

# Incentive Contracts and Worker Turnover in Search Equilibrium\*

Espen R. Moen<sup>†</sup> and Åsa Rosén<sup>‡</sup>

January 3, 2005

## Abstract

In this paper we analyze the interaction between long-term wage contracts and the turnover rate in the economy. Our starting point is that long-term incentive contracts with deferred compensation are warranted in order to motivate workers to undertake effort. On the other hand, deferred compensation may distort turnover decisions for experienced workers. We show that incentives systems based on deferred compensation become less attractive the higher is the overall turnover rate in the market. Furthermore, there exist feedback effects between the firms' choice of wage contracts and the turnover rate in the economy, and these feedback effects may lead to multiple equilibria: A low-turnover equilibrium where firms use deferred compensation to motivate workers, and a high-turnover equilibrium where they do not. Our model thus explains observed differences between countries like US and Japan and between regions like Silicon Valley and Route 128.

**Key Words:** Deferred Compensation, Incentive Contracts, Turnover, Multiple Equilibria.

**JEL Codes:** J41, J63.

---

\*We have benefited from comments by Denis Gromb and seminar participants at University of Oslo and the SED 2004 meeting in Florence.

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

<sup>‡</sup>Swedish Institute for Social Research, Stockholm University, S106-91 Stockholm, Sweden. E-mail: asa.rosen@sofi.su.se

# 1 Introduction

In this paper we analyze the interaction between firms' choice of wage contracts and the turnover rate in the economy. Our starting point is that long-term incentive contracts with deferred compensation are warranted in order to motivate workers to undertake effort. This may be because worker effort is difficult to measure accurately over a short period of time, or because worker effort may have consequences for future output. On the other hand, deferred compensation may distort turnover decisions; as future wages are set so as to motivate workers early in their careers and not in order to optimize worker turnover. In particular, workers may be reluctant to quit and lose the bonus waiting for them.

Turnover is modelled using the Diamond-Mortensen-Pissarides framework (Diamond 1982, Mortensen 1986, Pissarides 2000) with competitive wage setting as in Moen (1997). In our model, workers conduct on-the job search in order to find a job in which they are more productive and better paid than in their present position. Optimal turnover rates are obtained if a worker's current wage equals his productivity. With deferred compensation, wages for experienced workers exceed their productivity, and search effort is too low. We show that incentives systems based on deferred compensation become less attractive the higher is the overall turnover rate in the market. Furthermore, there exist feedback effects between the firms' choice of wage contracts and the turnover rate in the economy, and these feedback effects may lead to multiple equilibria: A low-turnover equilibrium where firms use deferred compensation to motivate workers, and a high-turnover equilibrium where they do not.

Our findings are interesting for several reasons. In a labour-contract perspective, our paper is the first to explore the relationship between deferred compensation and turnover, and show that there exists feedback effects between wage contracts and market conditions for employed workers. Furthermore, we find that short-term incentives and long-term incentives are substitutes. This finding is confirmed in a new empirical study (Moriones et al 2004), which finds that firms using deferred compensation to a smaller extent than other firms are using short-term bonuses as an incentive mechanism.

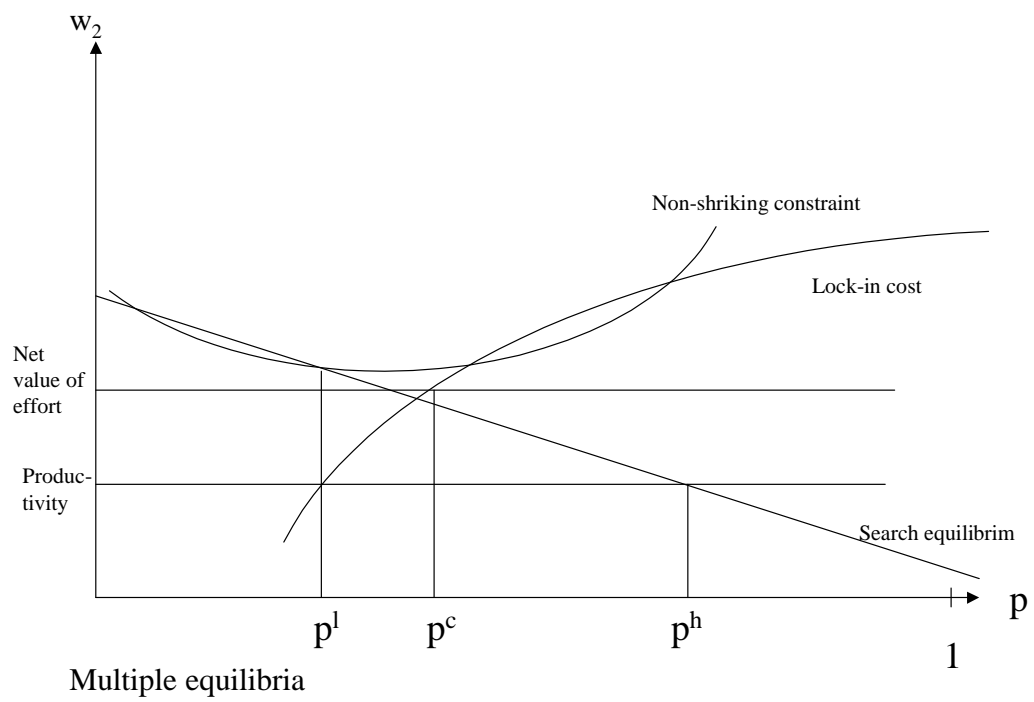
Our model may also give a new explanation for the huge variations in turnover rates between countries and regions. For instance, in the US and the UK the median tenures among employees were (in 1991) 3.0 and 4.4

years, while it was 7.5 years in Germany and 8.2 years in Japan. The fraction of workers with a tenure of less than one year were 28.8 percent in the U.S. and 18.6 % in the U.K., 12.8 % in Germany and only 9.8% in Japan (OECD 1993). Large differences in turnover rates also exist between regions in the same country. In Silicon Valley, turnover rates are extremely high. Along "Route 128" in Massachusetts, another prospering area with a well developed high-technology industry, turnover rates are much lower (Saxenian 1994). Our model predicts that the countries (regions) with low turnover to a larger extent use long-term wage contracts with deferred compensation (seniority-based wage, promotion etc.) and to a less extent use short-term bonus systems than do countries (regions) with a high turnover rate. Although we have not been able to find data that confirm this implication, it seems to be in line with popular conceptions of differences between US and Japan and between Silicon Valley.

To see why we may get multiple equilibria, note the following: If all the other firms in the market choose long-term wage contracts with deferred compensation, few firms open vacancies for employed workers, because the wages they have to pay in order to attract such workers are high. Subsequently, turnover rates are low, and the added costs of using long-term wage contracts due to distorted turnover decisions are small. We refer to this as low-turnover equilibrium. By contrast, if the other firms in the market choose short-term wage contracts, more firms will open vacancies for experienced workers, since they are easier to attract. The costs associated with distorted turnover decisions are thus higher. It follows that short-term wage contracts may be optimal. We refer to this as high-turnover equilibrium.

The figure illustrates the two equilibria. The horizontal axis represents the job finding probability  $p$ . The non-shirking constraint shows the wage necessary to induce long-term effort (hereafter effort) as a function of  $p$ . The search equilibrium line gives the job finding probability  $p$  as a function of the wage  $w_2$  for experienced workers in their existing jobs. The "lock-in costs" curve shows the costs associated with distorted turnover decisions as a function of  $p$  given that the wage satisfies the non-shirking constraint. The net-value of effort line shows the value of long-term effort less of effort cost (effort is high or low). The productivity line shows the worker's productivity in his current job.

The low-turnover equilibrium  $p^l$  is determined as the intersection between the non-shirking constraint and the search equilibrium line. For this low value



of  $p$ , the net value of effort exceeds the lock-in cost, and the firms implement a high effort level. The high-turnover equilibrium  $p^h$  is determined as the intersection between the worker productivity line and the search equilibrium line. For this high value of  $p$ , the lock-in cost that emerges if the firm sets a wage that satisfies the non-shirking constraint outweighs the gain from effort. The firm thus implements a low effort level and sets the wage equal to worker productivity to maximize the gain from on-the-job search.

An interesting issue concerns the welfare properties of the two equilibria. It turns out that the equilibria cannot be Pareto ranked. The state which gives rise to the highest welfare depends on parameter values.

Our model has several empirical implications. In low-turnover equilibrium, firms are inclined to give weaker incentives to short-term performance (less use of short-term bonuses), rely more on deferred compensation and invest more in firm-specific human capital than in the high-turnover equilibrium. Furthermore, entrepreneurship and venture capital may be more frequent in high-turnover equilibrium than in low-turnover equilibrium. The implications seem to be in line with popular conceptions of differences between US and Japan and between Silicon Valley and Massachusetts. The effects leading to multiple equilibria may be reinforced by other economic decisions.

There exists a related literature on multiple equilibria and labour market turnover. One branch of the literature relates multiple equilibria to adverse selection, see Chang and Wang (1995). With a high turnover rate, the average quality among the pool of workers that change jobs is high, and thereby also the wage they obtain in the new job. This makes turnover more attractive, and may lead to multiple equilibria. A similar argument is made in Acemoglu and Pischke (1998).

Saint-Paul (1995) shows that multiple equilibria with different levels of turnover may arise in a matching model with firing costs paid by the firm. After a negative productivity shock, workers start on-the-job search. If the job-finding rate is high, the firms will not fire the workers and thus save on firing costs. This makes it more attractive to open jobs in the first place. Moene and Wallerstein (1997) obtain multiple unemployment rates in a shirking model with search frictions and increasing returns to scale in the matching technology. If the unemployment rate is low, hiring costs are high, and firms hoard labour after a temporary firm-specific shock. As a result, the unemployment rate stays low.

Morita (2001) shows how multiple turnover rates may arise as a result

of firms' choice of production technology and learning-by-doing. There exist two types of technologies, and for one of the types turnover is more important than for the other type. If a firm chooses the "high turnover" technology, it may suffer from thin market effects if few other firms choose the same type of technology, and this may lead to multiple equilibria.

Within the search and matching literature, on-the-job search and worker turnover have been given considerable attention lately. Burdett and Mortensen (1998) use on-the-job search to explain wage differentials. Their model is extended in many directions, see for instance Manning (2003) and Burdett and Coles (2003). Mortensen and Pissarides (1994) and Pissarides (1994) study on-the-job search within the Diamond-Mortensen-Pissarides matching framework, while Moen and Rosén (2004) study turnover within the competitive search equilibrium framework. For a recent contribution on turnover and matching, we refer to Kiyotaki and Lagos (2004).

The paper is organized as follows. Section 2 describes the model. Section 3 characterises the high and low turnover equilibria. In section 4 the conditions for multiple equilibria are derived. Welfare is analyzed in section 5. In section 6 the model is extended to account for short-term effort, firm-specific human capital and entrepreneurship. Section 7 concludes.

## 2 The model

The model has two main ingredients: on-the-job search and deferred compensation. Before we go into the technical details of the model, we will discuss these two ingredients in some detail.

There are several reasons to believe that job-to-job movements are important for economic efficiency. First, it is a fact that a lot of job-to-job movements take place. For instance, Boeri (1999) finds that in Europe, between 48 and 70 percent of all gross hirings are job-to-job movements. Fallick and Fleischman (1999) find that in the U.S., 2,7 percent of workers change employers every *month*. Furthermore, job changes give workers a large pay-off. Topel and Ward (1992) find that "wage gains at job changes account for at least a third of early career wage growth" (for men). From a theoretical perspective, there are several arguments that may justify a high turnover rate, particularly among young employees. First, workers may try out several jobs to determine their comparative advantage (Johnson 1978). Second, workers

may do on-the-job search because of match specific productivity differences as in Jovanovic (1979). Third, a worker's relative productivity in different firms may change over time as he gains experience and expertise. Fourth, sectorial shocks to the economy may warrant a re-allocation of workers on firms. Finally, with technological progress, efficient dissemination of knowledge may require worker turnover as workers may learn from each other. Saxonian (1995) argues that dissemination of information through worker turnover is very important in high-tech areas as Silicon Valley. On the other hand, search frictions imply that a reallocation of workers on firms is costly, and hence that some workers may be stuck in jobs in which their productivity is relatively low.

Deferred compensation is profitable if effort is only observable with noise and workers have limited liability (the worst feasible outcome for the worker is to be fired). In this case, motivating workers through short-term contracts is costly as the firm has to pay a rent to the worker, as in the "shirking-model" of Shapiro and Stiglitz (1984). With deferred compensation, this rent may be reduced (Akerlof and Katz 1989 Lazear 1979, 1981). Lazaer also argues that deferred compensation implies that workers are paid more than their productivity at retirement age, and that this can rationalize mandatory retirement. Lazaer and Rosen (1981) and Malcomson (1984) show that tournaments can be used as motivators, which also implies that wages are deferred. Clearly, a firm may also want to use long-term incentives and deferred compensations as motivators if worker effort is only observable after a period of time. Let us give some examples: A seller may reduce service provision, which may have bigger long-term than short term effects on sales as the customer may consider to change supplier. The seller may also boost his sales in the short run by encouraging the buyer to buy an excessive amount. Again, this may come at the expense of a good customer relationship in the future. The quality of a product or a service may not be observable before a considerable amount of time has elapsed (Laffont and Tirole 1993 ch 4). Alternatively, effort may improve the quality of the pool of customers. For instance, a seller of loans may boost short term performance (the number of loans provided) by being less selective. Another example is that investments in new routines etc. may enhance productivity in the future.

As noted in Cahuc and Zylberberg (2004 p 353), there exist empirical studies that indicate that certain firms do in fact use deferred compensation. For instance, Lazear and Moore (1984) compare age-income profiles for tenured workers and for self-employ workers, for which there exists no

agency problems. They find that the returns to seniority is higher for tenured workers, and attribute this to deferred compensation. Katlikof and Gokhale (1992) study wages and productivity of more than 300000 workers in a Fortune 1000 firm. The identifying assumption for the productivity-age profile is that workers over their careers in the firm are paid a wage equal to their productivity in expected terms, and workers join the firm at different ages. For all categories of workers they find a substantial degree of deferred compensation. In particular, for 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) finds evidence that supports the theory of deferred compensation, as piece rate workers have negligible returns to seniority while workers that are not paid by piece rates have significant seniority effect. Finally, Dustmann and Meghir (2004) find substantial return to tenure.

Deferred compensation tends to imply that the discounted wage stream for experienced workers are higher than their discounted productivity stream, even if worker productivity peaks at this stage. Suppose there are no start-up costs for firms associated with hiring a person. Free entry of firms then implies that the discounted marginal value of the worker is equal to his discounted wage from that firm. If the compensation is deferred, this implies that at any point in time (except at the hiring point), the net present value of the wage stream exceeds that of his productivity stream. Hence, it is in the firm's interest that the worker quits. With start-up costs, the firm has to capitalize this cost, and the point in time in which the worker's expected wage stream exceeds his expected productivity stream is postponed.

We now turn to the details of the model. The economy consists of a continuum of workers and firms. The timing goes as follows:

1. Firms hire workers.
2. In period 1, workers choose effort  $e$ . We assume that  $e = 0$  or  $e = 1$ , and the cost of effort is  $c$ . Output this period is  $y_1 = y + e$ . The value of  $e$  can only be observed at the end of period 2.<sup>1</sup>
3. After period one, the worker decides whether to stay on in the firm or to quit.

---

<sup>1</sup>We assume that effort can be observed without noise at the end of period 2. Adding noise would further strengthen our results.



4. In the beginning of period two, employed workers do on-the-job search.
5. If on-the-job search is unsuccessful, he stays on in the firm. His output in period two is  $y_2 = y$ , and this is observable.
6. The period two wage can be made contingent on  $e$  if and only if the worker is employed in period 2 when  $e$  is observed.

In a more general setting, some dimensions of effort will be observable instantly, and can be rewarded by short-term bonuses. This extension will be analyzed in a later section.

As will be clear below, only "shirkers" (workers that do not exert effort in period 1) if any will quit after period one. We thus assume that the shirkers quit the job before the stage at which they do on-the-job search. This simplifies the algebra, but is not crucial for our results.

The most important assumption is that firms only pay a bonus to workers that are still employed when  $e$  is revealed. This may be rationalized in several ways. The first thing to note is that as firms have a large number of employees, stock options will be insensitive to the effort of any single employee. Thus, to incentivize workers by giving them stock options will be extremely costly (assuming limited liability). Furthermore, it may be hard for a worker to control that his former employer actually measures his long-term effect on productivity accurately without being in the firm. Thirdly, even if this control problem can be solved, including a payment based on period-two output when the worker has quit will, in a more general setting, lead to new distortions.

We will discuss the last point in some detail. Suppose the gain from  $e$  depends on the effort of the worker's predecessor. If the initial worker is residual claimant (or paid a share) of  $e$ , the firm and the new workers will not have (sufficient) incentives to fulfill the project. Thus, paying workers that quit on the basis of  $e$  is costly, as it distorts the incentives of his predecessor. We capture this by ruling out such payments entirely.<sup>2</sup>

---

<sup>2</sup>To formalise this point, suppose effort in period 1 consists of starting a project. In period 2, effort with effort cost  $c_2$  (observable) will increase the probability that the project is a success from  $1 - x$  to 1. Suppose  $x > c_2$ , so that inducing period 2 effort is efficient (the value of the project is 1).

Suppose the worker obtains a bonus  $w_2 = c$  if he quits the firm and the project succeeds. It is then in the firm's interest to induce period 2 effort iff  $(1 - c)x > c_2$ . Thus, period 2 effort will not be induced even if it is efficient if  $1 - c \geq c_2/x$ . In order to induce

Let  $w_1$  denote the wage in period 1,  $w_2$  the wage in period 2 if he stays in the firm, and  $w_p$  the wage obtained if search is successful. Let  $u_2(e)$  denote the expected income (utility) of a worker that does not quit to become self-employed in the beginning of period 2 (prior to on-the-job search). This utility can be written as

$$u_2 = \max_s sp(w_p - w_2(e)) + w_2 - g(s), \quad (1)$$

where  $s$  is the worker's search intensity and  $g(s)$  the cost of search. In what follows we assume that  $g(s) = s^2/2$ . The variable  $p$  reflects the probability of finding a job for a unitary search intensity. Both  $p$  and  $w_p$  are considered exogenous by individual workers and firms, but are determined endogenously in labour market equilibrium. Let  $U$  denote the utility of the worker in the beginning of period 1. Then

$$U(e) = w_1 - ce + \max[\bar{w}, u_2]. \quad (2)$$

where  $\bar{w}$  denote the (exogenous) pay-off to the worker if he quits after period 1.

## 2.1 Optimal wage contracts

The optimal wage contract  $[w_1, w_2, e^*]$  maximizes firm profits given the following conditions:

1. Worker participation constraint,  $U(e^*) \geq U^0$ , where the outside option  $U^0$  is endogenously determined in labour market equilibrium.
2. Incentive compatibility constraint:  $U(e^*) = \max_e U(e)$ .

Suppose first that firms want to implement high effort. Then the incentive compatibility constraint binds. Shirkers quit after period 1. Thus, a necessary condition for high effort is that the worker is better off exerting high effort and staying with the firm than shirking and quitting after period 1:

$$w_1 + \bar{w} \leq w_1 - c + u_2 \quad (3)$$

---

period two effort, the highest bonus that can be paid to the worker if he quits is such that  $(1 - w_2) = c_2/x$ . Thus, all the deferred payments cannot be handed out to a worker that quits, and our main result that deferred payments distort turnover is maintained.

where  $u_2$  is evaluated at  $e = 1$ .

Let  $\Omega$  denote the joint expected income to the worker-firm pair at the search stage. It follows that  $\Omega$  is given by

$$\Omega = sp(w_p - y) + y - g(s).$$

Optimal on-the-job search intensity  $s$  is the value of  $s$  that maximizes  $\Omega$ . However, the worker is the one that chooses  $s$  according to (1). It follows that the worker only maximizes  $\Omega$  if  $w_2 = y$ , in which case  $u_2 \equiv \Omega$ .

**Lemma 1** *Optimal on-the-job search requires that  $w_2 = y$ .*

In what follows, we assume that  $\Omega(y) - c < \bar{w}$ . It follows that if  $w_2 = y$ , workers shirk. Hence the following proposition must be true:

**Proposition 2** *Suppose  $\bar{w}$  satisfies the inequality stated above. Then high effort ( $e = 1$ ) is incompatible with efficient turnover (incompatible with  $w_2 = y$ ).*

Thus, in a high-effort equilibrium the firm has to increase  $w_2$  above marginal productivity  $y$  of the worker in order to incentivize the him, i.e., the firm has to defer compensation to the worker. This comes at a cost, as the turnover is distorted.

Let  $L$  denote the loss associated with an inefficient turnover (with  $s$  given by (1)). The loss is given by

$$L(w_2) \equiv \Omega(y) - \Omega(w_2). \quad (4)$$

We have that

$$\Omega(w_2) = \max_s [sp(w_p - w_2) + w_2 - g(s)] - (1 - ps)(w_2 - y). \quad (5)$$

In the appendix we show that as  $g(s) = s^2/2$ , we can write (see Appendix)

$$L(w_2; p) = p^2(w_2 - y)^2/2. \quad (6)$$

Note that  $L$  is increasing in both  $w_2$  and  $p$ .

As  $L$  is increasing in  $w_2$ , the incentive compatibility constraint binds, and the firm sets  $w_2$  such that  $u_2(w_2) = \bar{w} + c$ . The period 1 wage is then set

residually so as to satisfy the individual rationality constraint of workers. We assume free entry of firms, hence the period 1 wage can be written as

$$w_1 = y + \Omega(w_2) + e - u_2. \quad (7)$$

In what follows we assume that  $\Omega(y) - y > e - c > 0$ . Thus the value of effort is higher than the effort cost, but the net value is less than the (maximal) value of search.

## 2.2 The search market

The search market consists of experienced workers and firms which are particularly skillful in utilizing this experience. We refer to the firms as search firms. The productivity of an experienced worker in a search firm is  $y_p > y$ . The cost of opening a vacancy in a search firm is  $k$ . We assume that  $y_p - k$  is sufficiently high so that the search market is operating.

The value of a vacancy is given by the value of hiring a worker times the probability of finding one, that is

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

where  $q$  is the probability of obtaining a worker. The matching technology can be described as follows. Suppose  $u'$  workers search randomly for  $v'$  firms with vacancies. Then the matching function is captured by a reduced-form function  $q = \bar{q}(p)$ , relating the probability that a worker finds a job (per unit of search effort) and the probability of a vacancy is filled.<sup>3</sup> We require that  $0 < p, q < 1$  and that  $\lim_{p \rightarrow 1} \bar{q}(p) = 0$ .

Vacancies enter the market up to the point at which  $V = k$ , that is up to the point at which  $q(y_p - w_p) = k$ . The free entry condition defines a relationship between  $w_p$  and  $p$ ;  $w_p = w^{FE}(p)$ . It follows that the function  $w^{FE}(p)$  is decreasing in  $p$  and goes to minus infinity as  $p$  approaches 1 (since  $\lim_{p \rightarrow 1} \bar{q}(p) = 0$ ).

The equilibrium of the search market,  $(p, w_p)$ , can be modelled in different ways. We assume that the equilibrium maximizes the utility of the searching

---

<sup>3</sup>The probabilities  $p$  and  $q$  can be written as  $p = x(eu, v)/eu = x(1, \theta) = \tilde{p}(\theta)$  and  $q = x(eu, v)/v = x(1/\theta, 1) = \tilde{q}(\theta)$ . The matching technology can thus be summarised by a function  $q = \tilde{q}(\theta) = \tilde{q}(\tilde{p}^{-1}(p)) = \bar{q}(p)$ .

workers, and thus is defined as

$$\max_{w_p, p} u_2(w_p, p; w_2) \quad \text{S.T.} \quad V = k . \quad (8)$$

This equilibrium corresponds to the competitive search equilibrium derived in Moen (1997) and Acemoglu and Shimer (1999). The equilibrium can be the result of a market maker creating markets with different wages, or by wage advertisement by firms. The same equilibrium condition can be derived with wage advertisements with the urn-ball matching process (Montgomery 1991, Peters 1991, and Burdett et al. 2001). The equilibrium condition can also be derived with individual wage bargaining provided that a) the workers' outside option in the bargaining game is to continue working at wage equal to  $w_2$ , and b) that the Hosios condition is satisfied<sup>4</sup> (Hosios, 1990). From (8) it follows that we can write the equilibrium values of  $p$  and  $w_p$  as functions of  $w_2$ ;  $p = p^E(w_2)$  and  $w_p = w^E(w_2)$ . It is straight-forward to show that  $p^E(w_2)$  is strictly decreasing and  $w^E$  strictly increasing in  $w_2$ .

In our definition of equilibrium, given by equation (8), the wage  $w_2$  refers to the equilibrium wage in the market. We assume that a single firm takes the value of  $w_p$  and  $p$  as given, independently of how it sets  $w_2$ . Thus, if a firm sets a wage different from the equilibrium wage  $w_2$ , search firms cannot react by targeting jobs with a different wage attached to it towards this particular firm.

### 3 High-and low-turnover equilibria

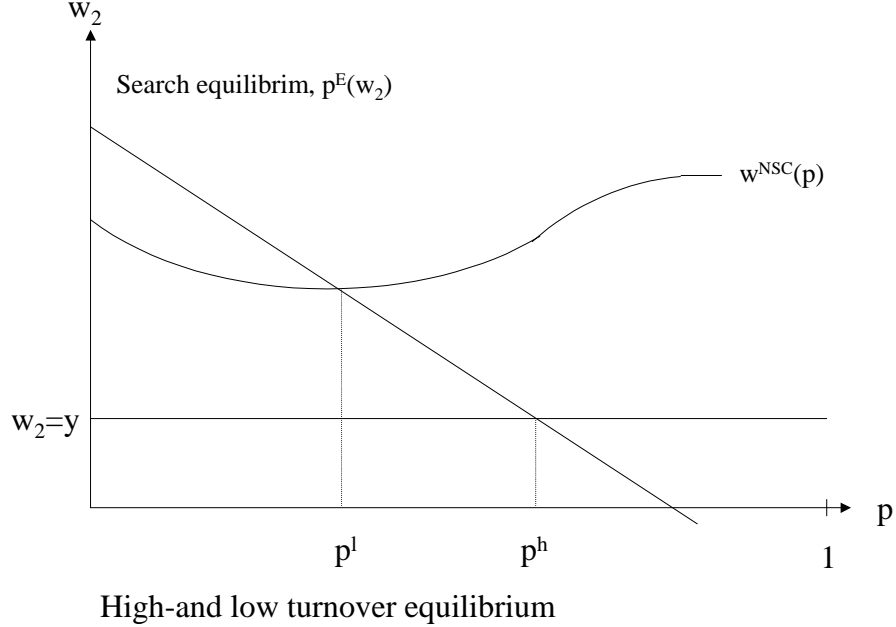
We first consider potential equilibria with deferred compensation. We referred to this as low-turnover equilibrium. Substituting the zero-profit condition of search firms into the incentive compatibility constraint (3) gives

$$\max_s [sp(w^{FE}(p) - w_2) - g(s)] + w_2 = \bar{w} + c, \quad (9)$$

which defines  $w_2$  as a function of  $p$ . We refer to this as the Non-Shiriking Condition, and write  $w_2 = w^{NSC}(p)$ . The NSC thus gives the wage  $w_2$  necessary to avoid shirking as a function of  $p$ . The low-turnover equilibrium  $(w_2^l, p^l, w_p^l)$  as follows:

---

<sup>4</sup>The Hosios condition states that the workers' bargaining power is equal to the absolute value of the elasticity of  $q$  with respect to the labour market tightness.



1. The NSC is satisfied:  $w_2^l = w^{NSC}(p)$
2. Competitive search equilibrium:  $p^l, w_p^l \in \max_{w_p, p} u_2(w_p, p; w_2)$  S.T.  $V = k$ .

The equilibrium is illustrated in figure 2

The equilibrium non-shirking constraint (NSC) obtains its minimum at the equilibrium point (see the appendix for a proof). At this point,  $u_2$  is maximized given  $w_2$ . As  $p$  increases,  $w_p$  decreases in order to satisfy the zero profit condition, and at some value  $p = p^{\max}$ , we will have  $w_p = w_2 = \bar{w} + c$ . At  $p = p^{\max}$  the turnover rate is zero and further increases in  $p$  does not influence  $w_2$ .

The next issue is whether the firm will find it in its interest to implement a high effort. The alternative for the firm is to set  $w_2 = y$  and implement  $e = 0$ . The gain by this strategy is that efficient turnover is obtained, and this increases firm profit through the workers' participation constraint. The

cost is the loss due to low effort, which is given by  $e - c$ . Above we defined the loss function related to inefficient turnover as a function of  $w_2$ , see equation (4). The equilibrium candidate  $(w_2^l, p^l, w_p^l)$  is an equilibrium if and only if

$$L(w_2^l, p^l) \leq e - c. \quad (10)$$

This equation will always be satisfied for sufficiently small values of  $c$ .

**Proposition 3** *For sufficiently low values of  $c$ , the low-turnover equilibrium exists.*

We now turn to the high-turnover equilibrium. In this equilibrium, the firms implement  $e = 0$  and obtain optimal turnover by setting  $w_2 = y$ , as explained above. The equilibrium candidate  $(w_2^h, p^h, w_p^h)$  is thus determined as follows:

1. Efficient turnover:  $w_2^h = y$
2. Competitive search equilibrium:  $p^h, w_p^h \in \max_{w_p, p} u_2(w_p, p; y)$  S.T.  $V = k$

This equilibrium is also illustrated in figure 2. For the high-turnover equilibrium to exist, the firms must prefer to implement  $e = 0$  to  $e = 1$ . In order to implement  $e = 1$ , a firm must satisfy NSC. For this not to be profitable, we must have that

$$L(w^{NSC}(p^h), p^h) \geq e - c. \quad (11)$$

For sufficiently high values of  $c$ , this condition is always satisfied.

**Proposition 4** *For sufficiently high values of  $c$ , the high-turnover equilibrium exists.*

**Proof.** Omitted ■

## 4 Multiple equilibria

In order to have multiple equilibria, both conditions (10) and (11) must be satisfied. We must thus have that

$$L(w^l, p^l) \leq e - c \leq L(w^{NSC}(p^h), p^h).$$

For this to hold, we must have that the loss  $L$  is higher for high values of  $p$  than for low values of  $p$ , which indeed is the case. More precisely, let  $L^{NSC}(p) \equiv L(w^{NSC}(p), p)$ . Recall that as  $L(w, p) = (w_2 - y)^2 p^2 / 2$  it is increasing in both arguments. Furthermore, we have already shown that  $w^{NSC}(p)$  is increasing in  $p$  on the relevant intervals. Thus,  $L^{NSC}(p)$  is increasing in  $p$  on the relevant intervals, and multiple equilibria exists.

**Proposition 5** *There exists an interval  $[\underline{c}, \bar{c}]$  for which the model exhibits multiple equilibria. One equilibrium has low turnover, high long-term effort, and deferred compensation. The other equilibrium has high turnover, no long-term effort, and no deferred compensation.*

**Proof.** This follows by construction if we let  $\bar{c}$  be the solution to the equation  $L^{NSC}(p^l) = e - \bar{c}$  and  $L^{NSC}(p^h) = e - \underline{c}$ . ■

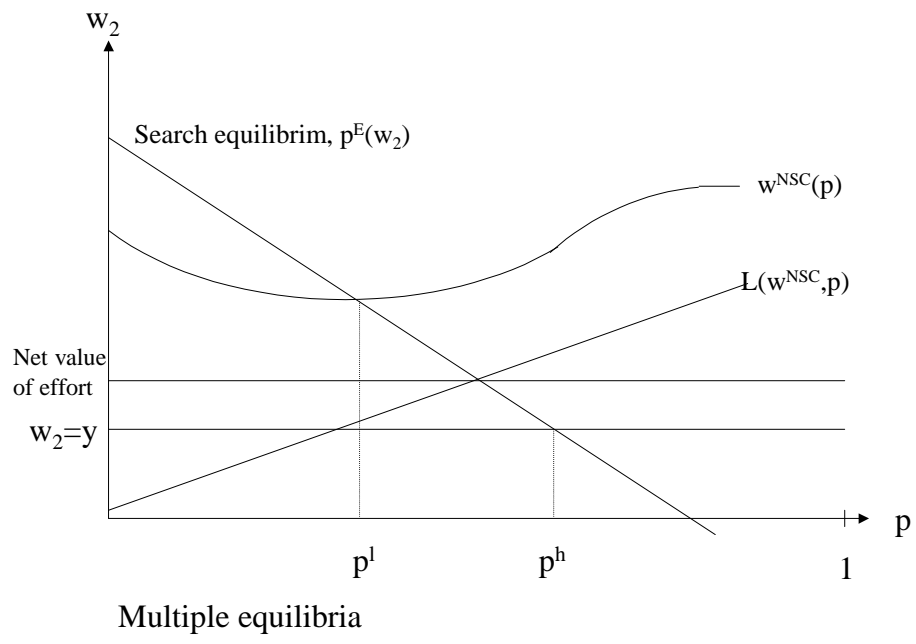
Multiple equilibria are illustrated in figure 3:

In this figure, the low-turnover equilibrium is determined as the intersection between the Non-Shirking Constraint  $w^{NSC}(p)$  and the search equilibrium line  $p^E(w_2)$ . At this point, the loss  $L$  due to low turnover is below  $e - c$ , hence it is optimal to implement high effort. The high-turnover equilibrium is defined as the intersection between the  $w_2 = y$  line and the search equilibrium line  $p^E(w_2)$ . At this point, the loss  $L$  of implementing a high effort caused by too low turnover exceeds the net value of effort  $e - c$ , hence it is optimal to implement low effort.

## 5 Welfare

In this section we analyze which of the equilibria that give rise to highest social welfare, (given that parameter values are such that multiple equilibria exist). As will be clear shortly, this depends on parameter values. Since we





assume free entry of firm, social welfare is simply defined by the expected utility of a worker,  $U^0$ . The starting point is to study the optimal value of  $p$  (that maximize the joint period two income  $\Omega(p)$ , given the free entry condition). In the regime without effort, where  $w_2 = y$ , it can be shown that  $p^h, w_p^h$  are optimal.<sup>5</sup> The less straightforward task is to determine the optimal value of  $p$  given that the NSC is satisfied. As will be clear shortly, that is not  $p^l$ .

Let  $\Omega^e(p)$  denote the value of  $\Omega(p)$  given that the worker is induced to exert high effort. Then the following holds

**Lemma 6** *Suppose  $\Omega^e(p^h) > \Omega^e(p^l)$ . Then the high-turnover equilibrium gives rise to higher welfare than the low-turnover equilibrium.*

**Proof.** From (2) and (7) follows that  $U^0(p^h) = y + \Omega(p^h, y)$  and that  $U^0(p^l) = y + e - c + \Omega^e(p^l)$ . In the high-turnover equilibrium  $L(p^h) = \Omega(p^h, y) - \Omega^e(p^h) > e - c$ . We have

$$U^0(p^l) = y + e - c + \Omega^e(p^l) < y + e - c + \Omega^e(p^h) < y + \Omega(y, p^h) = U^0(p^h).$$

As welfare depends on  $U^0$  only, the Lemma follows. ■

We want to study  $\Omega^e(p)$  in some more detail.

**Lemma 7**  *$\Omega^e(p)$  is increasing in  $p$  at  $p = p^l$  and decreasing in  $p$  at  $p = p^h$ .*

**Proof.** Consider  $p = p^l$ : Using equation (5) gives

$$\begin{aligned} \Omega^{e'}(p^l) &= \frac{d}{dp^l} \left( \max_s [sp^l(w_p - w_2) + w_2 - g(s)] - (1 - p^l s)(w_2 - y) \right) \\ &= w_2 - y > 0. \end{aligned}$$

Here we have used that  $\frac{ds}{dp^l} = 0$  and that the equilibrium in the search market as defined by (8) implies that  $\frac{d}{dp^l} (\max_s [sp^l(w^{FE}(p) - w_2) + w_2 - g(s)]) = 0$ .

Consider  $p = p^h$ : Using (4) and Lemma 6 gives that  $\frac{d}{dp} L(y, p^h) = \frac{d}{dp} [\Omega(y, p^h) - \Omega^e(p^h)] > 0$ . Due to the envelope theorem,  $\frac{d}{dp} \Omega(y, p^h) = 0$ , hence we must have that  $\frac{d}{dp} \Omega^e(p^h) < 0$ . ■

---

<sup>5</sup>See Moen and Rosén (2003) for a proof.

**Lemma 8** *Suppose  $s$  is exogenous. Then the high-turnover equilibrium always welfare dominates the low-turnover equilibrium.*

**Proof.** When  $s$  is exogenous, the optimal value of  $p$  is  $p^h$ , independently of  $w_2$  and therefore  $\Omega^e(p^h) > \Omega^e(p^l)$ . The above lemma then follows directly from lemma (8). ■

As before  $\bar{c}$  denotes the highest value of  $c$  for which the low-turnover equilibrium exists and we have the following Lemma.

**Lemma 9** *Suppose  $c$  is close to  $\bar{c}$ . Then the high-turnover equilibrium gives rise to a higher welfare than the low-turnover equilibrium.*

**Proof.** It follows that  $e - c$  is only marginally larger than  $L(p^l) = \Omega(y, p^l) - \Omega^e(p^l)$ . As  $U^0(p^l) = y + e - c + \Omega^e(p^l) \approx y + \Omega(y, p^l) < y + \Omega(y, p^h) = U^0(p^h)$ . ■

The opposite does not hold. It may be that if  $c$  is close to  $\underline{c}$ , the high-turnover equilibrium still dominates the low-turnover equilibrium. (Recall that  $\underline{c}$  denotes the lowest value of  $c$  for which the high-turnover equilibrium exists.) The reason is that it may still be that  $\Omega^e(p^h) > \Omega^e(p^l)$ . Thus, although inducing long-term effort will not reduce  $U^0$  at  $p = p^h$ , it may reduce  $U^0$  to move from  $p = p^h$  to  $p = p^l$ .

The latter result may indicate that the low-turnover equilibrium is never optimal. However, this is not the case. Consider the case in which  $s$  is either 0 or 1. Then it may well be that the worker will not search at all if offered a high-effort contract in the high-turnover equilibrium, while they will search in the low-turnover equilibrium.

**Lemma 10** *Suppose  $s$  takes two values 0 and 1 only, Then there exist values of  $e - c$  such that 1) there exists multiple equilibria, 2) the low-turnover equilibrium yields higher welfare than the high-turnover equilibrium welfare.*

**Proof.** The high-turnover will exist as long as  $\Omega(y, p^h) > e - c$  (which is satisfied by assumption). The low-turnover equilibrium yields higher welfare than the high-turnover equilibrium if  $\Omega(y, p^h) - \Omega(y, p^l) < e - c$ . Since  $\Omega(y, p^l) > 0$ , such values of  $e - c$  do exist. ■

## 6 Implications

In this section we study short-term effort, investments in firm specific human capital and supply of entrepreneurs.

### 6.1 Short versus long-term incentives

In this subsection we will elaborate slightly on the wage contracts, and introduce effort that is immediately observable to the firm in addition to effort that is observable at the end of period two. We will demonstrate that firms will provide stronger short-term incentives in the high-turnover equilibrium than in the low-turnover equilibrium. For a more careful treatment of wage contracts, we refer to Moen and Rosén (2004).

As before, we let  $e \in \{0, 1\}$  denote effort that is observable in period 2. Effort that is immediately observable is denoted by  $e_s$ . We assume that  $e_s$  is continuous. The disutility of effort is a function of total effort,  $\tilde{c}(e, e_s)$  where  $\tilde{c}$  is increasing and convex. Furthermore, we assume that that increasing effort along one dimension increases the marginal cost of effort along the other dimension. More specifically, we assume that  $c(e_s) = \tilde{c}(1, e_s) - \tilde{c}(0, e_s)$  is increasing in  $e_s$  (note that  $c$  corresponds to our previous effort cost, however, now it depends on  $e_s$ ).

In the equilibrium with  $e = 1$ , the firms will take into account that increased observable effort will make it harder to implement a high level of  $e$ ; the degree of deferred compensation will increase and thereby also the deadweights loss associated with distorted turnover decisions. Let  $w_2(e_s; p)$  denote the lowest wage that satisfy the I.C. constraint. In the low-turnover equilibrium, the firms are thus maximizing

$$\pi = y + e_s + e + \Omega(w_2(e_s; p)) - U^0 - \tilde{c}(e_s, 1).$$

From equation (5) it follows that  $\frac{ds}{dw_2} = -p$ . From the incentive compatibility constraint (3) and (5) it follows that

$$\frac{dw_2(e_s; p)}{de_s} = \frac{c'(e_s)}{1 - sp}$$

First order conditions for  $e_s$  can thus be written as

$$c(e_s) = 1 - \Omega'(w_2(e_s; p)) \frac{dw_2(e_s; p)}{de_s}$$

$$\frac{\partial \tilde{c}(e, e_s)}{\partial e_s} = 1 - \frac{pc'(e_s)}{1 - ps} < 1.$$

The next proposition follows:

**Proposition 11** *In low-turnover equilibrium, workers are given less than full incentives to exert short-term effort. Furthermore, the higher is the turnover rate  $p$ , the lower are the incentives to exert short-term effort*

The last part of the proposition is an artifact of the assumption that  $e$  is either 0 or 1.

In a high-turnover equilibrium with  $e = 0$ , the firm chooses  $e_s$  so as to maximize

$$\pi = y + e_s + \Omega(y) - U^0 - \tilde{c}(e_s, 0),$$

with first order condition  $\tilde{c}_{e_s}(e_s, 1) = 1$ . The firm thus chooses the first best level of observable effort (given that  $e = 0$ ).

We have thus shown the following proposition:

**Proposition 12** *In the high-turnover equilibrium, workers are given full incentives to exert short-term effort, that is,  $\tilde{c}_{e_s}(e_s) = 1$ . Thus, short-term incentives are stronger in the high-turnover equilibrium than in the low-turnover equilibrium.*

## 6.2 Firm-specific human capital

In this subsection we introduce firm-specific human capital. Not surprisingly, we find that firm-specific human capital investments are more profitable in low-turnover equilibrium than in high-turnover equilibrium. The reason for this is twofold. Firstly, with a low turnover the worker-firm pair is more likely to benefit from the investments. Secondly, investments in firm-specific human capital reduces the inefficiencies associated with deferred compensation in the low-turnover equilibrium.

Let  $a \in \{0, \bar{a}\}$  denote firm-specific human capital. The cost of investing  $a$  units is denoted by  $d$ . In a high-turnover equilibrium, firms will set  $w_2 = y + a$  in order to obtain optimal turnover. Hence firms invest in firm-specific human capital whenever

$$\Omega(y; p^h) < \Omega(y + \bar{a}; p^h) - da.$$

In the low-turnover equilibrium, the period 2 wage will be determined by the NSC whenever this gives a higher wage than  $y + a$ , otherwise the wage will be set equal to  $y + a$ . It follows that in the low-turnover equilibrium, firms invest in firm-specific human capital if and only if

$$\Omega(w_{a=0}^{NSC}, p^l) < \Omega(w_{a=1}^{NSC}, p^l) - da.$$

We are now able to show the following proposition:

**Proposition 13** *The return to firm specific human capital investments are higher in low-turnover equilibrium than in high-turnover equilibrium.*

**Proof.** The return to human capital in high-turnover equilibrium is given by  $\Delta^h(a) = \Omega(y + a; p^h) - \Omega(y; p^h)$ . Taking the derivative with respect to  $a$  gives

$$\Delta^{h'}(a) = (1 - p^h s^h).$$

In the low-turnover equilibrium. Firms set  $w_2(a) = \max[\hat{w}_2(a), y + a]$ , Let  $\Delta^l(a)$  denote the associated pay-off from firm-specific human capital. It follows that

$$\Delta^{l'}(a) = (1 - p^l s^l) > \Delta^{h'}(a),$$

for  $a \in \{0, \bar{a}\}$ . ■

### 6.3 Entrepreneurship and venture capital

Entrepreneurs are frequently former employees of firms in the same industry. Furthermore, entrepreneurs often need access to particular kinds of funding, for instance venture capital, for which the market may be thin. The matching process between venture capitalists and entrepreneurs can be similar to the search market described above, however, venture capitalists play

a more active role in searching for entrepreneurs relative to the situation in the employment search market.

For a potential entrepreneur, the shadow price of becoming an entrepreneur is continued employment. With deferred compensation, this shadow price is likely to be higher than without deferred compensation. Furthermore, when bargaining over terms of trade with a venture capitalist, the economic compensation of continued employment is likely to influence a potential entrepreneur's bargaining position. Thus, in low-turnover equilibrium with deferred compensation, entrepreneurship will be less attractive, because the shadow price, in terms of foregone wages, is high. This effect may also be reinforced as a lower number of entrepreneurs will enter the market. Furthermore, to the extent that the expected gains associated with entrepreneurship are large, this may tend to magnify the effects of deferred compensation: in low-turnover equilibrium, few venture capitalists enter the market, hence the loss of deferred compensation caused by reduced entrepreneurship is small. In the high-turnover equilibrium, by contrast, a large number of venture capitalists enter the market, and distortions associated with low entrepreneurial activity are large.

There exists a small literature on multiple equilibria in the supply of entrepreneurs. Other papers include Landier (2002), and Gromb and Scharfstein (2001), who explain cross-country differences in entrepreneurship by externalities. These papers obtain multiple equilibria because of differences in social norms and managerial talents. In contrast, we let firms choose between short-term and long-term incentive contracts.

## 7 Conclusion

In this paper we have analyzed the relationship between the firms' choice of wage contracts and the turnover rate among experienced workers. Furthermore, we have shown that feedback effects from the firms choice of wage contracts to the labor market for experienced workers may create multiple equilibria with different turnover rates. In low-turnover equilibrium, firms offer long-term contracts with deferred compensation, and workers exert long-term effort, while the opposite holds in the high-turnover equilibrium. It is tempting to think of low-turnover equilibrium as Japan and Germany and high-turnover equilibrium as US and UK. Alternatively, low-turnover equilibrium may be thought of as Massachusetts and low-turnover equilibrium as

Silicon Valley.

Our model has several empirical implications. In low-turnover equilibrium, firms are inclined to give weaker incentives to short-term performance (less use of short-term bonuses), rely more on deferred compensation and invest more in firm-specific human capital than in the high-turnover equilibrium. Furthermore, entrepreneurship and venture capital may be more frequent in high-turnover equilibrium than in low-turnover equilibrium. The implications seem to be in line with popular conceptions of differences between US and Japan and between Silicon Valley and Massachusetts.

## 8 Appendix

*Derivation of loss function*

With  $g(s) = s^2/2$  and equation (1) the worker's choice of  $s$  is given by

$$s = p(w_p - w_2).$$

It follows that

$$\begin{aligned}\Omega(w_2) &= p^2(w_p - w_2)^2/2 + w_2 - [1 - p^2(w_p - w_2)](w_2 - y) \\ &= \Omega(y) - p^2(w_2 - y)^2/2,\end{aligned}$$

and thus that  $L(w_2; p) = p^2(w_2 - y)^2/2$ .

$w^{NSC}(p)$  obtains its minimum at  $p = p_l$ .

Suppose there exists a point  $w'$  at the non-shirking line such that  $w' < w_2$ , and denote by  $p'$  the corresponding value of  $p$ . By the definition of the NSC, we must then have that  $u_2(w') = u_2(w_2)$ . Then it follows that  $u_2(w_2, p') > u_2(w', p')$  (since  $w_2 > w'$ ). But then  $p'$  cannot solve the maximization problem (8).

## References

Acemoglu, D. and Pischke, J-S (1998), "Why Do Firms Train?". Theory and Evidence". *Quarterly Journal of Economics* 113, 79-119.



- Acemoglu, D. and Shimer, R. (1999), "Efficient Unemployment Insurance", *Journal of Political Economy*, 107, 893-928.
- Akerlof, G.A., and Katz, L.F. (1989), "Workers' Trust Funds and the Logic of Wage Profiles". *Quarterly Journal of Economics* 104, 525-536.
- Aoki, M. (1990), "Toward an Economic Model of the Japanese firm". *Journal of Economic Literature*, 28, 1-27.
- Bayo-Moriones, A., Galdon-Sanchez, J.E., and M. Guell (2004), "Is Seniority-Based Pay Used as a Motivation Device? Evidence from Plant Level Data". IZA Discussion Paper no 1321.
- Burdett, K., and Coles, M.(2003), "Equilibrium Wage Tenure Contracts". *Econometrica* 71, 1377-1404.
- Burdett, K., and Mortensen, D.T. (1998), "Wage Differentials, Employer Size, and Unemployment". *International Economic Review* 39, 257-273.
- Burdett, K., Shi, S. and Wright, R. (2001), "Pricing and Matching with Frictions", *Journal of Political Economy*, 109, 1060-1085.
- Cahuc, P., and A. Zylberberg (2004), *Labour Economics*. The MIT Press, Cambridge, Massachusetts and London, England.
- Chang, C. and Wang, Y (1995), "A Framework for Understanding Differences in Labor Turnover and Human Capital Investments". *Journal of Economic Behaviour and Organization*, 28, 91-105.
- Diamond, P.A. (1982), "Wage Determination and Efficiency in Search Equilibrium", *Review of Economic Studies* 49, 217-227.
- Dustmann, C., and Meghir, C. (2004), "Wages, Experience and Seniority". *Review of Economic Studies*, forthcoming.
- Fallick, B.C., and Fleischman, C.A. (2001), "The Importance of Employer-to-Employer Flows in the U.S. Labor Market". Federal Reserve Board mimeo April.

- Fonseca, R., Lopez-Garcia, P., and P.A. Pissarides (2002), "Entrepreneurship, start-up costs and employment", *European Economic Review*, 45, 692-705.
- Gromb, D., and D. Scharfstein (2001), "Entrepreneurial Activity in Equilibrium". Miemo, MIT.
- Hamilton, B.H. (2000), "Does Entrepreneurship Pay? An Empirical Analysis of the Return to Self Employment". *Journal of Political Economy*, 108, 604-631.
- Hellman, Thomas (2001), "When do Employees Become Entrepreneurs?" Incentives, corporate strategy, and intellectual property rights. WP, Stanford University.
- Hosios, A.J. (1990), "On The Efficiency of Matching and Related Models of Search and Unemployment", *Review of Economic Studies*, 57, 279-298.
- Johnson, W.R. (1978), "A Theory of Job Shopping". *Quarterly Journal of Economics*, 92, 261-277.
- Jovanovic, B. (1979), "Job Matching and the Theory of Turnover". *Journal of Political Economy*, 87, 972-990.
- Kiyotaki, N., and Lagos, R. (2004), "A Model of Job and Worker Flows". Mimeo, London School of Economics and New York University.
- Kotlikoff, L.J., and Gokhale, J (1992). "Estimating a Firm's Age-Productivity Profile Using the Present Value of Workers' Earning". *The Quarterly Journal of Economics* 107, 1215-42.
- Laffont, J.J. and Tirole, J (1993), *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge.
- Landier, A (2002), "Entrepreneurship and the Stigma of Failure". Working paper, University of Chicago.
- Lazear, E.P (1979). "Why is there Mandatory Retirement?". *Journal of Political Economy*, 87, 1261-1264.
- Lazear, E.P. (1981), "Agency, Earnings Profiles, Productivity, and Hours Restrictions". *American Economic Review*, 71, 606-620.

- Lazear, E.P. and Moore, R.L. (1984). "Incentives, Productivity, and Labor Contracts". *The Quarterly Journal of Economics* 99, 275-296.
- Lazear, E.P. and Rosen, S. (1981), "Rank Order Tournaments as Optimum Labor Contracts". *Journal of Political Economy*, 89, 841-864.
- Malcomson, J. (1984), "Work Incentives, Hierarchy, and Internal Labor Markets" *Journal of Political Economy*, 92, 486-507.
- Manning, A. (2003), *Monopsony in Motion: Imperfect Competition in Labour Markets*. Princeton University Press, Princeton
- Moen, E.R. (1997), "Competitive Search Equilibrium". *Journal of Political Economy*, 105, 385-411.
- Moen, E. R., and Rosén, Å (2003), "Does Poaching Distort Training?", forthcoming, *Review of Economic Studies*.
- Moen, E.R. and Rosén, Å (2004), "Long-run versus short-run incentive contracts". Working paper.
- Moene, K.O., and Wallrestein, M. (1997), "Full Employment as a Worker Discipline Device." In *Property Relations, Incentives and Welfare*. Ed J. Roemer, London: MacMillan Press.
- Montgomery, J.D. (1991), "Equilibrium Wage Dispersion and Interindustry Wage Differentials", *Quarterly Journal of Economics*, 106, 163-179.
- Morita, H (2001), "Choice of Technology and Labour Market Consequences". *Economic Journal*, 111, 29-50.
- Mortensen, D.T (1986), "Job Search and Labour Market Analysis". In O.C. Ashenfelter and R. Layard (eds), *Handbook of Labor Economics* Volume 2, Amsterdam, North-Holland, 849-919.
- Mortensen, D.T., and Pissarides, C.A. (1994), "Job Creation and Destruction in the Theory of Unemployment" .. *Review of Economic Studies*, 61, 397-415.
- Pissarides, C.A (1994), "Search Unemployment with On-the-Job Search". *Review of Economic Studies* 61, 457-475.

- OECD (1993), *Employment Outlook*. ch 4. Paris
- Peters, M. (1991), “Ex Ante Price Offers in Matching Games: Non-Steady States”, *Econometrica*, 59, 1425-1454.
- Pissarides, C.A. (2000), *Equilibrium Unemployment Theory*. MIT press, Cambridge, Massachusetts.
- Saxenian, A.L. (1994), *Regional advantage: culture and competition in Silicon Valley and Route 128*. Cambridge, Mass.: Harvard University Press.
- Shapiro, C, and Stiglitz, J.E. (1984), “Equilibrium Unemployment as a Worker Discipline Device”, *American Economic Review*, 74, 433-444.