On-the-job search and moral hazard.*

Espen R. Moen† and Åsa Rosén‡

August 10, 2012

Abstract

We analyze the interaction between intertemporal incentive contracts and search frictions associated with on-the-job search. In our model, agency problems call for wage contracts with deferred compensation. At the same time workers do on-the-job search.

Deferred compensation improves workers’ incentives to exert effort but distorts their

---

*We would like to thank editor George-Marios Angeletos and three anonymous referees for valuable comments. We have also benefited from comments by Denis Gromb, Edgar Preugschat, Ketil Storsletten, Christopher Taber, and seminar participants at the University of Oslo, Georgetown University, Northwestern University, the University of Vienna, the 2011 SAM conference and the 2011 SED conference. We would also like to thank Øyvind Nilsen Aas for valuable research assistance. Financial support from the Swedish Research Council and the Norwegian Research Council are gratefully acknowledged. This paper is part of the research activities at the Centre of Equality, Social Organization, and Performance (ESOP) at the Department of Economics at the University of Oslo. ESOP is supported by the Research Council of Norway.

†Norwegian School of Management (BI) and CEPR. Box 580 N-1301 Sandvika, Norway. E-mail: espen.moen@bi.no.

‡Stockholm University (SOFI) and Oslo University (ESOP). Stockholm University, S106-91 Stockholm, Sweden. E-mail: asa.rosen@sofi.su.se.
on-the-job search decisions. We show that deferred compensation is less attractive when the value to the worker-firm pair of on-the-job search is high. Moreover, the interplay between search frictions and wage contracts creates feedback effects. If firms in equilibrium use contracts with deferred compensation, fewer firms with vacancies enter the on-the-job search market, and this in turn reduces the distortions created by deferred compensation. These feedback effects between the incentive contracts used and the activity level in the search markets can lead to multiple equilibria: a low-turnover equilibrium where firms use deferred compensation, and a high-turnover equilibrium where they do not. Furthermore, the model predicts that firms are more likely to use deferred compensation when search frictions are high and when the gains from on-the-job search are small.

**Key words:** On-the-job search, Moral Hazard, Deferred Compensation, Multiple Equilibria.

**JEL Codes:** J41, J63.
1 Introduction

There is broad evidence that on-the-job search and turnover are important features of the labor market. In addition, several studies indicate that wage contracts often have an intertemporal element, and that firms use deferred compensation to avoid agency problems and to motivate their workers. This paper studies the interplay between on-the-job search and deferred compensation and finds that multiplier effects can lead to multiple equilibria.

Our starting point is that reallocation of workers across firms is necessary to obtain an efficient allocation of resources, since experienced workers may have a comparative advantage in different tasks and different firms than inexperienced workers. To capture this, we set up an on-the-job search model of the Diamond-Mortensen-Pissarides type (Diamond, 1982; Mortensen, 1986; Pissarides, 1985). Workers live for two periods and search on the job between the periods. Efficient on-the-job search then requires that the senior workers’ wages in the original firm equal their productivity in that firm (as in Moen and Rosén, 2004). If wages are deferred, in the sense that the wages of senior workers exceed their productivity, workers will do too little on-the-job search. Since the workers’ participation constraint binds at the hiring stage, this cost is carried by the firm.

In addition, firms face agency problems created by imperfect monitoring. As in Shapiro and Stiglitz (1984), firms can only imperfectly monitor worker effort. We follow Lazear (1979, 1981) and allow firms to use deferred compensation to provide incentives for workers to exert effort. However, deferred compensation implies that a senior worker’s wage is above his or her productivity, so on-the-job search is distorted. We show that there are feedback effects between optimal wage contracts and the number of firms searching for employed workers.
Deferred compensation increases workers’ wage claims when bargaining with a new employer. In response, fewer firms enter the search market. Hence activity in the on-the-job search market falls, and the loss to the worker firm pair induced by distorted search falls. These feedback effects can lead to multiple equilibria: a high-effort /low-turnover equilibrium, in which firms use deferred compensation to motivate workers, and a low-effort high-turnover equilibrium, in which they do not.

We also characterize the parameter sets for which the two equilibria exist. We find that a high-effort/low-turnover equilibrium with deferred compensation is likely to exist when search frictions are high and productivity differences between firms are small.

Our model is relevant to industries where i) on-the-job search plays an important role in obtaining an efficient allocation of workers across firms and ii) deferred compensation is important to induce workers to exert effort. It seems that turnover is important economy-wide, as indicated by high turnover rates. Recent studies, such as that of Jolivet, Postel-Vinay and Robin (2006), document transitions from employer- to- employer, from employment - to -non-employment, and from non-employment to employment in 10 European countries and the U.S. The authors find that job-to-job transitions are of a similar magnitude to employment- to- non-employment (but vary across countries). Nagypál (2008) finds that for the U.S. labor market, about half of all separations from employers between 1994 and 2007 were to a new employer. Fallick and Fleischman (2004) find that nearly 40% of new jobs started between 1994 and 2003 were employer changes. Furthermore, the authors find that on-the-job search is of similar importance to off-the-job search in explaining the flow from employment to employment and from non-employment to employment. Jolivet et al (2006) structurally estimate a prototype on-the-job search model and find that it fits employment
and transition data reasonably well. The importance of worker turnover for economic efficiency is more difficult to assess. However, using a structural growth model with worker turnover, Lentz and Mortensen (2004) find that 53% of the growth in Danish manufacturing can be attributed to reallocation of workers from low- to high- productivity firms.

Several empirical studies provide evidence of deferred compensation. Medoff and Abraham (1980) find that pay increases with seniority, even when supervisors’ ratings of performance do not increase. Lazear and Moore (1984) compare age-income profiles for tenured and self-employed workers, for whom there exist no agency problems. The authors find that the returns to seniority are higher for tenured workers, and attribute this to deferred compensation. Kotlikoff and Gokhale (1992) compare the wages and productivity of more than 300,000 workers in a Fortune 1000 firm and find a substantial degree of deferred compensation for all categories of workers. In particular, managers’ productivity exceeds compensation by a factor of more than two at the age of 35, while the opposite is true at the age of 57. Barth (1997) documents that workers on piece-rate compensation schemes have negligible returns to seniority, while workers who are not paid piece-rate earn significant returns to seniority. Using personnel data from a large Italian firm, Flabbi and Ichino (2001) find that the effects of seniority on wages do not reflect higher productivity.

From a theoretical point of view, our paper introduces moral hazard into a model of on-the-job search. There is currently a small but growing literature on private information in search models. Moen and Rosén (2011) and Guerrieri (2008) introduce asymmetric information in competitive search equilibrium. Guerrieri, Shimer, and Wright (2010) analyze self-selection of heterogeneous workers in a search environment, and Rudanko (2009) and Menzio and Moen (2010) analyze optimal insurance with limited commitment in a search
context. We contribute to this literature by analyzing the relationship between (inter tempo ral) wage contracts and on-the-job search.

Also related are extensions of the Burdett-Mortensen model (Burdett and Mortensen, 1998), which allows for back-loading of wages, as in Burdett and Coles (2003) and Stevens (2004). The mechanism at play in these papers diverges from ours. In their models, search is inefficient, since it reduces the joint income of the incumbent firm and the employee. The employer discourages job quits by back-loading wages (but never to the extent that the wage is higher than the output). In our model, by contrast, on-the-job search is warranted, since it increases the joint value of the incumbent firm and the employee. Back-loading is used to motivate workers to exert effort and implies that wages for senior workers exceed output. Reduced on-the-job search then becomes a costly and unintended by product of such back-loading.

Our paper proposes a new explanation for differences in turnover between countries and between regions within a country (see Saxenian 1994 for documentation of the latter). Other explanations include differences in institutions, as in, for example Pries and Rogerson (2005), who argue that the differences in worker turnover between the U.S. and Europe can be explained by institutional factors. There also exist papers that analyze multiple equilibrium turnover rates. Acemoglu and Pischke (1998) develop a model where adverse selection may lead to multiplicity in quit rates. Related arguments are made in Chang and Wang (1995), Owan (2004), Saint-Paul (1995), and Moene and Wallerstein (1997). Morita (2001) shows how multiple turnover rates can arise as a result of firms’ choice of production technology and learning- by- doing. Our paper differs from this literature in several ways. First, in our model multiplicity is caused by interactions between the search market and incentive
contracts to mitigate worker moral hazard. Second, our paper is the only one that explicitly models on-the-job search as an equilibrium outcome in the presence of search frictions.

The paper is organized as follows. Section 2 sets up the model and defines equilibrium. Section 3 characterizes equilibrium, while Section 4 analyzes multiple equilibria. Section 5 studies additional implications of our model. Section 6 concludes.

2 Model and equilibrium

We study an overlapping generations model where workers live for two periods. Since there is no interaction between the generations, each generation can be studied in isolation. The economy consists of two types of firms: ordinary firms and specialized firms. All workers start their career in ordinary firms. After the first period they qualify for a job in a specialized firm, where their productivity is higher.\(^1\) All agents are risk neutral with a zero discount rate. The labor market for jobs in ordinary firms is Walrasian. The labor market for specialized jobs contains search frictions. Entry of firms in both markets implies zero profits in equilibrium.

In an ordinary firm, the first period can be divided into four stages, the contracting stage, the production stage, the remuneration stage, and the search stage. At the production stage, young workers choose effort level \(e\) and produce \(y_1 + e\) units of output, where \(e \in \{0, \bar{e}\}\) is

\(^1\)Several arguments support that turnover can be efficient. Workers may try out several jobs to determine their comparative advantage (Johnson, 1978) or because of match-specific productivity differences (Jovanovic, 1979). A worker’s relative productivity in different firms can also change over time as experience and expertise are gained (Moen and Rosén, 2004). Furthermore, sectorial shocks to the economy may warrant a re-allocation of workers. Finally, with technological progress, efficient dissemination of knowledge may require turnover, since workers learn from each other (Saxenian, 1994).
the effort level. The cost of effort is $ec, c \in (0, 1)$. We introduce a moral hazard problem that may call for deferred compensation, and we do so as simply as possible, by assuming a workers’s effort level can only be observed in the second period.\(^2\) At the contracting stage, ordinary firms offer the young workers a wage schedule $\omega = \{w_1, w_2(e)\}$, where $w_1 \in R$ denotes the wage in the first period and $w_2(e) : \{0, \tau\} \rightarrow R$ is the wage in the second period, given that the worker is still employed in that ordinary firm.\(^3\) At the remuneration stage, the young workers are paid their wages $w_1$ as specified in the contract. At the search stage, the young workers search for specialized jobs. If on-the-job search is successful, the worker leaves. If staying on in the ordinary firm in period 2, the worker produces $y_2$ in that period without any moral hazard problems and receives the contracted wage $w_2(e)$.

Worker search for a job in a specialized firm is costly. A search intensity $s$ implies an effort cost $\gamma s^2/2$. Specialist firms enter the search market at a cost $K$, which must be repeated each period the firm searches for workers. A worker produces $y_p > y_2$ in a specialized firm without any moral hazard problems, and wages $w_p$ are set by bargaining.\(^4\)

\(^2\)Deferred compensation may also be warranted if effort is observed the same period but with noise, see Moen and Rosén (2012).

\(^3\)We assume that firms do not pay workers who have left the firm. This may be because of the following reasons: a) It is hard to verify whether movers expended high effort in the first period. b) A firm’s reputation may suffer more from breaking a contract if the worker is still employed than if the worker has quit. c) Deferred compensation may reflect the (expected) gain from promotions and, as argued in Carmichael (1983), it may be easier for a firm to commit to promotions than to cash payments not associated with particular positions. e) It may be easier for a worker to retaliate in informal ways after a breach of contract if he is still employed than if he works in another firm.

\(^4\)The main results of this paper also hold under the less restrictive assumption that only some, rather than all, workers have a higher productivity in specialized firms, and that only those workers engage in search.
Matching takes place between the periods. The number of matches between searching workers and specialized firms is determined by a constant return to scale matching function \( x(sn, v) \), where \( n \) is the measure of searching workers, \( s \) is their average search intensity, and \( v \) is the measure of vacancies posted by specialized firms. We assume that the matching function is Cobb-Douglas, that is, \( x(sn, v) = A(sn)^\beta v^{1-\beta} \). Let \( p \) denote the probability of finding a job per unit of search intensity and \( q \) the expected number of applicants to a firm. It follows that

\[
p(\theta) = A\theta^{1-\beta}, \tag{1}
\]

\[
q(\theta) = A\theta^{-\beta}, \tag{2}
\]

where \( \theta = v/sn \).\(^5\) The probability of finding a job for a worker with search intensity \( s \) is \( sp \).

In equilibrium we require that \( sp \leq 1 \).\(^6\) For technical reasons, we allow \( q \) to be greater than one (hence a firm can attract and hire more than one worker).\(^7\)

Since the market for ordinary firms is competitive, an old worker can always be assured a wage equal to \( y_2 \). To retain an old worker who has not obtained a job offer in a specialized

---

\(^5\)One may think of our matching process as a reduced form of a matching process set in continuous time. The probability of finding a job may then be interpreted more broadly as a fraction of the available time the worker spends in the specialized firm.

\(^6\)Albrecht, Gautier, and Vroman (2006) note that coordination externalities associated with multiple applications can arise in a discrete setting if workers obtain more than one offer and must choose between them. However, the coordination externality disappears if the matching processes are set in continuous time, or if firms can make job offers to more applicants if the first applicant(s) turns down the offer (Kircher, 2009).

\(^7\)The alternative is to impose the constraint that \( x(u, v) \leq v \). While clearly doable, this is inconvenient for the existence and efficiency proofs.
firm, the wage contract must specify a wage \( w_2(e) \geq y_2 \). We refer to this as the worker’s interim participation constraint. As we will see shortly, the interim participation constraint does not bind if the worker exerts effort. If the worker does not exert effort, the firm is indifferent to retaining the worker at wage \( w_2 = y_2 \) or letting the worker go. We assume without loss of generality that the wage schedule also in this case satisfies the worker’s interim participation constraint. Thus, the expected utility of a worker is,

\[
u(\omega, e, s) = w_1 - ec - \gamma s^2 / 2 + spw_p + (1 - sp)w_2(e).
\]

The profit of an ordinary firm is

\[
\pi(\omega, e, s) = y_1 + e - w_1 + (1 - sp)(y_2 - w_2(e)).
\]

Wages in specialized firms are determined by Nash bargaining. The threat point of the worker is the second-period wage in ordinary firms. To avoid uninteresting technicalities, we assume that individual wages are unobservable by the specialized firm, which only knows the distribution of wages in the economy.\(^8\) Since we only consider pure strategy equilibria, all workers in equilibrium have the same fallback wage \( w_2(e^*) \), where \( e^* \) is the equilibrium effort level and constitutes the workers’ disagreement point. The disagreement point of the firms is zero. We assume that the Hosios condition is satisfied (Hosios, 1990), that is, wages are given by the Nash sharing rule, with the worker’s bargaining power equal to \( \beta \). Appendix F shows that the on-the-job search market maximizes the utility of searching workers subject

\(^8\)If \( w_2 \) were observable, this would imply that the current employer could jack up the wage and thereby the value of search for the employees. As shown by Shimer (2006), this can lead to an untractable equilibrium distribution of wages.
to the zero profit condition of firms. The wage in a specialized firm is given by

$$w_p = \beta y_p + (1 - \beta)w_2(e^*).$$  \hspace{1cm} (5)

The expected income to a specialized firm is

$$V = q(y_p - w_p)$$

$$= K,$$ \hspace{1cm} (6)

where the last equation follows from entry. From (1) and (2) it follows that $q = A^{\frac{1}{1-\beta}}p^{\frac{\beta}{1-\beta}}$. Substituting $q = A^{\frac{1}{1-\beta}}p^{\frac{\beta}{1-\beta}}$ and (5) into the zero profit condition (6) gives

$$p = A^{\frac{1}{\beta}}\left[\frac{(1 - \beta)(y_p - w_2(e^*))}{K}\right]^{\frac{1-\beta}{\beta}},$$ \hspace{1cm} (7)

which uniquely pins down $p$ as a function of $w_2$.

To ensure that the market for specialized firms is operating both with and without deferred compensation, we must make assumptions about the productivity differential between specialized firms and ordinary firms relative to the cost of opening a vacancy. More specifically, we make the following assumption.

**Assumption 1**

$$y_p - y_2 - \tau c > 0.$$ 

In addition, we must make parameter assumptions to ensure that the probability of finding a job, $sp$, is strictly lower than 1 in equilibrium.

**Assumption 2**

$$\gamma > A^{\frac{1}{\beta}}\left[\frac{(1 - \beta)(y_p - y_2)}{K}\right]^{\frac{2(1-\beta)}{\beta}}\beta(y_p - y_2).$$
Why the assumptions on the parameters take exactly these forms will be clear in the proofs of Lemmas 3 and 4.

Let \( u \) denote the expected utility of a young worker who enters the market. The optimal contract can be defined as follows:

**Definition 1** For given values of \( u, s, p, \) and \( w_p \), the optimal contract \((\tilde{\omega}, \tilde{e}, \tilde{s})\) is a wage schedule \( \tilde{\omega} = \{\tilde{w}_1, \tilde{w}_2(e)\} \), an effort level \( \tilde{e} \), and a search intensity \( \tilde{s} \) that solve

\[
\max \pi(\omega, e, s) \quad \text{subject to}
\]

1. **Incentive compatibility**

   \[
u(\tilde{\omega}, \tilde{e}, \tilde{s}) = \max_{e, s} u(\tilde{\omega}, e),
\]

2. **Interim participation**

   \[
   \tilde{w}_2(e) \geq \gamma_2, \quad e \in \{0, \tilde{e}\},
   \]

3. **Participation**

   \[
u(\tilde{\omega}, \tilde{e}, \tilde{s}) \geq u.
\]

We are now ready to define the equilibrium.

**Definition 2** The equilibrium is a contract \((\omega^*, e^*, s^*)\), a job finding rate \( p^* \), a wage \( w_p^* \) and a utility \( \overline{u}^* \) such that the following applies.

1. **The contract** \((\omega^*, e^*, s^*)\) is an optimal contract.

2. **Equilibrium in the search market**: \( w_p^* \) and \( p^* \) solve (5) and (7).

3. **Zero profit of ordinary firms**: \( \pi(\omega^*, e^*, s^*) = 0 \).
4. The equilibrium utility $\bar{u}$ is given by (3)

3 Characterizing equilibrium

Since the participation constraint $u \geq \bar{u}$ is binding, it follows from (3) and (4) that

$$
\pi = y_1 + y_2 + e(1 - c) + sp(w_p - y_2) - \gamma s^2/2 - \bar{u} = y_1 + y_2 + e(1 - c) + \Omega(s) - \bar{u},
$$

(8)

where

$$
\Omega(s) = sp(w_p - y_2) - \gamma s^2/2,
$$

(9)

is the value of search (the functional dependence on $p$ and $w_p$ is suppressed). Define

$\Omega_{\text{max}} = \max_s \Omega(s)$ and let $s_{\text{max}}$ denote the corresponding value of $s$. Then,

$$
\begin{align*}
    s_{\text{max}} &= \frac{p(w_p - y_2)}{\gamma}, \quad (10) \\
    \Omega_{\text{max}} &= \frac{p^2(w_p - y_2)^2}{2\gamma}. \quad (11)
\end{align*}
$$

Consider a worker’s search behavior. Incentive compatibility requires that $\tilde{s}$ maximizes $u(\tilde{w}, \tilde{e}, s)$, and from (3) the first order condition of this maximization problem reads

$$
\tilde{s} = \frac{p(w_p - w_2(\tilde{e}))}{\gamma}. \quad (12)
$$

By comparing (12) and (10) it follows that the worker maximizes the value of search $\Omega$ if and only if $w_2(\tilde{e}) = y_2$, in which case there is no externality on the firm from the worker’s search behavior. Define

$$
L = \Omega_{\text{max}} - \Omega(\tilde{s}). \quad (13)
$$
We refer to \( L \) as the *(deadweight) loss associated with inefficient search intensity* when \( w_2(\bar{e}) \neq y_2 \). The profit function (8) can now be written as

\[
\pi = y_1 + y_2 + e(1 - c) + \Omega^{\max} - L - \pi. \tag{14}
\]

Let \( D \equiv w_2(\bar{e}) - y_2 \) denote the *amount of deferred compensation* the worker receives. If the firm implements effort, \( D > 0 \).

**Lemma 1** *The loss \( L \) is a function of \( D \) and \( p \),

\[
L(D, p) = \frac{D^2p^2}{2\gamma}. \tag{15}
\]

**Proof.** See Appendix A □

The loss is increasing in the amount of deferred payment \( D \) and tightness \( p \) in the search market. The higher the value of \( D \), the further away the worker’s search intensity is from that which maximizes the value of search. The higher the value of \( p \), the more valuable the search intensity is, and the more it matters when the search intensity is too low. Note that the loss is independent of \( w_p \).

If the firm wants to implement high effort, \( \bar{e} = \bar{\tau} \), the contract must satisfy the incentive compatibility constraint \( u(\omega, \bar{e}, \bar{s}) \geq \max_s u(\omega, 0, s) \). We define a shirker as a worker who deviates and sets \( e = 0 \) when the contract prescribes \( e = \bar{e} \). The contract punishes shirkers as hard as possible, hence the interim participation constraint binds; \( w_2(0) = y_2 \). Let \( \overline{D} = w_2(\bar{e}) - y_2 \) denote the lowest amount of deferred compensation consistent with the incentive compatibility constraint for \( e = \bar{e} \). The following then holds.

**Lemma 2** *\( \overline{D} = \overline{D}(p, w_p; c) \) is implicitly defined by the expression

\[
\overline{D}p^2[\overline{D} - 2(w_p - y_2)] + 2\gamma\overline{D} = 2\gamma\bar{\tau}c. \tag{16}
\]

14
Furthermore, \( \bar{D} \) is strictly increasing in \( p, w_p \) and \( c \).

**Proof.** See Appendix B. ■

The value of \( \bar{D} \) must be high when the probability that the worker is there to pick up the bonus is low, that is, if \( p \) is high or \( w_p \) is high (since a high \( w_p \) means a high \( \bar{s} \)). If the cost of effort, \( c \), is high, \( D \) must also be high.

The structure of the equilibrium depends on whether firms implement effort or not. It is convenient to define two equilibrium candidates, a no-effort equilibrium candidate \((\omega^n, 0, s^n, p^n, w^n_p, \bar{\pi}^n)\) and an effort equilibrium candidate \((\omega^e, \bar{e}, s^e, p^e, w^e_p, \bar{\pi}^e)\). The no-effort (effort) candidate is defined in the same way as the equilibrium, with the restriction that the firms are forced to implement \( e = 0 \) \( (\bar{e} = \bar{e}) \).

Consider first a no-effort equilibrium candidate. The firms set \( w^n_2(\bar{\pi}) = w^n_2(0) = y_2 \) \( (D = 0) \), since this maximizes the value of search \( \Omega \). The zero profit constraint and (4) implies that \( w^n_1 = y_1 \). From (5) and (7), it follows that

\[
\begin{align*}
    w^n_p & = y_2 + \beta(y_p - y_2), \\
    p^n & = A^n(1 - \beta)(y_p - y_2)K^{1-\beta}. \\
\end{align*}
\]

(17)

(18)

Since the right-hand sides only contain exogenous variables, the existence and uniqueness of \( w^n_p \) and \( p^n \) follows directly from equations (17) and (18). Given \( w^n_p \) and \( p^n \), equations (10) and (11) uniquely determine \( s^n = s^{\text{max}}(p^n, w^n_p) \) and \( \Omega = \Omega^{\text{max}}(p^n, w^n_p) \), and the loss, \( L \), is zero.

From (14) and the zero profit condition, we have

\[
\bar{\pi}^n = y_1 + y_2 + \Omega^{\text{max}}(p^n, w^n_p),
\]

(19)
which uniquely determines $\bar{w}^n$.

**Lemma 3** The no-effort equilibrium candidate exists and is unique. The solution is such that $p^n > 0$ and $s^n p^n < 1$.

**Proof.** The first part of the Lemma (existence and uniqueness) was proved above. The last part is proved in Appendix C. ■

For the no-effort equilibrium candidate to constitute an equilibrium for the full model, firms cannot find it profitable to implement effort, that is, $\pi(\omega^n, 0, s^n) \geq \max_{\omega, s} \pi(\omega, \bar{e}, s)$ given the market parameters $(p^n, w^n, \bar{w}^n)$. If the firm implements effort, recall that it sets $w_2(0) = y_2$ and $w_2(\bar{e}) = y_2 + D(p^n, w^n_p)$. A deviating firm thus obtains a profit given by (from (14))

$$\pi = y_1 + y_2 + \bar{e}(1 - c) + \Omega^{\text{max}}(p^n, w^n_p) - L(D(p^n, w^n_p), p^n) - \bar{w}^n. \quad (20)$$

For the deviation to be strictly profitable, $\pi$ must be strictly positive. By inserting (19) into (20), it follows that the no-effort equilibrium candidate constitutes an equilibrium in the full model if and only if

$$L(D(p^n, w^n_p), p^n) \geq \bar{e}(1 - c), \quad (21)$$

that is, if the loss $L$ exceeds the gain from effort. If (21) is satisfied, we say that a *no-effort equilibrium* exists.
Consider then an effort equilibrium candidate. Firms set \( w_2(0) = y_2 \) and \( w_2^e(\bar{\tau}) = y_2 + D(p^e, w_p^e) \). From (5) and (7), it then follows that

\[
\begin{align*}
  w_p^e &= \beta y_p + (1 - \beta)(y_2 + \bar{D}^e), \\
  p^e &= A \frac{\log[(1 - \beta)(y_p - y_2 - \bar{D}^e)] - \beta}{K}, \\
  \bar{D}^e &= D(p^e, w_p^e).
\end{align*}
\]

Appendix D shows that (22) - (24) define \( w_p^e, p^e, \) and \( \bar{D}^e \) uniquely and that the solution is such that \( p^e > 0 \) and \( s^e p^e < 1 \). It follows that \( w_2^e(\bar{\tau}) = y_2 + \bar{D}^e \) is uniquely determined. Hence (10) and (11) uniquely determine \( s^e \) and \( \Omega^{\max} \).

From the profit equation (8) and the definition of the loss function (13) it follows that

\[
\pi^e(\bar{\tau}) = y_1 + y_2 + \bar{\tau}(1 - c) + \Omega^{\max}(p^e, w_p^e) - L(D(p^e, w_p^e), p^e) - \pi^e. \tag{25}
\]

Zero profits yields

\[
\bar{\pi}^e = y_1 + y_2 + \bar{\tau}(1 - c) + \Omega^{\max}(p^e, w_p^e) - L(D(p^e, w_p^e), p^e), \tag{26}
\]

which uniquely defines \( \bar{\pi}^e \). Equation (3) then defines \( w_1^e \) uniquely.

**Lemma 4** The effort equilibrium candidate exists and is unique. The solution is such that \( p^e > 0 \) and \( s^e p^e < 1 \).

For the effort equilibrium candidate to constitute an equilibrium in the full model, the firms cannot find it profitable to deviate and not implement effort. A necessary and sufficient condition for this is that \( \pi(\omega^e, \bar{\tau}, s^e) \geq \max_{\omega, s} \pi(\omega, 0, s) \), given the market parameters \((p^e, w_p^e, \bar{\pi}^e)\). A deviating firm that implements zero effort and sets \( D = 0 \) obtains a loss of
zero and a profit given by (from (14))

\[ \pi = y_1 + y_2 + \Omega_{\max}(p^e, w_p^e) - \bar{u}^e. \] (27)

Deviation is only strictly profitable if \( \pi \) is strictly positive. By substituting out \( \bar{u}^e \) in (27) by virtue of (26), it follows that the effort equilibrium candidate is an equilibrium of the full model if and only if

\[ L(D(p^e, w_p^e), p^e) \leq \bar{e}(1 - c). \] (28)

The equation states that the loss, \( L \), in the value of search due to deferred compensation is outweighed by the gain from effort. If (28) is satisfied, we say that an effort equilibrium exists.

**Proposition 1**  

a) There exists a threshold value \( c^a \) such that the no-effort equilibrium exists if and only if \( c \geq c^a \).

b) There exists a threshold value \( c^e \) such that the effort equilibrium exists if \( c \leq c^e \).

**Proof.** See Appendix E. ■

Note the difference between a) and b). The threshold \( c^a \) is the unique solution to (21) with equality, and the no-effort equilibrium exists if and only if \( c \geq c^a \). However, (28) with equality may have more than one solution. We have therefore defined \( c^e \) as the smallest solution to (28) with equality. Due to continuity, a smallest such value always exists. It follows that the effort equilibrium may also exist for values above \( c^e \) (hence there is no "only if"). The reason why (28) with equality does not necessarily have a unique solution is that a higher \( c \) implies a higher \( D \) (which increases \( L \)) as well as a lower \( p^e \), which tends to reduce \( L \). In Moen and Rosén (2012) we show that a sufficient condition for (28) with
equality having a unique solution (in terms of endogenous variables, unfortunately) is that
\[ \overline{D} \leq \beta(y_p - y_2)/(1 - \beta). \]

**FIGURE 1**

We illustrate the different equilibria in Figure 1. Appendix D shows that the generic expressions (22) and (24) define \( \overline{D} \) as a function of \( p \) only, \( \overline{D} = \overline{D}(p) \), and that
\[ \frac{d\overline{D}(p)}{dp} > 0. \] \hspace{1cm} (29)

This means that along a zero profit curve \((p, w_p(p))\), \( D \) is increasing in \( p \). Second, from the zero profit condition (7), we can write \( p \) as a function of \( \overline{D} \):
\[ p = p^{FE}(\overline{D}), \quad \frac{dp^{FE}(\overline{D})}{d\overline{D}} < 0. \] \hspace{1cm} (30)

The curve shows the job finding rate \( p \) that is consistent with the zero profit condition for specialized firms (or free entry condition) as a function of \( \overline{D} \). The higher the deferred
compensation is, the higher the Nash bargaining wage $w_p$ is. A high $w_p$ in turn leads to a lower entry and a lower $p$.

In the $p - D$ space, the effort equilibrium is obtained at the intersection of the two curves. The no-effort equilibrium is defined by $D = 0$ and $p^n = p^{FE}(0)$. The two equilibria are shown in Figure 1.

4 Multiple equilibria

The previous Section derives conditions under which effort and no-effort equilibria exist. An interesting issue is whether these can exist simultaneously. As the next Proposition shows, the answer is confirmatory:

**Proposition 2** The threshold values are such that $c^n < c^e$. The model exhibits multiple equilibria whenever $c \in [c^n, c^e]$. One equilibrium is characterized by high effort, low $p$, and deferred compensation, while the other is characterized by no effort, high $p$, and no deferred compensation.

**Proof.** See Appendix G. ■

FIGURE 2

The Proposition is illustrated in Figure 2. The no-effort equilibrium exists for $c \geq c^n$. The effort equilibrium exists for $c \leq c^e$. Multiple equilibria exist if $c \in [c^n, c^e]$. 
The intuition for multiplicity is as follows. Suppose we are in the no-effort equilibrium. Then \( w_2 \) is relatively low, since no firms defer compensation. Therefore many specialized firms enter the market and on-the-job search is valuable. If a firm deviates and defers compensation to implement high effort, this comes at a high cost (\( L \) is high). By contrast, in the effort equilibrium all firms defer wages; hence few specialized firms enter the market and the return from the workers’ on-the-job search is lower. A deviating firm that does not implement high effort and thus does not defer compensation only obtains a modest increase in the value of search, since there are relatively few specialized firms to search for anyway.

Put differently, when all the other firms use deferred compensation, the search market is "designed" for workers with a high second period wage, in the sense that the equilibrium maximizes the value of search for such workers. That means that there are few specialized firms and that they pay high wages. This is bad news for firms and workers that deviate and set wages equal to productivity, since they would get more out of the search market if wages were lower and the job finding rate higher (as in the no-effort equilibrium). This weakens incentives to deviate and not implement effort.

Note that the gain from on-the-job search depends on the number of specialized firms with vacancies entering the market, which again depends on the amount of deferred compensation obtained by other workers in the economy. Hence a feedback effect exists between the wage contracts given to the average worker in the market, and the gain from search for any individual worker. The gain from search again influences the firms’ choice of contracts, which closes the loop.

We want to analyze how the equilibrium configurations depend on the parameters of the model. For any parameter \( z \), let \( \mathcal{M}^z \) denote the set of permissible parameters for which
the effort equilibrium exists given that \( z = \bar{z} \). We say that an increase in \( z \) makes the effort equilibrium more likely if \( \mathcal{M}_z \supset \mathcal{M}_{z=\bar{z}} \) for any \( z_1 < z_2 \). Analogously, let \( \mathcal{M}_{z=\bar{z}} \) denote the set of parameters for which the no-effort equilibrium exists given that \( z = \bar{z} \) and define more likely accordingly.

**Proposition 3**  

i) An increase in the search frictions (decreased \( A \), increased \( K \) or increased \( \gamma \)) makes the no-effort equilibrium less likely and the effort equilibrium more likely.

ii) A decrease in the productivity difference between ordinary and specialized firms (decrease in \( y_p - y_2 \)) makes the effort equilibrium more likely and the no-effort equilibrium less likely.

**Proof.** See Appendix H.

A reduction in \( A \), or an increase in \( K \) or \( \gamma \), reflects that it becomes more costly to find a trading partner in the submarket of specialized firms, which can be interpreted as search frictions being more severe. The Proposition states that such a change will tend to favor the effort equilibrium. The same is true for a decrease in \( y_p - y_2 \). Higher search frictions and smaller productivity differentials between specialized and ordinary firms both imply that the search market is less important for the agents, hence distorting the workers’ search behavior becomes less costly. This makes the effort equilibrium more attractive relative to the no-effort equilibrium.

We also want to analyze the conditions under which multiple equilibria are "likely" to occur. This is difficult, since the model is highly nonlinear. However, we can derive a limit result. To do so, we assume that the value of search in the absence of deferred compensation is greater than the gain from effort, \( \Omega_\text{max} > \bar{\epsilon}(1 - c) \).

22
Assumption 3

\[ 2\gamma\bar{c}(1 - c) < A^2 \left[ \frac{(1 - \beta)(y_p - y_2)}{K} \right]^{2(1-\beta)} \beta(y_p - y_2) \]

From equations (11), (5), and (7), it follows straightforwardly that Assumption 3 implies that \( \Omega^{\text{max}} > \bar{c}(1 - c) \). By comparing Assumption 3 with Assumption 2, we see that the two requirements are consistent for all \( c \in [0, 1] \) if \( \bar{c} < 1/2 \).

Now suppose the cost of effort \( c\bar{e} \) converges to \( y_p - y_2 \), that is, toward the border defined by Assumption 1. We can then show the following result.

**Proposition 4** Suppose Assumptions 1-3 are satisfied. Then, if \( y_p - y_2 \) is sufficiently close to \( \bar{c}c \), the model exhibits multiple equilibria.

**Proof.** See Moen and Rosén (2012). ♦

When \( c\bar{e} \) is large compared with \( y_p - y_2 \), there is a large difference between \( p^e \) and \( p^n \). This implies that the difference in the deadweight loss of implementing effort between the no-effort equilibrium and the effort equilibrium candidates is large, which broadens the scope for multiple equilibria.

### 4.1 Welfare

Since the workers receive the entire economic surplus, the relevant welfare measure is the utility of workers entering the market, \( \bar{u} \). Recall from (26) and (19) that

\[
\bar{u}^n = y_1 + y_2 + \Omega^{\text{max}}(p^n, w_p),
\]

\[
\bar{u}^e = y_1 + y_2 + \bar{c}(1 - c) + \Omega^{\text{max}}(p^e, w_p) - L(p^e, D^e). \tag{31}
\]
Appendix F shows that the search market maximizes the income of searching workers, given
the zero profit condition of firms. Efficiency in the on-the-job search market is therefore
achieved when \( w_2 = y_2 \). The unconstrained efficient allocation thus requires that \( e = \bar{e}, \)
\( p = p^n \), and \( s = s^n \). It follows that both equilibria are inefficient: the no-effort equilibrium
because there is no effort, and the effort equilibrium because there is too little turnover.

Suppose the parameter constellation is such that multiple equilibria exist. We want to
explore whether the equilibria can be welfare ranked. The conclusion is negative: One cannot
generally show that one of the equilibria welfare dominates the other. The exception is if
the effort cost, \( c \), is close to the upper boundary \( c^e \) of the interval where the high-effort
equilibrium exists. In this case, we have (from equations (26) and (19) )

\[
\bar{w}^e \approx y_1 + y_2 + \Omega^{\text{max}}(p^e, w_p^e) < y_1 + y_2 + \Omega^{\text{max}}(p^n, w_p) = \bar{w}^n.
\]

Hence the low-effort equilibrium welfare dominates the high-effort equilibrium in this case.
To understand why, note that when \( c = c^e \), the net gain from effort in the effort equilibrium
exactly balances the loss associated with distortions in the workers’ search intensity. How-
ever, since wages are above the workers’ productivity, the search market does not maximize
the joint gain from search, since too few firms enter the market. Suppose, instead, that no
firms implement effort. Then the wages in specialized firms fall and more specialized firms
enter the market. At this point the search market maximise the joint gain from search; hence
this increasing the gain from search and thus also welfare.

However, similar logic does not hold for values of \( c \) in the interior of \((c^n, c^e)\). Given that
a firm implements effort, the joint income from on-the-job search, \( \Omega \), may be higher in the
effort equilibrium than for a deviator in the no-effort equilibrium, since the worker may have
very low search intensity in the latter case (zero if the worker’s second-period wage exceeds $w^n_p$). The cost of implementing effort in the no-effort equilibrium will then be very high, and the economy may be locked into the no-effort equilibrium even though the effort equilibrium is more efficient.

4.2 The Hosios condition

We assume that the Hosios condition is satisfied, so the search equilibrium maximizes the utility of searching workers given that firms break even (see Appendix F). Hence the search market in itself does not create inefficiencies. We now relax this assumption.

Suppose the workers’ bargaining power $\beta$ is lower than the value prescribed by the Hosios condition. Then wages in the no-effort equilibrium are too low and the job finding rate too high compared to the efficient values. Since deferred compensation increases the wages in specialized firms and reduces the job finding rate, this may now increase the value of search for a worker without deferred compensation, if the deviation from the Hosios condition is sufficiently large. In this case we conjecture that our multiple equilibria result may unravel. If $\beta$ exceeds the value prescribed by the Hosios condition, we conjecture that multiple equilibria still exist. Our conjectures are confirmed by numerical computations. We have computed the equilibrium for a series of parameter values, see Moen and Rosén (2012) for details. First we fix the matching function at $x(u, v) = u^{1/2}v^{1/2}$. In this case the model exhibits multiple equilibria for all permissible values of the workers’ bargaining power in all the computations we have conducted. However, when changing the matching function to $x(u, v) = u^{0.7}v^{0.3}$, the model does not exhibit multiple equilibria for low values
of the workers’ bargaining power $\beta$.

Our conjecture regarding welfare is analogous. Suppose again that the workers’ bargaining power is lower than the level the Hosios condition prescribes. From a welfare point of view, this will tend to favor the effort equilibrium over the no-effort equilibrium. This is also confirmed by our computational exercises. If $x(u, v) = u^{0.7}v^{0.3}$, the effort equilibrium welfare dominates the no-effort equilibrium for low values of the workers’ bargaining power $\beta$.

5 Implications

This Section discusses further implications and extensions of our model and compares them with empirical findings, where such findings exist.

5.1 The slope of the wage-tenure profile and turnover

An interesting issue is the relationship between the slope of the wage-tenure profile in firms and the industry turnover rate. In our model, all firms in the market offer the same wage contracts. However, since the model exhibits multiple equilibria, different markets can experience different outcomes even when the parameter constellations are identical. In markets where firms implement effort, wage profiles are steep and the turnover low, whereas the opposite is the case in the no-effort equilibrium. Hence, the model predicts a negative relationship between the steepness of the wage-tenure profile and turnover.

If the parameters differ between industries, we obtain similar results. Suppose the cost of implementing effort is higher in one industry than another. If firms implement effort in a low-cost industry but not in a high-cost industry, there is a negative relationship between
the slope of the wage-tenure profile and the turnover rate across the industries. If effort is implemented in both industries, there is less deferred compensation and hence more turnover in an industry with low effort costs. Again the model predicts a negative relationship between the steepness of the wage-tenure profile and the industry turnover rate in the economy.\footnote{The conclusions are not so clear for differences in other structural parameters. Suppose two industries differ in terms of matching efficiency $A$. Suppose both industries implement effort. The industry with the higher $A$ will have more deferred compensation. The turnover rate in this industry may or may not be higher than in the low-$A$ industry.}

To analyze within-industry differences, the model may be extended to allow for heterogeneous ordinary firms. Suppose ordinary firms differ in the cost of implementing effort (high or low). Suppose further that all employed workers search in the same search market and face the same value of $p$. If only low-cost firms implement effort, they have the steepest wage-tenure profiles. If both firm types implement effort, the high-costs firms have the steepest wage-tenure profiles. In both cases turnover will be lower in firms with steeper wage-tenure profiles.

Several papers explore the relationship between the slope of the wage-tenure profiles and turnover empirically. The general finding is that higher expected wage growth decreases turnover. Using establishment level U.S. survey data, Fairris (2004) finds that quits are lower when the wage-tenure profile is steeper and when seniority is used as a criterion for promotion. Galizzy and Lang (1998) study a sample of firms in Turin, Italy and find that expected wage growth and worker quits are negatively related. Using Belgian data, Leonard and Audenrode (1996) find that a steeper tenure-wage profile is associated with lower turnover, especially for white-collar workers.\footnote{In addition, Barth and Dale-Olsen (1999) study the relationship between the steepness of the wage-tenure profile and the turnover rate in the economy.}
These findings fit well with the predictions of our model. However, other theories may also explain a negative relationship between the slope of the wage-tenure contract and the turnover rate, such as firm-specific human capital (Becker, 1967) or learning-by-doing (Jovanovic, 1979).

There is also an empirical literature on the relationship between monitoring and pay. The Shapiro-Stiglitz model predicts that monitoring and pay level are substitutes, and most of the literature focuses on the level of pay rather than wage profiles.\textsuperscript{11} One exception is Sessions and Theodoropoulos (2008). Using matched employer-employee data for Britain, the authors find a negative relationship between the slope of the wage-tenure profile and the level of monitoring. Similarly Bayo-Moriones et al (2010) find that firms that base their wages partly on seniority are less likely to invest in monitoring devices. These findings indicate that monitoring and deferred compensation are substitutes.

5.2 When are the different equilibria more likely?

Our analysis predicts where we can expect to find deferred compensation. Our analysis concludes that the no-effort equilibrium is more likely when the value of the on-the-job search is high. This may indicate that deferred compensation is less attractive in dense areas, where there are a large number of potential firms. Silicon Valley is an prime example. Furthermore, the no-effort equilibrium is more likely if an experienced worker’s productivity profile and quits using Norwegian data. They find that a steeper seniority-profile lowers turnover. However, it is not possible to disentangle the effect of the steepness of the wage profile from the level of wages in their study.

\textsuperscript{11}This approach has been criticized by, for example, Allgulin and Ellingsen (2002).
varies little across firms. Almost trivially, this indicates that deferred compensation is more likely when productivity differences between firms are small. More interestingly, if a firm can offer a variety of job types, the probability that a good worker-job match can be found internally increases. Hence deferred compensation may be more attractive in large firms with efficient internal job markets than in small firms with thin internal job markets.

Finally, Proposition 4 indicates that multiple equilibria are more likely to occur if the cost of effort is relatively close to the output gap $y_p - y_2$. Hence, one can expect large differences between contract forms and turnover rates between countries in sectors where the gain from turnover is relatively modest compared to the effort cost.

5.3 Piece rate payment or deferred compensation?

So far we have focused solely on effort that can only be observed with a time lag and remunerated by deferred compensation. Other dimensions of effort or output may be readily observable and hence remunerated by short-term bonuses. In an extension of the model, we analyze the interplay between short-term and long-term incentives (available on request). Our underlying assumption is that the cost of effort is convex in the two dimensions. Thus, if a worker exerts higher short-term effort, this increases the cost of choosing the high long-term effort level $e$.

We show that as a result of this convexity, firms that implement high long-term effort will cut back on short-term incentives relative to firms that do not. The reason is that higher short-term effort increases the cost of long-term effort. Hence the amount of deferred compensation needed to implement long-term effort increases. This has a first order effect
on profits.

Our conclusion is that in the effort equilibrium, firms will be more restrictive in their use of short-term bonuses than in the no-effort equilibrium. Analogously, if firms in an industry are heterogeneous, say, regarding the value of long-term effort, firms that choose to implement long-term effort will use less incentive-powered short term bonus systems.

There is some evidence that suggests that firms that use deferred compensation to a lesser extent than other firms use short-term bonuses. Bayo-Moriones et al. (2010) study personnel practices in 734 industrial establishments in Spain. They find that firms that reward seniority are less likely to invest in monitoring devices and less likely to offer short term bonuses.\textsuperscript{12}

Our model predicts that firms in countries (or regions) with lower turnover rates rely more on long-term wage contracts with deferred compensation (seniority-based wages, promotions, etc.) and less on short-term performance-based systems, than firms in countries (regions) with higher turnover rates. This implication is in accordance with popular conceptions of the differences between the U.S. and Japan, and between Silicon Valley and Route 128 in Mass.\textsuperscript{13}

\textsuperscript{12}Another explanation for the findings in Bayo-Moriones et al. (2010) is that long- and short-term incentive schemes are alternative incentive mechanisms.

\textsuperscript{13}For evidence on differences in turnover rates and wage profiles between different regions in the U.S. see Saxenian (1994).
5.4 Firm-specific human capital

Firms’ choices of wage contracts will influence their incentives to invest in firm-specific human capital. If a firm implements effort, the turnover rate is reduced, and the firm has a higher probability of retaining workers in the second period. Hence the gains from firm-specific human capital investments increases.

Our model thus predicts that firms in countries/regions with different remuneration practices may exercise different strategies for investments in firm-specific human capital. Firms in countries or regions with deferred compensation and low turnover rates will invest more in firm-specific human capital than firms in countries or regions with less deferred compensation and higher turnover rates. If firms within a sector have different practices regarding deferred compensation, then those who offer higher deferred compensation and have lower turnover rates will invest more in firm-specific human capital.

6 Conclusion

This paper analyses moral hazard in a model of on-the-job search. Since worker effort is observed with a time lag, implementing effort requires that the wage contract includes deferred compensation. However, deferred compensation distorts the workers’ on-the-job search decisions, since it gives the workers too weak incentives to search on the job. Due to feedback effects between firms’ choices of wage contracts and entry to the on-the-job search market, multiple equilibria can emerge. In one equilibrium, firms offer incentive contracts with deferred compensation, which leads to high effort and low turnover rates. In the other
equilibrium, firms do not offer deferred compensation, which leads to low effort and high turnover.

Multiple equilibria can be viewed as an extreme form of multiplier effects. In our model, an increase say in the cost of opening specialized jobs, gives rise to multiplier effects: higher entry costs have a direct negative effect on the turnover rate for experienced workers, which again improves the firms’ incentives to defer compensation in order to implement effort. More deferred compensation reduces the incentives of specialized firms to enter the market further. In our setup, with identical firms, these multiplier effects lead to multiple equilibria. If we allow for more heterogeneity, for instance regarding the firms’ cost of implementing effort, the equilibrium may be unique, but the multiplier effects will still be present.

Our model contributes to a growing literature that incorporates private information into matching models of the labor market. Our model also sheds light on the observed differences in turnover rates between countries (the U.S. and Europe/Japan) and regions (Silicon Valley and Route 128 in Mass.).

Our model gives rise to several empirical implications. Equilibrium with deferred compensation is more likely to prevail in markets with large search frictions and small productivity differences between firms. Deferred compensation tends to reduce turnover, and this is supported by empirical findings. In addition, deferred compensation tends to occur in tandem with weak short-term incentives and high investments in firm-specific human capital.
Appendix A. Proof of Lemma 1

We have (using equation (12), when going from the second to the third equation),

\[ L(D, p) = \Omega^{\text{max}}(p) - \Omega(p, \bar{s}) = \frac{p^2(w_p - y_2)^2}{2\gamma} - \bar{s}p(w_p - y_2) + \frac{\gamma\bar{s}^2}{2} \]

\[ = \frac{p^2(w_p - y_2)^2}{2\gamma} - \frac{p^2(w_p - w_2)(w_p - y_2)}{\gamma} + \frac{p^2(w_p - w_2)^2}{2\gamma} \]

\[ = \frac{p}{2\gamma}((w_p - y_2) - (w_p - w_2))^2 = \frac{p^2}{2\gamma}(w_2 - y_2)^2 = \frac{p^2D^2}{2\gamma} \]

which completes the proof.

Appendix B. Proof of Lemma 2

Incentive compatibility w.r.t \( e = \bar{e} \) requires that \( u(\bar{\omega}, \bar{e}, \bar{s}) \geq \max_s u(\bar{\omega}, 0, s) \). From (3) and (12) it follows that the incentive compatibility constraint thus reads

\[ \frac{p^2(w_p - w_2)^2}{2\gamma} + w_2 - c\bar{e} \geq \frac{p^2(w_p - y_2)^2}{2\gamma} + y_2 \iff \]

\[ \frac{p^2(w_p - y_2 - D)^2}{2\gamma} + y_2 + D - c\bar{e} \geq \frac{p^2(w_p - y_2)^2}{2\gamma} + y_2 \iff \]

\[ p^2D [D - 2(w_p - y_2)] + 2\gamma D \geq 2\gamma c\bar{e}. \]

Thus \( \overline{D} \) is defined by

\[ p^2\overline{D} [\overline{D} - 2(w_p - y_2)] + 2\gamma \overline{D} = 2\gamma c\bar{e} \tag{B.2} \]

(ii) Differentiating (B.2) w.r.t \( \overline{D} \) and \( p \) gives

\[ (2p^2\overline{D} - 2p^2(w_p - y_2) + 2\gamma) d\overline{D} + 2p\overline{D} [\overline{D} - 2(w_p - y_2)] dp = 0, \]
which gives
\[
\frac{d\overline{D}}{dp} = \frac{-2p\overline{D} [\overline{D} - 2(w_p - y_2)]}{2p^2\overline{D} - 2p^2(w_p - y_2) + 2\gamma}.
\]

Using (B.2), we have
\[-2p\overline{D} [\overline{D} - 2(w_p - y_2)] = 2\gamma(\overline{D} - c\overline{e})/p > 0\]
and hence the numerator is positive. Since
\[2p^2\overline{D} - 2p^2(w_p - y_2) + 2\gamma > (p^2\overline{D} [\overline{D} - 2(w_p - y_2)] + 2\gamma\overline{D})/\overline{D}\]
and (using (B.2))
\[(p^2\overline{D} [\overline{D} - 2(w_p - y_2)] + 2\gamma\overline{D})/\overline{D} = 2\gamma c\overline{e}/\overline{D} > 0\]
the denominator is also positive, and therefore \(d\overline{D}/dp > 0\).

Differentiating (B.2) w.r.t to \(\overline{D}\) and \(w_p\), we have
\[
(2p^2\overline{D} - 2p^2(w_p - y_2) + 2\gamma) d\overline{D} - 2p^2\overline{D} dw_p = 0,
\]
or
\[
\frac{d\overline{D}}{dw_p} = \frac{2p^2\overline{D}}{2p^2\overline{D} - 2p^2(w_p - y_2) + 2\gamma}.
\]

We have already shown that the denominator is positive. Hence \(d\overline{D}/dw_p > 0\).

The claim that \(\overline{D}\) is increasing in \(c\) follows directly from the fact that the left-hand side of (16) is increasing in \(\overline{D}\) and the right-hand side is increasing in \(c\).

**Appendix C. Proof of Lemma 3**

From equation (18) and Assumption 1, we have \(p^n > 0\). From equations (10) and (18) it follows that the probability that a worker finds a job in a specialized firm, \(p^n s^n\), is given by
\[
p^n s^n = A^2 \left[ \frac{(1 - \beta)(y_p - y_2)}{K} \right]^{\frac{2(1 - \beta)}{\beta}} \frac{\beta}{\gamma} y_p - y_2 < 1,
\]
where the inequality follows directly from Assumption 2.
Appendix D. Claim that equations (22)- (24) have a unique solution

Much of the proof is given in the main text. We only have left to show that (22)- (24) uniquely determine $p^e$, $w_p^e$, and $w_2(\bar{c}) = y_2 + D^e$. The proof is constructed as follows. First, we show that equations (22) and (24) define $\overline{D}$ as a function of $p$, $\overline{D} = \tilde{D}(p)$. Second, we show that $\tilde{D}(p)$ is strictly increasing in $p$ and, third, that the equations $\overline{D} = \tilde{D}(p)$ and (23) have a unique solution.

1. Equations (22) and (24) uniquely define $\overline{D} = \tilde{D}(p)$

Rewrite (22) as (in generic form)

$$\overline{D} = \frac{w_p - \beta y_p}{1 - \beta} - y_2 = f(w_p).$$

Hence the two equations (22) and (24) can be condensed to

$$D(p, w_p) - f(w_p) = 0. \tag{D.1}$$

We first want to show that for any $p > 0$, (D.1) has a unique solution. Sufficient conditions for uniqueness are that $dD/dw_p < 1$, $df/dw_p > 0$, $D(p, w_p) - f(w_p) > 1$ for a (low) value of $w_p$, and $D(p, w_p) - f(w_p) < 0$ for a (high) value of $w_p$. To this end, first note that $f$ has the following two properties;

$$f(\beta y_p + (1 - \beta)y_2) = 0,$$

$$f'(w_p) = \frac{1}{1 - \beta} > 1.$$ 

For any given $p$, implementing effort requires that $D \geq \bar{c}$, hence, in particular

$$D(p, \beta y_p + (1 - \beta)y_2) > 0$$

Therefore $D(p, w_p) > 0 = f(w_p)$ for $w_p = \beta y_p + (1 - \beta)y_2$ (i.e., the wage in the no-effort equilibrium). Furthermore, $f(y_p) = y_p - y_2$ while $D(p, y_p) < y_p - y_2$ for all $p$. We have
just seen that $f'(w_p) = 1/(1 - \beta) > 1$. A sufficient condition for the existence of a unique solution is thus that $\partial \mathcal{D}(p, w_p) / \partial w_p < 1$, which we now show.

Incentive compatibility for optimal effort requires that $u(\tilde{\omega}, \tilde{\varphi}, s^e) \geq \max_s u(\tilde{\omega}, 0, s)$, which is satisfied with equality. Using (3), $w_2(\tilde{\varphi}) = y_2 + \mathcal{D}$, and $w_2(\tilde{\varphi}) = y_2$, we have

$$w_1 - c\varphi - \gamma(s^e)^2/2 + s^e p^e w_p + (1 - s^e p^e)(y_2 + \mathcal{D}) = w_1 - \gamma(s^s)^2/2 + s^s p^e w_p + (1 - s^s p^e)y_2,$$

where $s^e$ is the search intensity with effort and $s^s$ is the search intensity with no effort. Differentiating with respect to $\mathcal{D}$ and $w_p$ yields (because of the envelope theorem we can ignore changes in $s^s$ and $s^e$)

$$s^e p^e dw_p + (1 - s^e p^e)\frac{d\mathcal{D}}{dw_p} = s^s p^e dw_p$$

or

$$\frac{d\mathcal{D}}{dw_p} = \frac{s^s p^e - s^e p^e}{1 - s^e p^e} < 1,$$

provided that $s^s p^e < 1$. To show this, note that since $p^e < p^n$, it is sufficient to show that $s^s \leq s^n$. To this end, note that the no-effort equilibrium $(p, w_p)$ maximizes the value of search $p^2(w_p - y_2)^2/2\gamma$ (from (F.2)). Hence

$$p^n(w_p^n - y_2) > p^e(w_p^e - y_2).$$

But, it then follows directly from (12) that $s^s < s^n$. Hence we have shown that for any given $p^e$, the equations (22) and (24) have a solution, and hence we can write $\mathcal{D} = D(p, w_p(p) \equiv \tilde{\mathcal{D}}(p)$.

2. $\tilde{\mathcal{D}}(p)$ is strictly increasing in $p$. 

36
We want to show that \( \frac{d\tilde{D}(p)}{dp} > 0 \). To this end, first note that

\[
\frac{d\tilde{D}(p)}{dp} = \frac{dD(p, w_p(p))}{dp} = \frac{\partial D(p, w_p)}{\partial p} + \frac{\partial D(p, w_p)}{\partial w_p} \frac{dw_p}{dp}.
\]

Differentiate (22) to get

\[
\frac{dw_p}{dp} = (1 - \beta) \frac{d\tilde{D}}{dp},
\]

or

\[
\frac{d\tilde{D}(p, w_p(p))}{dp} (1 - (1 - \beta) \frac{\partial D(p, w_p)}{\partial w_p}) = \frac{\partial D(p, w_p)}{\partial p}.
\]

From Lemma 2 we know that the right-hand side is strictly positive, \( (\partial \tilde{D}(p, w_p)/\partial p > 0) \). Above we showed that \( \partial \tilde{D}(p, w_p)/\partial w_p < 1 \). The claim thus follows.

3. The equations \( \tilde{D} = \tilde{D}(p) \) and (23) have a unique solution.

From (23) we can write \( \tilde{D} \) as a function of \( p \), \( D = g(p) \), where

\[
g(p) = -p^{\frac{\alpha}{1-\beta}} A^{-\frac{1}{1-\beta}} \frac{K}{1-\beta} + (y_p - y_2),
\]

which is strictly decreasing in \( p \). Note that \( g(0) = y_p - y_2 \) (with \( D = y_p - y_2 \), it follows that \( w_2 = y_2 + D = y_p \)). By the definition of \( p^a \), we know that \( g(p^a) = 0 \).

Then consider \( \tilde{D}(p) \). From (16) it follows that \( \tilde{D}(0) = c\tilde{e} < y_p - y_2 \) (Assumption 1). It follows that the equation \( g(p) = \tilde{D}^e(p) \) has a unique solution \( p^e \in (0, p^a) \). Furthermore, from (12) it follows that \( s^e < s^a \), hence it follows that \( s^e p^e < s^a p^a < 1 \) (from Lemma 2) since \( p^e < p^a \).
Appendix E. Proof of Proposition 1

Proof: a) It is sufficient to show that equation (21) with equality has a unique solution for \( c \), that is,

\[
L(\overline{D}(p^n, w^n), p^n) - \overline{c}(1 - c) = 0
\]

has a unique solution \( 0 < c < 1 \). For \( c = 0 \), \( \overline{D} = 0 \) and hence \( L = 0 \), and the left-hand side is strictly negative. Furthermore, from Lemma 2 we know that \( \overline{D} \) is increasing in \( c \); hence the left-hand side of the equation is strictly increasing in \( c \). Finally, for \( c = 1 \), the last term is zero, and hence the left-hand is strictly positive. Part a) thus follows.

b) It is sufficient to show that equation (28) with equality has a solution for \( c \), that is,

\[
L(\overline{D}(p^e, w^e), p^e) - \overline{c}(1 - c) = 0
\]

(E.1)

has a solution \( 0 < c < 1 \). For \( c = 0 \), \( D = 0 \), and it follows that the lhs is strictly negative. For \( c = 1 \), the last term is zero, and the left-hand side is strictly positive. Due to continuity it follows that (E.1) has a solution and that there is a smallest solution, which we define as \( c^e \). The Proposition thus follows.

Appendix F. The search market maximizes the utility of searching workers

The problem of maximizing worker utility given the zero profit constraint, is written as

\[
\max_{s,p} sp(w_p - w^2) - \frac{\gamma s^2}{2} + w_2 \quad \text{s.t.} \quad A^{\frac{1}{1-\sigma}} p^{-\frac{\sigma}{1-\sigma}} (y_p - w_p) = K. \quad \text{(F.1)}
\]

The first order condition for optimal search intensity reads \( s = p(w_p - w^2)/\gamma \), which inserted yields
\[
\max_p \frac{p^2(w_p - w_2)^2}{2\gamma} + w_2 \quad \text{s.t.} \quad A\frac{1}{1-\beta}p^{-\frac{\beta}{1-\beta}}(y_p - w_p) = K. \tag{F.2}
\]

The associated Lagrangian is

\[
L = \frac{p^2(w_p - w_2)^2}{2\gamma} + w_2 - \lambda[A\frac{1}{1-\beta}p^{-\frac{\beta}{1-\beta}}(y_p - w_p) - K],
\]

with first order conditions

\[
\frac{\partial L}{\partial w_p} = 0 \Leftrightarrow \frac{p^2(w_p - w_2)}{\gamma} + \lambda A\frac{1}{1-\beta}p^{-\frac{\beta}{1-\beta}} = 0, \\
\frac{\partial L}{\partial p} = 0 \Leftrightarrow \frac{p(w_p - w_2)^2}{\gamma} + \lambda A\frac{1}{1-\beta}p^{-\frac{\beta}{1-\beta}1}(y_p - w_p). 
\]

Solving gives

\[
w_p = \beta y_p + (1 - \beta)w_2 = \beta y_p + (1 - \beta)(y_2 + D).
\]

Hence the equilibrium values \((p, w_p)\) maximizes the expected utility of the searching workers, given the zero profit constraint on opening vacancies.

**Appendix G. Proof of Proposition 2**

It is sufficient to show that \(L(D(p^e, w_p^e), p^e) < L(D(p^n, w_p^n), p^n)\) evaluated at \(c = c^e\), since this implies that \(c^e\) satisfies (21) with strict inequality. Hence it follows that \(c^n > c^e\).

From equation (15) we know that \(L(D(p^e, w_p^e), p^e) < L(D(p^n, w_p^n), p^n)\), evaluated at \(c = c^e\) if and only if \(D_e p^e < D_n p^n\). From (F.2) in Appendix F it follows that the no-effort equilibrium solves \(\max_{p,w_p} p(w_p - y_2)\) subject to the zero profit constraint, and hence \(p^e(w_p^e - y_2) \leq p^n(w_p^n - y_2)\).
Condition (16) can be written as

\[ p[2(w_p - y_2)p - pD] = 2\gamma(1 - \frac{\epsilon}{D}). \]  

(G.1)

Suppose \( D^e p^e \geq D^n p^n \). Since \( p^e(w^e_p - y_2) < p^n(w^n_p - y_2) \) and \( p^e < p^n \), it follows that

\[ p^e[2(w^e_p - y_2)p^e - p^e D^e] < p^n[2(w^n_p - y_2)p^n - p^n D^n]. \]

From (G.1) it thus follows that

\[ 2\gamma(1 - \frac{\epsilon}{D^e}) < 2\gamma(1 - \frac{\epsilon}{D^n}), \]

that is, \( D^e < D^n \). But then it follows that \( D^e p^e < D^n p^n \), a contradiction.
Appendix H. Proof of Proposition 3

1) Consider first an increase in $A$. From (21) it follows that a sufficient condition for the no-effort equilibrium to be more likely is that an increase in $A$ implies a strict increase in $L(D(p^n, w^n_p), p^n)$. From (16) it follows that $D$ only depends on $A$ through $p^n$ and $w^n_p$. From (17) and (18) it follows that $\frac{dw^n_p}{dA} = 0$ and $\frac{dp^n}{dA} > 0$. From Lemma 2 it then follows that $D^n$ strictly decreases in $A$. Hence an increase in $A$ makes the no-effort equilibrium more likely.

The proof that a reduction in $K$ makes the no-effort equilibrium more likely is analogous.

From (28) it follows that a sufficient condition for an increase in $A$ to make the effort equilibrium less likely is that $L(D(p^e, w^e_p), p^e)$ is strictly increasing in $A$. It is more convenient to show an equivalent claim, that $L(\tilde{D}^e(p^e), p^e)$ is strictly increasing in $A$. From (16) it follows that $D$ only depends on $A$ through $p^e$ and $w^e_p$. From (29) we know that $D = \tilde{D}^e(p)$, with $\frac{\tilde{D}^e(p)}{dp} > 0$. It is thus sufficient to show that $p^e$ is strictly increasing in $A$. Suppose not. Then $\tilde{D}^e = \tilde{D}^e(p^e)$ is strictly decreasing in $A$. Recall that $w^e_2 = y_2 + \tilde{D}^e$, which is then strictly decreasing. From (7) it follows that $p^e$ is increasing, a contradiction. The proof that a reduction in $K$ makes the no-effort equilibrium more likely is analogous.

Then consider an increase in $\gamma$. We claim that this is equivalent to reducing $A$. More specifically, we will show that if $\gamma$ increases from $\gamma_0$ to $\gamma_0 + \Delta \gamma$, there exists a $\Delta A > 0$ such that if $A$ simultaneously increases from $A_0$ to $A_0 + \Delta A$, the equilibrium is unchanged. It then follows that an increase in $\gamma$ from $\gamma_0$ to $\gamma_0 + \Delta \gamma$, given $A = A_0$, is equivalent to reducing $A$ from $A_0 + \Delta A$ to $A_0$, given that $\gamma = \gamma_0 + \Delta \gamma$. We have already shown the effects of the latter.

Suppose $\gamma$ increases from $\gamma_0$ to $\gamma_0 + \Delta \gamma$. Suppose there exists a change in $A$ such that
the equilibrium is unchanged. The search cost of workers must stay constant; hence (using (12))

\[ \frac{\gamma s^2}{2} = \frac{p^2(w_p - w_2)^2}{2\gamma} = \text{const}, \]

\[ p \sim \gamma^{1/2}. \]

From the free entry condition (7) it follows that \( p \sim A^{\frac{1}{3}} \). Hence the equilibrium remains unchanged if \( \gamma^{1/2} \sim A^{\frac{1}{3}} \) or \( A \sim \gamma^{\frac{1}{3}} \). Since the equilibrium is unique the claim follows.

2) It is sufficient to show that the result holds for an increase in \( y_p \). It is thus sufficient to show that the loss of implementing effort increases for both the effort- and no-effort equilibrium candidates. Consider first the effort equilibrium. From (B.1) we have

\[ \frac{p^2(w_p - y_2 - D)^2}{2\gamma} + D - c \bar{e} - \frac{p^2(w_p - y_2)^2}{2\gamma} \geq 0. \]

(H.1)

Denote the left-hand side of (H.1) by \( H \). Suppose \( dy_p = d\bar{D} \). From (22) it follows that \( dw_p = dy_p \), and from (7) \( dp = 0 \). It follows that

\[ dH = dy_p - \frac{p^2(w_p - y_2)}{\gamma} dy_p \]

\[ = dy_p - spdy_p \]

\[ > 0, \]

since \( sp < 1 \). Hence the incentive compatibility constraint is slack at this point. Let topscript \( e \) indicate equilibrium values before the shift. It follows that at \( p = p^e \), \( \bar{D}(p^e; y_p + dy_p) < D^e + dy_p \). At the same time, \( D = p^{FE-1}(p; y_p + dy_p) = D^e + dy_p \) (where \( p^{FE-1} \) is the inverse of the function \( p^{FE}(\bar{D}) \) defined by (30)). Hence, at \( p^e \), the \( p^{FE} \) curve is above the \( \bar{D}(p^e) \) curve after the shift. It follows that both \( p \) and \( \bar{D} \) increases, and hence so does the loss of implementing effort. The result then follows.
Consider, then, the no-effort equilibrium. From (17) and (18) it follows that $w_p$ and $p$ increases in $y_p$, and hence so does the loss $L(p, w_p)$. The result thus follows.

References


