

**DO GOOD WORKERS HURT BAD WORKERS—OR
IS IT THE OTHER WAY AROUND?***

BY ESPEN R. MOEN¹

Norwegian School of Management, Norway and CEPR, UK

I study the effect of worker heterogeneities on wages and unemployment in a directed search model. A worker's productivity in a given firm depends both on his type and on a worker-firm specific component. Firms advertise unconditional wage offers only. The resulting equilibrium is inefficient, with a too high wage premium for high-type workers, and too few high-type jobs. This reduces the welfare of high-type workers. My findings contrast with the findings in the literature on labor market segmentation, where it is argued that the existence of high-type workers forces down wages and reduces welfare for low-type workers.

1. INTRODUCTION

It is well known that the standard workhorse in economics, the competitive (Arrow-Debreu) model is not fully adequate for analyzing the labor market, as unemployment seems to be an inherent feature of this market whereas the competitive model assumes market clearing. Alternative models have therefore been developed to deal with unemployment, among which models of directed search probably are the closest to the competitive model.

The main idea behind directed search models is that firms may manipulate the number of applicants it obtains through the wage it advertises. The seminal directed search model is constructed on the basis of the so-called urn-ball matching process (Hall, 1979; Montgomery, 1991; Peters, 1997). There are large numbers of workers and firms in the market. In every period, firms advertise their vacancies and the wages attached to them, and unemployed workers send applications to these firms in an uncoordinated fashion. The number of applications for each vacancy is then random; hence some firms obtain many applicants per vacancy, whereas others obtain none. However, the probability that a vacancy will be filled depends on the advertised wage. Firms therefore face a trade-off between wage costs and search costs. Similar models are developed in a context in which the contact rate between workers and firms depends on aggregate variables; see Moen (1997), Acemoglu and Shimer (1999), and Mortensen and Pissarides (1999).

* Manuscript received January 2001; revised October 2001.

¹ I would like to thank Erik Grønn, Nobuhiro Kiyotaki, Chris Pissarides, Åsa Rosen, seminar participants at Centre for Economic Performance, London School of Economics, and finally two anonymous referees for valuable comments. Financial support from the Norwegian Research Council is gratefully acknowledged. Please address correspondence to: Espen R. Moen, Economics Department, Norwegian School of Management BI, Box 580, N-1301 Sandvika, Norway. Tel: +47 6755 7395. Fax: +47 6755 7675. E-mail: espen.moen@bi.no.

Since the directed search models are so close to the competitive model, their welfare properties are of great interest. The standard directed search model yields an efficient equilibrium allocation if workers are homogeneous. However, in a series of papers (see particularly, Lang and Dickens, 1992, 1993), Lang and Dickens find that even small worker heterogeneities may change the equilibrium dramatically and lead to an inefficient allocation of resources. They show this in a model involving two groups of workers. The workers in one of the groups are marginally less productive than the workers in the other group. A slightly better worker obtains a discrete advantage over slightly less productive workers applying for the same job, as he will always be preferred for that job. Dickens and Lang show that the resulting equilibrium is separating, and that in this separating equilibrium wages for low-productivity workers are forced down to the reservation wage of the high-productivity workers. The existence of the better type of workers thus reduces the welfare of the (marginally) less productive workers. Dickens and Lang refer to this as labor market segmentation, and use the model to analyze the discrimination (Lang et al., 1999) and the effects of minimum wages (Lang and Kahn, 1998).

The labor market segmentation rests on two assumptions. Firstly, that firms cannot advertise wages conditioned on worker type. Secondly, that each worker has exactly the same productivity in all matches so that if one firm prefers worker A to worker B, all other firms will surely do so as well. In this article, I keep the first assumption, as there may be limits to the extent to which firms may observe and condition wages on worker type. However, I alter the second assumption by introducing worker-firm specific productivities, implying that a worker's productivity may differ between firms. Firms offer unconditional wage offers, and hire the most productive applicants independently of their type.

As in the literature on labor market segmentation, I find that the resulting equilibrium is separating, with workers of different types searching in different markets. The reason is related to the fact that the expected productivity, conditioned on being above a certain level, is typically higher for high-type than for low-type workers. Firms therefore prefer to attract high-type rather than low-type applicants. Furthermore, as in the literature, the separation is inefficient in the sense that it reduces aggregate production in the economy. However, and contrary to the findings of Dickens and Lang, equilibrium wages for low-type workers are efficiently determined, whereas the high-type workers are paid a larger-than-optimal wage premium. This results in too few high-wage jobs, since firms are reluctant to open sufficiently many high-wage jobs. It is thus the existence of low-type workers that reduces the welfare of high-type workers, not the other way around, as in Lang and Dickens' model. Furthermore, a binding minimum wage will reduce the welfare of both low- and high-type workers. On the other hand, a tax on income exceeding the wages for low-type workers may increase output.

Dickens and Lang relate their findings to a substantial literature, in large part in sociology and industrial relations, on the labor market segmentation. In this literature, a shared belief is that the wage distribution is unfairly tilted against the poor (Taubman and Wachter, 1986). The model of Dickens and Lang gives an economic rationale for this. My analysis suggests that this rationale is less general than

that previously considered, and may not hold if we allow for a nondegenerated distribution of a given worker's productivity over different firms. On the other hand, my analysis confirms a second shared belief in the labor market segmentation literature, namely that too few high wage (primary sector jobs) are created in the market.

Introducing match-specific productivities in a model of directed search complicates the analysis considerably. To simplify the algebra, I therefore assume that each firm hires many workers simultaneously. The law of large numbers then eliminates the co-ordination failures associated with random matching. Although this assumption implies that my model differs somewhat from the models of Lang and Dickens (in addition to the fact that we allow for a match-specific productivity component), I believe the gains achieved by using a simpler model make this assumption worthwhile. Conceptually the assumption also makes sense, as the worker–firm specific productivity differentials in my model play much of the same role as the co-ordination problems caused by the random matching in the urn-ball process.

The outline of the model is as follows. Following this introduction, the model is laid out in Section 2 and the optimal solution derived in Section 3. Equilibrium is derived in Sections 4 and 5. In Section 6, I discuss implications and extensions of the model and discuss some important assumptions underlying the results. Section 7 concludes.

2. THE MODEL

There are two types of agents in the economy: workers and firms. All agents are risk neutral. There is a continuum of workers with measure L , where L is a large number. Firms are large in the sense that they hire many workers. Formally, each firm hires a continuum of workers with exogenous measure normalized to 1. This may reflect the optimal scale of operations.² As in Acemoglu and Shimer (1999), the economy is modeled as a one-shot game with the following time structure:

1. Firms enter the market.
2. Then they advertise their vacancies and the wages attached to them.
3. Each worker submits an application to one of the firms.
4. Firms decide which of the applicants to hire. Then production takes place, and workers are paid the wages advertised.

The assumption that each worker submits applications to a limited number of firms, normalized to one, is necessary in order to obtain unemployment in equilibrium. It reflects that workers have limited capacity with respect to producing applications, attending interviews, and so on.

The cost of entering the market is exogenous and given by K . The productivity of a given worker depends on two factors—a worker specific component a^k , and

² Equivalently, we may assume that the number of firms is exogeneously given whereas the number (measure) of workers in each firm is large and endogeneously given; see Moen (2001). This will give rise to exactly the same equilibrium.

a worker–firm specific component ϵ . The worker specific component a^k reflects ex ante heterogeneities, as some workers are generally more productive than others. We distinguish between high-type workers h and low-type workers l , and let $a^h > a^l$. The worker–firm specific component reflects the fact that a worker may fit in well in some firms and not in others. Ex ante (before the workers have been in contact with a firm) the worker–firm specific component ϵ is regarded as stochastic, drawn independently from the same distribution for both type of workers. Thus, workers who are identical ex ante may turn out to have different productivities in a given firm. The productivity of a worker of type k in firm i can thus be written as

$$z_i^k = a^k + \epsilon_i$$

for $k = h, l$. Let ϵ^{sup} and ϵ^{inf} denote the supremum and the infimum of the support of ϵ , respectively. I assume that $a^l + \epsilon^{\text{sup}} > K$ and that $a^h + \epsilon^{\text{inf}} < K$. As will be clear later, this ensures that some high-type workers are unemployed and some low-type workers employed in equilibrium. I denote the cumulative distribution of ϵ by G and its density by g , and make the standard assumption that G has an increasing hazard rate ($g/(1 - G)$ is increasing). Additional restriction on G will be given later. Total output of firm i is given by

$$\bar{z}_i = \int_{R_i} z_i d\bar{G}$$

where R_i is the productivity of the least productive employee in this firm and \bar{G} is the distribution of worker productivity among the employees. We refer to R_i as the productivity threshold and \bar{z}_i as the average productivity in firm i .

In what follows, I assume that a worker knows his type, but not his worker–firm specific productivity component in any firm, not even at the point in time when he is (eventually) offered a job. Firms, on the other hand, know their applicants' total productivity but not their types (or for other reasons cannot make wages or employment contingent on worker type). It follows that firms cannot advertise wages contingent on worker type, since they cannot observe it. Nor can they advertise wage contracts where wages depend on total productivity: Since workers cannot observe the worker–firm specific productivity part, a firm will always have incentives to claim that this component is low in order to save on wage costs. Thus, we assume that firms advertise unconditioned wage offers only, as in the papers on labor market segmentation. In the extension section, we characterize the equilibrium of the model with more general wage contracts.

3. OPTIMAL ALLOCATIONS

In this section, I derive the socially optimal allocation of resources in the model. To this end, I apply the standard method in economics, and assume that a planner makes all the decisions in the economy, facing the same resource and information constraints as the agents have when making their decisions.

Due to the law of large numbers, there are no coordination problems in the economy in the sense that some firms by chance obtain more applicants than other firms. Obviously, a planner will want to hire the most productive applicants, independent of type. For any given number of firms in the economy, applicants will be allocated to firms in such a manner that the productivity thresholds R_i in all firms will be equal. If not, the aggregate output would be increased by directing more applicants toward firms with lowest productivity thresholds.³ Denote this common productivity threshold by R . From the planner's perspective, we may therefore think of the economy as consisting of one firm, and that the problem is to determine the measure of jobs to open (we abstract away from integer problems arising from the fact that each firm has a fixed size). Let N denote the total measure of workers hired by the planner, and let N^h and N^l denote the number of high- and low-type workers hired, $N^l + N^h = N$. Aggregate output is then

$$(1) \quad W = N^l \bar{z}^l + N^h \bar{z}^h - NK$$

where \bar{z}^k denotes the average productivity among hired workers of type k . For simplicity, we set the shadow price of leisure equal to zero. The planner chooses the number of firms entering the market, and thereby the measure N of workers hired, so as to maximize the aggregate output. We refer to the solution to this maximization problem as the optimal allocation. In the Appendix, the optimal solution is derived formally. A heuristic argument is provided below.

Due to the law of large numbers, R is a deterministic function of N . The higher N is, the lower R will be, as the planner must hire workers with lower and lower worker-firm specific components ϵ . In optimum, N must be set such that R equals the entry cost K (a unit measure of workers with productivity R produces R).

A worker of type k has a productivity that exceeds K if and only if $\epsilon \geq K - a^k$. It follows that N^h and N^l are given by the equations

$$N^k = (1 - G(K - a^k))L^k$$

for $k = h, l$, where L^k is the measure of workers of type k .

The optimality condition is illustrated in Figure 1. Note that our assumption that $a^k + \epsilon^{\text{inf}} < K$ ensures that an optimal allocation of resources implies unemployment for both type of workers. Furthermore, our assumption that $a^l + \epsilon^{\text{sup}} > K$ ensures that some low-type workers are employed by the planner.

4. EQUILIBRIUM WITH ONE TYPE OF WORKERS

It is convenient to derive the equilibrium with only one type of workers first. I thus assume that $a^h = a^l = a$. The equilibrium of the model is derived by backward induction.

³ The measure of applicants to any firm is a deterministic function of the workers' search strategies. This will be clarified in the next section.

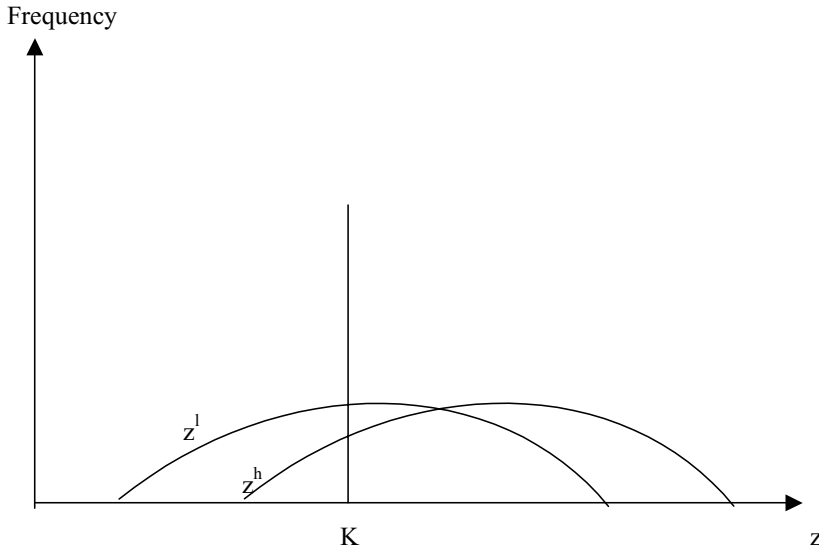


FIGURE 1

ILLUSTRATION OF THE OPTIMAL ALLOCATION. THE CURVES SHOW THE DISTRIBUTION OF z AMONG HIGH- AND LOW-TYPE APPLICANTS. THE PLANNER HIRES ALL WORKERS WITH PRODUCTIVITY ABOVE K

4.1. *Firms' Hiring Strategy (Stage 4)*. A firm collects all its applicants, inspects their productivity, and hires the most productive among them. Let L_i denote the measure of applicants to firm i . As clarified below, we will always have $L_i > 1$ in equilibrium. Due to the law of large numbers, the empirical distribution of the term ϵ among the applicants is equal to its theoretical distribution G . The threshold productivity R_i in this firm, that is, the productivity of the marginal worker, is thus given by

$$(2) \quad L_i(1 - G(R_i - a)) = 1$$

A potential problem may arise if $R_i < w_i$ (the wage), in which case the firm may be reluctant to fill all its vacancies. In the Appendix, I show that this never happens in equilibrium as long as the error term is not too important relative to the cost of creating jobs (more specifically, if $E[\epsilon | \epsilon \geq K - a] < 2K$ where K is the cost of opening a vacancy; see the Appendix). Even if this inequality is not satisfied, it is easy to show that it is always in the firm's ex ante interest to commit to filling all its vacancies. In what follows, we assume that R_i is given by (2).⁴

4.2. *Optimal Application Strategies (Stage 3)*. I refer to this stage as the application game. In this game, workers independently and simultaneously choose which firm to apply for given the number n of firms in the market and their

⁴ Moen (1999) characterizes the equilibrium when these assumptions are not satisfied.

advertised wages w_1, \dots, w_n . In order to simplify the presentation and avoid uninteresting technicalities, I focus on symmetric equilibria in which all workers choose the same strategy. A strategy of a worker is then a vector of probabilities $q = (q_1, \dots, q_n)$, where q_i is the probability that a worker will apply for firm i .

The expected income to a worker when applying for a job in firm i is

$$(3) \quad U_i = p_i w_i$$

where p_i is the probability of getting the job and where the income when unemployed is normalized to zero. When a worker randomizes over a set of actions, all actions must yield the same expected income. Thus, in equilibrium, *all firms that attract applicants must provide the same expected income to these applicants*, the law of one price in this context. We denote the equilibrium expected income to workers by U .

Due to the law of large numbers, it follows that $L_i = q_i L$ (the measure of workers in the market). It follows that R_i is a deterministic function of q , n , and (w_1, \dots, w_n) . Since a worker's productivity is given by $z = a + \epsilon$ we have that $p_i = 1 - G(R_i - a)$, and his expected income can be written as (from (3)):

$$(4) \quad U_i = (1 - G(R_i - a))w_i$$

The *indifference curve* of a worker shows combinations of w and R that give him the equilibrium expected income U . We write the indifference curve as $R = R^k(w; U)$, implicitly defined by (4). It follows that the slope of the indifference curve is given by

$$(5) \quad R_w(w; U) = \frac{1 - G(R - a)}{g(R - a)} \frac{1}{w} > 0$$

Since R is increasing in w and G has an increasing hazard rate, it follows from (5) that the derivative R_w is decreasing in w .

4.3. *Wage Determination (Stage 2)*. I assume that individual firms are small relative to the market and thus act as price-takers in the sense that they regard the equilibrium value of the expected incomes to workers as given, independently of their own wage policy. The (ex post) profit of a firm (entry costs excluded) is given by $\bar{z}_i - w_i$, where \bar{z}_i denotes the average productivity of that firm's employees. It follows that we can write \bar{z}_i as

$$(6) \quad \bar{z}_i = a + h(R_i - a)$$

where $h(R_i - a) \equiv E[\epsilon | \epsilon \geq R_i - a]$ is the expected value of the worker-firm specific term among the hired workers. It follows that h is an increasing function. Note also that the threshold productivity of a firm is independent of N_i . All firms thus

choose w so as to maximize profit per worker $\bar{z}_i - w_i$, i.e., to solve the problem (firm index skipped)

$$(7) \quad \max_w \bar{z} - w \quad \text{subject to} \quad R = R(w; U)$$

Denote the resulting maximum by π . It follows that we can write $\pi = \pi(U)$. Obviously, $\pi'(U) < 0$.

4.4. *Entry (Stage I)*. Firms will enter the market up to the point where the expected profit of entering is equal to the cost. We thus have that

$$(8) \quad \pi(U) = K$$

DEFINITION. An equilibrium of the model is a wage w and an expected income to workers U such that (i) w solves the profit maximization problem (7) and (ii) U is determined by the free entry condition (8).

The equilibrium is illustrated in Figure 2. It is straightforward to show that the equilibrium of the model exists provided that $a + \epsilon^{\text{sup}} > K$, and the proof is therefore omitted. Note that the structure of the equilibrium is almost recursive. Obviously, U is uniquely determined by (8). Given U , the advertised wage is determined by (7). The threshold productivity R is then determined by the

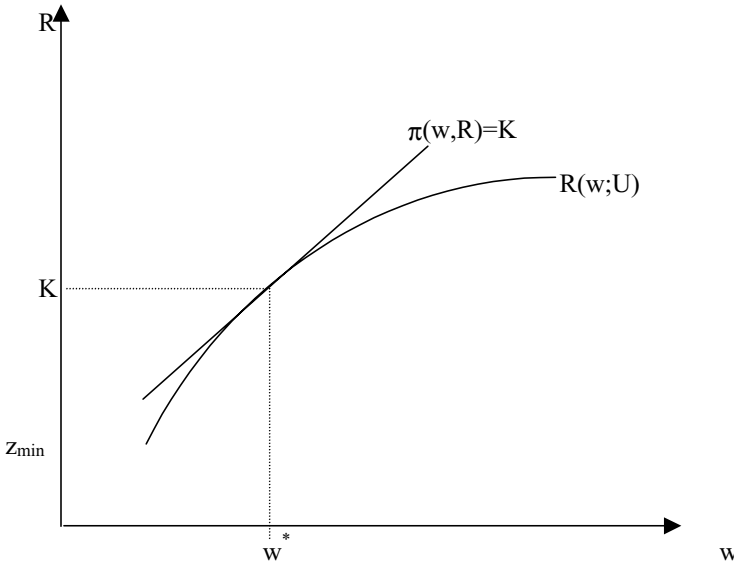


FIGURE 2

EQUILIBRIUM WITH ONE TYPE OF WORKERS, DETERMINED BY THE TANGENCY POINT BETWEEN THE ISO-PROFIT LINE $\pi = K$ AND THE INDIFFERENCE CURVE $R(w; U)$

indifference curve $R = R(w; U)$. For completeness, note that the aggregate capacity in the market is given by the equation $N = (1 - G(R - a))L$.

In the Appendix, we show that $\pi = R(w^*; U)$, where w^* denotes the profit maximizing wage. From (8), it thus follows that $R = K$ in all firms, and from the last section we thus know that the equilibrium is optimal

PROPOSITION 1. *The equilibrium with only one type of workers is optimal.*

By definition, $\pi = a + h(R - a) - w$. Combining this and the fact that $\pi = R = K$ thus gives $w = h(K - a) - (K - a)$. It follows that

$$(9) \quad \frac{dw}{da} = 1 - h'(K - a) > 0$$

Thus, an increase in the productivity term a increases wages if and only if $h'(K - a) < 1$. In words: If the productivity threshold in a firm increases by one unit, this increases the average productivity among the employees, but by less than one unit. This holds for most distributions and clearly for the uniform distribution. Actually, it is difficult to come up with a distribution with an increasing hazard rate that does not satisfy this condition. However, I have not been able to prove that such a distribution does not exist. In what follows, I assume that $0 < h' < 1$, in which case higher general productivity a leads to both higher employment and higher wages.

5. TWO TYPES OF WORKERS

Two types of workers are now allowed for. Let U^h and U^l denote the equilibrium values of U in the application game for high- and low-type workers, respectively. The following lemma then holds:

LEMMA 1. *The indifference curves of high- and low-type workers intersect at most once. If \widehat{w} denotes the intersection point, then $R^h(w; U^h) > R^l(w; U^l)$ for $w > \widehat{w}$ while $R^h(w; U^h) < R^l(w; U^l)$ for $w < \widehat{w}$.*

PROOF. By definition, $R^h(\widehat{w}; U^h) = R^l(\widehat{w}; U^l)$. Since, by assumption, the hazard rate $g(\epsilon)/(1 - G(\epsilon))$ is increasing in ϵ , it follows from (5) that $R_w^l(\widehat{w}) < R_w^h(\widehat{w})$. The indifference curve of high types thus crosses the indifference curve of low types from below. ■

The point is that high-type workers are more willing to trade off a low R in return for higher wages than are low-type workers. We continue to let \widehat{w} denote the intersection between the indifference curves for high- and low-type workers. The following lemma then follows almost directly:

LEMMA 2. *A firm that offers a wage above \widehat{w} attracts high-type workers only, while a firm that offers a wage below \widehat{w} attracts low-type workers only.*

To prove the lemma, consider a firm that advertises a wage $w' > \hat{w}$. The productivity threshold in this firm must be at least $R^h(w', U^h)$; otherwise, a high-type worker applying to this firm would obtain an expected income above the equilibrium income U^h , which is inconsistent with equilibrium in the application game. Since $R^l(w') < R^h(w')$, it follows that the expected income for low-type workers applying to a firm with wage w' must be strictly less than R^l . Similarly, firms advertising a wage below \hat{w} attract low-type workers only.

From (6) it follows that the expected productivity of a successful applicant of type k in a firm with productivity threshold R is $\bar{z}^k = a^k + h(R - a^k)$, $k = l, h$. Since $h' < 1$, it follows that *for a given productivity threshold R , firms will prefer to get high-type rather than low-type applicants*. Thus, a firm's profit as a function of the advertised wage is discontinuous at \hat{w} .

In any *pooling equilibrium*, there exists a wage w' such that if all firms but one advertise w' , it is also optimal for the last firm to advertise w' . It follows that a pooling equilibrium does not exist. If there was, a firm could increase its profit by increasing wages marginally and receive applications from high-type workers only, thus increasing its profit. The situation is described in Figure 3.

PROPOSITION 2. *The model has no pooling equilibrium.*

Any equilibrium of the model thus has to be separating, in the sense that some firms attract high-type workers only (the high-type submarket) while other firms

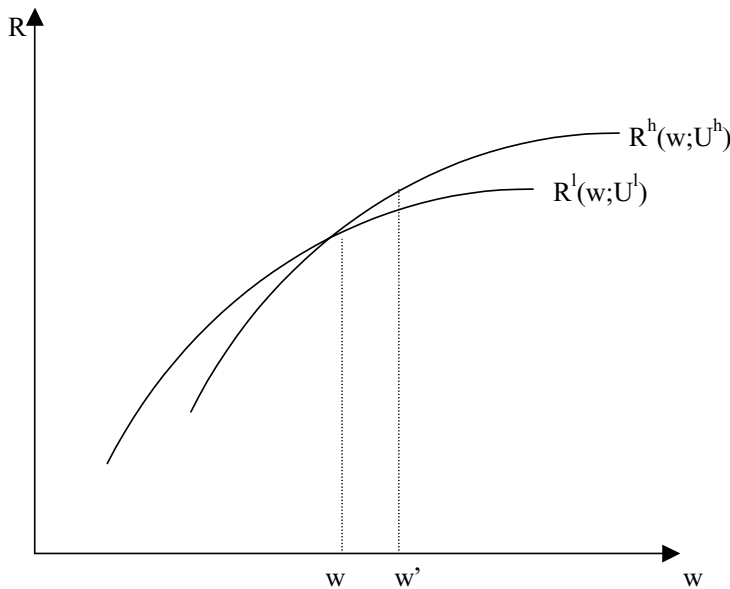


FIGURE 3

IF ALL FIRMS ADVERTISE THE WAGE w , A DEVIATING FIRM ADVERTISING $w' > w$ ATTRACTS HIGH-TYPE WORKERS ONLY

attract low-type workers only (the low-type submarket). Thus, two incentive compatibility constraints must be satisfied, ensuring that the low- and high-type workers have the incentive to apply for the submarket “assigned” to them. Intuitively, the incentive constraint for low-type workers will be binding. To see this, suppose that neither of the incentive constraints were binding. The equilibrium in the submarket for high- and low-type workers would then be as with one type of workers, with $a = a^l$ in the low-type submarket and $a = a^h$ in the high-type submarket. In both these submarkets, the productivity threshold would be equal to R . From equation (9), we know that equilibrium wages are increasing in a . Thus, *wages in the high-type submarket would be higher than in the low-type submarket, while the productivity threshold would be the same. Low-type workers would therefore be better off applying for high-wage firms.* In the Appendix, we show the following lemma in detail:

LEMMA 3. *In equilibrium, the incentive compatibility constraint for low-type workers binds, while it does not for high-type workers.*

Let w^l and w^h denote equilibrium wages for high- and low-type workers, respectively. It follows that the equilibrium in the low-type submarket is as posited in Section 4 (with one type of workers only). We thus know that w^l is determined as in Section 4 (with a replaced by a^l) and $R^l = K$. Let R^h denote the productivity threshold in high-wage firms. As the incentive compatibility constraint for the low-type workers binds, it follows that (w^h, R^h) is at the indifference curve of the low-type workers. Thus,

$$(10) \quad R^h(w^h; U^h) = R^l(w^h; U^l)$$

The zero-profit condition implies that $\pi^h = a^h + h(R^h - a) - w^h = K$, and together with (10), this equation determines R^h and w^h . Finally, the equilibrium value of U^h is determined by the indifference curve $R^h = R^h(w^h, U^h)$. The equilibrium can thus be defined as follows (a more formal argument is given in the Appendix):⁵

DEFINITION. A separating equilibrium of the model is a vector (U^l, U^h, w^l, w^h) , where U^l and w^l satisfy Equations (7) and (8), and w^h and U^h satisfy (10) and the zero-profit condition $\pi^h = K$.

The equilibrium is illustrated in Figure 4. Again it is easy to show that the equilibrium exists provided that $a^l + \epsilon^{\text{sup}} > K$ and that $a^h + \epsilon^{\text{inf}} < K$. In the Appendix, we show that the equilibrium is unique (in this proof we assume that $h'' < 0$). Furthermore, from this figure it is also evident that in equilibrium, the outcome to

⁵ Alternatively, we may let R^h be a choice variable to the firms, advertised together with the wage. The firm in the high-type submarket then maximizes profit given that (1) $R^h \leq R^h(w^h; U^h)$ (it attracts high-type applicants) and (2) $R^h \geq R^l(w^h, U^l)$ (incentive compatible constraint). As it will always be optimal to let both constraints bind, the resulting equilibrium is equivalent to the equilibrium presented in the text.

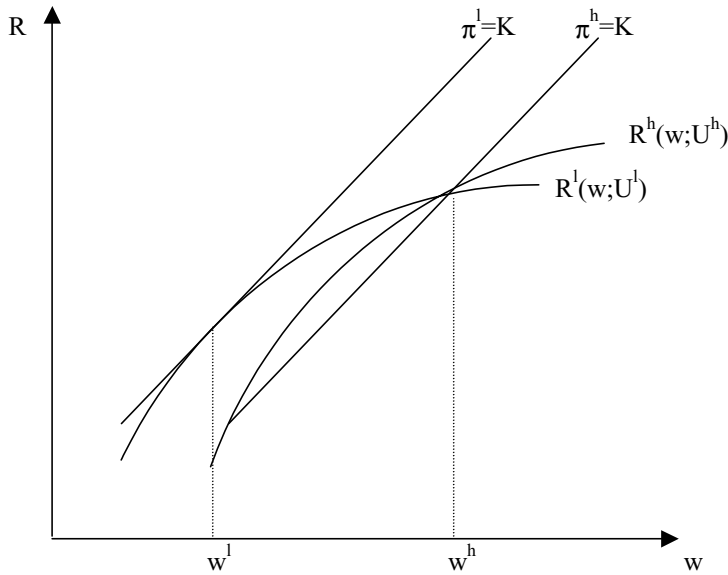


FIGURE 4

SEPARATING EQUILIBRIUM. THE WAGE FOR HIGH-TYPE WORKERS IS DETERMINED AT THE INTERSECTION BETWEEN THE INDIFFERENCE CURVE FOR LOW-TYPE WORKERS AND THE ISO-PROFIT CURVE FOR FIRMS ATTRACTING HIGH-TYPE WORKERS

high-type workers is maximized given the zero-profit condition for firms and the incentive compatibility constraint for low-type workers. However, as $R^h > K$ in this equilibrium, the equilibrium is not optimal in the sense that it maximizes the aggregate output.

PROPOSITION 3. *The separating equilibrium defined above is unique. The equilibrium allocation of resources is not optimal, as too few jobs for high-type workers are created compared with the socially optimal number.*

Note that it is the high-type workers who “suffer” from the inefficiencies. Although their wages are higher than the corresponding wages in a market with high-type workers only, this is more than compensated for by a higher unemployment rate. Note also that *although the unemployment level for high-type workers is higher than in the optimal solution, it may still be lower than for low-type workers*, since optimality requires that the unemployment rate among low-type workers is higher than among high-type workers.

Let us give some more intuition for our results. We start with the pooling result (consider Figure 3). Suppose all firms hired both types of workers at wage w in equilibrium. We would then have $R^h(w; U^h) = R^l(w; U^l)$ at this point, as firms hire the most productive workers independent of type. Since high-type workers on average are more productive than low-type workers (even conditioned on being

above the threshold productivity level), the firms make higher expected profits on high-type workers than on low-type workers.

Consider a firm that increases its wage slightly. This will increase competition for jobs in that firm, and the productivity threshold will increase slightly. Applications from high-type workers flow in, and the productivity threshold in this firm increases up to the point at which high-type applicants obtain their equilibrium expected income U^h . As low-type workers are more vulnerable to stiff job competition than high-type workers are, it is not worthwhile for them to apply at this point. The deviating firm therefore attracts high-type workers only, and thus makes a profit. This is not consistent with the equilibrium.

In any separating equilibrium, the free entry ensures that wages equal the average net productivity among the employees in a firm (job creation costs subtracted). Thus, a low-type worker who gets a job in a high-wage firm obtains a wage exceeding his expected net productivity in that firm, and this gives him an incentive to apply. We refer to this as the *productivity effect*. In order to obtain separation, the productivity effect must be balanced by a *job competition effect*; the competition for high-wage jobs must be stiffer and exactly balance the productivity effect for low-type workers. An efficient allocation of resources, by contrast, requires that there is no job competition effect, as the productivity threshold must be the same for both worker types. Finally, note that a high-wage firm that reduces its wage slightly will obtain low-type applicants only and lose money.

Lang and Dickens in their papers assume that there is a relatively small difference in productivity between the two types, that is, a^h is only slightly greater than a^l . This implies that the productivity effect is small. On the other hand, there is no worker–firm specific productivity term in their model. Thus, the job competition effect is extremely strong, as a high-type worker is always preferred to a low-type worker. As a result, it is the incentive compatibility constraint for the high-type worker that binds. Note that a similar exercise is uninteresting in our model: With no worker–firm specific component, there would be no unemployment as long as the worker productivity is above the cost of creating a job.⁶

Our assumption that firms hire many workers implies that the productivity threshold is constant. This may not be unrealistic for large firms which constantly hire workers, and which at a given point in time may adjust the number of hirings to the quality of their applicants. If we do not allow the law of large numbers to wash out uncertainty at firm level, the threshold productivity becomes stochastic, with a distribution that is hard to characterize. Our results will still hold if (1) high-type workers are more willing to accept harder job competition in return for higher wages than are low-type workers and (2) for a given R , the expected profit for a firm is higher when it attracts high-type applicants rather than low-type applicants. It is hard to derive precise conditions under which these two requirements will hold, though intuition suggests that they are likely to hold when the worker–firm

⁶ In order for the incentive compatibility constraint for low-type workers to be binding, we have to assume that $h' > 1$ (which means that in a market with only one type, equilibrium wages are *decreasing* in the worker productivity a). At the same time, we must retain the assumption that G has increasing hazard rate. As mentioned earlier, it is difficult to come up with such a distribution.

specific component is important relative to the type-specific component of worker productivity.

Our separating equilibrium is similar to the separating equilibrium in Rotchild and Stiglitz (1976). In contrast with the equilibrium in their model, the separating equilibrium in our model always exists, independently of the fraction of high- to low-type workers. The reason for this difference is that in my model, a firm that deviates and offers a lower wage attracts low-type workers only. In the model of Rotchild and Stiglitz, a deviating (insurance) firm obtains buyers of both types (a similar point is made in Inderst and Muller, 1999).

6. IMPLICATIONS AND EXTENSIONS

In this section, we first discuss policy implications of our model and its implications for the wage dispersion in the economy. Then we briefly study the consequences of having more than one production technology and discuss some important assumptions underlying the analysis.

6.1. Policy Implications. Since the equilibrium of the model does not maximize the aggregate output, there is scope for increasing output through economic policy. As the inefficiency problem is rooted in the fact that the incentive compatibility constraint for low-type workers bind in equilibrium, any policy which relaxes this constraint will reduce wages in the high-type submarket and may thereby potentially increase output.

A binding *minimum wage* increases wages in the low-type submarket. This reduces employment among low-type workers. Furthermore, as the equilibrium allocation in the low-type submarket is optimal, a minimum wage reduces welfare for low-type workers. This tightens their incentive compatibility constraint, as it becomes more attractive to apply for high-type jobs. Thus, wages for high-type workers increase and the distortions in the high-type submarket are amplified. A *maximum wage* may seem to make more sense, but as shown in Moen (2001), this may also reduce welfare unless it is sufficiently low so that it prevents separation completely. The argument is similar to the argument presented in Lang and Kahn (1998) when they discuss the effects of a *minimum wage* in a model with labor market segmentation.

Raising the *marginal tax rate* for high incomes will make it relatively less tempting for low-type workers to apply for high-wage firms. The point is illustrated in Figure 5, where I have assumed that the tax function is of the form $T = \max[0, t(w - w^l)]$ (so that a person pays taxes only on the part of his income in excess of the equilibrium wage w^l in the low-wage market). This shifts the low-type workers' indifference curve downward, reduces the wage w^h in the high-type market and increases output. As the marginal tax rate approaches one, the equilibrium allocation will converge to the output maximizing solution.

Finally, the output maximizing solution may be obtained by imposing the same wage (correctly set) in all firms. This may be a result of centralized bargaining. In this case, it would not be possible for firms to separate, and in the resulting pooling equilibrium the production threshold facing high- and low-type workers would be

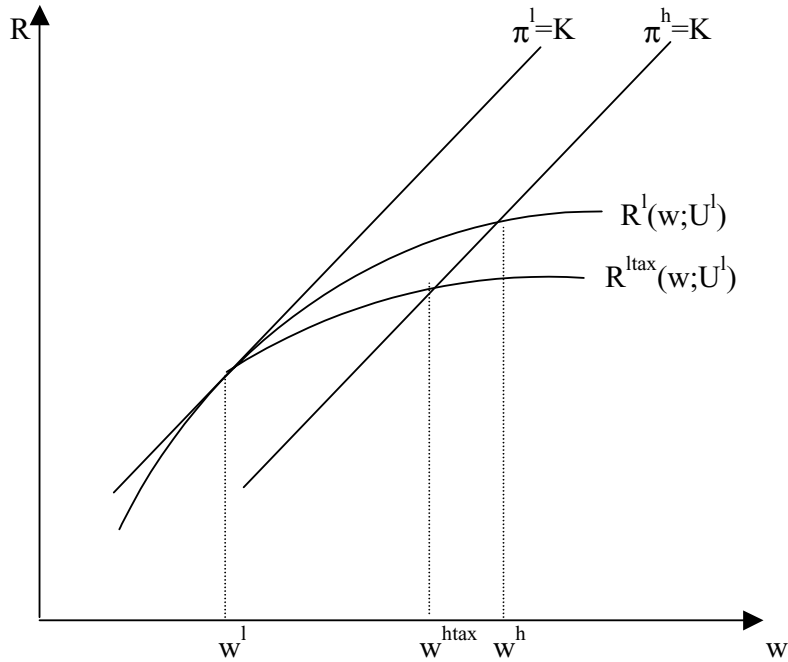


FIGURE 5

THE EFFECT OF TAXING INCOME EXCEEDING w^l . THE UPPER PART OF THE INDIFFERENCE CURVE OF LOW-TYPE WORKERS SHIFTS DOWN, AND THIS REDUCES THE EQUILIBRIUM WAGE FOR HIGH-TYPE WORKERS

the same. In Moen (2001), it is shown that if the fraction of low-type workers in the economy is sufficiently low, an optimally set centralized wage may actually be Pareto-improving. However, imposing a single wage on all workers would, of course, create distortions along dimensions not modeled here.

6.2. *Wage Dispersion.* It is easy to show that the equilibrium wage for high-type workers is continuous in a^h at $a^h = a^l$. Thus, if workers in one group are marginally more productive than the workers in the other group, this will not radically change the equilibrium relative to the situation with ex ante identical workers. On the other hand, the derivative of w^h does not exist at this point.

LEMMA 4. *The equilibrium wage w^h is continuous in a^h at $a = a^h$. However, we have that*

$$(11) \quad \lim_{a^h \rightarrow a^{l+}} \frac{dw}{da^h} = \infty$$

The proof is given in the Appendix. It follows that even small productivity differentials may lead to disproportionately large differences in wages. In Moen

(2001), similar results are shown for increases in the worker–firm productivity term. However, as should be clear by now, this reduces the welfare of high-type workers through the adverse effect on their employment rate.

6.3. *The Dual Economy.* Within the fields of sociology and economics, theories have been developed based on a two-tier picture of the labor market. The upper tier (the primary sector or the core) consists of well-paid jobs with high job security and low turnover, and with a queue of workers wanting jobs in these firms. The lower tier (the secondary sector or the periphery) consists of low-paid jobs with low job security and high turnover. Here wages are determined by market clearing. An economic rationale for dual labor markets based on efficiency wages is given in Bullow and Summers (1986). Saint Paul (1996) extends the model in various directions. In their model, monitoring is costly in the primary sector, and firms in this sector pay efficiency wages and workers are queuing to obtain jobs in the sector. In the secondary sector, monitoring is free and the market clears.

The mechanisms in this article may give rise to similar effects. Suppose, for instance, that the production technology in the primary sector is as described above. In the secondary sector, all workers are equally productive, and the labor market clears. The search market in the primary sector is modeled as above, and there is a cost associated with applying for a job in the primary sector. The equilibrium in this model is derived in Moen (2001) under parameter restrictions ensuring that it is unprofitable for primary sector firms to enter the market if they only attract low-type workers. As in the previous sections, wages in the high-tech sector have to be driven up to the point where low-type workers stop applying to the high-tech firms (due to stiff competition for jobs). In order to prevent low-type workers from applying, competition for primary sector jobs must be sufficiently high, and in order to obtain this, firms set wages high. Again it is clear that in equilibrium, wages in the primary sector are too high while the number of jobs in the sector is too low.

6.4. *More Complex Wage Contracts.* Suppose firms could advertise wages depending on productivity in our model. It is then straightforward to show that in equilibrium, workers would be paid their marginal product less a fee $\pi = K$ to the firm, and since all workers would accept the job if and only if $z \geq K$, the resulting equilibrium is efficient. It is also easy to demonstrate that efficiency can be obtained if firms attracting high-type workers make employment conditioned on worker type.⁷

The simplest way to introduce contingent wage contracts is through piece-rate pay, but this may induce substantial agency costs. In the absence of piece-rate payment, contingent wage contracts may be difficult to implement. Firstly, workers may not know their productivity in a given firm (as assumed in this article), and the firm in question then has an incentive to take advantage of this fact by claiming that

⁷ Shi (2001) derives an analogous result with urn-ball matching technology.

a given worker is less productive than he really is. Secondly, it may be hard to give an operational definition of a worker's productivity which is sufficiently accurate to serve as a basis for wage determination, and it may be harder still to communicate (and stick to) this definition to workers through a brief job advertisement. Thirdly, even if firms may rank their applicants (which is sufficient for the equilibrium with unconditional wage offers to exist), productivity may be costly or even impossible to quantify precisely.

What then about tying employment or salary to worker type? This should be relatively unproblematic if worker type refers to specific, observable characteristics, as formal education or years of experience. This article thus gives a rationalization as to why some firms require formal schooling, good grades, etc. from all employees, even if this is only imperfectly correlated with being productive in the specific firm. However, even after controlling for observable characteristics, there will probably still be a fair amount of *ex ante* heterogeneities left within the pool of workers a firm is hiring from. This may partly be for legal reasons: Suppose for instance that a is a person's IQ-score, while z is his ability to perform certain tasks in a firm. According to U.S. law, an employer may be allowed to give applicants a specific test that yields information about z , but he is usually not allowed to administer a general IQ test. Finally, even if worker type is observable, it may be illegal to use it in a wage contract if worker type is associated with sex, ethnicity, age, etc.

From an empirical point of view, there is if evidence that a worker's productivity is not fully reflected in his wage. Frank (1984) argues that in many firms, workers with large differences in productivities obtains the same wages. Furthermore, he shows that even in firms which offer piece-rate payment, the difference in productivity between workers by far exceeds the corresponding differences in wages.

7. CONCLUSION

In this article, I have constructed a directed search model with heterogeneous workers. Unemployment exists because a given worker has different productivities in different firms. I analyze the equilibrium of the model in a very simple and realistic setting in which firms advertise unconditional wage offers and hire the most productive workers that show up.

I find that the resulting equilibrium is not efficient, as wages in the submarket for high-type workers are too high and the number of job openings too low. Firms wanting to attract high-type workers are forced to offer too high wages and in turn are reluctant to hire sufficiently many workers at this high wage level. The submarket for low-type workers on the other hand is efficient. Thus, the existence of low-type workers in the market reduces the welfare of high-type workers but not the other way around. This contrasts one of the main two results in the literature on the labor market segmentation, namely that the existence of high-type workers reduces the welfare of low-type workers. However, my model gives support to the second main result in this literature, stating that too few primary sector (well paid) jobs will be created in equilibrium.

APPENDIX

A.1. *Derivation of the Optimal Solution.* The planner's problem can be expressed as maximizing

$$L^l \int_{R^l}^{\infty} zg(z - a^l) dz + L^h \int_{R^h}^{\infty} zg(z - a^h) dz - K(N^l + N^h)$$

with respect to R^l , R^h , N^l , and N^h , given the constraints

$$(A.1) \quad N^k = (1 - G(R^k))L^k$$

The associated Lagrangian is given by

$$L = L^l \int_{R^l}^{\infty} zg(z - a^l) dz + L^h \int_{R^h}^{\infty} zg(z - a^h) dz - K(N^l + N^h) \\ - \gamma^l (L^l (1 - G(R^l - a^l)) - N^l) - \gamma^h (L^h (1 - G(R^h - a^h)) - N^h)$$

First-order conditions for maximum can be written as

$$\frac{\partial L}{\partial R^k} = 0 \Rightarrow R^k = \gamma^k \\ \frac{\partial L}{\partial L^k} = 0 \Rightarrow \gamma^k = K$$

for $k \in \{l, h\}$. Combined, these equations yield

$$(A.2) \quad R^k = K$$

Equation (A.2), together with the constraints (A.1), determines the optimal aggregate allocation in the economy. It is trivial to show that the planner's maximization problem has an interior solution, and as the necessary condition (A.2) has a unique solution, it follows that it defines the global maximum.

A.2. *Characterization of When the Constraint $R(w) \geq w$ Does Not Bind.* Ex post profit π can be written as $E[z|z \geq R] - w$, and since $R = K$ in the unconstrained equilibrium, it follows from the zero-profit condition that $E[z|z \geq K] - w = K$. By inserting $w < K$ into this expression, we find that the constraint does not bind if and only if $E[z|z \geq K] > 2K$.

PROOF OF THE CLAIM THAT $\pi = R$. A worker is hired in a firm with productivity threshold R if and only if his worker-firm specific term exceeds $R - a$ in this firm. Define $\hat{\epsilon} = R - a$. A worker's indifference curve can then be expressed as $\hat{\epsilon}(w; U) = R(w; U) - a$, and obviously we must have that $\hat{\epsilon}_w(w; U) = R_w(w; U)$.

A firm's maximization problem can be written as

$$\max a + h(\hat{\epsilon}(w; U)) - w$$

Taking derivatives with respect to w gives the first-order condition

$$(A.3) \quad h'(\hat{\epsilon})\hat{\epsilon}_w(w; U) = 1$$

Since the distribution of ϵ given that $\epsilon \geq \hat{\epsilon}$ is given by $g(\epsilon)/(1 - G(\hat{\epsilon}))$, it follows that

$$h(\hat{\epsilon}) = \frac{\int_{\hat{\epsilon}}^{\infty} \epsilon g(\epsilon) d\epsilon}{1 - G(\hat{\epsilon})}$$

Taking derivatives gives

$$\begin{aligned} h'(\hat{\epsilon}) &= \frac{-\hat{\epsilon}g(\hat{\epsilon})(1 - G(\hat{\epsilon})) + g(\hat{\epsilon})\int_{\hat{\epsilon}}^{\infty} \epsilon g(\epsilon) d\epsilon}{(1 - G(\hat{\epsilon}))^2} \\ &= (-\hat{\epsilon} + h(\hat{\epsilon}))\frac{g(\hat{\epsilon})}{1 - G(\hat{\epsilon})} \end{aligned}$$

As $\hat{\epsilon}_w = R_w$, it follows from (5) that

$$h'(\hat{\epsilon})\hat{\epsilon}_w(w; U) = \frac{h(\hat{\epsilon}) - \hat{\epsilon}}{w}$$

which, together with (A.3) implies that

$$w = h(\hat{\epsilon}_i) - \hat{\epsilon}_i$$

Inserted into the profit function, this gives

$$\begin{aligned} \pi_i &= a + h(\hat{\epsilon}_i) - (h(\hat{\epsilon}_i) - \hat{\epsilon}_i) \\ &= a + \hat{\epsilon}_i \\ &= R_i \end{aligned}$$

which is what we wanted to prove. ■

PROOF OF LEMMA 3. We know that at least one of the constraints has to bind. Due to the single-crossing property, it follows that both constraints cannot bind. It is therefore sufficient to show that the constraint for high-type workers does not bind. Suppose it did. Then, in equilibrium, high-type workers would be indifferent between applying for high-type or for low-type firms. Consider a firm that increases its wage offer slightly above the low-type wage. This firm would then attract high-type workers only, and obtain a higher profit than when advertising the equilibrium

wage for low types. As this would be inconsistent with equilibrium, the proof is complete. ■

PROOF OF PROPOSITION 3. We will first show that the proposed equilibrium is unique. Then we will show that it actually constitutes an equilibrium.

That the equations defining the equilibrium in the low-type submarket have a unique solution is obvious, and therefore omitted. The profit to a firm, given that only high-type workers apply, as $\pi^h = \bar{z}^h(R^h) - w$. Inserting the indifference curve for low-type workers implies that we can write the profit function as $\pi^h(w) = \bar{z}^h(R^l(w)) - w$.

By construction, it follows that low-type workers cannot gain by deviating and applying to a high-type firm. We have already shown in the proof of Lemma 3 that the incentive-compatibility constraint for high-type workers does not bind; hence high-type workers will not gain by applying to a low-wage firm. If a high-wage firm deviates and advertises a wage marginally below w^h , it follows from Lemma 2 that this firm attracts low-type workers only; hence its profit falls.

Now consider a firm that deviates by offering a wage above w^h . Obviously, w^h is higher than the value of w that would prevail if the incentive-compatibility constraint for low-type workers did not bind. Furthermore, given that $h'' < 0$, it follows that the profit maximizing wage is increasing in U . Since U^h is lower than it would have been if the incentive-compatibility constraint for low-type workers did not bind, it follows that the deviating firm reduces its profit. As it is obvious that a low-wage firm cannot increase its profit by deviating, this completes the proof. ■

PROOF OF LEMMA 4. Write $a^h = a^l + \gamma$, and keep γ fixed, and write $w^h = w^h(\gamma)$. Suppose w^h is discontinuous in γ at $\gamma = 0$. Then there exists a δ such that for all $\gamma > 0$, $w^h \geq w^l + \delta$. Now consider a firm that offers a wage $w^l + \delta$ and attracts low-type workers. The profit of this firm can be written as $\pi(w^l + \delta, a^l) \approx \pi(w^l, a) + \pi_w \delta + \pi_{ww} \delta^2 = \pi(w^l, a) + \pi_{ww} \delta^2$, since $\pi_w = 0$ as w is optimally chosen. Since $\pi_{ww} < 0$, it follows that $\pi(w^l + \delta, a^l) < \pi(w^l, a)$.

Now consider a firm that offers the same wage, but attracts high-type workers. As these workers are high-type workers, the extra profit is given by $\gamma[1 - h'(R - a)] < \gamma$. Now choose γ such that $\gamma < \pi_{ww} \delta^2$. It follows that a firm offering $w^l + \delta$ and attracting high-type workers obtain less than $\pi(w^l, a^l) = K$. But this is inconsistent with equilibrium.

Let us then show Equation (11). The expected productivity of workers hired in high-wage firms can be written as $a^h + h(R^h - a^h)$. Since $R^h = R^l(w^h)$ (omitting U^l as an argument of R^l), we can write the zero-profit condition as $a^h + h(R^l(w) - a^h) - w^h = K$, which determines w^h . Taking derivatives with respect to a^h gives

$$1 + h'(R^l - a^h) \left[R_w^l \frac{dw^h}{da^h} - 1 \right] - \frac{dw^h}{da^h} = 0$$

or

$$(A.4) \quad \frac{dw}{da^h} = \frac{1 - h'(R^l - a^h)}{(1 - h'(R^l - a^h)R_w^l)}$$

By definition, w^l maximizes π^l , that is, maximizes $a^l + h(R^l(w) - a^l) - w^l$, with respect to w^l . The first-order condition for this maximization problem is given by $h'(R^l - a^l)R_w^l$. Due to continuity, it thus follows that the denominator goes to zero as $a^h \rightarrow a^{l+}$, and Equation (11) follows.

We want to show that $\frac{dw^h}{da^h} > 1 - h'(K - a^h)$ for all a^h . As the denominator in (A.4) is always less than one, it is sufficient to show that $1 - h'(R^l(w^h) - a^h) \geq 1 - h'(K - a^h)$, or that $h'(R^l(w^h) - a^h) \leq h'(K - a^h)$. Since $R^l(w^h) > K$, this follows from our assumption that $h'' < 0$. ■

REFERENCES

- ACEMOGLU, D., AND SHIMER, R., "Optimal Unemployment Insurance," *Journal of Political Economy* 67 (1999), 343–69.
- BULLOW, J., AND L. SUMMERS, "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination and Keynesian Unemployment," *Journal of Labor Economics* 4 (1986), 376–411.
- FRANK, R. H., "Are Workers Paid Their Marginal Products?" *American Economic Review* 74 (1984), 549–71.
- HALL, R. E., "A Theory of the Natural Rate of Unemployment," *Journal of Monetary Economics* 5 (1979), 153–69.
- INDERST, R., AND H. M. MULLER, *Competitive Search Markets with Adverse Selection* (University of Mannheim, 1999).
- JUHN, C., K. MURPHY, AND B. PIERCE, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy* 101 (1993), 410–42.
- LANG, K., AND W. T. DICKENS, "Labor Market Segmentation, Wage Dispersion, and Unemployment," NBER Working paper series No. 4073, 1992.
- , AND —, "Bilateral Search as an Explanation for Labor Market Segmentation and Other Anomalies," NBER Working paper series No. 4461, 1993.
- , AND S. KAHN, "The Effect of Minimum-Wage Laws on the Distribution of Employment: Theory and Some Evidence," *Journal of Public Economics* 69 (1998), 67–82.
- , M. MANOVE, AND W. T. DICKENS, "Racial Discrimination in Labor Markets with Announced Wages," Working paper, Boston University, 1999.
- MOEN, E. R., "Competitive Search Equilibrium," *Journal of Political Economy* 105 (1997), 385–411.
- , "How the Invisible Hand Create Excessive Wage Differentials," Discussion paper No. 9, Norwegian School of Management, 1999.
- , "Do Good Workers Hurt Bad Workers—or Is It the Other Way Around," Working paper, Norwegian School of Management, 2001.
- MONTGOMERY, J. D., "Equilibrium Wage Dispersion and Interindustry Wage Differentials," *Quarterly Journal of Economics* 106 (1991), 163–79.
- MORTENSEN, D. T., AND C. A. PISSARIDES, "New Developments in Models of Search in the Labor Market," in O. Ashenfelter, and D. Card, eds., *Handbook of Labor Economics*, Vol. 3b (Amsterdam: North-Holland, 1999), 2567–627.
- PETERS, M., "On the Equivalence of Walrasian and Non-Walrasian Equilibria in Contract Markets: The Case of Complete Contracts," *Review of Economic Studies* 64 (1997), 241–65.

- ROTCCHILD, M., AND J. E. STIGLITZ, "Equilibrium in Competitive Insurance Markets," *Quarterly Journal of Economics* 90 (1976), 629–50.
- SAINT PAUL, G., *Dual Labor Markets: A Macroeconomic Prospective* (Cambridge: MIT Press, 1996).
- SHI, S., "Frictional Assignment: 1. Efficiency," *Journal of Economic Theory* 98 (2001), 232–60.
- SHIMER, R., "Job Auctions," Working paper, Princeton University, 1999.
- TAUBMAN, P., AND M. L. WACHTER, "Segmented Labor Markets," in O. C. Ashenfelter, and R. Layard, ed., *Handbook of Labor Economics*, Vol. 2, Chapter 21 (Amsterdam: North-Holland, 1986) 1183–218.