers' characteristics. Second, there is nothing abnormal about a residual wage differential; rather, it is part of the equilibrium with fully rational players in a frictional labor market. Also, the residual wage differential is socially efficient, as shown below. Third, when similar workers get different wages, the ones who receive lower wage are not necessarily discriminated against. For example, where the productivity differential is small, the reverse wage differential is just a compensation to low-productivity workers for being ranked low in the selection. Despite the lower wage in this case, high-productivity workers are not discriminated against, because they are ranked the first for the job and they obtain higher expected wage.

The wage differential in this model also responds to the market condition in an interesting way. When there is a reverse wage differential as a result of a small productivity differential, an increase in the overall ratio of workers to firms, θ, increases the reverse wage differential. The explanation is as follows. When jobs become more scarce, workers value the employment probability more than the wage. Since high-productivity workers are given a higher employment probability through the ranking scheme, they are willing to take a larger wage cut to maintain this difference in the employment probability.

I now turn to the efficiency of the equilibrium. Since the equilibrium has unconventional features in wages, it is interesting to see whether the equilibrium is efficient under the constraint of the matching frictions. To examine efficiency, let me take aggregate output as the measure of social welfare. This measure is appropriate here because all agents are risk neutral.

Suppose that a fictional social planner tries to maximize aggregate output, subject to the same restrictions that matching frictions generate in the equilibrium. One of these restrictions is

that the firms cannot separate two identical workers. This restriction requires that the planner must treat all workers in the same group in the same way in the matching process. Thus, the planner can divide the firms into at most two groups, with each group potentially targeting a different group of workers. Another restriction is that the matching function in each group of workers and firms must be the same as in the equilibrium. To describe the matching function, let the two groups of firms be indexed by i, where $i = A, B$. Let group i firms be a fraction a_i of all firms, where $a_A + a_B = 1$. Let q_{ij} be the expected number of type j workers for each firm in group i, where $j = T, S$. Let R_i be the ranking of the workers by a firm in group i, where $R_i \in \{1, 0, \Phi\}$. As in the market equilibrium, the matching function facing the social planner is such that, in each group i, a firm receives one or more type j worker with probability $(1 - e^{-q_{ij}})$. A planner's allocation is $(a_i, q_{iT}, q_{iS}, R_i)_{i=A,B}$.

Expected output of a firm in group i is

$$
R_i \left[(1 - e^{-q_{iT}}) (1 + \delta) y + e^{-q_{iT}} (1 - e^{-q_{iS}}) y \right]
$$
$$
+ (1 - R_i) \left[(1 - e^{-q_{iS}}) y + e^{-q_{iS}} (1 - e^{-q_{iT}}) (1 + \delta) y \right].
$$

Re-arranging terms and weighting each group's output by the group's size, I can express expected output per firm in the economy as follows:

$$
\sum_{i=A,B} a_i \left\{ y \left[1 - e^{-(q_{iT}+q_{iS})} \right] + \delta y \left(1 - e^{-q_{iT}} \right) \left[R_i + (1 - R_i) e^{-q_{iS}} \right] \right\}.
$$

The planner chooses $(a_i, q_{iT}, q_{iS}, R_i)_{i=A,B}$ to maximize this output, subject to $q_{iT} \geq 0$, $q_{iS} \geq 0$ and the following (resource) constraints:

$$
a_A + a_B = 1, \quad a_A, a_B \in [0, 1],
$$

$$
a_A q_{Aj} + a_B q_{Bj} \leq \gamma_j \theta, \text{ for } j = T, S. \tag{4.1}
$$

Here $\gamma_T = \gamma$ and $\gamma_S = 1 - \gamma$. The first constraint is self-explanatory. The second constraint is the adding-up constraint (2.3) in the current context.

The following proposition holds and a proof is supplied in Appendix A:

Proposition 4.2. *The efficient allocation coincides with the equilibrium allocation described in Proposition 2.1.*

The reason why the equilibrium is efficient is similar to that in other directed search models, e.g., Moen (1997), Acemoglu and Shimer (1999) and Shi (2001, 2002). In particular, the directed search framework allows the firms to internalize the matching externalities. One way to see this is to recall the expressions for workers' expected market wages, E_T and E_S, which are given by (2.12) and (2.13). For each type of workers, the expected market wage is equal to the worker's marginal contribution to expected output which takes into account the worker's crowding-out on other workers' expected output. As a result, a worker's expected wage in the equilibrium is equal to the worker's social marginal value. To express this equality more formally, notice that a type j worker's social marginal value is equal to the Lagrangian multiplier of (4.1) in the planner's problem, denoted λ_j. Then, it can be verified that $\lambda_j = E_j$, for both $j = T$ and S. Because expected wages take into account each worker's crowding-out on other workers, they share the property of the efficient mechanisms described by Mortensen (1982, pp968-969) and Hosios (1990).

Let me make a few remarks on the efficiency result. First, it is expected wages, not actual wages, that serve the role of the efficient compensation scheme. To achieve efficiency, the market mechanism requires the use of not only prices (wages) but also the ranking of the workers. Second, separating the two types of workers in the matching process, completely or partially, is inefficient. Third, because the equilibrium is the only equilibrium, the wage differential (sometimes a reverse one) and the ranking are efficient.[9]

Now it is useful to contrast the results in my model with those in Lang et al. (2002). As mentioned before, the main difference is that Lang et al. restricts each firm to post only one wage. This restriction generates several important differences in results. First, complete separation is an equilibrium in Lang et al. but not in my model. Second, the workers who have a lower employment

---

[9]Although I have not examined entry by firms, it can be done easily (see the end of Section 5.1).

probability in my model can get higher wage than other workers when the productivity differential is small. This does not occur in Lang et al.. Third, the equilibrium in my model is inefficient, but not in Lang et al.. In particular, the separation of the two types of workers, the wage differential and the employment difference between workers in the model of Lang et al. are all inefficient.

5. Extension to Many Types of Workers

In this section I extend the model to incorporate many types of workers. One purpose of this analysis is to check whether the results in the simple model are robust. The other purpose is to examine how wages vary when the workers' productivity varies in a wide range.

5.1. The Equilibrium and Its Properties

Let there be $J \geq 2$ types of workers, which are indexed by $j = 1, 2, ..., J$. The productivity of a type j worker is y_j and each worker is of type j with probability γ_j, where $y_1 > y_2 > ... > y_J$ and $\sum_{j=1}^{J} \gamma_j = 1$. As before, I examine only the symmetric equilibrium, where all workers of the same type use the same strategy. In the presence of a large number of types of workers, it is difficult to establish the uniqueness of the equilibrium. Thus, I will only establish the existence of an equilibrium resembling the one in the simple model, i.e., an equilibrium in which all firms attract all types of workers and rank the workers according to productivity.

To examine this equilibrium, I need to examine deviations that do not attract all types of workers or do not rank the workers according to productivity. To do so, I introduce the following general notation. Use the notation $i \succ i'$ to mean that a firm ranks type i workers above type i' workers in the hiring process. For any integer K with $1 \leq K \leq J$, I use K to stand for both the number K and the ordered set $\{1, 2, ..., K\}$, where the ordering is $1 \succ 2 \succ ... \succ K$. As in the simple model, it is not optimal for a firm to give the same ranking to two types of workers. Thus, I will restrict attention to strategies, including deviations, that have strict ranking over the types of workers. A strategy is defined by three characteristics: a strictly ordered set of types, C, a vector of wages, w, and the associated vector of queue lengths, q.[10] The set C contains the types

---

[10]Strictly speaking, q should be treated as the workers' strategy, rather than a firm's. However, the use of q as

of workers that a firm attracts and the ordering of these types. It is useful to express C as

$$C = (j_k)_{k=1}^K, \ j_k \in J, \ K \leq J,$$

so that the ranking in C is according to the index k, with j_1 being ranked the first. The vectors w and q are sorted according to the ordering in C. Given (C, w, q), it is convenient to define

$$Q_{j_k} = \sum_{i=1}^{k} q_{j_i}. \tag{5.1}$$

Then, Q_{j_k} is the expected number of workers whom the firm attracts and whose ranking is higher than or equal to k in C. Set $Q_{j_0} = 0$.

An equilibrium consists of a strategy, (C, w, q), where the ordering in C is strict, and a vector of expected market wages, E, such that (i) given E, the strategy is optimal for each firm, and (ii) the strategy induces E.

The particular equilibrium which resembles the unique equilibrium in the simple model has the following additional features: (i) All firms use the same strategy; (ii) Every firm attracts all types of workers; and (iii) $C = J$, i.e., every firm ranks the workers according to productivity. With these features, each worker applies to all firms with the same probability, and so $q_j = \gamma_j \theta$ for all $j \in J$.

To characterize this equilibrium, suppose that each firm attracts all types of workers and ranks the workers according to the ordered set J. Then, with probability $e^{-Q_{j-1}}$, a firm receives no applicant whose productivity is higher than y_j. Thus, the probability with which the firm successfully hires a type j worker is $e^{-Q_{j-1}}(1 - e^{-q_j}) = e^{-Q_{j-1}} - e^{-Q_j}$. Similarly, a type j worker who applies to the firm will be selected with the following probability:

$$p_j = e^{-Q_{j-1}} G(q_j) = \left[e^{-Q_{j-1}} - e^{-Q_j} \right] / q_j. \tag{5.2}$$

Taking the vector E as given, each firm chooses (w, q) to solve the following problem:

$$max \ \pi = \sum_{j=1}^{J} \left(e^{-Q_{j-1}} - e^{-Q_j} \right) (y_j - w_j)$$

---

a firm's choice is convenient and it does not change the analysis. As it is clear from the simple model, a firm that intends to attract a type of workers must offer a combination of a wage and a queue length that gives the worker the expected market wage. Under this constraint, the firm's choice of w effectively determines q.

s.t. $\quad p_j w_j \geq E_j$ for all j. $\hfill (5.3)$

Because the firm attracts all types of workers in the conjectured equilibrium, then $q_j > 0$ for all j and (5.3) holds with equality. Thus,

$$w_j = E_j/p_j \text{ for all } j. \hfill (5.4)$$

In fact, as mentioned earlier, q is given as

$$q_j = \gamma_j \theta \text{ for all } j. \hfill (5.5)$$

Because $q_j > 0$, the first-order conditions of the above problem yield:

$$E_j = y_j e^{-Q_j} - \sum_{i=j+1}^{J} y_i \left(e^{-Q_{i-1}} - e^{-Q_i} \right), \quad \text{all } j. \hfill (5.6)$$

Notice that, since $G(q) > e^{-q}$ for all $q > 0$, then $p_j > e^{-Q_j}$ from (5.2) and $p_j w_j = E_j < p_j y_j$ from (5.6). That is, the wage rate specified by (5.4) is feasible for all j.

This compensation scheme generalizes (2.12) and (2.13). The right-hand side is a type j worker's expected contribution to output. The first term is a type j worker's contribution to output when the firm does not receive any applicant whose productivity is higher than or equal to y_j. The second term is expected output that a type j worker crowds out on the workers of lower productivity. Thus, the compensation scheme requires a worker's expected wage to be equal to the worker's expected output minus the worker's expected crowding-out on other workers' output. For this reason, the compensation scheme is socially efficient.

Proposition 5.1. *The strategy (J, w, q) and the vector E form an equilibrium if the ranking in J is strict and if (5.4), (5.5) and (5.6) are satisfied. In this equilibrium, $p_j > p_{j+1}$, $E_j > E_{j+1}$, $w_j < y_j$ and $w_j - y_j < w_{j+1} - y_{j+1}$ for all j. However, $w_j < w_{j+1}$ if $(y_j - y_{j+1})$ is small.*

A proof of this proposition appears in Appendix B. The intuition for why the particular (J, w, q) and E form an equilibrium is as follows. Consider a deviation by a single firm to (C, q^d, w^d), where $C = (j_k)_{k=1}^{K}$ for $1 \leq K \leq J$. First, any deviation that does not rank the

24

workers according to productivity can be improved upon. If there are two types of workers in C, say types j_s and j_{s+1}, whose relative ranking is opposite to productivity, then I can construct a further deviation that switches the rankings of these two types while maintaining the same wages and queue lengths for other types in C. Relative to the original deviation, the further deviation attracts more type j_{s+1} workers and fewer type j_s workers. Since type j_{s+1} workers are more productive than type j_s workers, the gain from the further deviation outweighs the loss, and so this further deviation improves upon the original deviation.

Second, any deviation that does not attract all types of workers can be improved upon. If the deviation does not attract type j^* workers, where $j_s < j^* < j_{s+1}$, then I can construct a further deviation that attracts type j^* workers and ranks type j^* between j_s and j_{s+1}. This further deviation gives the same wages and queue lengths to all the workers in the set C, except type j_{s+1}. In addition, the sum of the queue lengths of type j^* and type j_{s+1} workers in the further deviation is equal to the queue length of type j_{s+1} workers in the original deviation. Again, the gain from the higher productivity workers (type j^*) outweighs the loss from the lower productivity workers (type j_{s+1}), and so this further deviation improves upon the original deviation.

Notice that the further deviations in the above construction move the strategies toward the equilibrium strategy described in Proposition 5.1. They share with the equilibrium strategy the feature that every firm attracts all types of workers and ranks the workers according to productivity. By construction, however, all strategies with this feature are dominated by the strategy in the described equilibrium. Thus, the equilibrium described in Proposition 5.1 is indeed an equilibrium.

Because the equilibrium in this extended economy is of the same type as the one in the simple model, it is not surprising that it has similar properties. In particular, higher productivity workers have a higher employment probability and higher expected wage. Also, actual wages exhibit a differential opposite to the productivity differential when the productivity differential is small.[11]

---

[11] As in the simple model, expected wages in the equilibrium internalize the matching externalities. Thus, the equilibrium allocation is efficient.

The dependence of the actual wage on productivity can exhibit a number of different patterns. One is that the actual wage decreases in productivity, which occurs when the productivity differential $(y_j - y_{j+1})$ is small for all j. The opposite pattern is that the actual wage increases in productivity, which occurs when $(y_j - y_{j+1})$ is large for all j. Non-monotonic patterns can also arise. For example, if $(y_j - y_{j+1})$ is large for large j and decreases sharply as j decreases, then the actual wage increases first when productivity increases from the lowest level but decreases when productivity is sufficiently high.

The equilibrium wage distribution also depends on the pattern of productivity and the distribution of workers. To compute the wage distribution, recall that each type j worker is employed with probability p_j. Since the size of type j workers in the labor force is γ_j, the expected number of type j workers who are employed (at wage w_j) is

$$\gamma_j p_j = \frac{1}{\theta}\left(e^{-\theta \Gamma_{j-1}} - e^{-\theta \Gamma_j}\right), \text{ where } \Gamma_j = \sum_{i=1}^{j} \gamma_i.$$

The total number of employed workers of all types is $(1 - e^{-\theta})/\theta$, which is independent of the distribution of types. Therefore, the density of workers employed at wage w_j is[12]

$$f_j \equiv \frac{\gamma_j p_j}{\sum_{j=1}^{J} \gamma_j p_j} = \frac{\gamma_j p_j \theta}{1 - e^{-\theta}}. \tag{5.7}$$

Because $p_j > p_{j+1}$, then $f_j/\gamma_j > f_{j+1}/\gamma_{j+1}$. That is, the higher the productivity of a type of workers, the larger the fraction of these workers will be employed. If all types of workers have the same size in the labor force, then the density of wages is an increasing function of productivity. However, because wages are not necessarily an increasing function of productivity, the density of wages can be increasing, decreasing, or non-monotonic in wages. In the next subsection, I will provide some numerical examples to illustrate this distribution.

Before providing the numerical examples, let me introduce firms' entry into the market. Suppose that firms compete in setting up vacancies and the cost of setting up a vacancy is c. After the vacancies are set up, the firms play the recruiting game as described before. The expected

---

[12]Obviously, there are only a finite number of wage levels in the equilibrium distribution and so the phrase "density function" really means the frequency function.

profit of each firm from recruiting will depend on the overall market tightness, θ, and so let me denote it as $\pi(\theta)$. In the equilibrium, θ is determined by $\pi(\theta) = c$.

5.2. Numerical Examples

I use two examples to illustrate the wage distribution. In particular, these examples show that the model can generate a hump-shaped density of the wage distribution as the one documented in the literature (e.g., Mortensen, 2002). Since the model is not calibrated to the data, these examples only illustrative. In both examples, I will set $J = 20$, $y_1 = 100$, $c = 0.2y_J$ and $y_j = y_1(1 + j\Delta)^{1-j}$ for all j, where Δ will be determined by the choice of y_J through the requirement $y_J = y_1(1 + J\Delta)^{1-J}$. The two examples differ from each other in Δ and the distribution of workers in the labor force. Of course, the equilibrium tightness and wage distribution also differ in the two examples.

Example 5.2. $y_J = 98$ and $\gamma_j = j^{0.3} \left/ \sum_{j=1}^{J} j^{0.3} \right.$. *These imply* $\Delta = 5.319 \times 10^{-5}$ *and* $\theta = 0.819$.

In this example, the difference between the highest and the lowest productivity is small, and the distribution of workers' productivity in the labor force is a decreasing function. Figure 1 depicts the equilibrium in this example. In the upper panel, the variable on the horizontal axis is an increasing function of productivity. This panel shows that the employment probability is an increasing function of productivity, as the theory predicts. Also, because the productivity differential is small, the wage rate is a decreasing function of productivity.

The magnitude of this reverse wage differential is worth noting: Although the lowest productivity is only 2% lower than the highest productivity, the wage of the least productive workers is twice as much as the wage of the most productive workers. Despite this large reverse wage differential, high-productivity workers are rewarded properly as they receive higher expected wage. This is made possible by a positive differential in the employment probability that outweighs the reverse wage differential.

In Figure 1, I depict the wage density and the distributional density of workers' productivity. The wage density is hump-shaped. Because the wage rate is a decreasing function of productivity

in this example, the shape of the wage density implies that more workers with medium productivity are employed than workers with either very high-productivity or very low-productivity. There are very few low-wage workers in the equilibrium because these workers are high-productivity workers who are a small fraction of the labor force. There are also very few high-wage workers because these workers are low-productivity workers, who have a very low ranking and hence a very low employment probability.

The shape of the wage density sharply contrasts with the distribution of workers in the labor force, γ. In the lower panel of Figure 1, the plot of γ is the hypothetical density of wages when all workers of each type are employed. In contrast to the humped shape of the actual density, this hypothetical density is increasing. For any wage level, if the actual density exceeds the hypothetical density, the workers at that wage are employed at a rate higher than the average employment rate; if the actual density falls below the hypothetical density, the workers at that wage are employed at a rate lower than the average employment rate. Thus, in this example, low-wage workers are employed at a rate higher than the average rate while high-wage workers are employed at a rate lower than the average rate.

In the above example, the hump-shaped wage density comes with a reverse wage differential. However, this is not an inevitable prediction of the model, as the following example shows.

Example 5.3. $y_J = 50$ and $\gamma_j = [j\,(J + 11 - j)]^4 \Big/ \sum_{j=1}^{J} [j(J + 11 - j)]^4$. *These imply* $\Delta = 1.858 \times 10^{-3}$ *and* $\theta = 0.696$.

In contrast to Example 5.2, this example has a much larger difference between the highest and the lowest productivity. As a result, the wage rate is an increasing function of productivity, as depicted in the upper panel of Figure 2. Also, the distribution of workers in the labor force has a single peak at an intermediate level of productivity, rather than being a decreasing function of productivity. Notice that the wage differential is small, in contrast with the large difference in productivity. The large difference in productivity induces a large difference in the employment
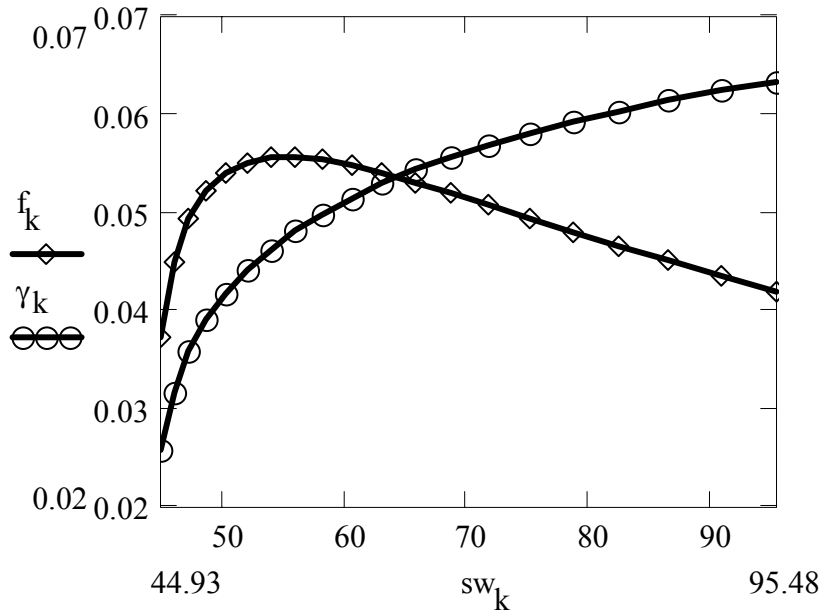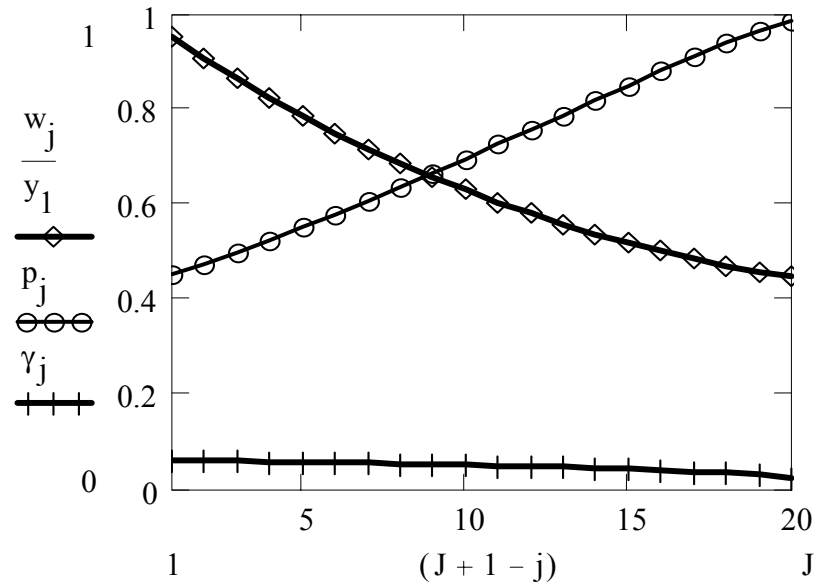
28

probability, and hence in the expected wage.

Figure 1. The equilibrium and the wage density in Example 1

j: worker's type, with a lower j corresponding to higher productivity;
$\frac{w_j}{y_1}$: wage of type j workers, normalized by y_1;
p_j: employment probability of a type j worker;
γ_j: fraction of type j workers in the labor force;
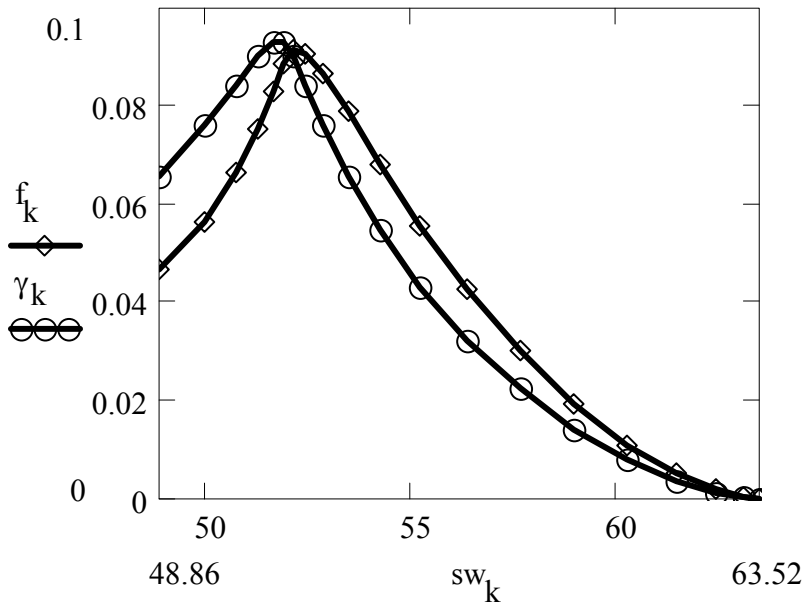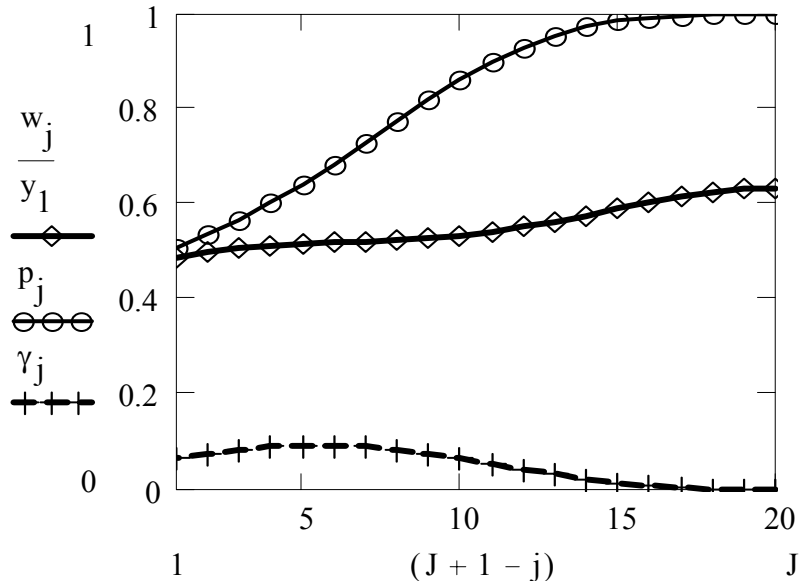sw: wages sorted in an ascending order; f: wage density.

29

Figure 2. The equilibrium and the wage density in Example 2

j: worker's type, with a lower j corresponding to higher productivity;

$\frac{w_j}{y_1}$: wage of type j workers, normalized by y_1;

p_j: employment probability of a type j worker;

γ_j: fraction of type j workers in the labor force;

sw: wages sorted in an ascending order; f: wage density.

The density of the wage distribution is still hump-shaped, as depicted in the lower panel of Figure 2. The hypothetical wage density when all workers are employed, γ, is also hump-shaped. In comparison with this hypothetical density, the equilibrium wage density peaks at a higher wage and a larger mass is distributed at higher wages. Thus, low-wage workers are employed at a rate lower than the average employment rate while high-wage workers are employed at a rate higher than the average rate. This result reflects the fact that high-wage workers in this example are high-productivity workers who are employed with a higher probability.

6. Conclusion

In this paper I construct a search model of a large labor market in which workers are heterogeneous in productivity and (homogeneous) firms post wages to direct workers' search. Each firm can rank the applicants and post different wages for different types of workers. Workers cannot coordinate their applications. In the unique (symmetric) equilibrium, high-productivity workers have a higher priority in employment and higher expected wages than low-productivity workers. The equilibrium is socially efficient. However, high-productivity workers receive higher actual wages than low-productivity workers only when the productivity differential is large. When the productivity differential is small, workers of higher productivity are paid lower wages. Moreover, this reverse wage differential increases as the productivity differential decreases and so it is strictly positive even when the productivity differential approaches zero.

These results suggest that actual wages may not be a good indicator of productivity. They also show that wage differentials among similar workers should not always be construed as discrimination, as they have often been viewed in the literature. Rather, the difference in workers' expected payoffs is a more reliable measure of discrimination. For example, when the productivity differential is so small that it appears statistically insignificant to an econometrician (but observable to the employers), the current model implies that the workers with slightly higher productivity obtain lower wage than other workers. Guided by the convention, the econometrician would conclude that high-productivity workers are discriminated against. This would be

misleading, because high-productivity workers enjoy a higher ranking and higher expected wage.

A natural question is whether the wage differential can persist in the long run. In a dynamic environment, the search model will have exogenous separation between matched firms and workers. A wage differential will continue to exist in the steady state, but it may not be reversely related to productivity when the productivity differential is small. For the reverse wage differential to survive in the steady state, exogenous job separation must be high. Finding out how high a job separation rate is needed is a quantitative exercise.

An extension to a dynamic economy will also generate some interesting predictions on the time path of wages. For example, consider the model with only two types of workers and suppose that the productivity differential between the two is small. Then, employed workers of high productivity are more likely to search on the job than employed workers of low productivity. This implies that high-productivity workers will have a steeper wage path, even though there is no learning-by-doing or human capital accumulation. Analyzing such on-the-job search is difficult in an environment with directed search. The reason is that, when offering a wage, a firm must take into account not only workers' tradeoff between the wage and the current employment probability, but also the tradeoff between the current wage and the probability of getting higher wage in the future. Perhaps a quantitative analysis can be conducted.

Appendix

A. Proof of Proposition 4.2

I can assume $a_A > 0$ and $a_B > 0$ without loss of generality. To see this, suppose that an allocation has $a_A = 1$, $a_B = 0$ and $(q_{iT}, q_{iS}, R_i)_{i=A,B}$. Then, an alternative allocation $(a_i^*, q_{iT}^*, q_{iS}^*, R_i^*)_{i=A,B}$ can be constructed as follows: $a_A^* = a_B^* = 1/2$, $q_{AT}^* = q_{BT}^* = q_{AT}$, $q_{AS}^* = q_{BS}^* = q_{AS}$, and $R_A^* = R_B^* = R_A$. This alternative allocation is equivalent to the original allocation, except that it re-labels a half of the firms in the original group A as group B.

Let λ_j be the Lagrangian multiplier of (4.1) in the planner's maximization problem. Then, the first-order conditions of q_{iT} and q_{iS} are as follows:

$$\lambda_T/y \geq e^{-(q_{iT}+q_{iS})} + \delta e^{-q_{iT}} \left[R_i + (1-R_i)e^{-q_{iS}} \right], \quad \text{`` = ''} \text{ if } q_{iT} > 0, \tag{A.1}$$

$$\lambda_S/y \geq e^{-(q_{iT}+q_{iS})} - \delta \left(1 - e^{-q_{iT}}\right)\left(1 - R_i\right)e^{-q_{iS}}, \quad \text{`` = ''} \text{ if } q_{iS} > 0. \tag{A.2}$$

The first-order condition of a_A (taking $a_B = 1 - a_A$ into account) is:

$$
\begin{aligned}
0 = {} & \delta \left(1 - e^{-q_{AT}}\right)\left[R_A + (1-R_A)e^{-q_{AS}}\right] - \delta \left(1 - e^{-q_{BT}}\right)\left[R_B + (1-R_B)e^{-q_{BS}}\right] \\
& + \left[e^{-(q_{BT}+q_{BS})} - e^{-(q_{AT}+q_{AS})}\right] - (q_{AT} - q_{BT})\lambda_T/y - (q_{AS} - q_{BS})\lambda_S/y.
\end{aligned}
\tag{A.3}
$$

The following features can be easily deduced from the planner's problem. First, from inspecting the objective function, I can infer that the choice $R_i = 1$ is efficient whenever $q_{iT} > 0$ and $q_{iS} > 0$. Second, it is efficient to utilize both types of workers, in the sense that $q_{Aj} + q_{Bj} > 0$ for both $j = T, S$. To see this, suppose $q_{AS} = q_{BS} = 0$, to the contrary. Then, (4.1) does not bind and so $\lambda_S = 0$. Because type S workers are not assigned to match with any firm, the ranking R is irrelevant for the allocation, and so R_i can be set to 1. In this case, (A.2) implies $0 \geq e^{-q_{iT}}$ for $i = A, B$, which is a contradiction. Similarly, the choices $q_{AT} = q_{BT} = 0$ are not efficient.

With the above features, the efficient allocation must be one of the following cases: (i) $q_{iT} > 0$ and $q_{iS} > 0$ for both $i = A$ and B; (ii) $q_{AT} > 0$ and $q_{AS} > 0$ but $q_{BT} = 0 < q_{BS}$; (iii) $q_{AT} > 0$ and $q_{AS} > 0$ but $q_{BT} > 0 = q_{BS}$; (iv) $q_{AT} > 0 = q_{AS}$ and $q_{BT} = 0 < q_{BS}$. (All other cases

33

are equivalent to these four cases, with the labels A and B being switched.) In the remainder of the proof, I will first show that case (i) is a solution to the planner's problem and it yields the same allocation as the equilibrium described in Proposition 2.1. Then I will show that cases (ii) through (iv) are not efficient.

Consider case (i) first. In this case, $R_i = 1$ for both $i = A$ and B. Also, the first-order conditions for q_{iT} and q_{iS} hold as equality for both $i = A$ and B. These first-order conditions and constraint (4.1) together yield $q_{AT} = q_{BT} = \gamma\theta$ and $q_{AS} = q_{BS} = (1-\gamma)\theta$. Thus, the two groups are identical, and the first-order condition for a is trivially satisfied. This allocation is the same as in the equilibrium described in Proposition 2.1.

In case (ii), $R_A = 1$. Also, (A.2) holds as equality for both $i = A$ and B, which yields $q_{AT} + q_{AS} = q_{BT} + q_{BS}$. In addition, (A.1) holds with equality for $i = A$. Combining the conditions (A.1) and (A.2) for $i = A$ yields $\lambda_T - \lambda_S = \delta y e^{-q_{AT}}$. Substituting these results into (A.3), I have:

$$0 = \delta\left[1 - (1 + q_{AT})e^{-q_{AT}}\right].$$

Because $1 > (1+z)e^{-z}$ for all $z > 0$, the right-hand side of the above condition is strictly positive. This is a contradiction, and so case (ii) is not a solution to the planner's problem.

In case (iii), $R_A = 1$, (A.1) holds as equality for both $i = A$ and B, and (A.2) holds as equality for $i = A$. Thus,

$$\lambda_T = y e^{-q_{AT}}\left(e^{-q_{AS}} + \delta\right), \quad \lambda_S = y e^{-(q_{AT}+q_{AS})},$$

$$q_{BT} = q_{AT} - \ln\left(\frac{\delta + e^{-q_{AS}}}{1+\delta}\right).$$

Substituting these conditions into (A.3) and simplifying, I have:

$$0 = -\ln\left(\frac{\delta + e^{-q_{AS}}}{1+\delta}\right) - \frac{q_{AS}}{1 + \delta e^{q_{AS}}}.$$

The right-hand side of the above condition is an increasing function of q_{AS} and, at $q_{AS} = 0$, its value is 0. Because $q_{AS} > 0$, then the above condition is violated. Thus, case (iii) is not a solution to the planner's problem.

Finally, consider case (iv). In this case, (A.1) holds as equality for $i = A$ and (A.2) holds as equality for $i = B$. Thus, $\lambda_T = (1 + \delta)ye^{-q_{AT}}$ and $\lambda_S = ye^{-q_{BS}}$. Substituting these results into (A.3) yields

$$0 = (1 + q_{BS})e^{-q_{BS}} + \delta - (1 + \delta)(1 + q_{AT})e^{-q_{AT}}. \tag{A.4}$$

Because no type S worker applies to a group A firm, R_A can be set to 1. Also, (A.2) holds as inequality for $i = A$. Substituting $\lambda_S = e^{-q_{BS}}$ and $R_A = 1$, this inequality implies $q_{BS} \leq q_{AT}$. Since the function $(1 + z)e^{-z}$ is decreasing for all $z > 0$, the result $q_{BS} \leq q_{AT}$ and the condition (A.4) imply $0 \geq \delta[1 - (1 + q_{AT})e^{-q_{AT}}]$. This cannot hold for $q_{AT} > 0$. Thus, case (iv) is not a solution to the planner's problem. **QED**

B. Proof of Proposition 5.1

Before proving that the strategy described in the proposition is an equilibrium, let me verify the properties of the equilibrium stated in the proposition. Using (5.2), I have

$$p_j - p_{j+1} = e^{-Q_j}\left(\frac{e^{q_j} - 1}{q_j} - \frac{1 - e^{-q_{j+1}}}{q_{j+1}}\right) > 0. \tag{B.1}$$

The inequality follows from the facts that $e^z - 1 > z > 1 - e^{-z}$ for all $z > 0$. From (5.6), I have

$$E_j - E_{j+1} = (y_j - y_{j+1})e^{-Q_j} > 0.$$

Using this result, (5.4) implies:

$$w_j - w_{j+1} = (y_j - y_{j+1})\frac{q_j}{e^{q_j} - 1} - E_{j+1}\left(\frac{1}{p_{j+1}} - \frac{1}{p_j}\right). \tag{B.2}$$

Because $q_j < e^{q_j} - 1$ and $p_j > p_{j+1}$, the above condition implies $w_j - w_{j+1} < y_j - y_{j+1}$ for all j. Thus, $w_j - y_j < w_{j+1} - y_{j+1}$ for all j. In turn, this implies that $w_j < y_j$ for all j if and only if $w_J < y_J$. The latter condition indeed holds, because (5.4) implies:

$$w_J = \frac{E_J}{p_J} = y_J\frac{q_J}{e^{q_J} - 1} < y_J.$$

The final property of the equilibrium is that $w_j < w_{j+1}$ if $(y_j - y_{j+1})$ is small. To show this, note from (B.1) that $(p_j - p_{j+1})$ is bounded strictly above zero even when $(y_j - y_{j+1})$ approaches zero. Thus, (B.2) implies that $w_j < w_{j+1}$ when $(y_j - y_{j+1})$ is sufficiently small.

Now I prove that there is no profitable deviation from the strategy described in the proposition. Consider a deviation by a single firm to the strategy (C, q^d, w^d), where $C = (j_k)_{k=1}^K$ and $1 \leq K \leq J$. Although the ranking in C may not necessarily be strict, I will consider only the case of strict ranking; that is, $j_1 \succ j_2 \succ ... \succ j_K$. A similar proof can be constructed to show that a deviation to having no priority between some elements in C is not profitable. With strict ranking in C, the deviation must satisfy the following properties for all $k \in K$:

$$\frac{e^{-Q^d_{j_{k-1}}} - e^{-Q^d_{j_k}}}{q^d_{j_k}} w^d_{j_k} = E_{j_k} \tag{B.3}$$

$$y_{j_k} e^{-Q^d_{j_k}} - \sum_{i=k+1}^{K} y_{j_i} \left(e^{-Q^d_{j_{i-1}}} - e^{-Q^d_{j_i}} \right) = E_{j_k}, \tag{B.4}$$

where $Q^d_{j_k}$ is defined similarly to (5.1), with q^d replacing q. The first condition is type j_k workers' participation constraint, and the second condition comes from the first-order condition of $w^d_{j_k}$. If one of these conditions is violated, then the deviation can be improved upon by a further deviation to the strategy that satisfies these conditions.

The deviation can differ from the equilibrium strategy in two ways: the ranking in C may not be according to workers' productivity and C may not include all elements of J. In Lemmas B.1 and B.2 below, I will show that if the deviation has either of these differences, then it can be improved upon by a further deviation that eliminates the difference. Thus, a profitable deviation must attract all types of workers and rank them according to productivity. By construction, the best strategy of this kind is the strategy proposed in the equilibrium.

Lemma B.1. *A deviation that does not rank workers according to productivity can be improved upon by another deviation that attracts the same types of workers as in the original deviation but that ranks the workers according to productivity.*

Proof. Consider the deviation (C, w^d, q^d) described above and suppose that the ranking in C is not according to productivity. Then, there exists a number $s = \max\{k \in K : y_{j_k} < y_{j_{k+1}}\}$. That is, type j_s and type j_{s+1} are the lowest ranked pair which exhibits a relative ranking

36

opposite to productivity. Consider a further deviation from this deviation $(C^*, q_{n_k}^*, w_{n_k}^*)_{k \in K}$, where $C^* = (n_k)_{k=1}^K$. This further deviation attracts the same types of workers as those contained in C but it reverses the ranking between type j_s and type j_{s+1} workers. That is, $n_k = j_k$ for all $k \notin \{s, s+1\}$, $n_s = j_{s+1}$ and $n_{s+1} = j_s$. Use q^* to construct Q^* similarly to (5.1). Furthermore, let this further deviation have the following properties:

(i) For all $k \notin \{s, s+1\}$, $w_{n_k}^* = w_{j_k}^d$ and (B.3) is satisfied with $(q_{n_k}^*, Q_{n_k}^*, w_{n_k}^*)$ replacing $(q_{j_k}^d, Q_{j_k}^d, w_{j_k}^d)$.

(ii) For $k \in \{s, s+1\}$, $(q_{n_k}^*, Q_{n_k}^*, w_{n_k}^*)$ solve the following problem:

$$max \sum_{k=s}^{s+1} \left(e^{-Q_{n_{k-1}}^*} - e^{Q_{n_k}^*} \right) \left(y_{n_k} - w_{n_k}^* \right)$$

subject to:

$$\frac{e^{-Q_{n_{k-1}}^*} - e^{-Q_{n_k}^*}}{q_{n_k}^*} w_{n_k}^* = E_{n_k}, \quad k = s, s+1, \tag{B.5}$$

$$q_{n_s}^* + q_{n_{s+1}}^* = q_{j_s}^d + q_{j_{s+1}}^d. \tag{B.6}$$

For $k \leq s - 1$, the two deviations have the same wages and both satisfy (B.3). Comparing the condition (B.3) for the two deviations and working from $k = 1$, I obtain $Q_{n_k}^* = Q_{j_k}^d$ for all $k \leq s - 1$, which implies $q_{n_k}^* = q_{j_k}^d$ for all $k \leq s - 1$. With (B.6), I have $Q_{n_{s+1}}^* = Q_{j_{s+1}}^d$. Using this result and again comparing the condition (B.3) for the two deviations, I have $Q_{n_k}^* = Q_{j_k}^d$ and hence $q_{n_k}^* = q_{j_k}^d$ for all $k \geq s + 2$. Thus, the only difference between the two deviations lies in the employment and wages of type j_s and type j_{s+1} workers.

To show that the further deviation improves upon the original deviation, substitute (B.5) and (B.6) for $(w_{n_s}^*, w_{n_{s+1}}^*, q_{n_{s+1}}^*)$ and use the facts that $n_s = j_{s+1}$, $n_{s+1} = j_s$, $Q_{n_{s-1}}^* = Q_{j_{s-1}}^d$ and $Q_{n_{s+1}}^* = Q_{j_{s+1}}^d$. Then, the objective function of the above maximization problem becomes:

$$\left(e^{-Q_{j_{s-1}}^d} - Q^{-Q_{n_s}^*} \right) y_{j_{s+1}} - q_{n_s}^* E_{j_{s+1}} + \left(e^{-Q_{n_s}^*} - Q^{-Q_{j_{s+1}}^d} \right) y_{j_s} - \left(q_{j_s}^d + q_{j_{s+1}}^d - q_{n_s}^* \right) E_{j_s}.$$

The first-order condition for $q_{n_s}^*$ yields:

$$0 = \left(y_{j_{s+1}} - y_{j_s}\right) e^{-Q_{n_s}^*} + \left(E_{j_s} - E_{j_{s+1}}\right).$$

From (B.4), I can derive the following relationship:

$$E_{j_s} - E_{j_{s+1}} = -(y_{j_{s+1}} - y_{j_s}) e^{-Q_{j_s}^d}. \tag{B.7}$$

The above two conditions imply $Q_{n_s}^* = Q_{j_s}^d$. Then

$$q_{n_s}^* = Q_{n_s}^* - Q_{n_{s-1}}^* = Q_{j_s}^d - Q_{j_{s-1}}^d = q_{j_s}^d.$$

The constraint (B.6) then yields $q_{n_{s+1}}^* = q_{j_{s+1}}^*$.

Let π^d be the firm's expected profit under the original deviation and π^* under the further deviation. Then

$$\pi^d = \sum_{k=1}^{K} \left(e^{-Q_{j_{k-1}}^d} - e^{-Q_{j_k}^d} \right) \left(y_{j_k} - w_{j_k}^d \right),$$

and π^* is given similarly. Subtract the two:

$$\pi^* - \pi^d = \sum_{k=s}^{s+1} \left[\left(e^{-Q_{n_{k-1}}^*} - e^{Q_{n_k}^*} \right) \left(y_{n_k} - w_{n_k}^* \right) - \left(e^{-Q_{n_{k-1}}^*} - e^{Q_{n_k}^*} \right) \left(y_{n_k} - w_{n_k}^* \right) \right].$$

Substituting the above results for (Q^*, q^*), the constraint (B.5) for the wages, and the relationship (B.7), I have

$$\pi^* - \pi^d = \left(y_{j_{s+1}} - y_{j_s}\right) e^{-Q_{j_s}^d} \left[e^{q_{j_s}} - \left(2 + q_{j_s}^d - q_{j_{s+1}}^d \right) + e^{-q_{j_{s+1}}^d} \right].$$

Since $e^z > 1 + z$ and $e^{-z} > 1 - z$ for all $z > 0$, the expression in $[.]$ above is positive. Because $y_{j_{s+1}} > y_{j_s}$, then $\pi^* > \pi^d$. **QED**

Lemma B.2. *A deviation that does not attract all types of workers can be improved upon by a further deviation that attracts all types.*

Proof. Let the original deviation be (C, w^d, q^d), where $C = (j_k)_{k=1}^K$ and $j_k \in J$. Because the deviation does not attract all types, then $K < J$. With Lemma B.1, I can assume that C ranks

38

the workers by productivity, i.e., $j_1 < j_2 < ... < j_K$. Also, since any deviation that does not satisfy (B.3) and (B.4) can be improved upon, the deviation must satisfy these conditions. Let $j^* = \max\{j \in J\backslash K\}$. There are two cases to consider: $j^* > j_K$ and $j^* < j_K$.

If $j^* > j_K$, construct the further deviation (C^*, w^*, q^*) as follows. Set $C^* = \{j_1, j_2, ..., j_K, j_{K+1}\}$, where $j_{K+1} = j^*$. For every $k \leq K$, set the wage $w^*_{j_k} = w^d_{j_k}$ and let the constraint (B.3) hold. For $k = K + 1$, let (B.3) and (B.4) hold. Then, $q^*_{j_k} = q^d_{j_k}$ for all $k \leq K$. This further deviation increases the firm's expected profit, provided $w^*_{j_{K+1}} < y_{j_{K+1}}$. The latter condition can be verified.

Now examine the case $j^* < j_K$. Let $s < K$ be such that $j_{s-1} < j^* < j_s$. From the original deviation, I construct a further deviation (C^*, w^*, q^*), where $C^* = (n_k)_{k=1}^{K+1}$ and n_k is given as:

$$n_k = \begin{cases} j_k, & \text{if } k \leq s-1 \\ j^*, & \text{if } k = s \\ j_{k-1}, & \text{if } k \geq s+1. \end{cases}$$

That is, the type j^* is inserted between the types j_{s-1} and j_s. Define $Q^*_{n_k} = \sum_{i=1}^{k} q^*_{n_k}$. Let the strategy in the further deviation satisfy the following properties:

(i) For all $k \leq s-1$, $w^*_{n_k} = w^d_{j_k}$ and (B.3) is satisfied with $(q^*_{n_k}, Q^*_{n_k}, w^*_{n_k}, K+1)$ replacing $(q^d_{j_k}, Q^d_{j_k}, w^d_{j_k}, K)$.

(ii) For all $k \in \{s, s+1, ..., K+1\}$, (B.3) and (B.4) are satisfied with $(q^*_{n_k}, Q^*_{n_k}, w^*_{n_k}, K+1)$ replacing $(q^d_{j_k}, Q^d_{j_k}, w^d_{j_k}, K)$.

As in the proof of Lemma B.1, property (i) implies $Q^*_{n_k} = Q^d_{j_k}$, $q^*_{n_k} = q^*_{j_k}$, and $w^*_{n_k} = w^d_{j_k}$ for all $k \leq s-1$. For $k \geq s$, recall that $n_{k+1} = j_k$. Subtract (B.4) for j_k in the original deviation and for n_{k+1} in the further deviation, I have:

$$0 = y_{j_k} D_{j_k} - \sum_{i=k+1}^{K} y_{j_i} \left(D_{j_{i-1}} - D_{j_i} \right), \text{ all } k \geq s$$

where $D_{j_k} = e^{-Q^*_{n_{k+1}}} - e^{-Q^d_{j_k}}$. Changing the index k to $k+1$ and subtracting the resulted equation from the above equation, I obtain $(y_{j_k} - y_{j_{k+1}})D_{j_k} = 0$. Since $y_{j_k} > y_{j_{k+1}}$, then $D_{j_k} = 0$ for all $k \geq s$. That is,

$$Q^*_{n_{k+1}} = Q^d_{j_k}, \text{ all } k \geq s.$$

39

In particular, $Q^*_{n_{K+1}} = Q^d_{j_K}$. Then,

$$Q^*_{n_{K+1}} - Q^*_{n_{k+1}} = Q^d_{j_K} - Q^d_{j_k}, \text{ all } k \geq s.$$

Working from $k = K - 1$ to $k = s$, the above equation yields

$$q^*_{n_{k+1}} = q^d_{j_k}, \text{ all } k \geq s + 1.$$

$$q^*_{n_s} + q^*_{n_{s+1}} = Q^*_{n_{s+1}} - Q^*_{n_{s-1}} = Q^d_{j_s} - Q^d_{j_{s-1}} = q^d_{j_s}.$$

Since (B.3) holds in both deviations, the above equalities imply

$$w^*_{n_{k+1}} = w^d_{j_k}, \text{ all } k \geq s + 1.$$

Thus, the only difference between the two deviations lies in type n_s and type n_{s+1} workers.

Let π^d be the firm's expected profit with the original deviation and π^* with the further deviation. Then

$$\pi^* - \pi^d = -\left(e^{-Q^d_{j_{s-1}}} - e^{-Q^d_{j_s}}\right)(y_{j_s} - w^d_{j_s}) + \sum_{k=s}^{s+1}\left(e^{-Q^*_{n_{k-1}}} - e^{-Q^*_{n_k}}\right)\left(y_{n_k} - w^*_{n_k}\right).$$

Substitute (B.3) for w^d and its counterpart for w^*, substitute the relationship (B.7), and use the above relationships between (Q^*, q^*) and (Q^d, q^d). Then,

$$\pi^* - \pi^d = \left(y_{n_s} - y_{n_{s+1}}\right)e^{-Q^*_{n_{s-1}}}\left[1 - (1 + q^*_{n_s})e^{-q^*_{n_s}}\right] > 0.$$

Thus, the original deviation can be improved upon by including the type $n_s = j^*$ in the set of workers to attract. This completes the proof of Lemma B.2 and hence of Proposition 5.1. **QED**

References

[1] Acemoglu, D. and R. Shimer, 1999, "Holdups and efficiency with search frictions," *International Economic Review* 40, 827-850.

[2] Altonji, J.G. and R.M. Blank, 1999, "Race and gender in the labor market," in O. Ashenfelter and D. Card, eds, *Handbook of Labor Economics*, Vol. 3c, 3143-3259, Amsterdam: Elservier.

[3] Black, D.A., 1995, "Discrimination in an equilibrium search model," *Journal of Labor Economics* 13, 309-334.

[4] Bowlus, A.J. and Z. Eckstein, 2002, "Discrimination and skill differences in an equilibrium search model," *International Economic Review* 43, 1309-1345.

[5] Burdett, K., Shi, S. and R. Wright, 2001, "Pricing and matching with frictions," *Journal of Political Economy* 109, 1060-1085.

[6] Hosios, A., 1990, "On the efficiency of matching and related models of search and unemployment," *Review of Economic Studies* 57, 279-298.

[7] Juhn, C., Murphy, K.M. and B. Pierce, 1993, "Wage inequality and the rise in returns to skill," *Journal of Political Economy* 101, 410-442.

[8] Julien, B., Kennes, J. and I. King, 2000, "Bidding for labor," Review of Economic Dynamics 3, 619-649.

[9] Lang, K., Manove, M. and W. Dickens, 2002, "Racial discrimination in labor markets with posted wage offers," manuscript, Boston University.

[10] Moen, E.R., 1997, "Competitive search equilibrium," *Journal of Political Economy* 105, 385-411.

[11] Mortensen, D., 1982, "Property rights and efficiency in mating, racing, and related games," *American Economic Review* 72, 968-979.

[12] Mortensen, D., 2002, "Wage dispersion: why are similar workers paid differently?" manuscript, Northwestern University.

[13] Peters, M., 1991, "Ex ante price offers in matching games: non-steady state," *Econometrica* 59, 1425-1454.

[14] Shi, S., 2001, "Frictional assignment. i. efficiency," *Journal of Economic Theory* 98, 232-260.

[15] Shi, S., 2002, "A directed search model of inequality with heterogeneous skills and skill-biased technology," *Review of Economic Studies* 69, 467-491.

[16] Shimer, R., 1997, "Do good guys come in first? How wage determination affects the ranking of job applicants," unpublished manuscript, Princeton University.