Available online at www.sciencedirect.com
SCIENCE DIRECT®

European Economic Review 49 (2005) 1525–1542

EUROPEAN ECONOMIC REVIEW

www.elsevier.com/locate/econbase

Competitive search markets with heterogeneous workers

Roman Inderst*

Department of Economics and Department of Accounting and Finance, London School of Economics, Houghton Street, London WC2A 2AE, UK

Received 12 December 2001; accepted 16 January 2003

Abstract

This paper extends the concept of competitive search markets to the case with heterogeneous workers. If offers can condition on workers' productivity as this depends on some formal qualification, the equilibrium is separating and efficient, though there is wage compression. If offers cannot condition on workers' productivity, the nature of the equilibrium depends on the size of the workers' productivity difference. For intermediate values of the productivity difference, separation is achieved only by paying high types an "efficiency wage" premium. We discuss several implications for wage dispersion and efficiency.

JEL classification: J41

Keywords: Search markets; Efficiency wages; Wage differential

1. Introduction

This paper explores wage setting in a model with heterogeneous workers and search frictions. In doing so, we follow recent developments in the search literature and apply a model of competitive search markets. In contrast to most previous search market approaches, this framework allows for the co-existence of separate search environments.

In our benchmark model, the workers' productivity can be verified as it is linked to years of schooling or some formal training. Offers can thus be made contingent on workers' types. In this case, our model extends previous work by Moen (1997), who considers homogeneous workers. We find that workers with different productivities search in separate environments. More productive workers have both higher wages

* Tel.: +44-20-7955-7291; fax: +44-20-7831-1840.
E-mail address: r.inderst@lse.ac.uk (R. Inderst).

0014-2921/$ - see front matter
doi:10.1016/j.euroecorev.2003.01.001

and shorter unemployment spells. Interestingly, wages are compressed, i.e., the wage difference is smaller than the productivity difference. The wage difference increases as the market becomes more efficient, i.e., as search costs decrease or market transparency increases.

If the workers' productivity is not (fully) determined by some formal training, offering contingent contracts may not be feasible. In this case, a separating equilibrium can only be supported if the high types' wage is sufficiently increased compared to the case with contingent contracts, i.e., if high types receive an "efficiency wage" premium. This is necessary to ensure that firms would rather dismiss a (deviating) low type instead of employing him at the conditions offered to high types. Interestingly, efficiency wages are only paid if the productivity difference between the two types takes on intermediate values. If the productivity difference is sufficiently small, only a pooling equilibrium exists, in which, by definition, both types receive the same wage. We argue that a comparison of the two regimes with contingent and non-contingent contracts may shed some light on how wages and wage differences are affected if the composition of skills or the nature of required skills change.

To our knowledge, search by heterogeneous workers has almost exclusively be considered in models that allow only for a single-market environment, i.e., employing the Diamond–Mortensen–Pissarides framework (see Pissarides, 1990). The few exceptions we are aware of are Lang and Dickens (1992, 1993), Moen (2000), and Inderst (2000). We comment on the relation to these papers in much detail below, when presenting our results. [1]

The rest of this paper is organized as follows. Section 2 introduces the search market environment. Section 3 studies the benchmark case of contingent contracts. In Section 4 we compare this to the case where contracts can no longer condition on workers' productivities. Section 5 concludes.

2. The search market environment

2.1. The matching technology

We consider an anonymous search market environment. The market operates in continuous time and is populated by a continuum of workers and firms. The number of open vacancies will be determined endogenously by a zero-profit condition. We assume that opening a new vacancy comes at the costs $k \geqslant 0$, which are incurred immediately. The productivity of a vacancy depends only on the productivity of the worker with which it is filled. There are two possible types of workers, indexed by $t \in T = \{l, h\}$, which we call low and high. Each type of worker is associated with a productivity y_t, for which we assume $0 < y_l < y_h$. The measure of workers in the

[1] Inderst and Müller (2002) explore competitive search markets with adverse selection. A somewhat related strand of the literature considers the case where firms do not announce wages but, instead, more complicated mechanisms such as auctions (e.g., McAfee, 1993; Peters and Severinov, 1997).

economy is normalized to one. The fraction $\mu^0 \in (0,1)$ of workers has type h. Workers and firms are risk neutral.

If a search market contains the measure u_t of workers of type t, the total measure $u := u_l + u_h$ of workers, and the measure v of firms, the measure of matchings that are formed per unit of time is given by the Cobb–Douglas matching technology $\xi[uv]^{1/2}$. The parameter $\xi > 0$ captures the transparency of the market and influences the speed with which matches are formed. Denote the tightness prevailing in the market by $\theta := v/u$. We then obtain for workers the transition rate $\xi\theta^{1/2}$ and for firms the transition rate $\xi\theta^{-1/2}$. While our qualitative results extend to more general matching technologies with constant returns to scale, we choose the Cobb–Douglas technology as it ensures uniqueness of an equilibrium.

A worker–firm pair starting production faces a constant probability rate s of job destruction. In this case, the worker starts to search anew, while the firm's position is removed.

2.2. Asset equations (Utilities from search)

Suppose the wage w prevails in some search market. As specified below, wages will be determined by a competitive process and not, as typically assumed in the search literature, by bilateral negotiations. We denote the utility of a job searcher by U. This is given by [2]

$$rU = z + \xi\theta^{1/2}(E - U), \tag{1}$$

where $z \geqslant 0$ denotes the unemployment income, $r \geqslant 0$ denotes the discount rate, and E denotes the expected utility if the worker meets a firm. Suppose for the moment that all matches are successful. Recall also that $s \geqslant 0$ denotes the exogenous risk of breakdown. We then obtain that E must satisfy

$$rE = w - s(E - U). \tag{2}$$

Substituting (2) into (1) allows to derive U as a function of θ and w:

$$U(w,\theta) := \frac{1}{r}\,\frac{(r+s)z + w\xi\theta^{1/2}}{r + s + \xi\theta^{1/2}}. \tag{3}$$

The value of a vacancy V is derived analogously. It satisfies

$$rV = -c + \xi\theta^{-1/2}(J - V), \tag{4}$$

where $c > 0$ denotes firms' search costs, while J equals firms' (expected) payoff if they employ a worker. If the probability of employing a high-type worker is given by μ, we obtain [3]

$$rJ = \mu y_h + (1-\mu)y_l - w - sJ. \tag{5}$$

[2] For an intuitive derivation of this and the following equations see Pissarides (1990).

[3] We follow much of the literature in assuming that the exogenous shock destroying the vacancy may only occur if the position is filled. Otherwise, we would have to transform (4) into $(r+s)V = -c + \xi\theta^{-1/2}(J - V)$, which would not lead to any qualitative change in results.

The value of a vacancy as a function of θ, w, and μ equals now:

$$V(w,\theta,\mu) := \frac{\xi\theta^{-1/2}[\mu y_h + (1-\mu)y_l - w] - c(r+s)}{(r+s)(r+\xi\theta^{-1/2})}. \tag{6}$$

As noted above, we have assumed that production takes place in all matches. This is only optimal for the worker if w exceeds his unemployment income z. Moreover, employing some type t is only optimal for firms if the expected profit $(y_t - w)/(r+s)$ does not fall short of the value of an open vacancy. (In equilibrium, the value of an open vacancy will be k as we will impose a zero-profit requirement for firms.) In what follows, these conditions will always apply whenever we make use of U and V.[4]

3. The case of contingent contracts

3.1. Market equilibrium

Following Moen (1997) and Acemoglu and Shimer (1999), we introduce the concept of a submarket.[5] A submarket is indexed by a positive natural number i and contains the measures of searching workers u_l^i, u_h^i and searching firms v^i. If the submarket opens up, it holds that $v^i > 0$ and $u^i = u_l^i + u_h^i > 0$. The distribution of types is denoted by $\mu^i = u_h^i/u^i$ and the tightness by $\theta^i := v^i/u^i$. We assume that all submarkets exhibit the same Cobb–Douglas matching technology. A submarket is associated with a fixed wage offer w^i, to which firms commit.

In this section, we further assume that wage offers can be made contingent on workers' types, implying that submarkets can discriminate between different workers. This assumption is reasonable if the worker's type can be documented by his years of schooling or another formal qualification. Below, we will consider the case where offers cannot be made contingent on types.

Denote the index set of all open submarkets by I and denote $U_t^* := \max_{i\in I} U(w^i,\theta^i)$. We require that firms and workers choose optimally whether and where to search.

(E.1) Workers and firms choose optimally whether to engage in search and in which submarket to do so. This gives rise to the following requirements:

(i) $V(w^i,\theta^i,\mu^i) = k$ holds for all $i \in I$.
(ii) $U(w^i,\theta^i) = U_t^*$ holds for all $i \in I$, $t \in T$ satisfying $u_t^i > 0$.

Note that, even though workers of some type t could benefit from entering some submarket i for which $u_t^i = 0$ and $U(w^i,\theta^i) > U_t^*$ hold, they can be prevented from doing so as offers are made contingent on workers' types.

[4] It is, however, straightforward to generalize the asset value equations, taking into account firms' and workers' optimal continuation decisions. For instance, we could re-write (5) such that $J(r+s)$ equals $\mu\max\{y_h - w, J\} + (1-\mu)\max\{y_l - w, J\}$.

[5] The idea that separate submarkets may form has also been used, though differently, in the literature on dual labor markets, e.g., in Saint-Paul (1996, Chapter 9).

Our following condition of competitiveness is equivalent to that in Mortensen and Pissarides (1998). They require that, in equilibrium, there is no scope for middlemen to profitably complete the market by opening up additional submarkets. A middleman is supposed to announce a pair (w,θ) and a set of admitted types. Potential entrants compare this opportunity to existing submarkets.[6]

(E.2) An equilibrium set of submarkets I must satisfy the following condition of competitiveness. There exists no $(w,\theta)\neq(w^i,\theta^i)$ and no type $t\in T$ such that $U(w,\theta)>U_t^*$ and $V(w,\theta,\mu)>k$ hold, where $\mu=1$ holds for $t=h$ and $\mu=0$ holds for $t=l$.

As the matching technology has constant returns to scale, we can require, without loss of generality, that all submarkets $i\in I$ have different characteristics (w^i,θ^i). Note, finally, that the restriction to a single-wage contract is without loss of generality. While we could also let firms offer a menu of contracts that condition on the verifiable type of the worker, in equilibrium only one contract would be chosen from the menu offered in a particular submarket. This follows as in equilibrium workers of different types will always be active in different submarkets, which we will show below.

3.2. Analysis

We abbreviate the analysis as it is largely analogous to that in Moen (1997). To ensure that production is efficient with both types, we make the following assumption.

Assumption 1. $(y_l-z)/(r+s)>k$.

From a worker's perspective, the zero-profit requirement for firms represents the following trade-off. To keep firms at zero profits, a higher wage must imply a shorter expected search time for firms, i.e., a longer expected unemployment spell for workers. By the requirement (E.2), this trade-off between a higher wage and longer expected search must be resolved *efficiently* in each submarket that opens in equilibrium. This is the case if the characteristics (w,θ) of some submarket in which type-t workers search solve the following program: Maximize workers' payoff $U(w,\theta)$ subject to firms' zero-profit constraint $V(w,\theta,\mu)=k$, where $\mu=1$ holds for $t=h$ and $\mu=0$ holds for $t=l$. Given our choice of the matching technology, this program has a unique solution for each type (see Proposition 1). For type t, we denote the respective wage by w_t^* and the respective tightness by θ_t^*.

We find that in equilibrium the two types must search in different submarkets.[7] Moreover, high-type workers both obtain a higher wage *and* find a job faster. These results follow immediately from the efficiency of the characteristics prevailing in all open submarkets. Efficiency dictates that the marginal rates of substitution between the

[6] The existence of a new submarket, which is opened up by deviating middlemen, does not change the composition of existing submarkets as it is supposed to be of negligible size.

[7] Note that this would also hold if firms could offer a menu of contracts. In that case, only one type would enter the respective submarket.

wage and the tightness, i.e., between the payoff in a successful match and the speed of matching, must be equal for workers and firms entering this submarket.

Suppose, for instance, that high-type workers would *not* match faster than low-type workers in equilibrium. For simplicity, suppose that low-type workers' search time is strictly lower, i.e., that $\theta_h < \theta_l$. By firms' zero-profit requirement and $y_h > y_l$, this implies that (i) high-type workers must receive a strictly higher wage, i.e., $w_h > w_l$, and that (ii) the profits of firms matching with high types are not higher, i.e., $y_h - w_h < y_l - w_l$. Note next that both firms and workers have a higher preference for matching faster in case their payoff from a match increases. In the proposed equilibrium, low-type workers receive a lower wage than low-type workers and match faster, while firms employing high-type workers realize a lower profit and match faster than firms employing low-type workers. Hence, holding wages constant, firms employing low-type workers have a stronger preference for matching faster than those employing high-type workers, while high-type workers have a stronger preference for matching faster than low-type workers. Clearly, this choice of submarket characteristics cannot be efficient.

A similar argument shows that high-type workers must not only match faster, but they must also obtain a higher wage.

Proposition 1. (i) *With contingent contracts there exists a unique equilibrium, in which workers with different productivities search in separate submarkets. High-type workers receive a higher wage and find a job faster, i.e.,* $w_h^* > w_l^*$ and $\theta_h^* > \theta_l^*$.

(ii) *The wage difference between high- and low-productivity workers* $w_h^* - w_l^*$ *is strictly smaller than the difference in productivities* $y_h - y_l$.

Proof. See the Appendix.

Note that assertion (ii) follows immediately from assertion (i) and the firms' zero profit requirement. To ensure that firms realize k in each submarket, $\theta_h^* > \theta_l^*$ implies that the wage difference $w_h^* - w_l^*$ must be strictly smaller than the difference in productivities $y_h - y_l$. It is interesting to note that this type of wage compression is achieved in an environment where (i) workers of different productivities are employed by different firms and (ii) wages are determined in a competitive environment and not, for instance, in bilateral negotiations.

The equilibrium described in Proposition 1 has the realistic feature that workers who are more productive and employed at higher wages have (on average) incurred a lower unemployment spell. Interestingly, this is not the case if workers are homogeneous and firms' vacancies have different productivities. As found in Moen (1997), firms with more productive vacancies will offer higher wages to attract more workers to the respective submarket, which reduces the time it takes to fill the respective vacancies. As a consequence, workers receiving a higher wage have previously experienced (on average) longer unemployment spells if only firms differ in productivities.

Inderst (2000) analyzes the competitive search market model with bilateral heterogeneity. If the productivity of a job is just the sum of the "productivities" of the respective firm and the respective worker, high-type workers still receive a higher wage and match faster than low-type workers. Moreover, in equilibrium high-type workers are matched with low-type firms and vice versa. This type of negative-assortative matching ensures that both high-type firms and high-type workers can match faster than the respective low types, while obtaining a higher payoff from matching. However, if workers' and firms' productivities are sufficiently complementary, high-type firms will match with high-type workers. In this case, workers receiving a high wage may face a longer expected unemployment spell than those receiving a low wage.

Recall that search is costly for various reasons. Until a match has formed, the two sides' joint flow utility $z-c$ is strictly lower than under employment. Moreover, waiting is costly as both firms and workers discount future wages and profits. We next analyze how the equilibrium is affected by the efficiency of the search market as measured both by the speed of matching, which is determined by the transparency parameter ξ, and the size of direct search costs c.

Proposition 2. *If the market becomes more efficient, i.e., if either firms' search costs c decrease or if the transparency parameter ξ increases, the following results hold*: (i) *both types of workers receive higher wages and find a job faster, i.e.,* w_t^* and θ_t^* *increase for both* $t=l,h$; (ii) *the wage gap* $w_h^* - w_l^*$ *widens.*

Proof. See the Appendix.

As the market becomes more efficient, wages increase and the duration of unemployment spells decreases. More interestingly, the wage difference between high- and low-type workers increases. In particular, as the market becomes increasingly transparent as $\xi \to \infty$, the wage for each type t approaches the upper boundary $y_t - k(r+s)$. The upper boundary is equal to the worker's productivity minus the rental price of the firm's vacancy.

In the US, wage dispersion seems to have increased over the last decades.[8] Wage inequality has risen both between groups (as identified by observable characteristics such as education, age, or work experience) and within groups. Recall that we allow offers to condition on workers' (observable) types. If we identify workers' types with different *skill* groups, wage dispersion should increase according to Proposition 2 in case labor markets become more efficient. Whether or not this has been the case is a debatable empirical question. In fact, according to some scholars, match efficiency seems to have deteriorated as documented by an outwards shift in the Beveridge curve.[9]

We next consider the case where workers' types are no longer determined by formal qualifications, which can be written into a job requirement. Instead, workers with a particular productivity can no longer be sorted "from a distance". This limits the ability for firms to directly condition offers on skill differences, which, as we show, may again increase the wage difference.

[8] Levy and Murnane (1992) provide a review of the literature.

[9] Nickell et al. (2002) discuss shifts in the Beveridge curve for OECD countries.

4. The case of non-contingent contracts

4.1. Market equilibrium

We assume now that contracts can no longer condition on workers' types. While firms can perfectly observe a worker's productivity once a match has formed and the worker has been interviewed, productivity is not fully captured by years of schooling or other formal training, which could be written as requirements into the job advertisement. While firms cannot make their offers contingent on workers' types, they still have the choice to employ workers at the advertised wage or to turn them down in order to search anew. As the firms' continuation value after dismissal is just k, a worker of type t will be turned down whenever it holds that

$$w \geqslant y_t - k(r+s), \tag{7}$$

i.e., whenever the firm's flow profits $y_t - w$ do not exceed the rental price of the vacancy.

With this observation we can now modify the equilibrium requirement (E.1).

(E$'$.1)

(i) $V(w^i, \theta^i, \mu^i) = k$ holds for all $i \in I$.
(ii) $U(w^i, \theta^i) = U_t^*$ holds for all $i \in I$, $t \in T$ satisfying $u_t^i > 0$. Additionally, $u_t^i = 0$ implies $U(w^i, \theta^i) \leqslant U_t^*$ if $w^i < y_t - k(r+s)$.

Hence, if workers of some type t do not enter a particular submarket i, one of the following two conditions must hold. Either it must not be more profitable to be employed under the respective conditions (w^i, θ^i), i.e., $U(w^i, \theta^i) \leqslant U_t^*$. Or, these workers would not be employed at the prevailing wage, i.e., $w^i \geqslant y_t - k(r+s)$. The latter issue of "incentive compatibility" is now also reflected in the following modification of requirement (E.2).

To modify (E.2), suppose again that middlemen specify the conditions (w, θ) for a new search environment. Firms entering this new submarket form rational expectations about the types of attracted workers. Workers of type t are not expected to enter if either their utility would fall short of their equilibrium utility, i.e., if $U(w, \theta) < U_t^*$, or if they would not expect to be employed, i.e., if $w \geqslant y_t - k(r+s)$. If both types of workers can be expected to enter the new submarket as this is profitable, we specify that firms expect the "fair" distribution μ^0 to prevail. We comment on this specification below.

(E$'$.2) There exists no $(w, \theta) \neq (w^i, \theta^i)$ with the following conditions:[10]

- If $w < y_l - k(r+s)$: $U(w, \theta) \geqslant U_t^*$ for some $t \in T$ and $V(w, \theta, \mu) > k$, where $\mu = 1$ is chosen for $U(w, \theta) \geqslant U_h^*$ and $U(w, \theta) < U_l^*$, $\mu = 0$ is chosen for $U(w, \theta) \geqslant U_l^*$ $U(w, \theta) < U_h^*$, and $\mu = \mu^0$ is chosen for $U(w, \theta) \geqslant U_h^*$ and $U(w, \theta) \geqslant U_l^*$.
- If $w \geqslant y_l - k(r+s)$: $U(w, \theta) \geqslant U_h^*$ and $V(w, \theta, 1) > k$.

[10] Note that we have already restricted consideration to wages $w < y_h - k(r+s)$, which can safely be done.

Note that a new submarket that is (weakly) attractive to both types $t = l, h$ will generate the distribution μ^0. We next comment in detail on this specification.

One alternative would be to take into account how much each type t gains in the new submarket compared to his respective equilibrium utility U_t^*. Workers who have more to gain should be more willing to choose the new submarket and should therefore make up a larger fraction of the resulting distribution of workers. One way to capture this, which is explored in Inderst and Müller (2002) for a different setting, is as follows. Instead of specifying both w and θ for a new submarket, only the wage offer w is chosen exogenously. For instance, instead of appealing to the existence of middlemen, we could think of firms starting to offer a different wage.[11] The new tightness θ and the new distribution μ are then both the outcome of an equilibrating process. If there was only a single type of workers t, the tightness θ would adjust until workers are indifferent between entering the new submarket or one of the already existing submarkets, i.e., until $U(w, \theta) = U_t^*$ prevails.[12] With heterogeneous types, we have to jointly determine the tightness θ and the distribution μ in case the offer satisfies $w < y_l - k(r + s)$. If the equilibrium utility of one type t is strictly higher, i.e., if $U_t^* > U_{t'}^*$, then only type t' enters the new submarket and the prevailing tightness θ satisfies $U(w, \theta) = U_{t'}^*$. Note that this implies $U(w, \theta) < U_t^*$. If both types have the same equilibrium utility, i.e., if $U_t^* = U_{t'}^*$ for $t' \neq t$, only the new tightness is still determined uniquely. In this case, it may again be reasonable to choose the distribution μ^0.

If we modified (E′.2) along these lines, our results would not change, i.e., the equilibrium set would not be affected. The intuition is as follows. In our model, the only way to sort between different types is to choose a wage that makes it optimal not to employ low-type workers at all. In particular, if both types are employed under some (new) offer w, both types have the same preferences over w and θ, i.e., over the payoff in a match and the speed of matching. We return to this issue in Section 5.

4.2. Analysis

If contingent contracts are no longer feasible, a separating equilibrium can only be supported if the wage prevailing in the high type's submarket ensures that low-type workers are not employed. Hence, the wage payed to high-type workers must not fall below some threshold w^D that satisfies by (7)

$$w^D := y_l - k(r + s).$$

As $y_l/(r + s)$ equals the expected value of production with a low-type worker, w^D exceeds the low type's wage under contingent contracts w_l^*, i.e., it holds that $w^D > w_l^*$. As a consequence, if the productivity difference $y_h - y_l$ is not too high, it must hold likewise that $w^D > w_h^*$. As w_h^* is chosen efficiently, raising the wage of high-type workers above w_h^* strictly reduces efficiency and, given the zero-profit condition for firms, the utility of high-type workers. By condition (E′.2), an equilibrium will then

[11] Again, the measure of firms opening up the new submarket by offering a different wage must be negligible so as not to influence the composition of existing submarkets.

[12] Of course, this only applies for wages $w > U_t^*$. For completeness, we could specify for offers $w \leqslant U_t^*$ that $\theta = \infty$.

only be separating if the resulting utility of high-type workers does not fall short of what they could realize in a pooling submarket. As pooling implies little losses if productivities are similar, we will only obtain separation if the productivity difference $y_h - y_l$ is sufficiently large.

We next characterize the equilibrium outcome. Subsequently, we discuss in detail the various cases.

Proposition 3. *With non-contingent contracts, there exists a (generically) unique equilibrium. Given a value y_l for the low type's productivity and a value μ^0 for the distribution of worker types, we find two thresholds $0 < \Delta_1 < \Delta_2$ on the productivity difference $y_h - y_l$ such that the following characterization applies*:

(i) *For $y_h - y_l \geqslant \Delta_2$, the equilibrium is identical to the case with contingent contracts.*
(ii) *For $\Delta_1 \leqslant y_h - y_l < \Delta_2$, the two types search in separate submarkets, where the low type receives the wage w_l^* and the high type the wage $w^D > w_h^*$.*
(iii) *For $y_h - y_l \leqslant \Delta_1$, the two types search in the same submarket.*

Proof. See the Appendix.

Take first the case of separating submarkets. For intermediate levels of the productivity difference, high-type workers are paid a higher wage than in the case with contingent contracts. Their wage is increased up to w^D so as to ensure incentive compatibility. We may refer to this wage premium as a kind of efficiency wage. Though various rationals have been proposed in the literature, Proposition 2 adds a new idea. (See Weiss (1980) and Layard et al. (1991) for overviews.) Previous models of efficiency wages under adverse selection have assumed that more able workers simply have a higher reservation value, implying that offering a higher wage allows firms to tap into a more productive pool of workers (see, e.g., Weiss, 1980; Malcomson, 1981).

Consider next case (iii) with low productivity differences. In this case, we find a pooling equilibrium. A pooling equilibrium can be supported as, in our framework, the only possibility to separate low- from high-type workers is to announce a sufficiently high wage, which makes it unprofitable to employ low-type workers. In particular, if both types will be employed for a given wage, they have the same preferences over w and θ. In Section 5, we compare this feature to the settings in Moen (2000) and Lang and Dickens (1992, 1993), which also consider heterogeneous workers.

4.3. Further implications

We next discuss further implications of Proposition 3. In doing so, we are mainly interested in the wage difference between high- and low-productivity types.

By Proposition 3, the wage difference between high- and low-productivity workers is no longer strictly increasing in the productivity difference. To see this, keep y_l fixed and increase y_h. For low differences $y_h - y_l$, all workers are employed at the same wage. As the productivity difference exceeds Δ_1, the wage of high-type workers jumps upwards, while that of low-type workers jumps downwards.[13] As we further increase the productivity difference by raising y_h, the high types' wage stays constant at the separating threshold w^D. The high types' wage starts to rise again as y_h surpasses the threshold Δ_2, at which point the wage difference again strictly increases.

A possibly interesting implication of Proposition 3 is the jump in the wage difference as the productivity difference becomes equal to Δ_1. If workers' skills do not derive primarily from verifiable formal training, a slight increase in the skills of the upper segment of workers may induce separation. This would be accompanied by a non-marginal decrease in wages of low-productivity workers and a non-marginal increase in wages of high-productivity workers. It is again tempting to speculate how this finding relates to the observed increase in the wage difference. According to Levy and Murnane (1992), much of this increase can be attributed to changes within groups, i.e., once most of the observable characteristics such as years of schooling have already been factored out. In our model, already a small increase in the productivity of high types can have large consequences precisely in the case where this productivity change is not related to differences in formal education, years of schooling, or other easy-to-verify characteristics.

According to our model, an increase in the wage difference can also arise if workers' skills become less formal and "hard". For instance, technological change may decrease the importance of formal training and previously acquired "hard" skills. Moreover, if workers are forced to change jobs and even industries more often as the whole economy undergoes profound changes, a worker's history of employment may become less informative about the worker's skills. As the worker's productivity becomes harder to verify "at a distance", our model would predict a sharp increase in the wage difference if $y_h - y_l$ lies between Δ_1 and Δ_2. At first glance, such a jump in the wage difference seems rather counterintuitive, given that firms can no longer offer wages contingent on formal skills. Recall, however, that an efficiency wage premium is paid to high-type workers precisely because of this new restriction.

Finally, if efficiency wages arise in equilibrium, it becomes essential that the firm can commit not to renegotiate the wage offer even if this would be mutually beneficial. In contrast to the case with contingent offers, both the firm offering the efficiency wage and a deviating low type would strictly prefer to renegotiate the wage downwards. This holds as the joint payoff from employment strictly exceeds the sum of the payoffs if the worker is not employed, i.e., the sum of k and U_l^*. Ensuring that the wage is not (re-)negotiated downwards even if this was mutually beneficial may thus require a strong commitment on behalf of firms offering (efficiency) wages to attract high-productivity workers. This may be easier to achieve in large firms where wages are centrally determined and may not be adjusted to individual cases. Similarly, agreements with a recognized unions may serve as a commitment. In a very informal way, we could thus conjecture that large and unionized firms will use their "commitment power" to attract high-productivity workers by paying the respective efficiency-wage premium. Larger and unionized firms would then be both more productive and pay higher wages.

[13] While Proposition 3 does not explicitly state that the wage obtained by low-type workers decreases at Δ_1, this is easily established.

These observations have, indeed, some empirical support. See, e.g., Booth (1995, Chapter 7) for evidence on unionized firms and Mellow (1982) for evidence on firm size effects.

If firms cannot commit not to renegotiate wages downwards in case this is mutually beneficial, we have shown in a previous version of this paper that a separating equilibrium is still feasible if workers' bargaining power (in the renegotiation game) is not too high. In this case, the wage w_h paid to high-type workers must further increase to ensure separation. This has the following intuition. An increase in w_h is again associated with a decrease in the tightness, implying a longer expected search time. While this harms both high-type workers and deviating low-type workers, only high-type workers can reap the gains from higher wages. In fact, the wage paid to deviating low-type workers will be the outcome of re-negotiations and will, therefore, be 18 independent of the initial offer w_h.

4.4. Efficiency

If job offers can be made contingent on workers' types, the equilibrium is efficient. I.e., the flow of vacancies opened up in the separating submarkets for low- and high-productivity workers ensures that the realized flow utility is maximized. For $y_h - y_l < \Delta_2$, this is no longer the case with non-contingent contracts. For intermediate values of the productivity difference, the tightness prevailing in the submarket for high-productivity workers is chosen inefficiently low. In other words, too few vacancies are opened up for workers with high skills. For low values of the productivity difference, only a single submarket opens up for both types. Compared to the efficient benchmark with contingent offers, too many vacancies are opened up for low-type workers and too few vacancies are opened up for high-type workers in this case.

The potential inefficiency in the case of non-contingent contracts raises the question whether and how a regulator could improve welfare. If the regulator acts as a middleman, we find that he can achieve efficiency by specifying two wages w_l, w_h and a subsidy σ. The subsidy is paid to firms who open up a vacancy that pays the wage w_h. We set $w_l = w_l^*$. In case this submarket attracts only low types, the zero-profit requirement for firms implies the tightness θ_l^*. Turn next to the submarket where the wage w_h is offered. If only high-type workers are attracted to this market, free entry of firms ensures that the tightness is equal to θ_h^* if

$$V(w_h, \theta_h^*, 1) = k - \sigma. \tag{8}$$

To ensure that only high-type workers are employed at the wage w_h, it must further hold that

$$y_l - w_h \geqslant (r+s)V(w_h, \theta_h^*, 1). \tag{9}$$

The pair (w_h, σ) is now determined by requiring that (9) holds with equality and that (8) is also satisfied. As is easily checked, these requirements uniquely pin down the wage w_h and the subsidy σ. In fact, substitution yields [14]

$$w_h = y_l - \frac{y_h - y_l}{r}\,\xi(\theta_h^*)^{-1/2} + \frac{c(r+s)}{r}.$$

To pay for the subsidy σ, the regulator could additionally levy a tax on the wage received by high types. Imposing an absolute tax burden of $w_h - w_h^*$ on high types will pay for the flow of subsidies. In this case, workers' utilities are equal to those obtained in the equilibrium with contingent contracts. For $\Delta_1 \leqslant y_h - y_l < \Delta_2$, this also constitutes a Pareto improvement compared to the market equilibrium with non-contingent contracts. While the utility of low-type workers is unchanged, the utility of high-type workers is strictly increased. In case $y_h - y_l < \Delta_1$ holds, we can still easily achieve a Pareto improvement by increasing the tax burden on high types and making a transfer to low types. We have thus obtained the following result.

Proposition 4. *For $y_h - y_l < \Delta_2$, the equilibrium with non-contingent contracts is not constrained efficient. In particular, the first-best outcome can be obtained by a regulator who can pay subsidies for creating vacancies that specify high wages. Additional transfers can then be used to achieve a Pareto improvement compared to the equilibrium with non-contingent contracts.*

4.5. Relation to the literature

We next discuss the relation of our model to two alternative approaches that consider the possibility of sorting between workers with different skills. Consider the case with non-contingent contracts and low productivity differences. According to Proposition 3, all workers search in the same market. A firm opening up a vacancy is not sure whether it will encounter a low- or a high-type worker. Suppose that the first encounter is with a low-type worker. Even though the firm would prefer to employ a high-type worker, it does not pay to wait. Instead, the firm is better off by employing the first applicant it meets. As a consequence, both types of workers face the same expected search time in the pooling submarket. This feature is the main difference to Moen (2000) and Lang and Dickens (1992, 1993), which we discuss next in more detail.

In Moen (2000), a finite number of firms faces a continuum of workers. Each firm has a mass of vacancies and offers a single wage. As there is a continuum of firms and only a finite number of workers, each applicant can perfectly predict the number of other applicants at each firm. Hence, there are no frictions related to search or matching. Moreover, the labor market closes after one period. The productivity of a given applicant depends both on his inherent type and on a random shock. The random component creates a risk of rejection for all workers. However, due to their skill-related productivity component, high-type workers are more likely to have a higher productivity than low-type workers. This creates the following possibility to sort between workers of different types. For a given posted wage, a firm will choose a particular cutoff level for the acceptable productivity of a hired worker. This threshold is increasing in the wage.

[14] Using Assumption 1 and $w_h^* < w^D$, it is easily checked that this wage is strictly positive. Indeed, it holds that $w_h > w_h^*$.

Given their higher probability of realizing a high productivity, high-type workers have more to gain from high wage offers. Using this possibility to sort between workers of different types, Moen finds that an equilibrium is always separating. Separation is achieved by paying workers with high skills a relatively high wage, which implies an inefficiently high probability of being rejected. Moen also refers to this as an efficiency wage premium.

The main difference between our analysis and that of Moen is that, in his model, the randomness of the match quality creates a "continuous" sorting variable. Any increase in the wage benefits high-type workers more (or hurts them less) than low-type workers. In contrast, as the productivity is only determined by workers' known type in our model, sorting is only achieved as the wage passes the threshold w^D, implying that high-type workers are still employed for sure whereas low-type workers are not employed for sure.

The probability of finding employment also becomes a sorting variable in Lang and Dickens (1992, 1993). (Moen (2000) provides a detailed discussion of the differences between his model and theirs.) There, firms make a single wage offer and applicants distribute themselves over firms. The presence of only a finite number of firms and workers can create co-ordination failure and thus frictions in the matching process.[15] Workers differ in their skills. The productivity at each firm depends only on a worker's known skills. As the model is set in discrete time (precisely, the market operates for only one period) and as there is co-ordination failure among applicants, a given firm may, with some probability, attract both high- and low-type workers. In this case, the firm will hire the applicant with the highest skills. Low-type workers will thus face a (much) lower probability of getting hired in case they apply to the same firms as high-type workers. This allows to support a separating equilibrium, in which high-type workers are paid higher wages.

5. Conclusion

This paper explores a competitive search market model with heterogeneous workers. We discuss both the case where wages can be made contingent on workers' types and the case where this is not feasible. This difference may reflect the difference between skills that can be documented by formal training or work experience and skills that are less easy to verify. We derive several implications for the wage difference between high- and low-productivity workers. In particular, high-type workers may receive an "efficiency wage" premium in the case of non-contingent contracts.

A natural limitation of the current analysis is its restriction to two types. Besides some conceptual issues, the extension to a continuum of types is straightforward in our first regime where offers can be made contingent on workers' productivities (Proposition 1). Under non-contingent offers, our results in Proposition 2 can only qualitatively survive with a continuum of types. It can be shown that, in equilibrium, a finite set

[15] More generally, co-ordination failure can arise if the cardinality of the set of workers and firms is the same. Inderst (1999) extends this to the case with homogeneous workers but multiple vacancies.

of submarkets opens up, covering each a convex (and non-overlapping) interval of types. With the exception of the submarket for the lowest interval of types, the wage in all submarkets is inefficiently high, i.e., it exceeds the wage that should be chosen to maximize flow utilities, given the prevailing average productivity.

Acknowledgements

A previous version of this paper was circulated under the title “Efficiency Wages in Competitive Search Markets with Adverse Selection” (first version 1998). I thank Dale Mortensen for drawing my attention to the model’s implication for the wage gap. The paper was improved by suggestions from three referees and the editor Klaus Schmidt.

Appendix

Proof of Propositions 1 and 2. For given μ, we abbreviate the expected profitability by $y^{\mathrm{E}} = \mu y_h + (1-\mu) y_l$. Define for all values $w < y^{\mathrm{E}} - k(r+s)$ the tightness $\theta^{\mathrm{F}}(w,\mu)$ by the requirement that firms realize zero profits, i.e., that $V(w, \theta^{\mathrm{F}}(w,\mu), \mu) = k$ holds. We obtain

$$\theta^{\mathrm{F}}(w,\mu) := \left[\xi \frac{y^{\mathrm{E}} - k(r+s) - w}{(r+s)(rk+c)}\right]^2. \tag{A.1}$$

By the equilibrium requirements (E.1) and (E.2), an equilibrium must be separating. For each type t, the respective submarket characteristics (w, θ) must maximize $U(w,\theta)$ subject to the requirement $V(w,\theta,\mu) = k$, where $\mu = 1$ holds for $t = h$ and $\mu = 0$ holds for $t = l$. Substituting from (A.1), the respective wage must maximize $U(w, \theta^{\mathrm{F}})$, while the respective tightness is equal to θ^{F}. Substitution of θ^{F} obtains

$$U(w,\theta^{\mathrm{F}}) = \frac{1}{r} \frac{(r+s)^2(rk+c)z + w\xi^2[y^{\mathrm{E}} - k(r+s) - w]}{(r+s)^2(rk+c) + \xi^2[y^{\mathrm{E}} - k(r+s) - w]}. \tag{A.2}$$

By Assumption 1, the maximum value of U will exceed z/r for both types. Inspecting (A.2) reveals that we can restrict consideration to values $z/r < w < y^{\mathrm{E}} - k(r+s)$. Uniqueness of a solution follows now from the fact that $U(w,\theta^{\mathrm{F}})$ is strictly quasiconcave over $z/r < w < y^{\mathrm{E}} - k(r+s)$. To see this, note that the sign of the derivative w.r.t. w is determined by the expression

$$D := \xi^2[y^{\mathrm{E}} - k(r+s) - w]^2 + (r+s)^2(rk+c)[y^{\mathrm{E}} - k(r+s) + z - 2w]. \tag{A.3}$$

Strict quasiconcavity follows as D is strictly decreasing in w for all $w < y^{\mathrm{E}} - k(r+s)$.

For what follows, it is convenient to denote the wage maximizing $U(w,\theta^{\mathrm{F}})$ for given μ and thus given y^{E} by w^*. Using the first-order condition, i.e., setting D equal to zero, and implicitly differentiating w.r.t. y^{E}, we obtain

$$\frac{\mathrm{d}w^*}{\mathrm{d}y^{\mathrm{E}}}=\frac{2\xi^2[y^{\mathrm{E}}-k(r+s)-w]+(r+s)^2(rk+c)}{2\xi^2[y^{\mathrm{E}}-k(r+s)-w]+2(r+s)^2(rk+c)}. \tag{A.4}$$

This shows that $0<\mathrm{d}w^*/\mathrm{d}y^{\mathrm{E}}<1$, which implies both $w_h^*>w_l^*$ and $w_h^*-w_l^*<y_h-y_l$. As θ_t^* is equal to the respective value of θ^{F}, it follows also that $\theta_h^*>\theta_l^*$.

We investigate next how w^* changes if the market becomes more efficient. We first analyze changes in the transparency ξ. Implicit differentiation of the first-order condition, i.e., of $D=0$, obtains

$$\frac{\mathrm{d}w^*}{\mathrm{d}(\xi^2)}=\frac{[y^{\mathrm{E}}-k(r+s)-w]^2}{2\xi^2[y^{\mathrm{E}}-k(r+s)-w]+2(r+s)^2(rk+c)}>0. \tag{A.5}$$

Hence, the optimal wage increases for all (expected) productivity levels. Substituting the first-order condition into (A.4), this yields

$$\frac{\mathrm{d}w^*}{\mathrm{d}y^{\mathrm{E}}}=1-\frac{1}{2}\,\frac{y^{\mathrm{E}}-k(r+s)-w^*}{w^*-z},$$

implying that w^* increases faster in the expected productivity if the (optimal) wage w^* is higher. As w^* is by (A.5) strictly increasing in ξ, we have thus shown that the difference $y_h^*-y_l^*$ is strictly increasing in ξ. An analogous argument can be used to show that $y_h^*-y_l^*$ is also strictly decreasing in c. This completes the proof of Propositions 1 and 2. □

Proof of Proposition 3. Observe first from the proof of Propositions 1 and 2 that w_h^* is continuous and strictly increasing in y_h, while it also satisfies $w_h^*>w^D$ for high y_h. Define Δ_2 by the requirement that $w_h^*=w^D$ holds. For $y_h-y_l\geqslant\Delta_2$, the assertion in Proposition 3 is then immediate from Proposition 1. Moreover, any separating submarket for low types must be characterized by w_l^* and θ_l^*. For the rest of the proof we assume that $y_h-y_l<\Delta_2$ holds. We proceed by proving a series of claims.

Claim 1. *If there is separation between high and low types, a single submarket opens up for high types where the wage satisfies $w_h=w^D$.*

Proof. By the observations in the main text, separation requires for the high types' wage $w_h\geqslant w^D$. Moreover, by (E′.2) the choice of w_h must maximize the high types' utility. Given $w_h^*<w^D$ and the strict quasiconcavity of $U(w,\theta^{\mathrm{F}}(w,1))$, which was proven in Proposition 1, there is a unique maximum at $w_h=w^D$. □

In case of separation, denote the respective utility of high types by

$$U^{\mathrm{S}}:=U(w^D,\theta^{\mathrm{F}}(w^D,1)).$$

Consider next the case where pooling submarkets operate.

Claim 2. *If high- and low-type workers search in the same submarket, only a single (pooling) submarket, which is uniquely determined, opens up.*

Proof. Note first that, whenever both types search in the same submarket in equilibrium, they realize the same payoff. By (E′.2), any submarket where the fraction of high types μ^i is smaller than μ^0 can thus not be supported in equilibrium as this would allow a profitable deviation. On the other hand, if there existed a submarket i such that $\mu^i > \mu^0$, there must also exist a submarket with $\mu^i < \mu^0$, which we have already ruled out.[16] Summing up, if there exists a pooling submarket, any submarket i must exhibit the same distribution μ^i, which must therefore be equal to μ^0. Finally, using the notation used to prove Propositions 1–2, (E′.2) implies that the wages offered in these submarkets must maximize $U(w, \theta^{\mathrm{F}}(w, 1))$. As shown previously, this program has a unique solution. □

We denote the workers' utility in a pooling equilibrium by U^{P}. The following assertion completes the proof of Proposition 3.

Claim 3. *Keeping y_l fixed, there exists a unique value Δ_1 satisfying $0 < \Delta_1 < \Delta_2$ such that a pooling equilibrium exists for $y_h - y_l < \Delta_1$ and a separating equilibrium exists for $y_h - y_l > \Delta_1$. Moreover,* for $y_h - y_l = \Delta_1$ *both equilibria exist.*

Proof. By comparing the utilities U^{S} and U^{P}, the following results follow immediately from (E′.2). If $U^{\mathrm{S}} > U^{\mathrm{P}}$ holds, there exists a unique equilibrium, in which workers of different types search in different submarkets. For $U^{\mathrm{S}} < U^{\mathrm{P}}$, there is a unique pooling equilibrium. And for the non-generic case where $U^{\mathrm{P}} = U^{\mathrm{S}}$ holds, both equilibria exist.

Note next that the difference $U^{\mathrm{P}} - U^{\mathrm{S}}$ is continuous in y_h, strictly negative at $y_h - y_l = 0$, and strictly positive at $y_h - y_l = \Delta_2$. We show that $U^{\mathrm{P}} - U^{\mathrm{S}}$ is strictly increasing, which proves existence of the asserted threshold Δ_1. Before providing a formal proof, note first that the strict monotonicity is intuitive. An increase in the high type's productivity y_h should increase the utility under separation U^{S} more than the utility in the pooling case U^{P}.

Recall that

$$U^{\mathrm{P}} - U^{\mathrm{S}} = \max_w U(w, \theta^{\mathrm{F}}(w, \mu^0)) - U(w^D, \theta^{\mathrm{F}}(w^D, 1)).$$

Note next that the derivative $\mathrm{d}[\max_w U(w, \theta^{\mathrm{F}}(w, \mu^0))]/\mathrm{d}y_h$ is strictly smaller than the derivative $\mathrm{d}[\max_w U(w, \theta^{\mathrm{F}}(w, 1))]/\mathrm{d}y_h$, where we have replaced μ^0 by the distribution $\mu = 1$. Hence, it remains to show that

$$\frac{\mathrm{d}[\max_w U(w, \theta^{\mathrm{F}}(w, 1))]}{\mathrm{d}y_h} < \frac{\mathrm{d}[U(w^D, \theta^{\mathrm{F}}(w^D, 1))]}{\mathrm{d}y_h} \tag{A.6}$$

[16] Note that the same rate of breakdown s applies to both matches of high- and of low-type workers.

holds over $y_l \leqslant y_h < y_l + \Delta_2$. This follows as (A.6) transforms to the requirement

$$\frac{w_h^* - z}{[(r+s)^2(rk+c) + \xi^2[y_h - w_h^* - k(r+s)]]^2} < \frac{w^D - z}{[(r+s)^2(rk+c) + \xi^2[y_h - w^D - k(r+s)]]^2},$$

which holds after substituting $w_h^* < w^D$ from $y_h - y_l < \Delta_2$. □

References

Acemoglu, D., Shimer, R., 1999. Holdups and efficiency with search frictions. International Economic Review 40, 827–849.

Booth, A., 1995. The Economics of the Trade Union. Cambridge University Press, Cambridge.

Inderst, R., 1999. On the foundation of competitive equilibrium with strategic buyers. Mimeo.

Inderst, R., 2000. Complementarities and assignment in complete market with frictions. Mimeo.

Inderst, R., Müller, H., 2002. Competitive search markets for durable goods. Economic Theory 19, 599–622.

Lang, K., Dickens, W.T., 1992. Labor market segmentation, wage dispersion and unemployment. NBER Working paper 4073.

Lang, K., Dickens, W.T., 1993. Bilateral search as an explanation for labor market segmentation and other anomalies. NBER Working paper 4461.

Layard, N., Nickell, S., Jackman, R., 1991. Unemployment: Macroeconomic Performance and the Labour Market. Oxford University Press, Oxford.

Levy, F., Murnane, R., 1992. U.S. Earnings levels and earnings inequality: A review of recent trends and proposed explanations. Journal of Economic Literature 30, 1333–1381.

Malcomson, J., 1981. Unemployment and the efficiency wage hypothesis. The Economic Journal 91, 848–866.

McAfee, P., 1993. Mechanism design by competing sellers. Econometrica 61, 1281–1312.

Mellow, W., 1982. Employer size and wages. Review of Economics and Statistics 64, 495–501.

Moen, E., 1997. Competitive search equilibrium. Journal of Political Economy 105, 385–411.

Moen, E., 2000. Do good workers hurt bad workers—or is it the other way around? Mimeo., Norwegian School of Management.

Mortensen, D., Pissarides, C., 1998. New developments in models of search in the labor market. Mimeo.

Nickell, S.J., Nunziata, L., Ochel, W., Quintini, G., 2002. The Beveridge curve, unemployment and wages in the OECD. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps. Princeton University Press, Princeton, NJ.

Peters, M., Severinov, S., 1997. Competition among sellers who offer auctions instead of prices. Journal of Economic Theory 75, 141–179.

Pissarides, C., 1990. Equilibrium Unemployment Theory. Basil Blackwell, Cambridge, MA.

Saint-Paul, G., 1996. Dual Labor Markets. MIT Press, Cambridge, MA.

Weiss, A., 1980. Job queues in the labor markets with flexible wages. Journal of Political Economics 88, 526–539.