Delegated Job Design*

Hans K. Hvide and Todd R. Kaplan†

December 5, 2005

Abstract

We develop a theory of delegation within organizations where agents are privately informed about whether they should be engaged in exploitation or in exploration activities. Excessive delegation lead agents to inefficiently herd into exploration in an attempt to boost their market value. The theory is consistent with both high-delegation practices and practices where agents are assigned to activities. Our main result is that an agent should be delegated more the weaker career concerns, a variable that is made endogenous through the firm's technology and its degree of transparency. The theory sheds light on empirical regularities that are previously unexplained, such as a positive relation between wages and delegation, and delegation being more prevalent in closed environments or environments with long-term employment contracts.

Keywords: Career concerns, Decentralization, New work practices, Sun Hydraulics, Wage inequality.

JEL#: C72, D23, D44, D82, J33, M12.

*We thank seminar participants at Aberdeen, Bristol, Carlos III, EARIE, Exeter, ESSLE 2003, Keele, London, Oslo, Paris, SOLE 2004, Stanford, and Tel-Aviv.

†Hvide: Norwegian School of Economics and Business, and CEPR. Kaplan: Department of Economics, University of Exeter.

1 Introduction

Who should decide which tasks or projects an employee engages in? The bureaucratic approach to job design is to first define the activities contained in the job slots and then to hire suitable employees to fill those slots, giving the employees limited leeway. The bureaucratic approach is thoroughly investigated in the principal-agent literature, which emphasizes centralized decision-making (Holmstrom & Milgrom, 1991, and Baker, 1992). An ample empirical literature (Osterman, 1994, Ichniowski et al., 1996, Baron & Kreps, 1999, Caroli et al., 2001, OECD 1999, and Rajan & Wulf, 2003) emphasizes decentralization of decision-making authority and autonomy, or delegated solutions to the job design problem. Two extreme but illustrating examples are the US companies Sun Hydraulics, where agents have "the right and responsibility to choose how they spend agents time" (Kaftan, 1984/1991) and Gore & Ass., producer of Gore-Tex© products, which encourages "maximum freedom for each employee" (Gore, 1990).

The purpose of the paper is to suggest a theory that addresses the question of why delegation occurs and what poses limits to delegation. The theory is based on the following trade-off: The upside to delegating more is that it enables the employee to use his private information when allocating effort between activities. For example, a worker may know better than his superiors whether he should be engaged in more routine tasks such as product updating or testing, or in more challenging tasks such as product innovation and design. Or a manager may know better than his superiors whether he has the skills to lead a new and bold research and development project.

The downside to delegating more is that the agent may use the freedom to reap private benefits through enhancing his market value. For example, if the most able agents in a hi-tech firm are engaged in product development, then less able agents engaging in product development may have better future prospects than if engaged in product updating, if the market views job description as an indicator of ability. Or if more able managers tend to initiate bolder projects, then less able managers may be tempted to act similarly.

How much should firms delegate given these two opposing effects? To address this question, we build upon a variation of the Roy (1951) model. Agents are of two possible types, privately known to them. In each firm there are two activities, the easy/exploitation activity and the difficult/exploration activity. An efficient allocation implies that low agents exploit and the high ability agents explore. By job design we mean the decision about which activity an agent should engage in (the activities are mutually exclusive).

When career concerns are weak (its determinants are discussed later) firms fully delegate the job design decision to the agent, and structure a compensation scheme so that agents do so efficiently. Such a scheme involves paying the low type a premium, i.e., pay above marginal productivity, to self-sort into the easy activity, since working in that activity will be a bad signal about ability. When career concerns are strong, however, full delegation would require a self-sorting premium so large that outside firms could cream-skim the high type agent, and a high degree of delegation would be unprofitable. Under strong career concerns, the principal assigns the agent to an activity, even if this entails that the agent's local knowledge is left unused. The principal assigns more the stronger career concerns, and more assignment implies a less efficient allocation of talent inside the firm. Therefore more delegation is associated with higher average wages. This agrees with recent empirical literature (Caroli & Van Reenen, 2001, Black et al., 2003, and Bauer & Brender, 2003) and contrasts with other theoretical literature such as Aghion & Tirole (1997).

We find that assignment will take the form of placing the agent in the difficult activity. The economic intuition for this is quite simple. The principal uses the degree of delegation as a tool in recruiting able agents. By restricting the agent to do the difficult activity, the firm avoids being stuck with low talent. Hence we obtain a theory that explains limits to delegation, but also how firms can use job design to attract talent.

The job design literature, Holmstrom & Milgrom (1991), Prendergast (1999), and Olsen & Torsvik (2000), among others, asks which combination of tasks should be included in the description of a job. Instead we focus on who should decide, the agent or the principal. Similarly, the job assignment literature, including Rosen (1982), Gibbons & Waldman (1999a) and Ortega (2001) considers settings with symmetric information between superiors and subordinates, a circumstance under which there would be no advantage of delegating. The same point applies to the literature on career concerns, as in Fama (1980) and Holmstrom (1982/1999). In the literature on adverse selection in labor markets (for example, Greenwald, 1986, Foster & Rosenzweig, 1996, and Acemoglu

& Pischke 1998) workers have private information prior to hiring, but ability is assumed to be revealed to the firm once the worker is hired. In contrast, firms in our setting face adverse selection both at the hiring stage and when allocating workers inside the firm.

The paper proceeds as follows. Section 2 introduces the basic model. Section 3 contains the simplified analysis without performance contracts. Section 4 contains a discussion of applications and modeling assumptions, and Section 5 concludes. The Appendix contains the analysis with optimal performance contracts.

2 The model

There is one industry with two active firms in each of two periods and one agent. Firms and the agent are risk neutral and there is no discounting. The agent has two possible types, $t \in T = \{L, H\}$. The agent knows his type while firms have the prior θ that the agent is H. In each (identical) firm there are two activities; the "easy" (E) and the "difficult" (D).[1] The two activities can be interpreted as different projects (like different R&D projects), different tasks (like product updating versus product development), or different methods in performing a certain task (like different types of analysis). Both types have productivity π^0 in the E activity. In the D activity, the low type has expected productivity π^L, and the high type has expected productivity π^H, where $\pi^L < \pi^0 < \pi^H$. We think of the E activity as having lower complexity and being more prone to standardization than the D activity. By job design we mean the decision about which activity the agent should specialize in.

Firms have a one-period lifespan and write one-period contracts.[2] The observed productivity of the agent in period i equals π_i. Contracts in each period specify a wage w and the degree of delegation d as a function of the observables, i.e., the history of the

[1] A quotation from March (1991, p. 71) illustrates well the distinction we have in mind: "Exploration includes things captured by terms such as search, variation, risk taking, experimentation, play, flexibility, discovery, innovation. Exploitation includes such things as refinement, choice, production, efficiency, selection, implementation, execution."

[2] Whether the first-period firms and the second-period firms are identical or not is immaterial for the analysis (as long as information among firms is symmetric and contracts are still one-period in length).

agent, his current-period performance π_i and his current-period activity.[3] The degree of delegation d is the key endogenous variable as to who makes the decision about what project the agent engages in, the principal or the agent (d is probability that the agent will decide).

In the basic analysis of Section 3, firms can not condition wages on the agent's output (think of this as a scenario where output measures are very noisy). This will highlight the role of delegation and activity choice in determining wages. First-period contracts must then be a wage w_1^D for the D activity, a wage w_1^E for the E activity, both non-negative, and the degree of delegation $d \in [0, 1]$. Since there is no incentives for misrepresentation in the second period, we assume for convenience that the agent chooses his efficient activity in the second period, so that a contract offer is simply a wage $w_2(I)$, where I is the information generated in the first period. In the Appendix, we solve a general model where pay can be conditioned upon performance.

The timing goes as follows. In the beginning of the first period, the firms simultaneously offer contracts $\{w_1^D,\ w_1^E, d\}$, and nature then chooses the type of the agent. The agent then learns his type. Having observed the contract offers and his type, the agent accepts to work for one of the firms. That firm is then free to improve its offer to its workers by proposing $\{w\text{'}_1^D,\ w\text{'}_1^E, d\text{'}\}$, where $\{w\text{'}_1^D, w\text{'}_1^E, d\text{'}\} \geq \{w_1^D, w_1^E, d\}$. This improved offer is valid only for the workers who accepted the original offer (workers can't switch firms at this stage). It is then determined whether the agent is delegated the decision of which activity to undertake or not (delegation occurs with probability d). The agent is then allocated to or chooses an activity, according to the outcome, and the wage is paid. The first period ends with all actions in the first period becoming public knowledge (the agent may still know his type privately). In the beginning of the second stage, the firms simultaneously offer the agent a contract conditional upon their mutual information.

A strategy for the first-period firms is a vector of initial offers $\{w_1^D,\ w_1^E,\ d\}$ and interim offers $\{w\text{'}_1^D,\ w\text{'}_1^E,\ d\text{'}\}$, and a strategy for the second-period firms is a wage $w_2(I)$. A strategy for the agent is a mapping from T and the contract offers to a choice of which contract to accept in the first period, a mapping from T and the renegotiated offer to

[3] We assume that a message game is too costly to implement. This assumption can be justified by intrinsic cost of communication (Dessein, 2002), or because the principal lacks commitment power.

which activity to choose (if given the choice), and a mapping from T and wage offers in the second period to a choice of which offer to accept. The agent maximizes his lifetime income. If the agent is indifferent between which firm to work for, we assume that he flips a fair coin to make a decision. We focus on Sequential Equilibria (SE).[4] That a strategy tuple is a SE means that: (i) the agent chooses a firm and activity optimally in each period, (ii) the firms anticipate the agent's behavior and chooses wage offers to maximize profits, (iii) the firm's posterior beliefs are formed using Bayes' rule whenever possible, and (iv) the equilibrium is a limit of one where beliefs are restricted to be interior. In terms of equilibrium selection, we focus on first-period initial wage offers that are "time-consistent" in that firms do not wish revise them in the interim. Informally speaking, by focusing on time consistent offers we exclude "bluff" offers.

Some comments to the setup. (i) The firms' option to raise any of the wages after the agent has accepted the contract is a natural requirement, since it allows for the firm and the agent to find a better contract once the agent has started in the firm. To focus on equilibria that are immune to such renegotiations (time-consistent) is fairly standard in the contracting literature (e.g., Krasa & Villamil, 2000). (ii) One may ask how the firm can commit to play essentially a mixed strategy in the interim (the decision as to whether delegation occurs). The most reasonable interpretation is that the firm hires several agents, so that we can interpret d as the fraction of agents that get the freedom to choose. (iii) The spirit of wage setting in the model is similar to compensating wage differentials (Rosen, 1986, and Stern, 2004), except that we have such differentials inside the firm rather than in the market. (iv) We have assumed that information about the agent is symmetric between the firms. Waldman (1983) and several later papers consider the possibility of asymmetric learning, where the inside firm learns more than the outside firm about the agent before the second period. The asymmetric learning case of our model was considered in the working paper version of the paper, and some of the results are referred to later.

[4] Employing Perfect Bayesian Equilibrium would yield identical results.

3 Analysis

First we focus on full-delegation, *separating equilibria*, where $d = 1$ and both types self-sort to their efficient activity in period 1. Then we analyze assignment equilibria, where $d < 1$ and the agent (with probability $1 - d$) is assigned to an activity. To break ties, we assume for convenience that the agent has lexicographic preferences over wages and activity: if the wage offers are such that he is monetary indifferent between the E activity and the D activity (taking into account the implicit incentives) then he prefers the efficient activity.[5]

3.1 Separating equilibrium

Proposition 1 *If $\pi^H - \pi^0 < \pi^0 - \pi^L$ then the unique time-consistent SE is a separating equilibrium with $d = 1$.*

Proof. In the conjectured equilibrium, first-period firms offer $\{w_1^E,\, w_1^D,\, d = 1\}$ in the beginning of the first stage, the agent chooses a firm randomly after learning his type, the firms do not adjust wages in the interim, and the agent chooses his efficient activity. In the second period, firms offer w_2^D (w_2^E) to if the agent was engaged in the D (E) activity in the first period. In the interim, the firm believes that the agent is H with probability θ if he accepted the offer. The second-period firms believe that the agent is H (L) with probability 1 if he was engaged in activity D (E) in the first period.

We first derive wages in the conjectured separating equilibrium, denoted by $\{w_1^E,\, w_1^D,\, w_2^E,\, w_2^D\}$, then prove existence, and finally uniqueness. For simplicity of exposition, we set $\theta = 1/2$ (the full expressions appear in the Appendix). If the agent is engaged in his efficient activity in the first period, his type is revealed before the second period. By Bertrand competition between second-period firms, he must be paid his marginal

[5] This tie-breaking rule is used quite extensively in the current section, where both types will be monetary indifferent between the wage contracts offered. This property of the equilibria constructed disappears with optimal performance contracts, as shown in the appendix. So the right interpretation of the model in the current section is as a convenient limit case of a more general model with imperfect monitoring.

productivity in the second period, i.e., $w_2^E = \pi^0$ and $w_2^D = \pi^H$.[6] For separation to occur in the first period, we have two IC conditions. The lifetime utility of a low type must be maximized from choosing E, i.e., $w_1^E + w_2^E \geq w_1^D + w_2^D$, and the lifetime utility for the high type must be maximized from choosing D, i.e., $w_1^D + w_2^D \geq w_1^E + w_2^E$. Using the second period wages, we therefore have that $w_1^E - w_1^D = \pi^H - \pi^0$ for both IC conditions to hold. A firm must make zero profits in expectation in the first period, i.e., $(\pi^H - w_1^D) + (\pi^0 - w_1^E) = 0$. The only first period wages consistent with self-sorting and zero profits are therefore $w_1^D = \pi^0$ and $w_1^E = \pi^H$. We have then derived wages assuming existence. We now prove existence of a separating equilibrium.

The IC conditions in the previous paragraph ensure that the agent does not have incentives to deviate from the conjectured equilibrium. The firms do not have incentives to renegotiate the offer in the interim, since this could only make harm (the agent self-sorting into the wrong activity) given their interim beliefs. It is also easy to see that second-period firms would not have incentives to deviate given their posterior beliefs.

We now check under which circumstances a firm would wish to deviate from $\{w_1^E, w_1^D, d = 1\}$, and specify off-equilibrium path beliefs. That a firm deviating by offering $d < 1$ cannot be profitable is shown in the proof of Proposition 2. Here we consider a firm deviating in wage offers. First note that a firm would never raise w_1^D or w_1^E because it would then attract both types, pay more for them and have them in less productive activities. (The firm earns zero profits in the conjectured equilibrium. Raising w_1^E will simply pay more for both types of workers going into task E and generating on average π^0 rather than $(\pi^0 + \pi^H)/2$. Raising w_1^D will pay more for both types of workers going into task D and generating $(\pi^L + \pi^H)/2$ rather than $(\pi^0 + \pi^H)/2$.) Also, a firm would not lower both w_1^D and w_1^E as it would then not attract any type.

Let us now suppose a firm deviates by lowering just w_1^D or w_1^E. Since we are using sequential equilibria, we need examine what happens when firm believes it will attract both types of workers with positive probability (though not necessarily in equal proportion). We first consider a deviation $\{w_1^E, w'^D_1\}$, where $w'^D_1 < w_1^D$. A high type accepting this offer would then work in E if w'^D_1 is not adjusted upwards in the interim. The firm

[6] The non-negativity of wages is enough to ensure this outcome. See Kaplan & Wettstein (2000) for details.

would revise $w\text{'}_1^D$ upwards in the interim if this could make the high type "switch" to D in a profitable way (without affecting the low type). Denote the revised offer for $w\text{"}_1^D$. The extra compensation required to make the high type pick D rather than E would be the agent's loss of career gains from choosing D, i.e., $w_2^E - w_2^D = \pi^0 - \pi^H$. Since this expression is negative, the firm can pay the agent a lower wage in the D activity than in the E activity and still make him switch. The cheapest way to make this occur will be to set $w\text{"}_1^D = \pi^0 - \pi^H + w_1^E$. But this means that $w\text{"}_1^D = \pi^0 = w_1^D > w\text{'}_1^D$. In other words, by revising its wage offer in the interim, the firm can both reduce the wage bill and improve productivity. Hence $w\text{'}_1^D$ cannot be a time-consistent deviation.

Let us now consider a firm deviating by offering $\{w\text{'}_1^E, w_1^D\}$, where $w\text{'}_1^E < w_1^E$. Suppose again that the firm believes that after a deviation it will attract both types with a positive probability. A low type accepting this offer would then work in D if $w\text{'}_1^E$ is not adjusted upwards. The firm would revise $w\text{'}_1^E$ upwards in the interim if this could make L "switch" to D in a profitable way (without affecting the high type). Denote the revised offer for $w\text{"}_1^E$. The extra compensation required to make the low type pick E rather than D would be the agent's loss of career gains from choosing D, i.e., $w_2^D - w_2^E = \pi^H - \pi^0$. The cheapest way to make the agent willing two switch would be to set $w\text{"}_1^E - w_1^D = \pi^H - \pi^0$ (which implies that $w\text{"}_1^E = w_1^E$). Hence, a firm would prefer to raise the wage for E if the extra wage compensation is less than the productivity improvement,

$$\pi^H - \pi^0 < \pi^0 - \pi^L. \tag{1}$$

If (1) holds, then deviating by offering $w\text{'}_1^E$ would not be time consistent and a separating equilibrium exists. On the other hand, if (1) does not hold, a firm can profit by deviating through (credible) cream-skimming, and a separating equilibrium cannot exist.

Now uniqueness. Suppose first that there exists an alternative separating equilibrium. For self-sorting to occur in the first period, it is necessary that ($w_1^E = \pi^H - k$, $w_1^D = \pi^0 - k$) for the agent's IC conditions to hold (the separating equilibrium above corresponds to $k = 0$). Suppose there exists a separating equilibrium with $k > 0$. But in that case there would exist ϵ such that a firm could deviate by offering $w_1^E = \pi^H - k + \epsilon$, $w_1^D = \pi^0 - k + \epsilon$ and attract the agent with probability 1 (instead of with probability 1/2) and make a gain. Hence there is a unique separating equilibrium. To see that there cannot exist an assignment equilibrium is relegated to the proof of Proposition 2. ■

In a separating equilibrium a firm sets $d = 1$, hence the agent is hired and then fully delegated the job design. A high type gets a pay less than marginal productivity in the first period but sends a good signal about his abilities, that has monetary value through a higher pay in the second period. For low type, the equilibrium outcome is the reverse. To be willing to reveal low ability by choosing the easy activity rather than herding in with the high type, a low type must be compensated by a high wage in the first period, due to the inferior career prospects from doing so.[7] The information rent earned by the low type creates a potential incentive for firms to deviate in order to attract only the high type, by holding the offer w_1^D fixed and reducing w_1^E. However, when it is sufficiently inexpensive for firms to make the low type choose the easy activity instead of the difficult activity in the interim, by raising the offer w_1^E at that point, such cream-skimming is not time consistent (credible), and a separating equilibrium exists.

The proposition illustrates how delegation can boost productivity through improved allocation of human resources. If a firm instead of delegating assigned to the E activity or to the D activity without consulting the agent, it would obtain an expected production of $\max\{\theta\pi^H + (1-\theta)\pi^L, \pi^0\}$. On the other hand, the expected production in separating equilibria equals $\theta\pi^H + (1-\theta)\pi^0$. The extra production (and, by zero profits, wages) in separating equilibria reaches its maximum for an intermediate value of $\theta = (\pi^0 - \pi^L)/(\pi^H - \pi^L)$.[8]

To see that the existence of a separating equilibrium is linked to career concerns, we measure the strength of career concerns as the wage differential $w_2^D - w_2^E$. When $w_2^D - w_2^E$ is low, it is cheap to revise the offer w_1^E upwards in the interim, to make the low type choose the easy rather than the difficult activity. Therefore, a separating equilibrium is more likely to exist the weaker career concerns. We can further link the strength of the career concerns to the technology of the firm. A natural measure of the *returns to*

[7] In the Appendix we show that with optimal performance contracts the low type does not necessarily make a higher wage than the high type in the first period.

[8] We see this since for large θ, we have $\theta\pi^H + (1-\theta)\pi^L > \pi^0$. The extra production is then $(1-\theta)(\pi^0 - \pi^L)$, which is higher the smaller the θ. For small θ, we have $\theta\pi^H + (1-\theta)\pi^L < \pi^0$. The extra production is then $\theta(\pi^H - \pi^0)$, which is higher the larger the θ. Thus, the extra production is maximized for θ such that $\theta\pi^H + (1-\theta)\pi^L = \pi^0$.

ability is $\pi^H - \pi^0$. This is the difference in productivity between the low type and the high type under an efficient allocation. On the other hand, $\pi^0 - \pi^L$ measures the cost of misallocating a low type to the difficult activity. We coin this "Peter's cost". Therefore, a separating equilibrium is less likely to exist the higher returns to ability and the lower Peter's cost. We can note that an increase in the returns to the exploitation activity, π^0, both decreases the returns to ability and increases the Peter's cost, so that such an increase unambiguously increases the degree of delegation. This will be used later.

The full delegation in a separating equilibrium differs radically in spirit from the assignment and job design literatures, where firms direct agents to do specific activities rather than delegating the choice. Since $w_2^D - w_2^E > 0$ in a separating equilibrium, we can interpret this as the difficult activity endogenously becoming the "prestige" activity. In contrast to Aghion & Tirole (1997) the benefit a agent enjoys from delegation is therefore *endogenous*, in that it depends on the equilibrium sorting into a prestigious and non-prestigious activity.

3.2 Assignment equilibria

We now examine a firm's delegation policy when the condition in Proposition 1 does not hold.

Proposition 2 *If $\pi^H - \pi^0 > \pi^0 - \pi^L$, then the unique time-consistent SE is an assignment equilibrium where*

(i) The agent is assigned to the difficult activity with probability $1 - d$.

(ii) Delegation d decreases in the returns to talent $\pi^H - \pi^0$ and increases in the Peter's cost $\pi^0 - \pi^L$.

Proof. We first derive wages assuming the existence of an equilibrium where the worker is assigned to the D activity, then show the existence of such an equilibrium, and finally prove uniqueness. Since the on- and off-equilibrium path beliefs are analogous to those in Proposition 1, we leave out the discussion of these. We start by determining the wages $\{w_1^D, w_2^D, w_1^E, w_2^E\}$ in an assignment equilibrium with a given d, and then determine d. For convenience, we change variables so that the degree of assignment equals f $(\equiv 1-d)$.

In that case, the probability of an agent being H conditional on activity becomes (recall that assigned workers are in D and all H types given a choice will also choose D),

$$\begin{aligned} \theta^D &= \frac{\theta}{\theta + f(1-\theta)} \\ \theta^E &= 0 \end{aligned} \tag{2}$$

Since the firms must earn zero profits in the second period and only the L type will be engaged in E, we can determine the second period wages as,

$$\begin{aligned} w_2^D &= \theta^D \pi^H + (1-\theta^D)\pi^0 \\ w_2^E &= \pi^0 \end{aligned} \tag{3}$$

By self-sorting in the first period, we must as in a separating equilibrium have that $w_1^E - w_1^D = w_2^D - w_2^E$, where the right hand side is a function of f. Using equations (2) and (3), we find that,

$$\begin{aligned} w_2^D - w_2^E &= \frac{\theta}{\theta + f(1-\theta)}\pi^H + \frac{f(1-\theta)}{\theta + f(1-\theta)}\pi^0 - \pi^0 \\ &= \frac{\theta}{\theta + f(1-\theta)}(\pi^H - \pi^0) \end{aligned} \tag{4}$$

$\{w_1^D(f), w_2^D(f), w_1^E(f), w_2^E(f)\}$ is then determined from (3), (4), and first-period firms making zero profits. We now claim that f^* is determined by

$$(\pi^0 - \pi^L) - \frac{\theta}{\theta + f(1-\theta)}(\pi^H - \pi^0) = 0 \tag{5}$$

Since the left hand side increases in f, this expression uniquely determines f^*.[9] Note that (ii) follows immediately from inspection. We now prove the claim in Steps 1-4.

Step 1. Suppose that both firms offer $\{w_1^D(f), w_2^D(f), w_1^E(f), w_2^E(f), f\}$ where $f \in (0,1)$. In the interim, there is a potential gain from decreasing f (increasing delegation), since this improves allocation. Label the renegotiated f by $f'(< f)$. The gain from setting $f' = 0$ (setting f' to an intermediate value is treated in the same manner) would be the productivity improvement from making the low type switch, i.e., $\pi^0 - \pi^L$. The cost would

[9] If there does not exist $f^* \in (0,1)$ that solves (5) then the equilibrium must be pooling. In the following we assume that there exists an interior f^*.

be the extra compensation needed to make him switch, i.e., $w_2^D - w_2^E$. Using (4), the net gain from setting $f' = 0$ in the interim, $NG(f)$, equals

$$NG(f) \equiv (\pi^0 - \pi^L) - (w_2^D - w_2^E) = (\pi^0 - \pi^L) - \frac{\theta}{\theta + f(1-\theta)}(\pi^H - \pi^0) \qquad (6)$$

$NG(f)$ increases in f and $NG(f^*) = 0$ by (5).

Step 2. Consider $f > f^*$. From (6), we would then have $NG(f) > 0$. In words, $f > f^*$ would not be a time-consistent deviation since the firm would benefit from allowing full delegation in the interim.

Step 3. Suppose that the other firm sticks to $\{w_1^D(f^*), w_1^E(f^*), f^*\}$. Setting $f < f^*$ would attract the low type with probability 1 and cannot be a profitable deviation.

Step 4. Since $f < f^*$ would not be profitable and $f > f^*$ would not be time consistent, $f = f^*$ is the only candidate equilibrium value of f^*.

Having shown that f^* is the only candidate equilibrium value for f, we need to show that profitable wage deviations are not possible (deviations that combines adjusting w or f are treated in the same manner). Analogous to the proof of Proposition 1, the critical deviation is $w\text{'}_1^E < w_1^E(f^*)$. But by the same argument as in Proposition 1, such a deviation can only be time consistent if the gain from moving a low type in the interim exceeds the extra compensation needed. This will be the case if $NG(f) > 0$, but this cannot hold by the definition of f^*.

We now prove that there cannot be assignment equilibrium where the number of slots in activity D is restricted. If the number of slots in D is restricted, there are two possibilities. First, it can be the case that both types prefer D. In that case, the proportion of types should be the same in both activities. If this happens, there are no career concerns since no information inferred by activity choice. Because of this, the firm can induce the high type to switch from E to D, by paying the same wage in D as in E. Such a scheme would increase productivity without increasing costs. So in equilibrium, it cannot be the case that both types wish to engage in D. The second possibility is that the low type wishes to engage in E, while the high type wishes to engage in D. In that case, total wages must be equalized across activities. But then, the firm can increase profits by increasing delegation, and allow the agent to engage in D rather than E (since only the high type would wish to move). This occurs since both the wage in D is lower than in E (the probability of H is higher in D than in E) and the productivity of the high type is higher in D. Hence a situation where the slots in D are rationed cannot be an equilibrium.

To prove uniqueness, first observe that by Proposition 1 there cannot exist a separating equilibrium, since $\pi^H - \pi^0 > \pi^0 - \pi^L$. To see that there cannot exist other equilibria where the worker is assigned to E, suppose that there exists such an assignment equilibrium with $f < f^*$. But in that case a firm could profitably deviate by increasing f (it would attract only the high type, since a low type prefers a low f). This deviation would be time consistent since $f < f^*$. On the other hand suppose that there exists an assignment equilibrium with $f > f^*$. Then both firms would have incentives to lower f in the interim, and hence this deviation would not be time consistent. ■

In assignment equilibria the firm is forced to act as a traditional principal, restricting the activities possible for the agent, and a centralized solution to the job design problem emerges endogenously.[10] The economic intuition for an assignment equilibrium is quite simple. The principal uses the degree of delegated job design as a tool in recruiting able agents. By restricting the agent to do the difficult activity (with a high probability) the firm avoids being stuck with low talent. We can most clearly see this point by assuming that one of the firms deviates by offering the same wages but offer a higher degree of delegation. The high type would not be affected by such a deviation, while the low type would realize that he would get the preferred activity (E) with a higher probability. Therefore, the low type would attend the deviating firm with probability 1, resulting in negative profits for the deviator. On the other hand, offering a lower degree of delegation would not be a credible deviation, as the firm would have incentives to increase the degree of delegation in the interim. Hence we obtain a theory that not only explains why agents are given limited delegation, but also why firms would opt for “bold” projects in order to attract talent.

Note that if firms do not assign workers - and set equal wages for the two activities, the low type would imitate the high type and herd into the more prestigious activity, resulting in an efficiency loss. This argument highlights the role of firms in our model; firms do not exist for traditional reasons such that better ability to bear risk or because of

[10] Note also that an alternative interpretation of rationing equilibria is that of job rotation; all interested workers are allowed to do the easy task, but only a certain amount of time.

the need to coordinate different lines of production, but to adjust wages (and the degree of delegation) to ensure a second-best allocation of workers.

4 Discussion

The simple yet powerful insight that we obtain is when career concerns are weak, a worker's private benefit to misrepresentation is small and firms opt for a liberal delegation practice, and when the career concerns are strong, this private benefit is large and firms opt for the traditional emphasis on centralized, top-down job design. We note that this contrasts with Fama (1982) and Holmstrom (1982/1999), which emphasize the beneficial incentive effects of career concerns on reducing agency costs. In the present setting, career concerns increases agency costs and necessitates limits to delegation. In this section, we consider some applications and then discuss the robustness of our theoretical results.

While the early literature on decentralization of authority to workers and new work practices tended to focus on documenting a variety of facets of the phenomenon, more recently the empirical literature has attempted to link degree of delegation to wage levels (Caroli & Van Reenen 2001, Black et al, 2003, Bauer & Brender, 2003). A robust finding from this strand of the literature is a positive relation between degree of delegation and wage levels, controlling for a variety of worker and firm characteristics. While established theories of organizational delegation, such as Aghion & Tirole (1997), does not predict a positive relationship between delegation and wages, one does emerge in our paper for the simple reason that increased delegation will be associated with a more efficient allocation of talent inside the firm, leading to higher wages.[11,12]

[11] In the present model, firms make zero profits so that the agent gets all the surplus created. Obviously, there is a wide range of other sharing rules that share this positive relationship between delegation and wages.

[12] In Aghion & Tirole (1997), agents get (exogenous) private benefits from delegation, which drives the relationship. Other papers along these lines include Baker et al., (1999), and Zabojnik (2002). There is a largely independent literature on delegation that considers other motives than private information, such as reducing managerial overload (Aoki 1986), costly writing of contracts (Marschak & Reichelstein, 1998), and delegation as a commitment device (Fershtman and Judd, 1986). Other papers with private information as an ingredient in the delegation choice includes Laffont & Martimort (1998) on the costs

The last decades have seen some major changes in employment relations, such as the increased use of IT, increased wage inequality, and increased delegation and decentralization of decision-making in firms (see the survey by Bresnahan et al., 2002).[13] With firm-level data, Caroli & Van Reenen (2001) and Bresnahan et al. (2002) find that increased use of IT and increased delegation has arrived in a cluster, suggesting a causal link between cheaper IT and increased delegation. Suppose that stronger and easier to handle computers increases the productivity (of a given worker) in exploitation activities (such as quality testing and updating, information collection and simple analysis connected with e.g., sales), but has less of a productivity augmenting effect in exploration activities associated with conceptualization and the exploration of new paths (as way of analogue, the cheaper use of computers probably had a much stronger productivity augmenting effect on research in applied mathematics such as branches of physics and engineering, than in pure mathematics such as algebraic topology). Interpreted in terms of the model, the increased use of IT increases π^0, decreases the returns to ability and the degree of career concerns, and therefore should lead to increased delegation.[14]

Another application of the model is to understand cross-sectional differences in degree

---

of communication and collusion between agents, and Faure-Grimaud et al. (2002) on the equivalence between centralization and delegation in a Laffont-Martimort type of setting. These papers explore the conditions under which the revelation principle implies that centralization and delegation are equivalent.

[13] Bresnahan et al. (2002) argues that the trend towards more IT and more delegation have been associated with a third trend, towards more wage inequality. This trend can also be accommodated by the framework if we introduce "unskilled" workers into the framework.

[14] Consider two candidate explanations. One reasonable effect of cheaper IT technology is to lower the cost of communication between different layers of the organization, e.g., through the use of E-mail. But from arguments as in e.g., Dessein (2001), such a decreased cost of communication should lead to less delegation, not more, since a decreased cost of communication makes centralization of decision-making authority cheaper. Another candidate explanation is the following. Suppose that increased use of IT makes tasks more complex (Bresnahan et al., 2002). A reasonable consequence of such added complexity is that performance evaluation becomes harder. But as shown in the appendix, and in a related setting by Prendergast (2002), worse information about performance should lead to less delegation, not more. There are other alternative explanations. For example, a transition from mainframes to personal computers implies a lesser need for coordination at the managerial level. Or, the increased use of IT has lead both to more complex tasks *and* improved performance evaluation, for example through supervisors increased use of IT.

of delegation. For example, by any reasonable measure of degree of delegation, typical bureaucracies, public or private, delegate less than hi-tech firms.[15] While some of the differences in delegation level might be traced back to differences in firm size (a greater need for coordination in large firms and hence less delegation), some large companies such as Microsoft (Herbold, 2002) still have a high degree of delegation. Our argument goes as follows. In bureaucracies, there are well-defined ladders for job titles and salaries. An agents' placement on these ladders is visible to the outside market. In contrast, in hi-tech firms job titles and salaries are more opaque. As an illustrating example, in Sun Hydraulics job titles did simply not exist and the pay policy is extremely covert (Baron & Kreps, 1999). This implies that employees with the same degree of delegation in bureaucracies have *stronger* career concerns than employees in hi-tech firms because of this increased visibility of the choices they make. Employees in bureaucracies therefore should be delegated less than their counterparts in hi-tech firms, a prediction that is consistent with the available evidence.

One of the potential extensions of the model is extending the contract duration. We can model this by assuming that first-period productivities equal $\alpha\pi$ and second period productivities equal $(1-\alpha)\pi$, where $\alpha \in (0,1)$ and α is a measure of contract duration; α close to 1 would correspond to long-term contracts (the analysis of the previous sections corresponds to the case $\alpha = 1/2$). In a separating equilibrium, the second period wages equal $w_2^E = (1-\alpha)\pi^0$ and $w_2^D = (1-\alpha)\pi^H$. If a firm attempts to cream-skim in the first period, then in the interim it will get a productivity gain of $\alpha(\pi^0 - \pi^L)$ from raising the wage for the easy activity, while the necessary compensation equals $(1-\alpha)(\pi^H - \pi^0)$. A separating equilibrium therefore exists if the first effect dominates the second, or if $\alpha(\pi^0 - \pi^L) > (1-\alpha)(\pi^H - \pi^0)$. Clearly this equation is more likely to hold the higher α. The interpretation is straightforward; a higher α makes career concerns less of an issue, and more delegation can be sustained in equilibrium. This insight can be applied to understand differences between labor markets in the US and Japan. As described in an extensive literature surveyed by Aoki (1986), Japanese firms, compared to the US, have the job security of long-term employment relationships and delegate significantly more to

[15] This is based on evidence from case studies of firms such as Dell, Gore-Tex, Hewlett-Packard, Microsoft, and Sun Hydraulics.

their workers than American firms.[16] For other evidence supporting the same idea, Rajan & Wulf (2003) consider pay and organizational structure in 300 large US companies, and find that companies with more long-term compensation (stocks, options) delegate more to lower level managers.[17]

Robustness. To assess the robustness of our theoretical results, let us briefly discuss the role of the most important modeling assumptions. First, what would happen if the agent learns about his activity match only after starting with a firm, but before choosing activity? In that case, there would be no adverse selection at the hiring stage, and a higher degree of delegation would be sustainable in equilibrium. Second, one could argue that there are other private benefits than career prospects that may contribute to an employee desiring prestige activities, such as sense of importance or recognition by peers and friends. Such effects will work in a similar manner to career concerns, and tend to limit the degree of delegation. Third, suppose that we allow for effort costs that are observable and contractible. If the easy activity has a lower (first best) effort cost than the difficult activity, the effective (net) π^0 would increase relative to the effective π^H and π^L. A higher degree of delegation would occur. If the cost of effort differ for the two types, this creates the possibility of the firm screening the two types, which would work in the same qualitative manner as performance contracts.

[16] Our model can be extended to explain this difference of contract length between countries (or industries) by formalizing the ideal contract length as a trade-off between benefits from increased delegation and costs from being stuck with undesirable matches.

[17] We can also consider the effect of human capital acquisition between the periods, reflecting the idea that the agent becomes more productive with time. Suppose that second period productivities equal $h\pi$, where $h > 1$. This will have exactly the same effect on delegation as increasing the length of the second period, i.e., delegation will decrease. Therefore, first period wages will decrease due to inferior allocation inside the firm, while second period wages will increase due to improved productivity in that period. Empirical work has shown that worker (nominal) wages typically increase over time (see Gibbons & Waldman, 1999b). Adding human capital acquisition is one of the possibilities to ensure that separating equilibria have this property (performance contracts is another, as shown in the Appendix).

5 Conclusion

We have developed a theory of delegation within organizations where agents are privately informed about whether they should be engaged in exploitation or exploration activities. The theory is based on two simple components, private information and career concerns. In the model constructed, these components were shown to, depending on technological and other economic factors, produce equilibria with a varying degree of worker delegation. The stronger career concerns, the less delegation. The less delegation, the less efficient allocation of talent inside the firm, and the lower wages. This basic finding is consistent with evidence and not easily explainable by existing theories. In addition, the theory seems consistent with evidence on the relation between contract length and delegation.

While we have considered several ways to vary the degree of career concerns, we have not considered varying degrees of private information. In the corporate finance literature, a low liquidity of a stock is commonly seen as an indicator of a high level of information asymmetry between different traders. Analogously, we might interpret a low turnover in an industry as an indication of a high level of asymmetric information (obviously other factors such as the degree of firm-specific human capital would also play a role). We would therefore predict that the higher turnover rates, the lower degree of delegation. This conjecture seems to have empirical support from Aoki (1986) and Black & Lynch (2000) and deserves to be investigated further.

6 References

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

Aghion, P. & J. Tirole (1997). Formal and Real Authority in Organizations. *Journal of Political Economy*, 105, 1-29.

Aoki, M. (1986). Horizontal vs. Vertical Information Structure of the Firm. *American Economic Review*, 76, 971-983.

Baker, G. P. (1992). Incentive Contracts and Performance Measurement. *Journal of Political Economy.*

Baker, G. P., R. Gibbons & K. J. Murphy (1999). Informal Authority in Organizations. *Journal of Law, Economics and Organization*, 15, 56-73.

Baron, J. E. & D. M. Kreps (1999). Strategic Human Resources: Frameworks for General Managers. *John Wiley Press.*

Bauer, T. & S. Bender (2003). Flexible Work Systems and the Structure of Wages: Evidence from Matched Employer-Employee Data. IZA working paper #353, University of Bonn.

Black. S & L. M. Lynch (2001). How to compete: the impact of workplace practices and information technology on productivity. *Review of Economics and Statistics.*

Black, S., L. M. Lynch & A. Krivelyova (2003). How workers fare when employers innovate. NBER WP# 9569.

Bresnahan, T. F, E. Brynjolfsson, L. M. Hitt (2002). Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Quarterly Journal of Economics*, 117, 339-76.

Caroli, E., N. Greenan & D. Guellec (2001). Organizational Change and Skill Accumulation. *Industrial and Corporate Change*, 10, 481-506.

Dessein, W. (2002). Authority and Communication in Organizations. *Review of Economic Studies*, 69, 811-838.

Fama, E. (1980). Agency Problems and the Theory of the Firm. *Journal of Political Economy*, 88, 288-307.

Faure-Grimaud, A., J.-J. Laffont & D. Martimort (2002). Collusion, Delegation and Supervision with Soft Information. USC Center for Law, Economics & Organization, Research Paper No. C02-9. Forthcoming, *Review of Economic Studies.*

Fershtman, C. & K. Judd (1986). Equilibrium Incentives in Oligopoly. *American Economic Review*, 77, 927-940

Foster, A. & M. R. Rosenzweig (1996). Comparative Advantage, Information and the Allocation of Workers to Tasks: Evidence from an Agricultural Labour Market. *Review of Economic Studies*, 63, 347-74.

Gibbons, R. & M. Waldman (1999a). A Theory of Wage and Promotion Dynamics Inside Firms. *Quarterly Journal of Economics*, 114, 1321-58.

Gibbons, R. & M. Waldman (1999b). Careers in Organizations: Theory and Evidence.

Handbook of Labor Economics, Vol. 3, 2373-2437.

Gore, W. L. (1990). Freedom versus Structure: The Lattice Organization. In: *Handbook for Creative and Innovative Managers*, edited by R. L. Kuhn.

Greenwald, B. (1986). Adverse Selection in the Labor Market. *Review of Economic Studies*, 53, 325-47.

Herbold, R. J. (2002). Inside Microsoft. *Harvard Business Review*, 80, 72-79.

Holmstrom, B. (1982/1999). Managerial Incentive Problems: A Dynamic Perspective. *Essays in Economics and Management in Honor of Lars Wahlbeck* (Helsinki: Swedish School of Economics). Reprinted (with an added abstract) in *Review of Economic Studies*, 66, 1999.

Holmstrom, B. & P. Milgrom (1991). Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics and Organization*, 7, 24-52.

Hvide, H. K. & T. Kaplan (2003). Delegated Job Design. CEPR working paper 3907.

Kaftan, C. (1984/1991). Sun Hydraulics Corporation A, B, and C. *Harvard Business School Case Studies*.

Kaplan, T. and D. Wettstein (2000). Mixed-Strategy Equilibria with Constant-Returns-To-Scale Technology under Bertrand Competition, *Spanish Economic Review,* 2(1): 65-71

Krasa, S. and A. P. Villamil (2000). Optimal Contracts when Enforcement is a Decision Variable. *Econometrica*, 68, 119-134.

Laffont, J.-J. & D. Martimort (1998). Collusion and Delegation. *Rand Journal of Economics*, 29, 280-305.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, 2, 71-87.

OECD (1999). Employment Outlook. OECD, Paris.

Olsen, T. & G. Torsvik (2000). Discretion and Incentives in Organizations, *Journal of Labor Economics*, 18, 377-404.

Ortega, J. (2001). Job Rotation as a Learning Mechanism. *Management Science*, 47, 1361-1370.

Prendergast, C. (1996). A Theory of Responsibility in Organizations. *Journal of Labor Economics,* 13, 387-400.

Prendergast, C. (2002). The Tenuous Trade-off Between Risk and Incentives. *Journal of Political Economy*, 110, 1071-1102.

Rajan, R. G. & J. Wulf (2003). The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies. Working paper, GSB University of Chicago.

Rosen, S. (1982). Authority, Control and the Distribution of Earnings. *Bell Journal of Economics*, 13, 311-23.

Rosen, S. (1986). The Theory of Equalizing Differentials. *Handbook of Labor Economics*, Vol. 1, 641-92.

Roy, R. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers,* 3, 135-46.

Stern, S. (2004). Do Scientists Pay to be Scientists? *Management Science*, 50, 835-53.

Waldman, M. (1983). Job Assignments, Signalling, and Efficiency, *Rand Journal of Economics*, 15, 255-67.

Zabojnik, J. (2002). Centralized and Decentralized Decision-Making in Organizations. *Journal of Labor Economics*, 20.

7 Appendix: optimal performance contracts

Here we solve the more general model where the agent's type can depend upon performance measures.

There are two possible output levels in the D activity, π^{low} and π^{high}, where a low (high) type has probability P_L (P_H) of obtaining π^{high}. The expected output for the low (high) type equals π^L (π^H), and therefore $P_L < P_H$. Output in the E activity is as before independent of ability, and with mean π^0. We denote the performance of the agent in the first period by π.

Firms now offer (and renegotiate) performance dependent contracts $\{w_1^D(\pi), w_1^E(\pi), d\}$. Timing is as before with the addition that performance is becomes public knowledge after the first period. [18] For simplicity, we assume that output is not contractible in the

[18]Since workers reveal their type in a separating equilibrium, the conditions for existence of such an equilibrium do not depend upon performance being observable to the outside firm or not. The rationing

second period (this will have no qualitative bearing on our results). As before, we focus on time-consistent SE.

Since performance in E is independent of ability, an agent that chooses E will therefore be offered a fixed salary denoted by F, i.e., $w_1^E(\pi) = F$. If the agent is engaged in D, he gets a wage contingent on whether the output is π^{low} or π^{high}, denoted by B_L and B_H respectively. To avoid trivial forcing contracts, we assume that the agent has limited liability, so that F, B_L, B_H must be non-negative.

We assume that the low type breaks ties by choosing the E activity in the first period. If the agent had a different tie-breaking, we would get identical results but firms would maximize profits by offering $\sup(0, F)$ instead of F to the low type in the first period to ensure self-sorting. Our tie-breaking rule avoids such non-essential technicalities.

7.1 Separating equilibria

In a separating equilibrium, agent type is revealed to both firms before the second period and the second period wage must be π^H for the high type π^0 for the low type. To induce self-sorting as cheap as possible, optimal contracts must have $B_L = 0$, and we can therefore write B_H simply as B. Denoting the lifetime utility for a type i agent choosing activity j in the first period for U_i^j, we then have,

$$U_H^D = P_H B + \pi^H \tag{7}$$

The first term $P_H B$ is the expected wage in the first period, and the second term π^H is the wage in the second period, if the agent chooses the D activity (remember in a separating equilibrium only the high type chooses this activity). On the other hand, the utility for the low type for choosing the E activity equals,

$$U_L^E = F + \pi^0 \tag{8}$$

Where F is the fixed wage in the first period and π^0 is what he gets in the second period. We have two IC conditions for a separating equilibrium,

$$P_H B + \pi^H \geq F + \pi^0 \tag{IC1}$$

$$F + \pi^0 \geq P_L B + \pi^H \tag{IC2}$$

equilibrium would also have the same qualitative features but slightly different wages.

(IC1) is the self-sorting constraint for the high type, and (IC2) is the self-sorting constraint for the low type.[19]

If $F > \pi^0$, then (IC2) binds.[20] In that case, we can determine F as,

$$F = P_L B + \pi^H - \pi^0 \tag{9}$$

This implies (by $P_L < P_H$ and IC2) that the high type strictly prefers the D activity in a separating equilibrium, thus (IC1) holds as well. The first-period zero profit condition is,

$$\theta\pi^H + (1-\theta)\pi^0 = \theta B P_H + (1-\theta)F \tag{10}$$

The left hand side is the expected productivity of the firm, and the right hand side is the total wage bill. (IC2) and the zero profit conditions then determine the equilibrium values of F and B, denoted by F^* and B^*, as

$$\begin{aligned} B^* &= \frac{\pi^0 - (1-2\theta)(\pi^H - \pi^0)}{\theta(P_H - P_L) + P_L} \\ F^* &= \frac{\theta(P_H + P_L)(\pi^H - \pi^0) + P_L\pi^0}{\theta(P_H - P_L) + P_L} \end{aligned} \tag{11}$$

We can note that the solution for the simplified model in Section 3 is obtained by setting $P_L = P_H = P$ and $\theta = 1/2$, in which case $F = \pi^H$ and $PB = \pi^0$.

To have the same type of separating equilibrium as before, where the low type is paid above marginal productivity to self-sort, we need that $F^* > \pi^0$.[21] From (11), this occurs whenever $P_L\pi^H/P_H + \pi^H - \pi^0 > \pi^0$. However, with the opposite inequality, $P_L\pi^H/P_H + \pi^H - \pi^0 < \pi^0$, we get $F^* < \pi^0$ from (9), which clearly cannot occur in (separating) equilibrium, since a firm would make a profit no matter who shows up in the E activity. In that case, there exists a separating equilibrium with $F^* = \pi^0$ and $B^* = \pi^H/P_H$, that

[19] There is a π^H on RHS of (IC2), since the wage in the second period cannot be contingent on second-period output and should not be contingent upon first-period output since both firms believe the worker in the D task is high even if his performance is low. Again, assuming second period output is contractible does not qualitatively change results.

[20] If not, a firm can offer a contract with a lower F and obtain only the high ability workers. This firm would not have incentive to later raise the low ability worker's wage since such worker would already have incentive to self-sort.

[21] The liability constraint, $B^* \geq 0$, is satisfied whenever $\theta > \frac{1}{2} - \pi_0/(\pi_H - \pi_0)$. Hence a low θ is an additional reason to get rationing, but here we assume that θ is sufficiently high.

is both types get (expected) wage equal to marginal productivity in both periods, which is a qualitatively different separating equilibrium from that obtained previously.[22] To examine additional conditions for existence of a separating equilibrium where the low type is paid a premium to self-sort, that is when $F^* > \pi^0$, we now consider the possibility of cream-skimming.

Suppose one of the firms deviates by offering a low wage for the easy activity (in an attempt to cream-skim). This firm will have incentives to renegotiate this offer after the agent has chosen which firm to work for, by raising the wage for E such that $w_1^E = F$, if the production gain exceeds the wage compensation loss. The extra compensation needed to induce the low type to switch contracts equals $\pi^H - \pi^0$, that is the wage loss in period 2 from being revealed as having being the low type. It will pay to make this compensation only if the productivity improvement exceeds the extra compensation, that is

$$\pi^H - \pi^0 < \pi^0 - \pi^L \tag{12}$$

When this no cream-skimming condition holds, a separating equilibrium exists, which is analogous to the case without performance contracts (equation 1). By combining the no cream-skimming condition and the condition $P_L\pi^H + P_H(\pi^H - 2\pi^0) > 0$, we see that a separating equilibrium of the type considered in the main text, where the low type is compensated to self-sort, exists whenever (12) holds and $\frac{\pi^H}{P_H} > \frac{\pi^L}{P_L}$. Since this condition always holds for $P_L = P_H$, the essential requirement for this type of separating equilibrium is that the difference $P_H - P_L$ is not too great, or in other words that the monitoring technology is not too precise, which is an intuitively appealing result. Let us summarize.

Proposition 3 *When the no cream-skimming condition (12) holds and the monitoring technology is not too precise, a separating equilibrium exists where the low type is paid above marginal productivity in the first period. When monitoring is precise, a separating equilibrium exists where both types are paid their marginal productivity. In both types of separating equilibria, the agent is fully delegated the job design decision, and a high type strictly prefers the difficult activity.*

[22] This solution will satisfy (IC2) if $2\pi_0 \geq P_L\frac{\pi_H}{P_H} + \pi_H$, which is the same condition that determines when our candidate F^* is less than π_0. Thus, we can get a separating equilibrium for this case.

Let us now see what happens if a separating equilibrium does not exist due to the cream-skimming threat.

7.2 Assignment equilibria

In an assignment equilibrium, an agent that is engaged in E in the first period will be low type with probability 1, and will therefore get the wage π^0 in the second period. For an agent that chooses D, the wage in the second period will depend on the fraction of high types in D and on whether the agent obtained a bonus or not. Recall the assumption that pay can only be conditioned on performance in the first period, and hence that the agent simply gets his expected productivity in the second period.

Let θ_H (θ_L) be the probability that an agent with high (low) performance is the high type, and as before, let f be the probability that the low type is assigned to D, while a fraction $1-f$ $(=d)$ are allowed to choose freely, and hence choose E. Then,

$$\begin{aligned} \theta_H &= \frac{\theta P_H}{\theta P_H + (1-\theta) f P_L} \\ \theta_L &= \frac{\theta(1-P_H)}{\theta(1-P_H) + (1-\theta) f (1-P_L)} \end{aligned} \tag{13}$$

Furthermore, let w_2^H (w_2^L) be the second period wage for an agent with a high (low) performance in the first period. Then,

$$\begin{aligned} w_2^H &= \theta_H \pi^H + (1-\theta_H)\pi^0 \\ w_2^L &= \theta_L \pi^H + (1-\theta_L)\pi^0 \end{aligned} \tag{14}$$

$w_2^H > w_2^L$ since the high type has a better chance of getting a bonus than the low type. We now have the IC conditions for an assignment equilibrium,

$$P_H(B + w_2^H) + (1-P_H)w_2^L \geq F + \pi^0 \tag{IC3}$$

$$F + \pi^0 \geq P_L(B + w_2^H) + (1-P_L)w_2^L \tag{IC4}$$

(IC3) is the self-sorting constraint for the high type in an assignment equilibrium, and (IC4) the self-sorting constraint for the low type. As with a separating equilibrium, if $F > \pi^0$ and (IC4) were not binding, a firm can improve profits by lowering F and getting the low type with a smaller probability. Hence we can determine F as,

$$F = P_L(B + w_2^H) + (1 - P_L)w_2^L - \pi^0 \tag{15}$$

Since (IC4) binds, (IC3) becomes redundant (by $P_L < P_H$ and $w_2^L < w_2^H$), and the high type must strictly prefer D also in a assignment equilibrium. The first period zero profit condition is,

$$\theta\pi^H + (1-\theta)(1-f)\pi^0 + (1-\theta)f\pi^L = \theta BP_H + (1-\theta)(1-f)F + (1-\theta)fBP_L. \tag{16}$$

The left hand side is the expected productivity of the firm, and the right hand side is the total wage bill. The first term on the left hand side is the productivity of the high type, the second term is the productivity of the low type in E, and the third term is the productivity of the low type in D. The right hand side gives the corresponding wages for those three groups of agents. The third equilibrium condition is that firms should be indifferent between shifting the low type (i.e., decreasing f) on the margin, i.e., that $\pi^0 - \pi^L = F - P_L B$. Again, the productivity improvement from shifting the low type is on the left hand side, and the required extra compensation on the right hand side. We now have five endogenous variables, F, B, f, w_2^L, and w_2^H, and five equations, the no-shifting equation, zero profits, (IC2), and the equations determining w_2^L, and w_2^H. This system has a unique solution equal to,

$$\begin{aligned} B^* &= \frac{\theta(\pi^H - \pi^L) + \pi^L}{\theta(P_H - P_L) + P_L} \\ F^* &= \pi^0 + \frac{\theta(P_L\pi^H - P_H\pi^L)}{\theta(P_H - P_L) + P_L} \\ f^* &= \frac{\theta P_H(P_L(\pi^H - \pi^0) + \pi^L - \pi^0)}{P_L(1-\theta)(P_L(\pi^H - \pi^0) + 2\pi^0 - \pi^H - \pi^L)} \end{aligned} \tag{17}$$

The degree of assignment f^* can be seen to decrease in π^0 and increase in π^H and in π^L. Moreover, f^* increases in θ and in P_H, and is ambiguous to changes in P_L. A self-sorting premium is paid to the low type ($F^* > \pi^0$) whenever $\frac{\pi^H}{P_H} > \frac{\pi^L}{P_L}$, which is the same condition on monitoring as described above.[23] To see that there cannot be assignment in the case of perfect monitoring technology, that is when $P_L = 0$ and $P_H = 1$,

[23] If $(2\pi_0 - \pi_H)/\pi_H \leq P_L/P_H \leq \pi_L/\pi_H$, the (IC4) constraint may not be binding and as before we must have $F^* = \pi_0$.

observe that the denominator of f^* goes to 0 when P_L approaches zero. By solving for $f^* = 0$, we get that assignment occurs whenever $P_L > \frac{\pi^0 - \pi^L}{\pi^H - \pi^0}$, from which it follows that $\pi^H - \pi^0 > \pi^0 - \pi^L$ must hold to get assignment, as shown before. We can then summarize.

Proposition 4 *If a separating equilibrium does not exist, there exists an assignment equilibrium where the agent is assigned to the D activity with probability f. In such an equilibrium, the low type is paid a premium to be willing to self-sort, and the high type strictly prefers the D activity to the E activity. Moreover, the degree of assignment decreases in* π^0 *and increases in* π^H *and in* π^L.

The introduction of contractible measures of individual performance thus strengthens the qualitative insights of the paper in the following sense: With optimal performance contracts, we can still get assignment, a low type agent is paid a premium to be willing to self-sort, and moreover a high type agent strictly prefers the D activity to the E activity, provided that the monitoring technology is not too precise. In other words our line of argument is not dependent on the double indifference condition in the previous sections, nor on individual performance not being contractible. More generally, if other screening mechanisms are available, but are imperfect due to for example measurement costs, then job design gives information about ability, and we get the interaction of private information and career concern effects that has been our focus.