Discrimination in the Equilibrium Search Model with Wage-Tenure Contracts

Fang, Zheng and Sakellariou, Chris

February 2010

Discrimination in the Equilibrium Search Model with Wage-Tenure Contracts

Zheng Fang[1] and Chris Sakellariou
Division of Economics
Nanyang Technological University
Singapore, 639798
Email: fang0031@ntu.edu.sg

Abstract: We extend the Burdett and Coles (2003) search model with wage-tenure contracts to two types of workers and firms and derive the equilibrium earnings distributions for both types of workers, by means of which we succeed in predicting many stylized facts found in empirics. For example, we find that at the same wage level, majority workers almost always experience a faster wage increase than the minority workers; minority workers have a higher unemployment rate; discriminating firms make lower profit than non-discriminating firms and offers to minority workers by non-discriminating firms are consistently superior to those provided by discriminating firms etc. Besides, we find a similar result to the classical discrimination theory that the average wage of the majority workers, though higher in most cases, can be smaller than their counterpart's wage when the fraction of discriminating firms is small and the degree of recruiting discrimination and disutility are mild. We also show that in a special case of CRRA utility function with the coefficient of relative risk aversion approaching infinity, our model degenerates to Bowlus and Eckstein (2002).

Key words: discrimination, wage gap, equilibrium search, wage-tenure

[1]Corresponding Author

1. Introduction

The traditional discrimination literature--the taste-based theory of discrimination (Becker, 1971; Borjas and Bronars, 1989) and statistical discrimination (Aigner and Cain, 1977) are often subject to criticism on the grounds that prejudice cannot possibly be sustained in the long run due to the assumption of a competitive market.[2] Discrimination in the search framework, on the other hand, does not share this problem. For example, Black (1995) studies discrimination in an equilibrium search model where a cost is incorporated in the job search process and discriminating firms are assumed to hire only majority workers. He shows that the wage minority workers receive is less than the wage of their counterparts in the presence of prejudice; and it increases with the proportion of the minority workers in the labor market. However, his model cannot predict wage dispersion among equally productive workers. To overcome this weakness, Bowlus and Eckstein (2002) construct a discrimination framework based on Burdett and Mortensen (1998) in which on-the-job search implies a non-degenerate wage distribution among identical workers. Nonetheless, it also inherits some defects of Burdett and Mortensen (1998), i.e., inefficient equilibrium results and the unrealistic constant wage assumption (noted in Burdett and Coles (2003) and Stevens (2004)). So, in order to generate wage dispersion among similar workers and derive a more realistic model, we use the Burdett and Coles (2003)'s general equilibrium search model with wage-tenure contracts as the framework to explore the implications of discrimination. We are, then, able to identify differences in the patterns of wage dynamics for both types of workers resulting from discrimination, which to the best of our knowledge, has not been explored before.

In what follows, we will outline the discrimination search model with wage-tenure contracts and some equilibrium results. To discuss the effect of discrimination on labor market outcomes, we introduce two types of workers and firms: (1) majority workers A and minority workers B; (2) discriminating firms D and non-discriminating firms N.[3] Workers are assumed to be identical except for their appearance. Firms who experience a disutility from hiring minority workers recruit them at a slower rate. So, for type A workers firms are homogenous while for type B workers they are heterogeneous. In this paper, discrimination is associated with 3 parameters: the fraction of D-firms, the degree of recruiting discrimination and the disutility taste D-firms have when hiring type B workers, all of which are assumed to be exogenously determined.[4]

Our model belongs to a class of random search models.[5] Firms post tenure-based contracts for both types of workers, recruit workers and pay wages specified in the contracts. Workers, both unemployed and employed search for jobs randomly, accept the offers which arrive at an exogenous rate if and only if the expected lifetime value from the new offer is higher than the current one. Firms cannot fire workers or counter-offer workers' outside offers.

---

[2]See Cain (1986) for a good review.

[3]In this paper, majority and minority workers are only representations of different groups of people, say male and female (gender), or white and black (race). They do not necessarily indicate group size.

[4]As shown in the equilibrium, a firm, no matter the discriminating firm or non-discriminating firm, will design different optimal wage contracts for type A and B workers. However, this plausibly "discriminating" result is not what we mean by discrimination in this paper. Instead, we focus on the effect of discriminating recruitment and distaste some firms have on the equilibrium outcomes of the labor market.

[5]Shi (2009) build a directed search model for wage-tenure contracts. However, the incorporation of discrimination in the directed search model with wage-tenure contracts is too complicated to derive a tractable equilibrium solution.

We show that in equilibrium the optimal contract for B provided by N-firms are uniformly better than that provided by D-firms. Though, by offering a high tenure-wage N-firms extract a lower profit from each type B worker, they can hire far more B workers who are willing to stay for a longer period so that the total profit B workers have created in N-firms exceeds that in D-firms. In addition, since both firms make the same profit from type A workers, the total profit is also higher in N-firms than D-firms. This is a general finding encountered in the discrimination literature (see, for example, Becker, 1971; Black, 1995).

Another finding is that the range of discriminating wages is positively related to the fraction of D-firms and inversely related to the degree of recruiting discrimination. Specifically, the fewer D-firms are there in the labor market, the lower the upper bound of discriminating wages would be. Similarly, the more severe the recruiting discrimination, the lower the upper bound would be. Implications as to the lower bound of the discriminating wages are simply reversed. We also find that at the steady state, the lowest wage A-workers are willing to accept is smaller than the lower bound of B's wages only because A-workers can expect a faster wage increase and a larger probability of getting a new offer than their counterparts. However, both are smaller than the unemployment insurance.

The sign of the mean wage gap between type A and B workers is uncertain. If D-firms don't hire B workers at all, it has been proved that on average A earn more than B. However, in a general case, minority workers may have a higher average wage than the majority workers when only a few discriminating firms with weak distaste and recruiting discrimination are there in the labor market.[6] But as shown in the numerical exercises, the odd is quite small. Also found in the numerical section is that in almost all cases, A's wage increases faster than B's, a result that can only be obtained in our discrimination model where the wage is not constant but increases with tenures.

Finally, we point out that our model is a generalization of Bowlus and Eckstein (2002).[7] In a special case of CRRA utility function with the coefficient approaching zero, our model degenerates to Bowlus and Eckstein (2002) and reaches the same equilibrium results. In addition, a sticky floor effect that the wage differential decreases along the wage distribution is found in this case.

This paper makes two contributions to the literature. First, we construct a discrimination model in the search framework with wage-tenure contracts and derive an equilibrium which not only shows the difference in wage distributions but also in wage dynamics. More importantly, our model succeeds in predicting some stylized facts in the labor market, such as a higher unemployment rate of the minority workers; that Whites enjoy faster wage increase along tenures compared to Blacks; that male workers are associated with a wider wage range than female workers; that in most cases, the discriminated group has a lower average wage; and that a sticky floor effect is mainly documented in some Asian countries; etc.

The next section sets up the model and discusses workers' and firms' optimal decisions. Section 3 characterizes the equilibrium solutions and section 4 shows the equilibrium properties. In section 5, we

[6]A similar result is found in Becker(1971) who shows that in a neoclassical framework, equally productive workers, though discriminated by employers in the labor market, may not be paid less if there is only a small fraction of employers with prejudice. Aigner and Cain (1977) also find in one special case that discriminated workers earn more on average.

[7]Strictly speaking, our model is a generalization of the simplified version of Bowlus and Eckstein (2002) because in their model productivity and birth/death rate vary between the two types of workers and offer arrival rates for the employed and the unemployed are also different.

explore the implications using a special case of utility function, the CRRA utility function. When the coefficient of relative risk aversion approaches infinity, the optimal wage-tenure contracts degenerate to a constant wage and our discrimination search model with wage-tenure contracts degenerate to a variant of Bowlus and Eckstein (2002). Further, to facilitate comparisons of the average wages and their dynamics, we carry out numerical exercises in section 6. Section 7 concludes and points out possible future research. All the proofs of the propositions and expressions are given in the appendix.

2. The Model

In this section, we extend the equilibrium search model with wage-tenure contracts (Burdett and Coles 2003) to two types of workers and firms.

2.1 The Environment

Consider an economy consisting of two types of workers and firms. The total population is $n,$ among which the majority workers (type A) are $(1-\theta)n$ and the minority workers (type B) are $\theta n.$ Among all the firms in the labor market, a fraction σ has a distaste for minority workers, denoted by D; and *(1-σ)* are non-discriminating firms denoted by N. Workers are assumed to be equally productive (productivity level P), and have utility function $u(w)$, where $u' > 0, u'' < 0$. They are finitely lived, with a death rate δ. To balance the population, it's assumed that birth rate equals death rate and the newly born people enter the labor force immediately as unemployed. Unemployed workers can obtain an insurance compensation b per instant. Workers--both employed and unemployed--search for better opportunity to maximize their expected lifetime utility.

On the other hand, firms post the wage-tenure contracts and hire workers to maximize their profits. The wage-tenure contract is denoted by $w(t)$, where t denotes tenure—the duration a worker has stayed in the firm. Suppose the offer arrival rate is the same for both employed and unemployed workers. Specifically, it is λ for A; for B, the offer arrival rate from N-firms is λ while from D-firms is $(1-k)\lambda$, where $k \in (0,1)$ shows the degree of recruiting discrimination.[8] The larger k is, the more severe the discrimination. D-firms experience a disutility d from hiring B which enters the profit function directly. Therefore, the instantaneous profit from a B worker who has stayed in the D-firm for a duration t is: $P - w_B^D(t) - d$.

In addition, assume firms cannot fire workers but workers can quit for a better job without suffering any punishment from the previous employer. Furthermore, time preferences of workers and firms are assumed to be zero for the purpose of simplicity. It is also assumed no recalls in the process.

2.2 Workers' Optimal Decision

[8]Differences in search intensity can also account for the differences in job arrival rates. But in this paper, since we assume that both types of workers exert the same level of effort in looking for jobs, the differences in job arrival rates reflect the degree of recruiting discrimination. Indeed, the existence of recruiting discrimination against blacks and women are widely documented, see, for example, Goldin and Rouse (2000), Bertrand and Mullainathan (2004) and Pager et al.(2009).

Let $V_i\left(t\middle|\widetilde{w_i^j}\right)$ be the expected lifetime utility of a type i ($i = A, B$) worker who has tenure t under the wage-tenure contract $\widetilde{w_i^j}$ and use an optimal quit strategy in the future, where $\widetilde{w_i^j}$ denotes the wage-tenure contract a type i worker has signed with firm j ($j = D, N$). $F_A(V_0)$, $F_B^N(V_0)$ and $F_B^D(V_0)$ are the offer distributions for A and B where superscripts N, D denote non-discriminating and discriminating firms respectively. Here V_0 is the starting expected lifetime value of the offer. Thus, the offer distribution measures the proportion of firms who provide workers an starting expected lifetime value no greater than V_0. Since all firms treat A the same, there is no difference in the offer distributions for A provided by N -or D -firms. Let $\underline{V_A}(\overline{V_A})$ denote the infimum (supremum) of the support of F_A and $\underline{V_B^j}$ $(\overline{V_B^j})$ the infimum (supremum) of the support of F_B^j where $j = N, D$.

First consider the situation of employed workers. The standard Bellman equations for employed type A and type B workers are:

$$0 = u\big(w_A(t)\big) - \delta V_A(t|\widetilde{w_A}) + \lambda \int_{V_A(t|\widetilde{w_A})}^{\overline{V_A}} [V_0 - V_A(t|\widetilde{w_A})] dF_A(V_0) + \frac{dV_A(t|\widetilde{w_A})}{dt} \tag{1a}$$

$$0 = u\big(w_B(t)\big) - \delta V_B\left(t\middle|\widetilde{w_B^j}\right) + (1-\sigma)\lambda \int_{V_B\left(t\middle|\widetilde{w_B^j}\right)}^{\overline{V_B^N}} \max\{0, \left[V_0 - V_B\left(t\middle|\widetilde{w_B^j}\right)\right]\} dF_B^N(V_0)$$

$$+\sigma(1-k)\lambda \int_{V_B\left(t\middle|\widetilde{w_B^j}\right)}^{\overline{V_B^D}} \max\{0, \left[V_0 - V_B\left(t\middle|\widetilde{w_B^j}\right)\right]\} dF_B^D(V_0) + \frac{dV_B\left(t\middle|\widetilde{w_B^j}\right)}{dt} \tag{1b}$$

Note that A receives an offer at rate λ, whereas B has a probability of $(1-\sigma)\lambda$ receiving an offer from N -firms and a probability of $\sigma(1-k)\lambda$ receiving an offer from D -firms. The optimal quit strategy implies that they will quit and accept the new offer if and only if its starting value is greater than their current expected lifetime value.[9] The last term in both equations calculates the instantaneous change in the expected lifetime value.

Similarly, we can get the Bellman equations for unemployed workers of both types:

$$0 = u(b) - \delta V_{AU} + \lambda \int_{V_{AU}}^{\overline{V}_A} [V_0 - V_{AU}] dF_A(V_0) \tag{2a}$$

$$0 = u(b) - \delta V_{BU} + (1-\sigma)\lambda \int_{V_{BU}}^{\overline{V_B^N}} [V_0 - V_{BU}] dF_B^N(V_0) + \sigma(1-k)\lambda \int_{V_{BU}}^{\overline{V_B^D}} [V_0 - V_{BU}] dF_B^D(V_0) \tag{2b}$$

The offer provided by firms should be no less than the unemployed lifetime value V_U; otherwise, no workers would be hired. Therefore, $\underline{V_A} \geq V_{AU}$ and $\underline{V_B^j} \geq V_{BU}$ where $j = D, N$.

2.3 Firms' Optimal Decision

[9]Since the relationship between the current expected lifetime value and the supremum of offers from $N(D)$ firm is not clear yet, the maximum between zero and instantaneous change if accepted the offer makes sure the non-negativity and economic meaning. Intuitively, the current value should always be smaller than $\overline{V_B^N}$, but may or may not be smaller than $\overline{V_B^D}$ which means the first max is trivial.

The optimization problem faced by firms is to choose two wage-tenure contracts, one for A and the other for B, to maximize the total expected profit at the steady state. To begin with, we need to derive the expressions of total expected profit for each firm.

Since the quit rate of a type A worker who has stayed t periods under the wage-tenure contract $w_A(t)$ is $\lambda(1-F_A(V_A(t|\widetilde{w_A}))$, the survival probability of such a worker is:

$$\psi_A(t|\widetilde{w_A}) \triangleq \exp\left\{-\int_0^t \left[\delta+\lambda\left(1-F_A\big(V_A(s|\widetilde{w_A})\big)\right)\right] ds\right\} \tag{3a}$$

Similarly, the survival probability of B is:

$$\psi_B\left(t\middle|\widetilde{w_B^J}\right) \triangleq \exp\left\{-\int_0^t \left\{\delta+(1-\sigma)\lambda\left[1-F_B^N\left(V_B\left(s\middle|\widetilde{w_B^J}\right)\right)\right]+\sigma(1-k)\lambda\left[1-F_B^D\left(V_B\left(s\middle|\widetilde{w_B^J}\right)\right)\right]\right\} ds\right\} \tag{3b}$$

Let $G_A(V)$ denote the steady state proportion of A who have an expected lifetime utility less than or equal to V (including the unemployed); and correspondingly, $G_B(V)$ for B. Thus, at the steady state, a firm posting an offer V can recruit $[\lambda G_A(V)(1-\theta)n]$ type-A and $\lambda G_B(V)\theta n$ (if N-firm) or $\lambda(1-k)G_B(V)\theta n$ (if D-firm) type-B workers. The steady state profits of N and D firms are then functions of the wage-tenure contracts:

$$\Omega^{\mathrm{N}}\left({V_0}^A, {V_0}^B\right) =$$

$$\lambda G_A\left({V_0}^A\right)(1-\theta)n\int_0^\infty \psi_A(t|\widetilde{w_A})[P-w_A(t)]\,\mathrm{dt}+\lambda G_B\left({V_0}^B\right)\theta n\int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^N}\right)[P-w_B^N(t)]\,\mathrm{dt} \tag{4a}$$

$$\Omega^{\mathrm{D}}\left({V_0}^A, {V_0}^B\right) =$$

$$\lambda G_A\left({V_0}^A\right)(1-\theta)n\int_0^\infty \psi_A(t|\widetilde{w_A})[P-w_A(t)]\,\mathrm{dt}+\lambda(1-k)G_B\left({V_0}^B\right)\theta n\int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^D}\right)[P-w_B^D(t)-d]\,\mathrm{dt} \tag{4b}$$

In each equation, the first part is the profit from A and the second part is the profit from B. The integration calculates the expected profit that each worker brings to the firm; the part before the integration measures the steady state number of workers hired at given offers. So, the multiplication reflects the firms' expected profit from each type of workers. As both firms treat A equally, profit earned from A is the same between firms in equilibrium.

To derive the optimal decisions of firms, we need to solve the profit maximization problems. Due to additivity, we can solve separately for A; B in N-firms and B in D-firms. Each sub-problem can be solved in two steps:

(i) Conditional on the offer chosen, the optimal wage-tenure contract solves[10]:

[10]It turns out to be an optimal control problem when the two control conditions are rewritten in the form of differential equations with starting values $\psi_i\left(0\middle|\widetilde{w_i^J}\right)=1$ and $V_i\left(0\middle|\widetilde{w_i^J}\right)={V_0}^{ij}$.

$$\underset{w_i^j(t)}{\text{Max}} \int_0^\infty \psi_i\left(t\middle|\widetilde{w_i^j}\right)\left[P - w_i^j(t)\right] \mathrm{dt}$$

$$\text{s.t } \psi_i\left(t\middle|\widetilde{w_i^j}\right) \text{ satisfies (3)}$$

$$V_i\left(t\middle|\widetilde{w_i^j}\right) \text{ satisfies (1)}$$

(ii) The optimal offer solves:

$$\text{Max}_{V_0{}^{ij}}\, G_i\left(V_0{}^{ij}\right) \int_0^\infty \psi_i\left(t\middle|\widetilde{w_i^j}\right)\left[P - w_i^j(t)\right] \mathrm{dt}$$

$$\text{s.t } w_i^j(t) \text{ solves (i)}$$

where $i = A, B; j = N, D$.

When it comes to type B workers in D-firms, the disutility taste d should be further subtracted from $P - w_i(t)$.

3 Equilibrium

Since worker A faces homogenous firms in the labor market, the market equilibrium outcomes for this sub-problem are exactly the same as specified in Burdett and Coles (2003). To facilitate the discussion, we replicate the results in proposition 3.1:[11]

Proposition 3.1 Given $\underline{w_A} > 0$ and $F_A(V)$ is increasing and continuously differentiable, there exists a unique market equilibrium in type-A's labor market. At the steady state equilibrium, the baseline salary scale satisfies:

$$\frac{P - \overline{w_A}}{P - \underline{w_A}} = \left(\frac{\delta}{\delta + \lambda}\right)^2 \tag{5}$$

$$u\left(\underline{w_A}\right) = u(b) - \frac{\sqrt{P - \underline{w_A}}}{2} \int_{\underline{w_A}}^{\overline{w_A}} \frac{u'(x)dx}{\sqrt{P - x}} \tag{6}$$

The optimal wage-tenure contract follows the dynamic path**:**

$$\frac{dw_A}{dt} = \frac{\delta(P - w_A)}{u'(w_A)} \int_{w_A}^{\overline{w_A}} \frac{u'(x)dx}{\sqrt{(P - \overline{w_A})(P - x)}} \tag{7}$$

The earnings distribution is given by[12]:

[11] Refer to Burdett and Coles (2003) for detailed proof; or, see the proof for Proposition 3.3 in the appendix for an outline.
[12] The earnings distribution is not described in Burdett and Coles (2003), but can be easily derived.

$$K_w^A(w) = \frac{\delta}{\lambda}[\sqrt{\frac{P-w_A}{P-w}} - 1] \tag{8}$$

And the unemployment rate is:

$$u_A = \frac{\delta}{\lambda+\delta} \tag{9}$$

Baseline salary scale is a succinct way to describe all the equilibrium solutions. For any starting expected lifetime value V_0 from the support of offer distribution F_A, there exists a point t_0 such that $V_0 = V_s(t_0)$ where the subscript s denotes baseline. So the wage-tenure contract with a starting value V_0 can be expressed as $w(t|V_0) = w_s(t + t_0)$; that is, any equilibrium wage-tenure contract can be found on the baseline salary scale starting with a specific point t_0. In this paper, we suppress s-subscript for simplicity of representation. The optimal decision implied in the proposition is: a firm can set any wage between $[\underline{w_A}, \overline{w_A})$ as the starting wage offer and backload it as described in the optimal wage-tenure dynamic (3); the total profit from A will be the same across firms no matter which wage-tenure contract they choose. Since $\frac{dw_A}{dt}$ is positive, the optimal wage increases with tenure and the upper limit of the increment is $\overline{w_A}$. Obviously, the wage support for type A workers can be solved by combining (5) and (6), from which the earnings distribution (8) can be derived. The unemployment rate is also given for the purpose of comparisons later.

Next, we solve the steady state equilibrium for B. To begin with, we show in proposition 3.2 that the optimal offer for B provided by D-firms is uniformly smaller than that provided by N-firms. Detailed Proofs of all the propositions in the paper are given in the appendix.

Proposition 3.2: Let ${V_{0N}}^B$ denote the optimal offer for B given by N-firms and ${V_{0D}}^B$ the optimal offer provided by D-firms; then we have ${V_{0N}}^B \geq {V_{0D}}^B$.

Proposition 3.2 simplifies the subsequent analysis substantially. As ${V_{0N}}^B \geq {V_{0D}}^B$, equations (1b) and (3b) can be rewritten for B in N- and D- firms separately. Specifically, the Bellman equation for B workers working in N-firms is reduced to:

$$0 = u\left(w_B^N(t)\right) - \delta V_B^N\left(t\middle|\widetilde{w_B^N}\right) + (1-\sigma)\lambda \int_{V_B^N\left(t\middle|\widetilde{w_B^N}\right)}^{\overline{V_B^N}} \left[V_0 - V_B^N\left(t\middle|\widetilde{w_B^N}\right)\right] dF_B^N(V_0) + \frac{dV_B^N\left(t\middle|\widetilde{w_B^N}\right)}{dt} \tag{10}$$

For those working in D-firms the Bellman equation becomes:

$$0 = u\left(w_B^D(t)\right) - \delta V_B^D\left(t\middle|\widetilde{w_B^D}\right) + (1-\sigma)\lambda\left[EV_B^N - V_B^D\left(t\middle|\widetilde{w_B^D}\right)\right]$$

$$+\sigma(1-k)\lambda \int_{V_B^D\left(t\middle|\widetilde{w_B^D}\right)}^{\overline{V_B^D}} \left[V_0 - V_B^D\left(t\middle|\widetilde{w_B^D}\right)\right] dF_B^D(V_0) + \frac{dV_B^D\left(t\middle|\widetilde{w_B^D}\right)}{dt} \tag{11}$$

Similarly, survival probabilities of B workers who are employed by N-firms and D-firms change from 3(b) to:

$$\psi_B\left(t\middle|\widetilde{w_B^N}\right) = exp\left\{-\int_0^t [\delta + (1-\sigma)\lambda(1 - F_B^N\left(V_B\left(s\middle|\widetilde{w_B^N}\right)\right))]ds\right\} \tag{12}$$

$$\psi_B\left(t\middle|\widetilde{w_B^D}\right) = exp\left\{-\int_0^t \left[\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda(1 - F_B^D\left(V_B\left(s\middle|\widetilde{w_B^D}\right)\right))\right] ds\right\} \quad (13)$$

This makes disentanglement of the sub-problems for B workers in N-and D-firm possible. The following proposition describes the equilibrium outcomes in the labor market for B. The crucial step in the proof is to define $G_B^D(V_0)$ and $G_B^N(V_0)$ to replace $G_B(V_0)$. Let $G_B^D(V_0), V_0 \in [\underline{V_B^D}, \overline{V_B^D})$ be the proportion of B who have an expected lifetime value no greater than V_0 in all type-B workers excluding those working in N-firms and $G_B^N(V_0), V_0 \in [\underline{V_B^N}, \overline{V_B^N})$ be the proportion of type-B with expected lifetime value no greater than V_0 in all type-B workers. Then, the proof of the equilibrium outcomes could fit nicely in that of Burdett and Coles (2003). Moreover, through constructing the overall $G_B(V_0)$ from $G_B^D(V_0)$ and $G_B^N(V_0)$, we show that the lower bound of the starting wage in N-firms is the upper limit of starting wages offered by D-firms.

Proposition 3.3: Given $\underline{w_B} > 0$ and $F_B^N(V)$, $F_B^D(V)$ is increasing and continuously differentiable, there exists a unique market equilibrium in the labor market for type-B workers. At the steady state equilibrium, the baseline salary scale for worker B satisfies:

$$\frac{P-\overline{w_B^D}-d}{P-\underline{w_B^D}-d} = \left(\frac{\delta + (1-\sigma)\lambda}{\delta + (1-\sigma k)\lambda}\right)^2 \quad (14)$$

$$u\left(\underline{w_B^D}\right) = u(b) - \frac{\sqrt{P-\underline{w_B^D}-d}}{2}\int_{\underline{w_B^D}}^{\overline{w_B^D}} \frac{u'(x)dx}{\sqrt{P-x-d}} \quad (15)$$

$$\underline{w_B^N} = \overline{w_B^D} \quad (16)$$

$$\frac{P-\overline{w_B^N}}{P-\underline{w_B^N}} = \left(\frac{\delta}{\delta + (1-\sigma)\lambda}\right)^2 \quad (17)$$

The dynamics of baseline salaries are:

$$\frac{dw_B^D}{dt} = \frac{(\delta + (1-\sigma)\lambda)(P-w_B^D-d)}{u'(w_B^D)}\int_{w_B^D}^{\overline{w_B^D}} \frac{u'(x)dx}{\sqrt{\left(P-\overline{w_B^D}-d\right)(P-x-d)}} \quad (18)$$

$$\frac{dw_B^N}{dt} = \frac{\delta(P-w_B^N)}{u'(w_B^N)}\int_{w_B^N}^{\overline{w_B^N}} \frac{u'(x)dx}{\sqrt{(P-\overline{w_B^N})(P-x)}} \quad (19)$$

The earnings distribution is:

$$K_w^B(w) = \begin{cases} \frac{\delta}{(1-\sigma k)\lambda}\left[\sqrt{\frac{P-\underline{w_B^D}-d}{P-w-d}} - 1\right], & if\ w \in [\underline{w_B^D}, \overline{w_B^D}] \\ \frac{\delta+(1-\sigma k)\lambda}{(1-\sigma k)\lambda}\sqrt{\frac{P-\overline{w_B^N}}{P-w}} - \frac{\delta}{(1-\sigma k)\lambda}, & if\ w \in [\underline{w_B^N}, \overline{w_B^N}] \end{cases} \quad (20)$$

The unemployment rate is:

$$u_B = \frac{\delta}{\delta + (1-\sigma k)\lambda} \tag{21}$$

This proposition shows that D-firms can set any starting wage between $[\underline{w_B^D}, \overline{w_B^D})$ and then backload the wage using the rule described in (18). Profit from type-B workers is the same across the discriminating firms. Similarly, N-firms can determine any starting wage between $[\underline{w_B^N}, \overline{w_B^N})$, increase the wage with tenure as described in (19) and make the same profit as other N-firms. One point to note is that although $\underline{w_B^N} = \overline{w_B^D}$, $\underline{V_B^N} \neq \overline{V_B^D}$. Rather, employees hired in N-firms with a payment $\underline{w_B^N}$ have a higher expected lifetime value than the high-earners in D-firms, i.e., $\underline{V_B^N} > \overline{V_B^D}$; because workers with $\underline{w_B^N}$ can expect an immediate increase in the payment while those approaching $\overline{w_B^D}$ cannot.

Second, from the expression of unemployment rate (21), we can see that disutility d has no effect on u_B; and it is always higher than A 's unemployment rate given in (9) if and only if there is discrimination in the labor market ($\sigma k \neq 0$). Note that if any of the two indicators equals to zero, there would be no discriminating firms existing in the labor market.

Third, we can easily get the maximized total profit earned by a D-firm

$$\Omega^{\mathrm{D}} = \lambda(1-\theta)n\frac{P-\overline{w_A}}{\delta} + \lambda(1-k)\theta n\delta\frac{P-\overline{w_B^D}-d}{[\delta+(1-\sigma)\lambda]^2} \tag{22}$$

and that by a N-firm:

$$\Omega^{\mathrm{N}} = \lambda(1-\theta)n\frac{P-\overline{w_A}}{\delta} + \lambda\theta n\frac{P-\overline{w_B^N}}{\delta} \tag{23}$$

Substituting $\overline{w_B^N}$ in (23) with (17), and replacing $\underline{w_B^N}$ with $\overline{w_B^D}$, the difference in profits in N- and D- firm is:

$$\Omega^{\mathrm{N}} - \Omega^{\mathrm{D}} = \lambda\theta n\frac{\delta[k(P-\overline{w_B^D})+(1-k)d]}{[\delta+(1-\sigma)\lambda]^2} > 0 \tag{24}$$

This is a general finding in the discrimination literature. We can see that, though D-firms extract a lot from a single B-worker by paying a lower wage, the total profit is less than that in N-firms; because the negative effect of smaller employment and higher quit rate in D-firm outweights the positive effect of a lower wage. Besides, the disutility taste D-firms have towards workers B widens the profit gap further. It is easy to see that, the larger θ, n, k and d is, the larger the gap.[13] This indicates that having more minority workers in the labor market places the discriminating firms in a worse situation; and, the more prejudiced the discriminating firms are, the higher the loss they will bear.[14]

---

[13]Though values of k and d also influence $\overline{w_B^D}$ in the expression of profit difference, the negative correlation between k, d and $\overline{w_B^D}$ (which to be shown in section 4) will enhance the positive relationship between k, d and the profit gap.

[14]As the taste theory of employer discrimination (Becker, 1971) shows, the discriminating firms have to bear the cost of their distaste for the minority workers. Nevertheless, in a competitive environment the conclusion is often

In the next section, we focus on a more intriguing issue: comparisons of the steady state wages and the dynamics between both types of workers.

4 Equilibrium Properties

First, let's consider the impact of 3 parameters associated with discrimination on B workers' equilibrium wages. We can solve for the discriminating wage bounds from (14) and (15) and non-discriminating wage bounds from (16) and (17). Obviously, they are functions of productivity P, unemployment insurance b, birth-death rate δ, normal offer arrival rate λ and three discrimination indicators (σ, k, d). The comparative statics yield:

(1) $\frac{\partial \underline{w_B^D}}{\partial \sigma} < 0 \quad \frac{\partial \overline{w_B^D}}{\partial \sigma} > 0$

(2) $\frac{\partial \underline{w_B^D}}{\partial k} > 0 \quad \frac{\partial \overline{w_B^D}}{\partial k} < 0 \quad \frac{\partial \overline{w_B^N}}{\partial k} < 0$

(3) $\frac{\partial \overline{w_B^D}}{\partial d} < 0 \quad \frac{\partial \overline{w_B^N}}{\partial d} < 0$

Besides, we can prove:

(4) $\underline{w_B^D} \leq b$; and $\overline{w_B^D} > b$ if $b \leq \frac{3}{4}(P - d)$

Property (1) shows that the higher the proportion of D-firms in the market, the wider the range of discriminating wages will be. Moreover, the range extends in both directions. On the contrary, the degree of recruiting discrimination has an opposite effect; severe discrimination in the hiring process would lead to a narrowing of the discriminating wage range which converges to the unemployment insurance (which is implied by property (4)). The highest non-discriminating wage also decreases as the recruiting discrimination increases. Finally, the disutility of D-firms is negatively related to the upper bound of both the discriminating wages and non-discriminating wages for type B workers.

It's interesting to observe that the lowest acceptable wage is lower than the unemployment insurance. This result is unique within the search model with wage-tenure contract literature. In Burdett and Mortensen (1998) where firms set a constant wage rather than a wage-tenure contract, the lowest acceptable wage is the unemployment insurance b when the offer arrival rate is the same for both the employed and the unemployed. Under the wage-tenure framework, however, workers are willing to work at a wage lower than the unemployment insurance only because they can expect an immediate increase in the payment. In fact, the expected lifetime value at the lowest wage is virtually equal to that at the status of unemployment.

Comparing the wage range of type A and B results in property(5):

(5) $\underline{w_A} < \underline{w_B^D} \leq b < \overline{w_B^D} < \overline{w_A}$

criticized, as it is not persisting in the long-run equilibrium. In this paper, the issue disappears due to implicit assumptions of the frictional labor market and exogenously given fraction of (non)discriminating firms in the model.

A's lowest acceptable starting wage is less than the lowest starting wage for B because first, worker A's wage increases with tenure more quickly than B's; second, compared to B, A is more likely to get a new and better job offer in the labor market. The upper bound of A's wages being higher than that of their counterpart is within expectation. Discriminating firms are unlikely to set too high a wage due to their disutility tastes.

Next, to see whether our model can predict the findings in empirical studies that female workers earn less than male workers on average (or the black earn less than the white) even though the productivity characteristics such as experience, education and training are controlled for, we derive the mean wages of both types of workers from (5), (8), (14), (16), (17) and (20), which gives:[15]

$$Ew_A = \int_{\underline{w_A}}^{\overline{w_A}} w dK_w^A(w)$$

$$= \overline{w_A} + \frac{\delta}{\lambda}(\overline{w_A} - \underline{w_A}) - \frac{2\delta}{\delta + \lambda}(P - \underline{w_A}) \tag{25}$$

$$Ew_B = \int_{\underline{w_B^D}}^{\overline{w_B^N}} w dK_w^B(w)$$

$$= \overline{w_B^N} + \frac{\delta}{(1-\sigma k)\lambda}(\overline{w_B^N} - \underline{w_B^D}) - \frac{2}{1-\sigma k}\left[\frac{\delta\sigma(1-k)}{\delta+(1-\sigma k)\lambda}\left(P - \underline{w_B^D} - d\right) + \frac{(1-\sigma)[\delta+(1-\sigma k)\lambda]}{\delta}(P - \overline{w_B^N})\right] \tag{26}$$

Note that the unemployed workers are not included in the calculation.

If only N-firms hire type B workers, then "minority workers receive lower wages than workers not facing discrimination" (Black, 1995) as long as there are discriminating firms in the labor market ($\sigma \neq 0$).[16] Stated in property (6), that is:

(6) If $k = 1$ and $\sigma \neq 0$, then $Ew_A > Ew_B$.

However, this finding cannot be generalized. In the numerical example, we will show that if D-firms can hire B ($0 \leq k < 1$), the average type B worker might be able to earn slightly higher wage than A. Discussions concerning the comparison of the average wages between A and B are deferred in section 6 as it is almost impossible to get any conclusions without a particular utility function form or the parameters' values.

At last, we derive the wage quantiles of worker A and B:

[15]See Altonji and Blank (1999) for an overview; Blau and Kahn (1996, 1997, 2000, 2003) for discussions on gender wage differentials across countries; and Smith (1993), Chandra (2000) and Antecol and Bedard (2004), among others for discussions on racial gender gap.

[16]In Black (1995), the wage of minority workers increases with their proportion. However, here the fraction of workers doesn't enter any equilibrium outcomes. There are two reasons for this. First, in Black (1995) the number of discriminating firms decreases as the minority workers increase, while in this paper σ is assumed to be exogenous. Second, the search friction varies in the two models. Unlike Black (1995) in which an exogenous cost is imposed on the search process, this paper presents a random search model where the friction is embodied in the finite job arrival rate. Thus wage dispersion resulted from random search makes comparisons more complicated than Black (1995)'s case.

$$w_A^q = P - (P - \underline{w_A})\left(\frac{\delta}{\delta + \lambda q}\right)^2 \tag{27}$$

$$w_B^q = \begin{cases} P - d - \left(P - \underline{w_B^D} - d\right)\left(\frac{\delta}{\delta + (1-\sigma k)\lambda q}\right)^2 & if\ 0 \le q < \frac{\delta(1-k)\sigma}{(1-\sigma k)[\delta + (1-\sigma)\lambda]} \\ P - \left(P - \overline{w_B^N}\right)\left(\frac{\delta + (1-\sigma k)\lambda}{\delta + (1-\sigma k)\lambda q}\right)^2 & if\ \frac{\delta(1-k)\sigma}{(1-\sigma k)[\delta + (1-\sigma)\lambda]} \le q \le 1 \end{cases} \tag{28}$$

Comparisons between the wage quantiles will be shown in a special case and numerical exercises later. We will see how the wage disparity changes along the wage distribution. Can our model predict glass ceilings or sticky floors commonly found in empirics, though we are aware that the skewness of the earnings distribution (equations (8) and (20)) do not fit the data well?[17]

5. A special case

In this section, a special case of the CRRA utility function: $u(w) = \frac{w^{1-\gamma}}{1-\gamma}$ $(\gamma \to \infty)$ is considered.[18] Tractable equilibrium solutions that are derived from proposition 3.1 and 3.3 and the special CRRA utility function can shed more light on the labor market with discrimination. Proposition 5 below summarizes the equilibrium results in this special case.

Proposition 5: Given that both types of workers have the same CRRA utility function: $u(w) = \frac{w^{1-\gamma}}{1-\gamma}$ with $\gamma \to \infty$, the following statements hold:

(1) The optimal strategy of firms is to set fixed wages instead of the wage-tenure contracts, i.e., $\frac{dw_i^j}{dt} = 0$.

(2) The wage bounds are:

$$\underline{w_A} = b \qquad \overline{w_A} = P - \left(\frac{\delta}{\lambda + \delta}\right)^2 (P - b)$$

$$\underline{w_B^D} = b \qquad \overline{w_B^D} = \underline{w_B^N} = P - d - \left(\frac{\delta + (1-\sigma)\lambda}{\delta + (1-\sigma k)\lambda}\right)^2 (P - b - d)$$

$$\overline{w_B^N} = P - \left(\frac{\delta}{\delta + (1-\sigma)\lambda}\right)^2 d - \left(\frac{\delta}{\delta + (1-\sigma k)\lambda}\right)^2 (P - b - d)$$

And the relationships among these bounds are $\underline{w_A} = \underline{w_B^D} = b < \overline{w_B^D} < \overline{w_B^N} < \overline{w_A}$.

(3) A's earnings distribution first order dominates B's earnings distribution, i.e., $K_w^A < K_w^B$ for all w.

[17] A growing literature studying the wage differential across distributions has emerged in recent decades. Glass ceilings, that the relative wage gap increases with quantile are commonly documented in developed countries like Sweden (Albrechet et al. 2003) and most European countries (Arulampalam et al. (2007)) while sticky floors that the relative wage gap reaches its maximum at the lower tail of the distribution are mainly found in Asian areas such as Singapore, the Philippines (Sakellariou (2004a; 2004b), Thailand (Fang and Sakellariou (2010), Vietnam (Pham and Reilly, 2007) and China (Chi and Li(2007)).

[18] Another two special cases, risk neutral and log utility functions are of interest as well since tractable solutions may be derived from partial differential equations with initial value conditions. We leave this as future work.

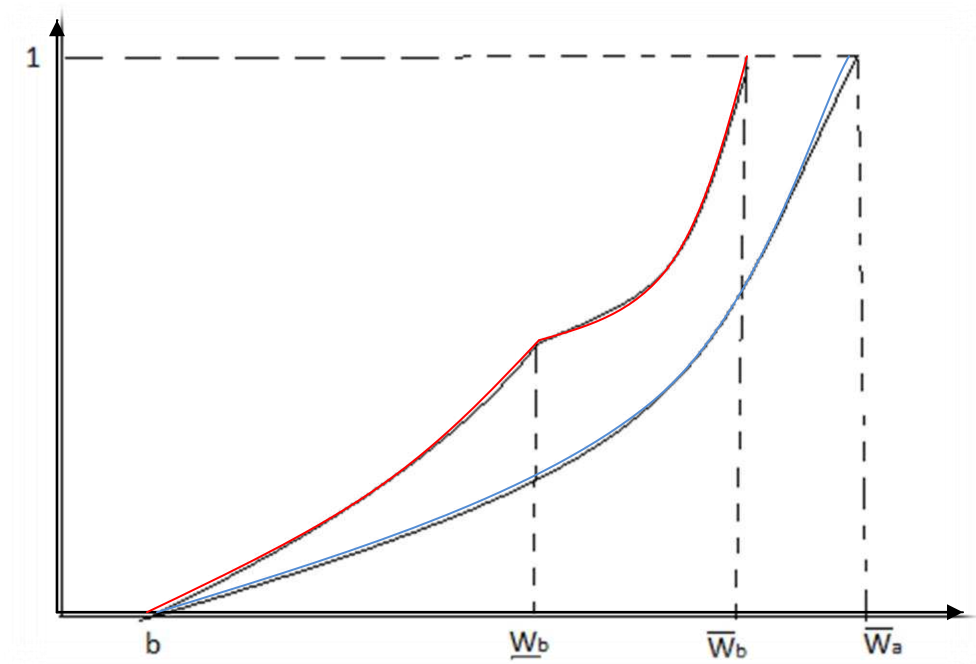
(4) $Ew_A > Ew_B$ and the mean wage gap increases with (σ, k, d).

(5) $w_A^q > w_B^q$ and the difference in wages at qth quantile decreases with q.

As $\gamma \rightarrow \infty$, workers are infinitely risk averse; thus the optimal wage contract is constant wages, i.e., $\frac{dw_i^j}{dt} = 0$. The equilibrium search model with wage-tenure contracts then degenerates to Burdett and Mortensen (1998) and the discriminating wage-tenure equilibrium search model degenerates to a simplified version of Bowlus and Eckstein (2002).[19] Figure 1 describes the earnings distributions for both types of workers and obviously A's cumulative earnings distribution first order dominates B's distribution. From first order dominance, properties (4) and (5) are directly obtained. In addition, the same reservation wages between A and B result from the assumption that the offer arrival rate is invariant between the employed and unemployed workers. The upper wage limit for B is less than A's because of the existence of the three non-zero discrimination parameters (σ, k, d).

Figure 1: Earnings distributions

------------ type A worker ------------ type B worker



Moreover, the larger (σ, k, d) is, the smaller B's average wage is. Since (σ, k, d) does not enter type A worker's wage, the average wage gap increases as (σ, k, d) increases. This conclusion is in line with the empirical findings. For example, Charles and Guryan (2008) plot the black-white wage gap against

[19] Bowlus and Eckstein (2002) extend Burdett and Mortensen (1998)'s model to discuss the contributions of discrimination and skill differences to the wage gaps. In their paper, the offer arrival rate is assumed to be different between the employed and the unemployed and therefore unlike what we get in this special case, the reservation wage is larger than the unemployment compensation

prejudicial attitude and find a wider gap at regions where many people will not vote for the black candidate for presidency or are against interracial marriages.

In addition, (5) indicates a "sticky floor" effect since the (relative) wage gap is decreasing along the distribution under the assumption that both types of workers possess the same productivity. This special case seems to show that countries with high risk averse population are very likely to experience a sticky floor effect, a hypothesis definitely requiring more rigorous analysis and empirical evidence. We leave it as one possible direction for future research.

6 Numerical Example

As mentioned in section 4, it is interesting to examine the effect of the three discrimination-relevant parameters on the difference in the mean wages between type A and B workers. We assume in the section that all workers have the same CRRA utility function. The parameter values we choose are $P = 300, b = 100$, $\lambda = 0.03$ and $\delta = 0.003$.[20] If the coefficients of relative risk aversion are $0.4, 0.9, 1.4$ and 1.9, equation (25) gives that A's average wages are 272.1134, 273.3307, 275.3025 and 276.8115 respectively. It seems that the more risk averse workers are, the higher the average wage they would earn.

For worker B, we vary the values of (σ, k, d) to see how the mean wage changes accordingly. Results are presented in table 1 in which the first panel fixes d and k, and changes the share of discriminating firms σ; the second panel changes the recruiting discrimination k and keeps the other two measures unchanged; the third one modifies disutility taste d given certain values of σ and k. The findings are as follows. First, the mean wage of type B worker decreases in σ and d, but increases in γ while the relationship with k is uncertain. Second, the fraction of D-firms plays a key role in the average wage; the other three parameters, though matter to some extent, have only limited influence on the wage outcomes. Third, if only D-firms exist in the labor market (see the case $\sigma = 1$ in Panel 1), the wage gap is very large; however, the gap will drop dramatically when N-firms begin to appear. In addition, Panel (2) indicates that the wage gap does not change much even when D-firms are forbidden to discriminate in hiring (see $k = 0$); on the other hand, what appears to be against expectation is that severe discrimination in recruitment leads to higher average wage for B and hence smaller wage gap (see $k = 0.9$). However, one should realize that this does not mean type B workers are better off because only a few will be hired in this situation and the overall welfare of type B workers is in fact jeopardized.

Finally, compared to A's average wage, the numbers in Table 1 are almost consistently smaller, which accords with the common sense that discriminated workers have lower average wage. But, there are some exceptions. For example, when $\sigma = 0.2$ and $\gamma = 0.9$, B's average wage is 273.3993 in Panel 1, a little larger than A's mean wage 273.3307. This implies that when there is only mild discrimination against the minority workers, the discriminated group may earn more than the non-discriminated group. This is because convex earnings distribution indicates more workers distributed at high wages. If the number of N-firms is sufficiently large, only a few B receive lower discriminating wages and some B workers even

[20] The values are borrowed from Bowlus and Eckstein (2002) where P is identified from the mean weekly earnings among black and white males who worked full-time between 1985 and 1988; b is the minimum weekly wage and λ and δ are estimates of unemployment rate and death rate.

get wages higher than their counterparts. So the average values of B's wages can exceed A's average wage in rare cases. Indeed, similar results have appeared in the discrimination literature. Becker (1971) shows that in a competitive labor market, a wage differential occurs if and only if the fraction of discriminating firms is large enough. Aigner and Cain (1977) demonstrate in the case where the mean productivities are the same but variances are different that for less skilled workers the discriminated-against workers have a higher average wage than their counterparts. In this paper, not only have we obtained this surprising result, but derived wage dispersions among equally productive workers which cannot be achieved in the first two types of discrimination in the literature.

Table 1: The mean wage of type B workers

(1) $d = 80$ $k = 0.2$

	$\gamma = 0.4$	$\gamma = 0.9$	$\gamma = 1.4$	$\gamma = 1.9$
$\sigma = 0.2$	272.1293	273.3993	274.2985	274.9699
$\sigma = 0.4$	266.8115	268.1293	269.0627	269.7598
$\sigma = 0.6$	257.9216	259.2913	260.2616	260.9865
$\sigma = 0.8$	239.8846	241.3105	242.3210	243.0759
$\sigma = 1.0$	199.2173	200.7044	201.7586	202.5463

(2) $d = 80$ $\sigma = 0.5$

	$\gamma = 0.4$	$\gamma = 0.9$	$\gamma = 1.4$	$\gamma = 1.9$
$k = 0.0$	263.3952	264.9373	265.9374	266.6517
$k = 0.3$	263.3267	264.4216	265.2213	265.8328
$k = 0.6$	263.9390	264.4312	264.8563	265.2231
$k = 0.9$	265.6882	265.7274	265.7662	265.8044

(3) $\sigma = 0.5$ $k = 0.5$

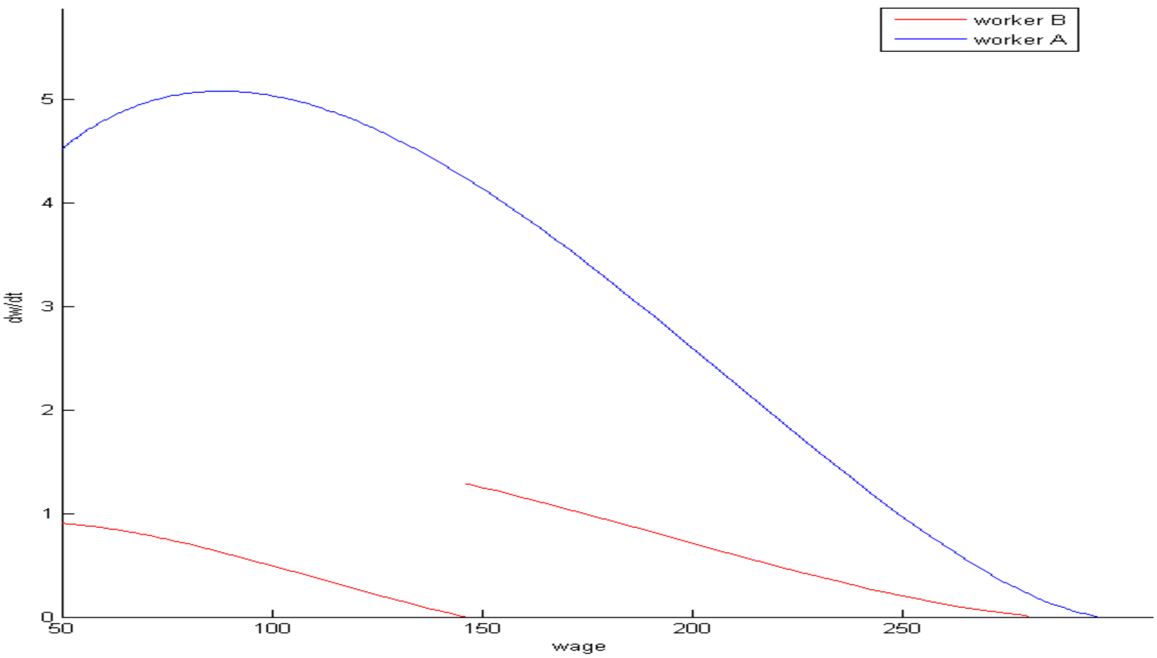
	$\gamma = 0.4$	$\gamma = 0.9$	$\gamma = 1.4$	$\gamma = 1.9$
$d = 10$	267.9042	269.3757	270.4330	271.2300
$d = 80$	263.6369	264.3334	264.9009	265.3683
$d = 150$	259.8387	259.9760	260.1040	260.2230

Next, we discuss the difference in wage dynamics between both types of workers. To be representative, we choose a most realistic case where $\gamma = 0.9$, $k = 0.8$, $\sigma = 0.8$ and $d = 80$ and a special case in which B's mean wage exceeds that of type A worker (See figure 2).[21]

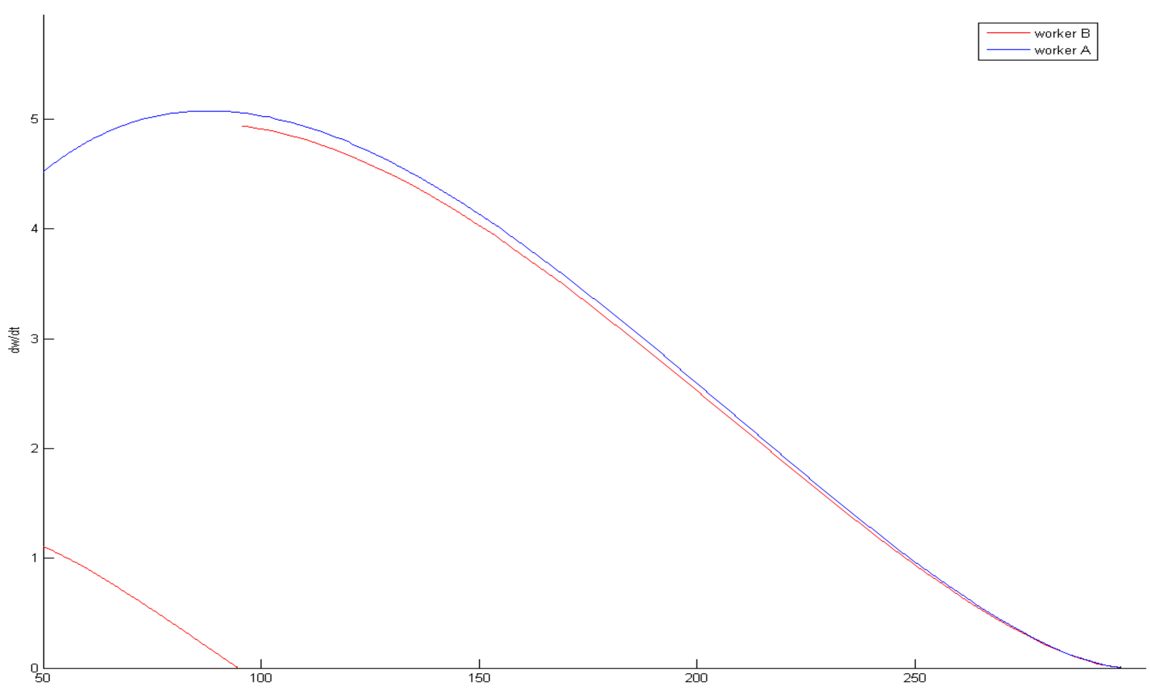
[21] Given those values, the simulated average wages for A and B are 273.3307 and 227.6730 respectively, very close to 273.9 and 230.96 derived from real data (Bowlus and Eckstein, 2002).

Figure 2: Wage Dynamics

(a) Severe discrimination



(b) Mild discrimination



There are several points worth noting. First, the slope of the wage-tenure contract is positive, meaning that the wage will increase with tenure. Second, for type A workers, the increase accelerates at the beginning, and slows down gradually; on the other hand, for type B workers the increasing rate drops from the very beginning. Besides, it is found that the slope of A's wage-tenure contract is always larger than B's in D-firms. However, as to the slope of wage-tenure contracts designed for type B workers by N-firms, it can be very close to type A's slope under mild discrimination, or even exceed that (figure b); while under severe discrimination where many firms have disutility taste towards type B workers and offer them less job opportunities, N-firms will not have an incentive to offer a sufficiently attractive contract for B and hence the increasing rate is much smaller than type A's slope (figure a).[22]

Through the numerical example, we can clearly see and compare mean wages, wage dynamics and other aspects of interest.[23] One surprising result highlighted in this paper is that given the same productivity, the discriminated group could earn more than the non-discriminated group on average if discrimination in the labor market is not severe.

7 Conclusions

In this paper we construct a discrimination search model with wage-tenure contracts based on Burdett and Coles (2003) and succeed in predicting many stylized facts found in empirics using this model framework. For example, we show that discriminating firms earn less than non-discriminating firms, the unemployment rate is higher for minority workers than majority workers and earnings distributions for both types of workers vary. In addition, we show that for minority workers, the tenure-based wage in discriminating firms is invariably smaller than that in non-discriminating firms; and it increases much slower compared to the majority workers' wage. Finally, the finding we would like to emphasize is that the mean wage gap is not positive under all circumstances. In very scarce cases of mild discrimination, equally productive and discriminated workers may get a higher average wage. Some other implications of discrimination on the labor market outcomes are discussed as well.

In future research, the assumption of same productivity among workers can be relaxed. It will not affect the form of equilibrium results, however the discussion on the comparisons of wage-tenure contracts will become considerably complicated as analytical solutions may be impossible. Nevertheless, one may empirically identify and estimate parameters in the structural model and hence would be able to get the contributions of productivity differences vs. discrimination in the wage differentials between workers. Wage data by tenure could be used in such future research. In addition, one can incorporate the free entry condition to endogenize the equilibrium firm numbers in the labor market. Under this assumption, total profit should decline to zero; but an additional assumption on the fraction of potential discriminating firms will have to be added.

[22] The slopes of the wage-tenure contracts are compared at the same wage level.
[23] Wage differentials across quantiles could also be obtained in the numerical example, but not discussed here due to its failure to be representative.

References

Aigner, D. J. and G. G. Cain (1977). "Statistical Theories of Discrimination in Labor Markets." Industrial and Labor Relations Review **30**(2): 175-187.

Albrecht, J., A. Björklund, et al. (2003). "Is There a Glass Ceiling in Sweden?" Journal of Labor Economics **21**(1): 145-177.

Altonji, J. G. and R. M. Blank (1999). "Race and Gender in the Labor Market" in Handbook of Labor Economics , vol. 3c, edited by O. Ashenfelter and D. Card: 3143-3259. Amsterdam: Elsevier.

Antecol, H. and K. Bedard (2004). "The Racial Wage Gap: The Importance of Labor Force Attachment Differences across Black, Mexican, and White Men." Journal of Human Resources XXXIX(2): 564-583.

Arulampalam, W., A. L. Booth, et al. (2007). "Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution." Industrial and Labor Relations Review **60**(2): 163-186.

Becker, G., Ed. (1971). The Economics of Discrimination 2nd ed., Chicago: University of Chicago Press.

Bertrand, M. and S. Mullainathan (2004). "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." The American Economic Review **94**(4): 991-1013.

Black, D. A. (1995). "Discrimination in an Equilibrium Search Model." Journal of Labor Economics **13**(2): 309-334.

Blau, F. D. and L. M. Kahn (1996). "Wage Structure and Gender Earnings Differentials: An International Comparison." Economica **63**: S29-S62.

_______ and _______ (1997). "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s." Journal of Labor Economics **15**(1): 1-42.

_______ and _______ (2000). "Gender Differences in Pay." Journal of Economic Perspectives **14**(4): 75-99.

_______ and ______ (2003). "Understanding International Differences in the Gender Pay Gap." Journal of Labor Economics **21**(1): 106-144.

Borjas, G. J. and S. G. Bronars (1989). "Consumer Discrimination and Self-Employment." Journal of Political Economy **97**: 581-605.

Bowlus, A. J. and Z. Eckstein (2002). "Discrimination and Skill Differences in an Equilibrium Search Model." International Economic Review **43**(4): 1309-1345.

Burdett, K. and D. T. Mortensen (1998). "Wage Differentials, Employer Size, and Unemployment." International Economic Review **39**(2): 257-273.

________ and M. Coles (2003). "Equilibrium Wage-Tenure Contracts." Econometrica **71**(5): 1377-1404.

Cain, G. G. (1986). "The Economic Analysis of Labor Market Discrimination: A Survey." in Handbook of Labor Economics , vol. 1, edited by O. Ashenfelter and R. Layard: 693-781. Amsterdam: Elsevier.

Chandra, A. (2000). "Labor-Market Dropouts and the Racial Wage Gap: 1940-1990." American Economic Review **90**(2): 333-338.

Charles, K.K. and J. Guryan (2008). "Prejudice and Wages: An Empirial Assessment of Becker's *The Economics of Discrimination*." Journal of Political Economy **116**(5): 773-809.

Chi, W. and B. Li (2008). "Glass ceiling or sticky floor? Examining the gender earnings differential across the earnings distribution in urban China, 1987-2004." Journal of Comparative Economics **36**(2): 243-263.

Fang, Z. and C. Sakellariou (2011). "A Case of Sticky Floors: Gender Wage Differential in Thailand." Forthcoming in Asian Economic Journal.

Goldin, C. and C. Rouse (2000). "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians." American Economic Review **90**(4): 715-741.

Pager, D., B. Western, et al. (2009). "Discrimination in a Low-Wage Labor Market: A Field Experiment." IZA DP No. 4469.

Pham, H. T. and B. Reilly (2007). "The Gender Pay Gap In Vietnam, 1993-2002: A Quantile Regression Approach." Journal of Asian Economics **18**(5): 775-808.

Sakellariou, C. (2004a). "The Use of Quantile Regressions in Estimating Gender Wage Differentials: A Case Study of the Philippines." Applied Economics **36**(9): 1001-1007.

Sakellariou, C. (2004b). "Gender-Earnings Differentials Using Quantile Regressions." Journal of Labor Research **25**(3): 457-468.

Shi, S. (2009). "Directed search for equilibrium wage-tenure contracts." Econometrica **77**(2): 561-584.

Smith, J. P. (1993). "Affirmative Action and the Racial Wage Gap." American Economic Review **83**(2): 79-84.

Stevens, M. (2004). "Wage-Tenure Contracts in a Frictional Labour Market: Firms' Strategies for Recruitment and Retention." The Review of Economic Studies **71**(2): 535-551.

Appendix

Proof of proposition 3.2

Proof:

Since ${V_{0N}}^B$ and ${V_{0D}}^B$ are offers chosen by N-and D-firms to maximize their respective profit flow at the steady state, it implies

$$\lambda G_B\left({V_{0N}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^N}\right)[P - w_B^N(t)]\,\mathrm{dt}$$
$$\geq \lambda G_B\left({V_{0D}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^D}\right)[P - w_B^D(t)]\,\mathrm{dt}$$

and

$$(1-k)\lambda G_B\left({V_{0D}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^D}\right)[P - w_B^D(t) - d]\,\mathrm{dt}$$
$$\geq (1-k)\lambda G_B\left({V_{0N}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^N}\right)[P - w_B^N(t) - d]\,\mathrm{dt}$$

Note that $\widetilde{w_B^j}$ $(j = N, D)$ is the wage-tenure contract designed to deliver the offer, so it's a function of ${V_{0j}}^B$. The two inequalities then imply:

$$G_B\left({V_{0N}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^N}\right)[P - w_B^N(t)]\,\mathrm{dt} - G_B\left({V_{0N}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^N}\right)[P - w_B^N(t) - d]\,\mathrm{dt}$$
$$\geq G_B\left({V_{0D}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^D}\right)[P - w_B^D(t)]\,\mathrm{dt} - G_B\left({V_{0D}}^B\right)\theta n \int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^D}\right)[P - w_B^D(t) - d]\,\mathrm{dt}$$

If we define:

$$\Psi\left({V_0}^B\right) \triangleq dG_B\left({V_0}^B\right)\int_0^\infty \psi_B(t|\widetilde{w_B})\,\mathrm{dt}$$

Then the above inequality is:

$$\Psi({V_{0N}}^B) \geq \Psi({V_{0D}}^B)$$

Because,

$$\Psi'\left({V_0}^B\right) = d\frac{\partial G_B\left({V_0}^B\right)}{\partial {V_0}^B}\int_0^\infty \psi_B(t|\widetilde{w_B})\,dt + dG_B\left({V_0}^B\right)\int_0^\infty \frac{\partial(\psi_B(t|\widetilde{w_B}))}{\partial {V_0}^B}dt > 0$$

due to the increasing property of $G_B\left({V_0}^B\right)$ and $\psi_B(t|\widetilde{w_B})$ with respect to ${V_0}^B$, we have ${V_{0N}}^B \geq {V_{0D}}^B$.

Proof of proposition 3.3

Proof:

(1) First consider the optimal wage-tenure contract designed for B-workers by discriminating firms. Given the starting offer V_0, the wage-tenure function solves:

$$\max_{w(t)>0}\int_0^\infty \psi_B\left(t\middle|\widetilde{w_B^D}\right)[P - w_B^D(t) - d]\,\mathrm{dt}$$

where $$\dot{\psi_B} = -[\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda(1 - F_B^D(V_B^D))]\psi_B \qquad \text{(A1)}$$

$$\dot{V_B^D} = \delta V_B^D - u\left(w_B^D(t)\right) - (1-\sigma)\lambda[EV_B^N - V_B^D] - \sigma(1-k)\lambda\int_{V_B^D}^{\overline{V_B^D}}[x - V_B^D]dF_B^D(x) \qquad \text{(A2)}$$

with starting values $\psi_B(0) = 1;\ V_B^D(0) = V_0$

To solve the dynamic optimization problem, define the Hamiltonian:

$$H = \psi_B[P - w_B^D(t) - d] - \Lambda_\psi\left[\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda\left(1 - F_B^D(V_B^D)\right)\right]\psi_B$$

$$+\Lambda_V\left[\delta V_B^D - u\left(w_B^D(t)\right) - (1-\sigma)\lambda[EV_B^N - V_B^D] - \sigma(1-k)\lambda\int_{V_B^D}^{\overline{V_B^D}}[x - V_B^D]dF_B^D(x)\right]$$

Where Λ_ψ, Λ_V are costate variables with respect to ψ_B and V_B^D.

The necessary conditions are:

$$H_w = -\psi_B - \Lambda_V\, u'\left(w_B^D(t)\right) = 0 \qquad \text{(A3)}$$

$$\dot{\Lambda_\psi} = -H_\psi = -[P - w_B^D(t) - d] + \Lambda_\psi[\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda(1 - F_B^D(V_B^D))] \qquad \text{(A4)}$$

$$\dot{\Lambda_V} = -H_V = -\Lambda_V\left[\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda[1 - F(V_B^D)]\right] - \Lambda_\psi\sigma(1-k)\lambda F_B^{D\prime}(V_B^D)\psi_B \qquad \text{(A5)}$$

And the two differential equations ψ_B and V_B^D should satisfy (A1), (A2).

Integrate (A4) with the integrating factor ψ_B yields:

$$\Lambda_\psi\psi_B = \int_t^\infty \psi_B\left(s\middle|\widetilde{w_B^D}\right)[P - w_B^D(s) - d]\,\mathrm{ds} + \mathrm{C}_1$$

Define the expected future profit flow from tenure period t onwards as:

$$\Pi_B^D\left(t\middle|\widetilde{w_B^D}\right) = \int_t^\infty \frac{\psi_B\left(s\middle|\widetilde{w_B^D}\right)}{\psi_B\left(t\middle|\widetilde{w_B^D}\right)}[P - w_B^D(s) - d]\,\mathrm{ds}$$

Then,

$$\Lambda_\psi = \Pi_B^D\left(t\middle|\widetilde{w_B^D}\right) + \frac{\mathrm{C}_1}{\psi_B\left(t\middle|\widetilde{w_B^D}\right)}$$

Since it's an autonomous control problem, the optimized Hamiltonian is zero, i.e., $H = 0$. Substituting Λ_ψ, Λ_V in H out yields:

$$0 = [P - w_B^D(t) - d] - \left\{\Pi_B^D\left(t\middle|\widetilde{w_B^D}\right) + \frac{\mathrm{C}_1}{\psi_B\left(t\middle|\widetilde{w_B^D}\right)}\right\}\left[\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda\left(1 - F_B^D(V_B^D)\right)\right]$$

$$-\frac{1}{u'\left(w_B^D(t)\right)}\left[\delta V_B^D - u\left(w_B^D(t)\right) - (1-\sigma)\lambda[EV_B^N - V_B^D] - \sigma(1-k)\lambda\int_{V_B^D}^{\overline{V_B^D}}[x - V_B^D]dF_B^D(x)\right]$$

Therefore, C_1 has to be zero to make Π_B^D bounded. Thus $\Lambda_\psi = \Pi_B^D\left(t\middle|\widetilde{w_B^D}\right)$ and (A4) turns to be

$$\frac{d\Pi_B^D\left(t\middle|\widetilde{w_B^D}\right)}{dt} = -[P - w_B^D(t) - d] + \Pi_B^D\left(t\middle|\widetilde{w_B^D}\right)[\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda(1 - F_B^D(V_B^D))] \qquad \text{(A6)}$$

And (A2), (A6) and $H = 0$ give:

$$\frac{dV_B^D\left(t\middle|\widetilde{w_B^D}\right)}{dt} = -u'\left(w_B^D(t)\right)\frac{d\Pi_B^D\left(t\middle|\widetilde{w_B^D}\right)}{dt} \qquad \text{(A7)}$$

Integrate (A5) with the integrating factor $\frac{1}{\psi_B}$ and substitute Λ_ψ with Π_B^D yields:

$$\frac{\Lambda_V}{\psi_B} = -\int_0^t \Pi_B^D \sigma(1-k)\lambda F_B^{D\prime}(V_B^D)ds + C_2$$

To Substitute Λ_V in (A3) using the above expression and differentiate with respect to t, we get:

$$-\frac{u''(w_B^D)}{u'(w_B^D)^2}\frac{dw_B^D(t)}{dt} = \sigma(1-k)\lambda F_B^{D\prime}(V_B^D)\Pi_B^D \tag{A8}$$

In addition, the transversality condition implies $\lim_{t\to\infty} V_B^D\left(t\middle|\widetilde{w_B^D}\right) = \overline{V_B^D}$.

(2) Next, we present the equilibrium results in terms of baseline wage.

If the solution to the above optimization problem with $V_0 = \underline{V_B^D}$ is taken as the baseline, then for any starting offer $V_0 \in [\underline{V_B^D}, \overline{V_B^D})$, there exists t_0 such that $V_s^{BD}(t_0) = V_0$. So, the optimal wage contract of any firm and all the equilibrium solutions could be expressed in terms of the baseline. For example, $w_B^D(t|V_0) = w_s^{BD}(t_0+t)$, $V_B^D\left(t\middle|\widetilde{w_B^D}\right) = V_s^{BD}(t_0+t)$ and $\Pi_B^D\left(t\middle|\widetilde{w_B^D}\right) = \Pi_s^{BD}(t_0+t)$. Then, it's easy to derive $w_s^{BD} \uparrow \overline{w_B^D}$ and $\Pi_s^{BD} \uparrow \overline{\Pi_B^D}$. Further, from (A2) we can obtain $\overline{V_B^D} = \frac{u\left(\overline{w_B^D}\right)+(1-\sigma)\lambda EV_B^N}{\delta+(1-\sigma)\lambda}$; and from (A6), we get $\overline{\Pi_B^D} = \frac{P-\overline{w_B^D}-d}{\delta+(1-\sigma)\lambda}$.

Let u_B denote the unemployment rate, d_B denote the share of B workers employed in D-firms and n_B the share employed in N-firms. The flow conditions imply

$$\delta = u_B(\delta + (1-\sigma)\lambda + \sigma(1-k)\lambda);$$
$$u_B\sigma(1-k)\lambda = d_B(\delta + (1-\sigma)\lambda);$$
$$(1-n_B)(1-\sigma)\lambda = n_B\delta$$

So, the unemployment rate is:

$$u_B = \frac{\delta}{\delta + (1-\sigma k)\lambda}$$

And the employment rate of type B workers in D-firms and N-firms are:

$$d_B = \frac{\delta\lambda\sigma(1-k)}{[\delta+(1-\sigma k)\lambda][\delta+(1-\sigma)\lambda]}; \qquad n_B = \frac{(1-\sigma)\lambda}{\delta+(1-\sigma)\lambda}$$

Let $G_B^D(V_0), V_0 \in [\underline{V_B^D}, \overline{V_B^D})$ be the proportion of B workers who have an expected lifetime value no greater than V_0 in all the B workers excluding those working in N-firms. Then $G_s^{BD}(t)$ is the corresponding baseline expression which satisfies:

$$G_s^{BD}(0) = \frac{u_B}{u_B+d_B} = \frac{\delta+(1-\sigma)\lambda}{\delta+(1-\sigma k)\lambda} \tag{A9}$$

and the flow condition for B workers employed in D firms with salary point greater than t:

$$[\delta + (1-\sigma)\lambda]\big(1 - G_s^{BD}(t)\big) = \frac{dG_s^{BD}(t)}{dt} + G_s^{BD}(t)\sigma(1-k)\lambda\big(1 - F_s^{BD}(t)\big) \tag{A10}$$

As every D-firm makes the same profit from B-workers at the equilibrium, and $G_s^{BD} \to 1$, $\Pi_{\mathrm{s}}^{\mathrm{BD}} \to \overline{\Pi_B^D}$, from the profit function:

$$\Omega_B^D = \lambda(1-k)G_s^{BD}(t)\theta n(1-n_B)\Pi_{\mathrm{s}}^{\mathrm{BD}}(\mathrm{t})$$

we can get:

$$G_s^{BD}(t)\Pi_{\mathrm{s}}^{\mathrm{BD}}(\mathrm{t}) = \frac{P-\overline{w_B^D}-d}{\delta+(1-\sigma)\lambda}.$$

So,

$$\frac{dG_s^{BD}}{dt}\Pi_{\mathrm{s}}^{\mathrm{BD}} + \frac{d\Pi_{\mathrm{s}}^{\mathrm{BD}}}{dt}G_s^{BD} = 0$$

Then substituting out $\frac{dG_s^{BD}}{dt}$ and $\frac{d\Pi_{\mathrm{s}}^{\mathrm{BD}}}{dt}$ using (A6) and (A10) and combining it with (A10) yields:

$$G_w^{BD} = \sqrt{\frac{P-\overline{w_B^D}-d}{P-w_B^D-d}}$$

$$\Pi_w^{BD} = \frac{1}{\delta+(1-\sigma)\lambda}\sqrt{\left(P-\overline{w_B^D}-d\right)(P-w_B^D-d)}$$

Putting the expression of G_w^{BD} into (A9) thus gets equation (10) in the proposition, i.e.,

$$\frac{P-\overline{w_B^D}-d}{P-\underline{w_B^D}-d} = \left(\frac{\delta+(1-\sigma)\lambda}{\delta+(1-\sigma k)\lambda}\right)^2.$$

The offer distribution could be derived from (A6), (A7), (A8) and the expression of Π_w^{BD}:

$$1 - F_w^{BD} = \frac{\delta+(1-\sigma)\lambda}{\sigma(1-k)\lambda}\left[\sqrt{\frac{P-w_B^D-d}{P-\overline{w_B^D}-d}} - 1 - \frac{1}{2u'(w_B^D)}\int_{w_B^D}^{\overline{w_B^D}}\frac{u'(x)dx}{\sqrt{\left(P-\overline{w_B^D}-d\right)(P-x-d)}}\right] \tag{A11}$$

Further, $V_s^{BD}(0) = V_{BU}$ at the equilibrium.

Since,

$$\frac{dV_s^{BD}(0)}{dt} = u(b) - u(\underline{w_B^D})$$

which is derived from the baseline expression of (A2) at $V_B^D = V_s^{BD}(0)$ and the Bellman equation for unemployed B workers; and,

$$\frac{dV_s^{BD}(0)}{dt} = \frac{\sqrt{P-\underline{w_B^D}-d}}{2}\int_{\underline{w_B^D}}^{\overline{w_B^D}}\frac{u'(x)dx}{\sqrt{P-x-d}}$$

which could be derived from substitutions using (A6), (A7), (A11) and the expression of Π_w^{BD}; we can derive another relationship between the bounds of the support of discriminating wages, i.e., equation (11):

$$u\left(\underline{w_B^D}\right) = u(b) - \frac{\sqrt{P-\underline{w_B^D}-d}}{2}$$

Besides, the dynamics of baseline tenure-wages (equation (14)) could be easily derived from (A8), (A11) and Π_w^{BD} expression.

(3) By the same token, we can get the equilibrium outcomes for B workers in the non-discriminating firms. Following the same procedures, we can prove that (15) holds. However, the support of the non-discriminating wages is somewhat different in the derivation.

Let $G_B^N(V_0), V_0 \in \left[\underline{V_B^N},\ \overline{V_B^N}\right)$ be the proportion of B workers (including the unemployed) who have an expected lifetime value no greater than V_0. Then, for the baseline expression, we have :

$$G_w^{BN} = \sqrt{\frac{P-\overline{w_B^N}}{P-w_B^N}}\ .$$

So, the overall proportion of type B workers (including the unemployed) who earn less than or equal to w at the steady state is:

$$G_w^B(w) = \begin{cases} \frac{\delta}{\delta+(1-\sigma)\lambda} G_w^{BD}, & if\ w \in [\underline{w_B^D}, \overline{w_B^D}] \\ G_w^{BN}, & if\ w \in [\underline{w_B^N}, \overline{w_B^N}] \end{cases}$$

Since $G_w^B\left(\overline{w_B^D}\right) = G_w^B\left(\underline{w_B^N}\right)$ and $G_w^B(w)$is monotonically increasing, $\overline{w_B^D}$=$\underline{w_B^N}$. Further, as:

$$G_s^{BN}(0) = \frac{\delta}{\delta+(1-\sigma)\lambda}$$

we can get:

$$\frac{P-\overline{w_B^N}}{P-\underline{w_B^N}} = \left(\frac{\delta}{\delta+(1-\sigma)\lambda}\right)^2$$

Thus, (12) (13) are proved.

(4) Finally, we derive the earnings distribution of type B workers.

Given G_w^{BD} and G_w^{BN}, the earning distributions of B workers in the D-and N- firms at the steady state are:

$$K_w^{BD} = \frac{u_B+d_B}{d_B}\left[G_w^{BD}(w) - \frac{u_B}{u_B+d_B}\right]$$

And:

$$K_w^{BN} = \frac{1}{n_B}\left[G_w^{BN}(w) - (u_B + d_B)\right]$$

So, the overall earning distribution is:

$$K_w^B(w) = \begin{cases} \frac{d_B}{d_B+n_B} K_w^{BD}, & if\ w \in [\underline{w_B^D}, \overline{w_B^D}] \\ \frac{d_B}{d_B+n_B} + \frac{n_B}{d_B+n_B} K_w^{BN} & if\ w \in [\underline{w_B^N}, \overline{w_B^N}] \end{cases}$$

Substituting the expressions of G_w^{BD}, G_w^{BN}, u_B, d_B and n_B inside, gives equation (16).

Proof of Equations (20), (21)

Proof

As shown above:

$$\Omega_B^D = \lambda(1-k)G_s^{BD}(t)\theta n(1-n_B)\Pi_{\mathrm{s}}^{\mathrm{BD}}(\mathrm{t}) = \frac{\lambda(1-k)\theta n\delta}{[\delta + (1-\sigma)\lambda]^2}\left(P - \overline{w_B^D} - d\right)$$

Similarly,

$$\Omega_B^N = \lambda G_s^{BN}(t)\theta n \Pi_{\mathrm{s}}^{\mathrm{BN}}(\mathrm{t}) = \lambda\theta n \frac{P - \overline{w_B^N}}{\delta}$$

Profits from A are:

$$\Omega_{\mathrm{A}} = \lambda G_s^A(t)(1-\theta)n\Pi_{\mathrm{s}}^{\mathrm{A}}(\mathrm{t}) = \lambda(1-\theta)n\frac{P - \overline{w_A}}{\delta}$$

So, $\Omega^{\mathrm{D}} = \Omega_{\mathrm{A}} + \Omega_B^D$ and $\Omega^{\mathrm{N}} = \Omega_{\mathrm{A}} + \Omega_B^N$.

Proof of properties (1)-(6) in section 4

Proof

First, let's consider properties (1)-(3).

From equation (11), taking partial derivatives with respect to (σ, k, d) yields:

$$A\frac{\partial \underline{w_B^D}}{\partial \sigma} + B\frac{\partial \overline{w_B^D}}{\partial \sigma} = 0\ ;$$

$$A\frac{\partial \underline{w_B^D}}{\partial k} + B\frac{\partial \overline{w_B^D}}{\partial k} = 0\ ;$$

$$A\frac{\partial \underline{w_B^D}}{\partial d} + B\frac{\partial \overline{w_B^D}}{\partial d} = \frac{1}{4}C.$$

Where:

$$A = \frac{u'\left(\underline{w_B^D}\right)}{2} - \frac{1}{4\sqrt{P - \underline{w_B^D} - d}}\int_{\underline{w_B^D}}^{\overline{w_B^D}} \frac{u'(x)dx}{\sqrt{P-x-d}}$$

$$B = \frac{u'\left(\overline{w_B^D}\right)}{2}\sqrt{\frac{P-\underline{w_B^D}-d}{P-\overline{w_B^D}-d}} > 0$$

$$C = \frac{1}{\sqrt{P-\underline{w_B^D}-d}}\int_{\underline{w_B^D}}^{\overline{w_B^D}} \frac{u'(x)dx}{\sqrt{P-x-d}} - \sqrt{P-\underline{w_B^D}-d}\int_{\underline{w_B^D}}^{\overline{w_B^D}} \frac{u'(x)dx}{(P-x-d)^{3/2}} < 0$$

Similarly, partial differentiation of equation (10)gives:

$$\frac{\partial \underline{w_B^D}}{\partial \sigma} = {\eta_1}^2 \frac{\partial \overline{w_B^D}}{\partial \sigma} - \eta_2 (P - \underline{w_B^D} - d) \; ;$$

$$\frac{\partial \underline{w_B^D}}{\partial k} = {\eta_1}^2 \frac{\partial \overline{w_B^D}}{\partial k} + \eta_3 (P - \overline{w_B^D} - d);$$

$$\frac{\partial \underline{w_B^D}}{\partial d} = {\eta_1}^2 \frac{\partial \overline{w_B^D}}{\partial d} + {\eta_1}^2 - 1.$$

Where:

$$\eta_1 = \frac{\delta + (1-\sigma k)\lambda}{\delta + (1-\sigma)\lambda}; \; \eta_2 = \frac{2\lambda(\delta+\lambda)(1-k)}{[\delta + (1-\sigma)\lambda][\delta + (1-\sigma k)\lambda]} \text{ and } \eta_3 = \frac{2\sigma\lambda}{\delta + (1-\sigma k)\lambda}$$

Substituting them into the first group of equations, results in:

$$(A{\eta_1}^2 + B)\frac{\partial \underline{w_B^D}}{\partial d} = \frac{1}{4}C{\eta_1}^2 + ({\eta_1}^2 - 1)B = \frac{\eta_1(\eta_1 - 1)}{2}\left[u'\left(\overline{w_B^D}\right)(\eta_1 + 2) - u'\left(\underline{w_B^D}\right)\eta_1\right]$$

$$(A{\eta_1}^2 + B)\frac{\partial \overline{w_B^D}}{\partial d} = \frac{1}{4}C - ({\eta_1}^2 - 1)A < 0$$

Since $\underline{w_B^N} = \overline{w_B^D}$, the partial derivative with respect to (σ, k, d) in (13) yields

$$\frac{\partial \overline{w_B^N}}{\partial k} = \left(\frac{\delta}{\delta + (1-\sigma)\lambda}\right)^2 \frac{\partial \overline{w_B^D}}{\partial k} \; ;$$

$$\frac{\partial \overline{w_B^N}}{\partial d} = \left(\frac{\delta}{\delta + (1-\sigma)\lambda}\right)^2 \frac{\partial \overline{w_B^D}}{\partial d} \; ;$$

$$\frac{\partial \overline{w_B^N}}{\partial \sigma} = \left(\frac{\delta}{\delta + (1-\sigma)\lambda}\right)^2 \frac{\partial \overline{w_B^D}}{\partial \sigma} - \frac{2\delta^2\lambda}{[\delta + (1-\sigma)\lambda]^3}\left(P - \overline{w_B^D}\right).$$

$$= \frac{2\delta^2\lambda}{[\delta + (1-\sigma)\lambda]^3}\left[\frac{P - \overline{w_B^D} - d}{\left(\frac{\delta + (1-\sigma)\lambda}{\delta + (1-\sigma k)\lambda}\right)^2 \frac{B}{A} + \delta + (1-\sigma k)\lambda} - \left(P - \overline{w_B^D}\right)\right]$$

So, $\frac{\partial \overline{w_B^N}}{\partial k}$ and $\frac{\partial \overline{w_B^N}}{\partial d}$ have the same sign as $\frac{\partial \overline{w_B^D}}{\partial k}$ and $\frac{\partial \overline{w_B^D}}{\partial d}$. But the sign of $\frac{\partial \overline{w_B^N}}{\partial \sigma}$ is uncertain.

Next, prove property (4).

From (11) we get : $u\left(\underline{w_B^D}\right) \le u(b)$. Thus, $\underline{w_B^D} \le b$ because of the increasing property of $u(w)$.

In proving the other side by contradiction, let's assume: $\overline{w_B^D} \le b$, then the integrated variable satisfies $\underline{w_B^D} \le x \le \overline{w_B^D} \le b$. So we have:

$$\int_{\underline{w_B^D}}^{\overline{w_B^D}} \frac{u'(x)dx}{\sqrt{P-x-d}} \le \int_{\underline{w_B^D}}^{b} \frac{u'(x)dx}{\sqrt{P-b-d}} = \frac{\sqrt{P - \underline{w_B^D} - d}}{2\sqrt{P-b-d}}\left[u(b) - u(\underline{w_B^D})\right]$$

If $b \le \frac{3}{4}(P-d)$, then $\underline{w_B^D} > 4b - 3(p-d)$. Thus, $\frac{\sqrt{P - \underline{w_B^D} - d}}{2\sqrt{P-b-d}} < 1$ and $\int_{\underline{w_B^D}}^{\overline{w_B^D}} \frac{u'(x)dx}{\sqrt{P-x-d}} < u(b) - u(\underline{w_B^D})$ which violates equation (11). Therefore, the assumption is false and we have proved $\overline{w_B^D} > b$ if $b \le \frac{3}{4}(P-d)$.

As for property (6), if $k = 1$, equation (24) is reduced to

$$Ew_B = \overline{w_B^N} + \frac{\delta}{(1-\sigma)\lambda}(\overline{w_B^N} - \underline{w_B^N}) - \frac{2\delta}{\delta+(1-\sigma k)\lambda}(P - \underline{w_B^N})$$

where $\underline{w_B^N}$ and $\overline{w_B^N}$ satisfy:

$$u\left(\underline{w_B^N}\right) = u(b) - \frac{\sqrt{P-\underline{w_B^N}}}{2}\int_{\underline{w_B^N}}^{\overline{w_B^N}} \frac{u'(x)dx}{\sqrt{P-x}}$$

and:

$$\frac{P-\overline{w_B^N}}{P-\underline{w_B^N}} = (\frac{\delta}{\delta+(1-\sigma)\lambda})^2.$$

The only difference in the system of equations compared with those for type A workers is the offer arrival rate, i.e., $\sigma = 0$ for type A while $\sigma > 0$ for type B.

Let $k \triangleq \frac{\delta}{\delta+(1-\sigma)\lambda}$, after some algebra the mean wage could be rewritten as:

$$Ew = P - (k^3 + k^2 + k)(P - \underline{w})$$

From the system of equations about $\underline{w}$ *and* $\overline{w}$, we can get:

$$\frac{\partial \underline{w}}{\partial k} = \frac{u'(\overline{w})(P - \underline{w})}{A + \frac{u'(\overline{w})k}{2}}$$

where,

$$A = \frac{u'(\underline{w})}{2} - \frac{1}{4\sqrt{P-\underline{w}}}\int_{\underline{w}}^{\overline{w}} \frac{u'(x)dx}{\sqrt{P-x}} > 0.$$

So,

$$\frac{\partial Ew}{\partial k} = -(3k^2 + 2k + 1)(P - \underline{w}) + (k^3 + k^2 + k)\frac{\partial \underline{w}}{\partial k}$$

$$= \frac{(P-\underline{w})}{A+\frac{u'(\overline{w})k}{2}}\left[\frac{u'(\overline{w})(k-k^3)}{2} - A(3k^2 + 2k + 1)\right] < 0$$

where the last inequality holds due to:

$$\frac{u'(\overline{w})(k-k^3)}{2} - A(3k^2 + 2k + 1) < \frac{u'(\overline{w})(k-k^3)}{2} - \frac{u'(\underline{w})(3k^3+2k^2+k)}{2} < 0$$

In addition, as k is increasing in σ, we get $\frac{\partial Ew}{\partial \sigma} < 0$. So the proposition is proved.

Proof of proposition 5

Proof

(1) and (2) can be directly derived from proposition 3.1 and proposition 3.3. $\overline{w_B^N} < \overline{w_A}$ because

$$\overline{w_A} - \overline{w_B^N} = (P - b)\left[\left(\frac{\delta}{\delta+(1-\sigma k)\lambda}\right)^2 - \left(\frac{\delta}{\lambda+\delta}\right)^2\right] + d\left[1 - \left(\frac{\delta+(1-\sigma)\lambda}{\delta+(1-\sigma k)\lambda}\right)^2\right]\left(\frac{\delta}{\delta+(1-\sigma)\lambda}\right)^2 > 0$$

Next, consider the comparison of earning distributions.

Since

$$K_w^A = \frac{\delta}{\lambda}[\sqrt{\frac{P-b}{P-w}} - 1],$$

$$K_w^B = \begin{cases} \frac{\delta}{(1-\sigma k)\lambda}[\sqrt{\frac{P-b-d}{P-w-d}} - 1], & if\ w \in [\underline{w_B^D}, \overline{w_B^D}] \\ \frac{\delta}{(1-\sigma k)\lambda}[\frac{\delta + (1-\sigma k)\lambda}{\delta + (1-\sigma)\lambda}\sqrt{\frac{P-\underline{\underline{w_B^N}}}{P-w}} - 1], & if\ w \in [\underline{w_B^N}, \overline{w_B^N}] \end{cases}$$

and

$$\frac{\delta}{\lambda} < \frac{\delta}{(1-\sigma k)\lambda}, \quad \sqrt{\frac{P-b}{P-w}} < \sqrt{\frac{P-b-d}{P-w-d}}, \quad \sqrt{\frac{P-b}{P-w}} < \frac{\delta + (1-\sigma k)\lambda}{\delta + (1-\sigma)\lambda}\sqrt{\frac{P-\underline{\underline{w_B^N}}}{P-w}}$$

we can get $K_w^A < K_w^B$ for all w.

Putting the wage expressions into (23) and (24), we get:

$$Ew_A = \frac{\lambda}{\lambda+\delta}P + \frac{\delta}{\lambda+\delta}b,$$

$$Ew_B = \frac{(1-\sigma k)\lambda}{\delta + (1-\sigma k)\lambda}P + \frac{\delta}{\delta + (1-\sigma k)\lambda}b - \rho d$$

Where,

$$\rho = \frac{\delta\{[\delta + (1-\sigma)\lambda](1-\sigma k)\lambda + (1-k)(1-\sigma)\sigma\lambda^2\}}{(1-\sigma k)\lambda[\delta + (1-\sigma)\lambda]^2} - \frac{\delta}{\delta + (1-\sigma k)\lambda} > 0.$$

Therefore, $Ew_A - Ew_B = \frac{\lambda\delta k\sigma}{(\delta+\lambda)[\delta + (1-\sigma k)\lambda]}(P-b) + \rho d > 0.$

Through tedious calibration, we can get the comparative statics of $Ew_A - Ew_B$: $\frac{\partial[Ew_A - Ew_B]}{\partial(\sigma,k,d)} > 0$.

Define $\Delta_q = w_A^q - w_B^q$. From (25) and (26), we have

$$\Delta_q = \begin{cases} d + (P-b-d)\left(\frac{\delta}{\delta+(1-\sigma k)\lambda q}\right)^2 - (P-b)\left(\frac{\delta}{\delta+\lambda q}\right)^2 & if\ \ q \in [0, q_0) \\ d\left(\frac{\delta[\delta+(1-\sigma k)\lambda]}{[\delta+(1-\sigma)\lambda][\delta+(1-\sigma k)\lambda q]}\right)^2 + (P-b-d)\left(\frac{\delta}{\delta+(1-\sigma k)\lambda q}\right)^2 - (P-b)\left(\frac{\delta}{\delta+\lambda q}\right)^2 & if\ \ q \in [q_0, 1] \end{cases}$$

Taking partial derivative with respect to q yields

$$\frac{\partial\Delta_q}{\partial q} = \begin{cases} -2(P-b-d)\frac{\delta^2(1-\sigma k)\lambda}{(\delta+(1-\sigma k)\lambda q)^3} - 2(P-b)\frac{\delta^2\lambda}{(\delta+\lambda q)^3} \\ -2d\frac{\delta^2\delta[\delta+(1-\sigma k)\lambda]^2[\delta+(1-\sigma)\lambda](1-\sigma k)\lambda}{([\delta+(1-\sigma)\lambda][\delta+(1-\sigma k)\lambda q])^3} - 2(P-b-d)\frac{\delta^2(1-\sigma k)\lambda}{(\delta+(1-\sigma k)\lambda q)^3} - 2(P-b)\frac{\delta^2\lambda}{(\delta+\lambda q)^3} \end{cases} < 0$$

So, property (5) is proved.