Delegated Job Design*

Hans K. Hvide[†] and Todd R. Kaplan[‡]

October 3, 2003

Abstract

Why do firms delegate job design decisions to workers and what are the implications of such delegation? We develop a private-information based theory of delegation where delegation provides a more efficient allocation of talent inside the firm, but at the cost that low ability workers must be compensated to self-sort. Career concerns limit the effectiveness of delegation: when returns to ability or market observability of job content are high, the compensation needed to get low-ability workers to self-sort is high, and firms limit delegation to avoid cream-skimming of the high-ability workers. We investigate implications for how misallocation of talent within firms may occur, the optimal design of incentive contracts, and which decisions are more likely to be delegated to subordinates.

Keywords: Career concerns, Delegation, Discretion, Peter Principle, Sun Hydraulics.

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

*We thank numerous seminar participants, and in particular Bengt Holmström, Eirik G. Kristiansen, Ed Lazear, David de Meza, Trond Olsen, and Mike Waldman for discussions and comments to earlier versions of the paper. The usual disclaimer applies.

[†]Department of Finance and Management Science, Norwegian School of Economics and Business, and CEPR.

[‡]School of Business and Economics, University of Exeter.

1 Introduction

The traditional method of job design, as evidenced by hiring procedures in government bureaucracies, is to first define the activities contained in the job slots and then to hire suitable workers (or to reallocate existing workers) to fill those slots. This gives workers limited discretion in designing their job. In recent years, this bureaucratic, top-down, solution to the job design problem has been challenged. For example, the engineering company Sun Hydraulics gives employees "the right and responsibility to choose how they spend their time," and Gore & Ass., the producer of Gore-Tex© products, encourages "maximum freedom for each employee" (as described by Baron & Kreps, 1999). While these two examples are extreme, the delegated job design practices of Sun and Gore are part of mainstream managerial thinking as evidenced by a burgeoning empirical literature (e.g., Caroli et al., 2001, OECD 1999, and Rajan & Wulf, 2003). This literature documents the widespread use of practices such as job rotation, matrices, and self-monitoring groups, which all may be seen as increased delegation and flexibility used by firms when designing jobs.[1]

Why do firms delegate job design to workers? Why do different firms or industries practice different degrees of delegation? Several aspects may be relevant. For example, delegation can reduce managerial overload. Or workers simply enjoy the freedom implied by delegation and are willing to take a pay cut to obtain it, as may be the case in academics.

While these issues may be important, we wish here to develop a theory of delegation with worker private information and career concerns as the key ingredients. To motivate, workers may have private information about whether they are creative or not, a characteristic that is notoriously difficult to capture with for example personality tests. Or workers may have better knowledge of customer tastes than the manager. Career concerns mean that the job design decision today affects a worker's welfare tomorrow. For example, if the

[1] For example, Caroli et al. (2001) states: "With more decentralized firms and more small businesses the organizational picture of western economies is changing. This is to be contrasted with the previously dominant scheme, based on a Taylorist tradition, which emphasized the advantage of setting precise norms and closely monitoring workers through their specialization in conception and execution activities." (p. 482).

most able workers in a hi-tech firm are engaged in product development, then low ability workers also engaged in product development may have better future prospects than low ability workers engaged in say product updating, since the market may (correctly) view job description as an indicator of ability. Or if the leading analysts use a new and complex market analysis tool, then other analysts may have better future prospects if they also use this tool. One example is stock market analysts during the recent dot-com boom, who started using valuation techniques based more on vague estimates of growth prospects rather than perhaps more precise estimates based on current cash flow and systemic risks. Analysts that did not approve of these new techniques were viewed as out of date.[2]

The basic tension stems from two effects of private information. On one hand, private information favors delegation since workers are better equipped to know what they should do, or how to do it. On the other hand, private information means that workers may have incentives to engage in wasteful signalling activities under delegation, to reap private benefits. For example, less able workers may engage in product development to herd in with the high ability workers and improve their own future prospects. Or less able analysts may use an analysis tool currently in mode, to give the market the impression that they are of high ability.

How much should firms delegate given these two opposing effects? To anticipate, the main cost from delegation is that low ability workers need to be compensated to self-sort efficiently, that is to choose activities with low returns to ability. When career concerns are less important (its determinants are discussed later), an internal labor market emerges where a firm sets a small premium for such activities coupled with a high degree of delegation. When career concerns are more important, the required self-sorting premium to low ability workers becomes high. In that case, outside firms can cream-skim the high ability workers, and a high degree of delegation would be unprofitable. To avoid the cream-skimming problem, the firm needs to reduce the premium necessary to compensate the low workers to self-sort. It obtains this by reducing the career concerns through limiting delegation and instead assigning workers to activities.

Our main message is therefore that when career concerns are weak, firms can opt for a liberal delegation practice, and when the career concerns are strong, firms should opt

[2]History showed of course that there were good reasons to be suspicious of these valuation methods.

for the traditional emphasis on centralized, top-down job design. This message stands in contrast to Holmstrom (1982/1999), which emphasizes the beneficial incentive effects of career concerns on agency costs. In the present setting, career concerns is the dark force that creates agency costs and necessitates limits to delegation.

Our model is based on a simple version of the classic Roy (1951) model. Workers are of two possible ability levels, and there are two activities, the "easy" and the "difficult". These activities may be thought of as different tasks or as different methods in performing the same task. An efficient allocation of workers occurs when the low (high) ability workers specialize in the easy (difficult) activity. By job design we mean the decision about which activity a worker should specialize in. There are two periods. In the first period, firms offer one-period contracts to the workers, which specify degree of delegation and pay for the different activities, and workers choose which firm to work for. Before the second period, the firms make offers to each worker conditional on their knowledge about ability, and workers accept the highest offer.

Let us summarize the main results. Under some circumstances there exist an efficient, separating, equilibrium where firms fully delegate the job design decision to workers, and a compensation scheme is structured so that workers do so efficiently. Such a scheme involves paying the low ability workers a premium to self-sort, i.e., pay above marginal productivity. A separating equilibrium resembles play in companies such as Sun Hydraulics and Gore & Ass., in that job design to a large extent is decided by the employees rather than by a manager.

When a separating equilibrium does not exist, there exist an assignment equilibrium where only a fraction of employees (which may be equal to zero and hence encompasses pooling) are delegated the job design decision, and the remaining fraction of employees have their jobs assigned by the manager. A rationing equilibrium with a high degree of assignment resembles play in bureaucracies, with little or no delegation, while a rationing equilibrium with a moderate degree of assignment resembles play in "typical" firms, where only a fraction of workers are delegated the job design.

Which equilibrium occurs depends on the strength of workers career concerns. When returns to ability are high, it is more tempting for a low ability worker to imitate a high ability worker (since the future wage differential is high between workers that are probably high ability and workers that are probably low ability) and the premium required to make low ability workers self-sort is high. Such a high premium makes the threat of cream-skimming the high workers stronger, and as a result workers will be delegated less in equilibrium. Likewise, when the market observability of worker activity is higher, the compensation needed to make low workers self-sort becomes higher, and cream-skimming is more of a threat. In essence, stronger career concerns implies that workers are delegated less decision-making authority. In addition, a high cost of misallocating a low worker (the "Peter's cost") will make the firm more willing to pay the self-sort premium, which results in a higher degree of delegation.

The limits to delegation in rationing equilibria imply that the workers' private information is not used efficiently, and a misallocation of workers therefore occurs. One may think that the greatest source of misallocation arising from assignment would be able workers that are not permitted to do the difficult activity. It turns out, however, that the inefficiency invoked by optimal behavior of firms in our model is the opposite: low workers are assigned to the difficult activity. This result accords with the Peter principle,[3] in that misallocation occurs due to workers being allocated to activities above their competence level (rather than able workers being occupied below their competence level).

Given the concerns that a high degree of delegation can make job design a (wasteful) signalling activity, one would expect that the degree of delegation and the degree of misallocation of labor input would be positively related. However, when taking into account the contractual response by firms to the signalling motive, that is in equilibrium, we find that more delegation is associated with *less* misallocation. Hence while it may be true that more delegation leads to more misallocation for a given firm at the margin, the hypothesis we obtain for a cross-section of firms is a negative relation between the degree of delegation and the degree of misallocation.

Theories of delegation have focused on situations where only one decision can be delegated. Our results give some hints as to *which* decisions will be delegated to agents. These will be activities where the returns to talent are low, or where the market observability is low. In short, activities that are less prone to harmful signalling activity will be delegated.

[3] The Peter's principle (Peter & Hull, 1969) states that in a hierarcy, employees are promoted to their incompetence level.

For example, since work method of a worker is probably less observable than task choice, we can conjecture that firms are more prone to delegate work method decisions than task choice decisions.

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, where monitoring and production technology are prime determinants, and how to give incentives such that workers undertake those activities. There are two main differences between this literature and the current paper. There is a technological difference in that we consider a situation where the firm attempts to make workers specialize efficiently, while the job design literature considers settings where firms attempt to make workers spread their effort on several tasks. More importantly, due to lack of worker private information, the job design literature has no notion of attempting to draw on worker's competence in designing jobs. Similarly, the job assignment literature, which includes Rosen (1982) and Gibbons & Waldman (1999a), considers settings where workers and firms have symmetric information at the hiring stage, circumstances under which there would be no advantage of delegating the activity choice decision. The same point applies to the literature on career concerns, as in Holmstrom (1982/1999).

The delegation literature, which spans areas in political economy, monetary economics, industrial organization, and economics of organization, has emphasized other motivations of delegation than private information, such as reducing managerial overload (Milgrom & Roberts, 1992, Aoki 1986), costly writing of contracts (Marschak & Reichelstein, 1998), delegation as a commitment device (Fershtman and Judd, 1986), or that workers may have private benefits from delegation which induces harder work (Aghion & Tirole, 1997, Baker et al., 1999, and Zabojnik, 2001). Prendergast (2002) considers the interaction between delegation and incentive contracts when worker private information may justify delegation. There are some clear differences to the present paper. For example, Prendergast's setting is static, which excludes career concerns, and the principal puts limits to delegation because she may be well-informed about which project the worker should attend to. In contrast, we take the degree of information asymmetry as given. Other papers with private information as an ingredient in the delegation choice includes Laffont & Martimort (1998) on the costs 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. Dessein (2002) shows that the equivalence no longer holds in a setting where the principal cannot commit to a reward scheme as a function of the agent's messages.

The paper proceeds as follows. Section 2 introduces the model. Section 3 considers the case where firms have symmetric information about worker ability at all stages, and Section 4 considers the case with asymmetric information before bidding in the second period. In Section 5, we relate our theory to some evidence, discuss misallocation of talent within firms, discuss which decisions are more likely to be delegated to subordinates. Section 6 extends the model to allow for performance contracts, and Section 7 concludes.

2 The basic model

Here we first describe the technology and contracts of the model, and then the timing.

2.1 Technology and contracts

There is a continuum of workers and several firms, for simplicity taken to equal two. Each worker privately knows whether he has either low or high ability, while only the share of high ability workers, $\theta \in (0,1)$, is publicly known.[4] In each firm, there are two possible activities for a worker; the "easy" and the "difficult", denoted by E and D. Both workers have productivity π^0 in the E activity. In the D activity, however, the low type has productivity π^L, and the high type has productivity π^H. We confine attention to the case where it is efficient that high workers are allocated to D and that low workers are allocated to E, that is when $\pi^L < \pi^0 < \pi^H$. We emphasize that the different "activities" can be interpreted as different tasks (e.g. product development vs. product updating), or they can be interpreted as different work methods in doing the same task (e.g., using an old or a new work method in designing products).

[4] Papers that support this notion of asymmetric information in labor markets include Acemoglu & Pischke's 1998 study of apprenticeship in Germany, and Foster & Rosenzweig (1996) on allocation of workers to tasks in agricultural Philippines.

We assume for the moment that measures of performance is sufficiently noisy to preclude the use of individual or group performance contracts. Contract offers must then simply consist of one wage w_1^D for the D activity, one wage w_1^E for the E activity, and the degree of delegation d.[5] The variable d may alternatively be viewed as the probability of a worker given full delegation once inside the firm, or the probability of a given worker being offered a full-delegation contract. In Section 6, we extend the our analysis to allow for individual performance contracts, to analyze the interaction between the implicit incentives created by career concerns and the explicit incentives created by performance contracts.

All workers and firms are risk neutral and have discount factors equal to one. We assume that if the wage offers are such that a worker is indifferent between doing the E activity or the D activity (taking into account the implicit incentives) and he is delegated the choice, then he will choose the efficient activity. This may be due to an (unmodeled) option plan or ownership share, or alternatively due to increased job satisfaction in the efficient activity. The equilibria we construct will use this tie-breaking rule quite extensively, since both low and high type workers will be indifferent between the wage contracts offered. One may therefore suspect that we obtain knife-edge results. The robustness of our results are discussed in Section 6 (performance contracts), where indifference only holds for the low type.

2.2 Timing

In the first period, workers are born knowing their ability level, and firms compete in attracting workers. Firms can only commit to contracts lasting one period. A firm offers workers a contract $\{w_1^D,\ w_1^E, d\}$. Given the offers, workers accept an offer, which binds them to a firm for one period.

Importantly, before workers engage in production (but after they have chosen which firm to work for) a firm has the option to *raise* any of the wages $\{w_1^D,\ w_1^E\}$ offered, and allow workers to switch activities. In other words, firms can commit to not lowering

[5] It may seem odd that an offer by a firm is a vector of wages, rather than just a wage. However, we can interpret the vector as reflecting differences in overtime payment or fringe benefits between the two possible tasks.

wages or to delegate less, but may in the interim choose to raise one of the wages and allow more delegation. This is a natural requirement, because both the firm and workers would (weakly) prefer such a reneged contract.[6] Although such wage raises do not occur in equilibrium, it will turn out to have an impact on equilibrium, through affecting which $\{w_1^D, w_1^E, d\}$ combinations that can credibly be offered. Workers are then either assigned to an activity or delegated the choice, and finally production takes place.

After the first period, the firms bid for the workers according to a first-price sealed-bid auction and workers accept the highest offer.[7] The expected wages in the second period conditional on activity in the first period are denoted by w_2^D and w_2^E.

3 Symmetrically informed firms

We now consider the benchmark case when the inside and the outside firm are symmetrically informed about the first period activity of a worker before bidding in the second period. This allows us to explore the role of technology (the π's) in delegation.[8]

First we focus on *separating equilibria*, where $d = 1$ and both types of employees work on their appropriate activity in period 1, and then on *rationing equilibria*, where $d < 1$ and (a fraction of) employees work on the wrong activity in period 1. Note that there is no incentive for worker misrepresentation in the second period, and hence inefficiencies, if they occur, do so in the first period.[9]

3.1 Separating equilibrium

We now analyze the separating equilibrium that occurs when workers self-sort to their appropriate activity.

[6] In technical terms, we are imposing the criterion of renegotiation-proofness.

[7] Under symmetric information between firms before the bidding, virtually any auction will give the same equilibrium wages. Under asymmetric information the auction plays a role. This is discussed in Section 4.

[8] Throughout, we view the production technology of the firm as exogenous. A richer analysis would take into account that such technology can be endogenous.

[9] The model can be extended to cover an arbitrary number of periods, in which case there can be inefficiencies in all periods except the last one.

When the sorting is efficient at time 1, both firms know a worker's ability before the second period. We therefore have that,

Remark 1 *Given that a separating equilibrium is played,*
$w_2^E = \pi^0$, $w_2^D = \pi^H$.

Proof. If workers allocate themselves efficiently, ability is revealed and bidding in the second period implies the remark. ■

A worker that chooses the difficult activity in the first period enjoys better career prospects than a worker that chooses the easy activity in the first period, since the outside firm learns the ability of the workers.

In a separating equilibrium, the difficult activity endogenously becomes the "prestige" activity. A worker that chooses the easy activity in the first period must therefore be compensated, due to the inferior career prospects. The following proposition describes the contracts and wage setting in separating equilibria in more detail.

Proposition 1 *A separating equilibrium has the following properties:*
(i) The job design is fully delegated to workers, and workers separate efficiently.
(ii) High workers have a steeper wage profile than low workers across the two periods.

Proof. The argument is simple. In order for a low worker to choose the right activity in the first period, the lifetime utility for a low worker for choosing the E activity must be at least as high as the lifetime utility for choosing the D activity,

$$w_1^E + w_2^E \geq w_1^D + w_2^D \tag{1}$$

Applying the same argument for a high worker, such a worker chooses the right activity if and only if,

$$w_1^D + w_2^D \geq w_1^E + w_2^E \tag{2}$$

Combining (1) and (2), we get that a separating equilibrium must have,

$$w_1^E + w_2^E = w_1^D + w_2^D \tag{3}$$

(3) is the double indifference condition referred to in Section 2.1. The only way to ensure an efficient allocation of workers is to set wages such that (3) holds, and allow workers to choose their activity. Hence workers are given full delegation in a separating equilibrium.

We showed in Remark 1 that $w_2^E = \pi^0$, $w_2^D = \pi^H$. It then follows from (3) and zero profits (i.e., $\theta\pi^H + (1-\theta)\pi^0 = \theta w_1^D + (1-\theta)w_1^E$) that $w_1^D = \pi^0 + (2\theta - 1)(\pi^H - \pi^0)$ and $w_1^E = \pi^H + (2\theta - 1)(\pi^H - \pi^0)$. Hence low (high) workers are paid weakly more (less) than their marginal productivity in both periods in a separating equilibrium. It also follows that high workers have a steeper wage dynamics than a low worker in a separating equilibrium. ■

In a separating equilibrium, a firm sets no limit to entry in any of the activities, and we can interpret this equilibrium as a situation where workers are hired and then fully delegated the job design choice.[10] To be willing to reveal low ability by choosing the easy activity, low workers must be compensated by a high wage in the first period. This implies that the wage profile of high workers is steeper than the wage profile of the low workers.

A separating equilibrium identifies the advantages of delegation in this economy. If firms instead assign workers to the E activity or to the D activity 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 θ $(= (\pi^0 - \pi^L)/(\pi^H - \pi^L))$.[11] 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 $\theta\pi^H + (1-\theta)\pi^L = \pi^0$. Furthermore, separating equilibria are "fair", in that the lifetime utility of low and high workers are equal. This follows from the fact that $w_1^E + w_2^E = w_1^D + w_2^D$ must hold for workers to sort into their efficient activity in the first period.

We now explain the conditions for existence of a full delegation, separating, equilib-

[10] Instead of delegating the job design choice, a separating equilibrium could be implemented by workers reporting their type and the principal assigning workers. However, low ability workers might fear that the firm would use the report against them, through various measures of discrimination, and delegation may be a cheap way for the firm to commit to not exploiting the information. Another reason for strictly prefering delegation could be costs of communication. Dessein (2002) contains an analysis that goes much further into this point.

[11] A similar result holds if the firm assigns workers randomly, in which case production would equal $\theta\frac{\pi^H + \pi^0}{2} + (1-\theta)\frac{\pi^L + \pi^0}{2}$.

rium. A natural measure of the *returns to ability* in this economy is $\pi^H - \pi^0$. This is the difference in productivity between the workers under full information, and also the difference in actual productivity in a separating equilibrium. On the other hand, $\pi^0 - \pi^L$ measures the cost of misallocating a low worker to the difficult activity. We will denote this by the "Peter's cost".

Remark 2 *A separating equilibrium is more likely to exist the lower returns to ability and the higher Peter's cost.*

Proof. Denote the wages in a separating equilibrium by $\{w_1^E,\ w_1^D,\ w_2^E,\ w_2^D\}$. As shown before, these wages are uniquely determined by bidding in the second period, the self-sorting constraint (3) and first period zero profits for firms. We now check under which circumstances these wages are consistent with optimal behavior by firms in the first period.

First notice that a firm would never raise w_1^D or w_1^E because it would then attract both type of workers and since the initial equilibrium has zero profits for the firm, this would result in a loss. Likewise, a firm cannot gain from lowering w_1^D because the high workers would then prefer the other firm. We therefore need to consider deviations where firms attempt to cream-skim by lowering w_1^E and keeping w_1^D constant. Suppose therefore that firm 1 sticks to the wage schedule $\{w_1^E,w_1^D\}$ and firm 2 deviates by offering the wage schedule $\{w\text{'}_1^E,w_1^D\}$, where $w\text{'}_1^E < w_2^E$. If firm 2 could commit to such a schedule, it would attract a share of the high workers while a disproportionate share of the low workers would choose firm 1. Consequently firm 2 would run a profit, since high workers are paid less than their marginal productivity.

Suppose that a low worker, after an offer $w\text{'}_1^E$, by mistake chose to work for firm 2 (in which case he would choose the D activity). If the probability of a mistake is positive, firm 2 would wish to revise $w\text{'}_1^E$ after the workers have chosen which firm to work for. This would make some workers choose E rather than D, and improve the allocation of workers inside the firm (and profits).[12] Denote this revised offer for $w\text{"}_1^E$. The extra compensation required to make this low worker prefer E to D would be the loss of career gains from

[12]If there are costs of adjusting the wage upwards, the probability of mistake needs to be correspondingly greater than zero.

choosing D, i.e. $w_2^D - w_2^E = \pi^H - \pi^0$. (The firm would choose to increase the wage by the minimum amount, so that $w\text{"}_1^E = w_1^D + (w_2^D - w_2^E) = w_1^E$.) The productivity gain from making a low worker choose E instead of D is $\pi^0 - \pi^L$. Hence, a firm would prefer to raise the wage for E to $w\text{"}_1^E$ (from $w\text{'}_1^E$) if the extra wage compensation is less than the productivity improvement or if,

$$\pi^H - \pi^0 < \pi^0 - \pi^L. \tag{4}$$

Cream-skimming by offering $w\text{'}_2^E$ would not be credible if (4) holds. Consequently, there exists a separating equilibrium when (4) holds. On the other hand, when (4) does not hold, a firm can profit by deviating through (credible) cream-skimming, and a separating equilibrium cannot exist. Hence a separating equilibrium is more likely to exist the lower returns to talent (the lower the left hand side of equation 4) and the higher the Peter's cost (the higher the right hand side of equation 4). ■

In a separating equilibrium, firms pay low workers a premium above their marginal productivity in the first period, to make such workers self-sort. This creates a potential incentive for firms to deviate in order to attract only high workers, by holding the offer w_1^D fixed and reducing w_1^E. When it is sufficiently inexpensive for firms to make low workers 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 credible, and a separating equilibrium exists. When returns to ability are low, $w_2^D - w_2^E$ is low, and it is cheap to revise the offer w_1^E upwards to make low workers choose the easy rather than the difficult activity. Hence when returns to ability are low, cream-skimming cannot be credible and a separating equilibrium exists. On the other hand, when returns to ability are high, $w_2^D - w_2^E$ is high, and it is expensive to revise the offer w_1^E upwards to make low workers switch from the difficult to the easy activity in the interim. Therefore, a separating equilibrium is less likely to exist the higher returns to ability and the lower Peter's cost.

The full delegation in a separating equilibrium differs radically in spirit from the assignment and job design literatures, where firms direct workers to do specific activities rather than delegating the choice. That high-delegation practices inside firms are common on a wide basis are indicated by the pioneering study of Osterman (1994), which reports on the degree of employee discretion in 875 US companies (with 50+ employees). Osterman finds that 45% of employees have complete or large discretion over the choice of work method. This is captured well by the model, recall that we can interpret delegation as both on which job to undertake and which work method to use. Section 5.1 contains a wider discussion of the relation between our work and empirical evidence on delegation.

Empirical work has shown that worker (nominal) wages and wage dispersion typically increase over time (see Gibbons & Waldman, 1999b, for an excellent overview of the careers in organization literature). In an older version of the paper, we showed that separating equilibria have these properties given that we accommodate a degree of human capital acquisition between the two periods and that the effort cost (such as hours on the job) is higher for completing the difficult activity than the easy activity.[13]

The literature on adverse selection in labor markets (for example, Greenwald, 1986, Foster & Rosenzweig (1996), and Acemoglu & Pischke 1998) implicitly assumes that the workers ability is revealed to the firm once hired, in contrast to in our setting. Due to this difference in assumptions, firms in our setting face not one but *two* adverse selection problems: at the hiring stage and when workers allocate inside the firm. To illustrate the second adverse selection problem, suppose a firm simply decided not to assign the workers - and set equal wages for the two activities. In that case, low-quality workers would imitate high-quality workers and herd into the more prestigious activity (to obtain a higher future compensation), resulting in an efficiency loss.

This argument highlights the role of firms in our model; firms do not exist because they can bear risk better than workers, or because they are needed to coordinate different lines of production, but to adjust wages (and the degree of delegation) to ensure a second-best allocation of workers. This point will be illustrated in the next section.

3.2 Rationing equilibria

We now consider the delegation policy of a firm when cream-skimming is a viable threat and a separating equilibrium consequently does not exist.

[13]These properties also make separating equilibria consistent with the low ability workers making lower wages than high ability workers in both periods. Briefly, human capital acquisition will ensure that the wage profile of both types of workers are increasing, and extra hours required to finish the difficult task ensures that the high ability workers will earn more in both periods.

Proposition 2 *(i) When a separating equilibrium does not exist, there exists a rationing equilibrium, where only a fraction of workers are delegated the job design decision, and the remaining fraction of workers is assigned to the difficult activity.*

(ii) There does not exist a rationing equilibrium where any workers are assigned to the easy activity.

(iii) The fraction of workers that are assigned increases in the returns to talent $\pi^H - \pi^0$ *and decreases in the Peter's cost* $\pi^0 - \pi^L$*.*

Proof. We show at the end of this proof that a rationing equilibrium must involve the E activity slots being rationed and the slots in the D activity being freely available.

We start by determining the equilibrium wages $\{w_1^D, w_2^D, w_1^E, w_2^E\}$ for a given level of such rationing, and then determine the degree of rationing. We then show the proposed equilibrium is renegotiation-proof.

Suppose that the degree of rationing equals f so that a fraction f of the low workers are forced into the difficult activity. In that case, the fraction of high workers in the two activities become,

$$\begin{aligned} \theta^D &= \frac{\theta}{\theta + f(1-\theta)} \\ \theta^E &= 0 \end{aligned} \tag{5}$$

Since the firms must earn zero profits in the second period, we can determine the second period wages as,

$$\begin{aligned} w_2^D &= \theta^D \pi^H + (1-\theta^D)\pi^0 \\ w_2^E &= \pi^0 \end{aligned} \tag{6}$$

By self-sorting and zero profits in the first period, we can derive the difference in first period wages as $w_1^E - w_1^D = w_2^D - w_2^E$. We are now able to characterize this wage difference as a function of f.

Using equations (5) and (6), we find that,

$$\begin{aligned} w_2^D - w_2^E &= \frac{\theta}{\theta + f(1-\theta)}\pi^H + \frac{f(1-\theta)}{\theta + f(1-\theta)}\pi^0 - \pi^0 \\ &= \frac{\theta}{\theta + f(1-\theta)}(\pi^H - \pi^0) \end{aligned} \tag{7}$$

The net gains from inducing one low worker to switch in the interim, NG, equals

$$NG(f) = \pi^0 - \pi^L - (w_2^D - w_2^E) = \pi^0 - \pi^L - \frac{\theta}{\theta + f(1-\theta)}(\pi^H - \pi^0) \tag{8}$$

The first two terms are the productivity improvement from a switch and the third term is the compensation needed to make a worker switch. Since changing f in the interim will affect the compensation necessary to induce workers to switch, the net gains from moving more than one worker needs to take into account that the second period wages are a function of f. The gain from moving workers (the first two terms) is independent of f, and the cost from moving workers (the third term) is convex and decreasing in f. There is thus a unique f^* such that $NG(f^*) = 0$. For all $f < f^*$, we have $NG(f) > 0$ which gives incentive to allow low-ability workers to switch to the E activity and thereby increase f. Once $f = f^*$, there is no longer such an incentive since $NG(f) < 0$ for all $f > f^*$. Hence a firm would not wish to increase delegation (decrease f) in the interim, and the proposed equilibrium is renegotiation-proof.

By looking further at the condition where $NG(f^*) = 0$, we have

$$\pi^0 - \pi^L = \frac{\theta}{\theta + f^*(1-\theta)}(\pi^H - \pi^0) \tag{9}$$

From this we see that f^* is increasing in $\pi^H - \pi^L$ and increasing in $\pi^0 - \pi^L$. Hence, the equilibrium degree of rationing increases in the returns to ability and decreases in the cost of misallocating a low worker.[14] We have then proved (i) and (iii).

We now prove (ii), that there cannot be rationing 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 wish to work in D. In that case, the proportion of workers should be the same in both jobs. If this happens, there are no career concerns since no information inferred by activity choice. Because of this, the firm can induce a high worker 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 of workers wish to work in D. The second possibility is that the low type wishes to work in E, while the high type workers wish to work in D. In that case, total wages must be equalized across activities. But then, the firm can increase

[14] If there does not exist a solution for f^* on (0,1), the equilibrium must be separating or pooling.

profits by allowing a higher fraction of workers in D, by allowing workers to move from E to D (since only high workers would wish to move). This occurs since both the wage in D is lower than in E (the fraction of high workers in D is higher than in E) and productivity of high workers is higher in D. Hence a situation where the slots in D are rationed cannot be an equilibrium. ■

When the returns to ability are high or the Peter's cost is small, full delegation implies that the wage difference $w_2^D - w_2^E$ would be high, the low workers would require a high premium in the first period to separate, and credible cream-skimming by the other firm would make full delegation unprofitable. To avoid cream-skimming, a firm therefore assigns (some) workers, and thereby reduces the compensation required by the low workers that self-sort to the easy activity.[15] In rationing equilibria the firm is in effect 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.

The intuition for why there cannot be a rationing equilibrium where the number of slots in the difficult activity is that if the D slots were rationed, firms could increase productivity without increasing the costs of compensation, by letting more (high) workers do the difficult activity.[16]

A higher returns to ability or a lower Peter's cost imply a higher degree of rationing. Therefore, rationing equilibria capture both firms with a low degree of delegation, as in government bureaucracies, and more typical firms, where a certain fraction of employees are delegated the choice of specialization. When the measures are sufficiently extreme, there can exist *pooling equilibria* where all workers are assigned.

It is somewhat surprising that rationing equilibria imply an assignment to the difficult activity, not to the easy activity. We can add plausibility to this result by considering an example. A frequent complaint about bureaucracies is that too many persons are employed in middle-level administrative/management positions, rather than working on more customer-oriented, clerical activities (the Peter principle). We can interpret admin-

[15] 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.

[16] If the production technology were such that the simple task must be done (as with the product catalogue of Sun Hydraulics), a high degree of rationing in equilibrium implies that separate workers, without the option to switch to the difficult task, must be hired to do the easy task.

istration/management as the difficult activity and the clerical activity as the easy activity. Our results then provide an argument for why there are too many employees at the management level: a more efficient allocation would make it too easy for outside firms to cream-skim high-quality employees.[17]

In the extension of that point, observe that the argument behind rationing equilibria can provide a limit to the effectiveness of organizational reforms in the public sector, an issue continuously debated in many countries. In the short run, public sector bureaucracies might be able to keep the same level of production by downsizing and delegating more to the retained workers. However, such a policy would induce low future wages for those that reveal themselves as having lower ability, and to compensate these workers their current wage would have to be raised. This, in turn would create incentives for outside (perhaps private) firms to cream-skim, by offering worse conditions for low ability workers than the public sector would do. In the longer run this process could lead to the public sector being be drained of its talent, in that the fraction of low ability workers, paid above their marginal product, would become high. Hence a certain amount of misallocation in the public sector can be desirable.

What would happen if workers learn about the match with particular activities only after entering a firm, but before choosing activities? In that case, there would be no adverse selection at the hiring stage, since cream-skimming is not a viable strategy, and a higher wage gap would be sustainable in equilibrium. Workers learning about the match with activities after they enter the firm thus would support more delegation. This argument might be relevant for explaining why consulting firms hiring workers at the bottom level often give such workers, after an initial general training, a relatively high degree of discretion in deciding which industries to specialize in.

4 Asymmetrically informed firms

We now consider the case when the inside and the outside firm have asymmetric information about which activities the worker undertook before the bidding in the second

[17] Recent papers that discuss the Peter's principle within hierarchical models include Fairburn & Malcolmson (2001) and Lazear (2001).

period.[18] This will shed light on the relation between a firm's transparency and its delegation practices.

To fix ideas, we can think of the degree of asymmetric information as determined by the extent to which job titles and salaries are precise or diffuse. In this respect, Sun Hydraulics lies at one end of the spectrum by not having job titles for its employees, and a very covert pay policy, while bureaucracies, with well-defined job titles, job descriptions, and salary ladders, being at the other end. Since transparency can be part of organizational design, however, it can less obviously be considered as exogenous than the productivity parameters. At the end of the section we therefore discuss some possible determinants of transparency, and justify taking this variable as exogenous.

Before the second period, the *inside firm* (a worker's first period employer) is assumed to be fully informed about the activity a worker was engaged in. The *outside firm* (the competitor of a worker's first period employer), however, receives some private, imprecise, signal about it. Given their information, the inside firm and the outside firm compete for the workers before the second period. We assume that the bidding follows a first-price sealed-bid auction; each firm gives an offer to a worker, in ignorance of the other firm's offer, and the worker accepts the highest offer.[19]

We model the outside firm's information, which is private, about the activity a worker was engaged in (or his wage) in the first period as an independent realization of a random variable X, which can take two values, E and D. If the worker is in activity D, then

[18] Waldman (1983) consider job assignment when employers know more about the abilities of their workers than other firms do. Such private information may give employers incentives to hide their able workers, by e.g., delaying promotion. However, Waldman (1984) considers the case when the employer and the worker are equally well informed about the ability of the worker, excluding the main issues of the present paper.

[19] The first-price sealed-bid auction is realistic for situations where firms may bid in turn, but where workers have no way of verifying an offer made by one firm to the other firm. Hence firms make secret or unverifiable offers to workers, so that a worker cannot start a "bidding war" by documenting an offer from the other firm. This bidding setup ensures that there will be positive a turnover rate between the two periods (and a higher turnover rate for low workers). Hence there will be a lemons problem in equilibrium, but not to the extent that trade breaks down. The more standard sequential bidding structure of Greenwald (1986) is unable to explain equilibrium turnover without assuming 'utility shocks', i.e., an urge to change employer even if the inside firm offers a higher wage, in contrast to our approach.

$X = D$ with probability 1, and if the worker is in activity E, then $X = E$ with probability $p \in [0, 1)$, and $X = D$ with probability $1 - p$. This signal structure is assumed purely for convenience; our results are robust to a variety of ways to model the auction.[20] The signal precision p, or outside visibility, is common knowledge. When $p = 1$ the inside firm and the outside firm are symmetrically informed, as in the previous section.

The asymmetric information case is more complex than the symmetric information case due to the richer structure of the bidding equilibrium before the second period. We start out by solving for the bidding equilibrium given that a separating equilibrium is played.

Remark 3 *Given that a separating equilibrium is played,*
(i) $\pi^0 < w_2^E < w_2^D < \pi^H$.
(ii) $w_2^D - w_2^E$ *increases in* p.

Proof. See Appendix A. ■

Remark 3 gives the more important properties of the mixed-strategy Nash Equilibrium of the bidding game between the inside firm and the outside firm. Recall that the inside firm bids conditional on the true type of each worker (since ability is revealed to the inside firm in a separating equilibrium), and the outside firm bids conditional on the signal X. The intuition for part (i) is the by now familiar one: a worker that chooses the difficult activity in the first period enjoys better career prospects than a worker that chooses the easy activity in the first period, since the outside firm (partially) learns the ability of the worker. The intuition for part (ii) is that a more informative signal means that more is learned by the outside firm about the ability type of a worker before the second period, and there will be a more intense competition for a worker that chooses the difficult activity in the first period. Hence the wage difference in the second period increases in the degree

[20] We make this assumption to ensure that the outside firm makes zero profits, which makes the auction solution simpler. Our results in Section 4 are robust to letting the private signal structure being symmetric, and to the information received by the outside firm being public. Since it is not obvious what the actual 'rules' of bidding games in labor markets are, we should emphasize that our modeling choice is one of convenience; any bidding setup where Remark 3, part ii), holds would work, which would be the case e.g., in certain hybrid versions of first-price auctions and the auction considered by Greenwald (1986).

of outside observability of activity choice in the first period.

Property (ii) of Remark 3 is the main building block to the next result. To anticipate, the immediate implication of part (ii) is that a firm must pay low ability workers a higher wage in the first period to be willing to sort, the higher p. This in turn makes the threat of cream-skimming stronger since the necessary compensation for choosing the easy activity. Therefore, a separating equilibrium is less prone to exist the higher p. To alleviate the wage difference between low and high ability workers in period 2 (and thus the potential for cream-skimming in the first period) a firm rations the slots in one of the activities. Hence, we get an equilibrium where the firm designs the job for some workers.

Proposition 3 *(i) A separating, full-delegation, equilibrium is more likely to exist the lower the p and always exists for $p = 0$. (ii) When a separating equilibrium does not exist, there exists a rationing equilibrium where the slots in the easy activity are rationed. (iii) The fraction of workers that are assigned in rationing equilibria increases in the signal precision p.*

Proof. See Appendices A & B. ■

We see that a separating equilibrium is more likely to exist the lower transparency. Public sector units normally have job titles (and individual salaries) that are on a clear-defined ladder, and are hence relatively informative about the type of work individuals do. Part of the reason for this is probably that coordination costs from having obscure job titles may be high in a larger organizations, but equally important there will be political regulations promoting transparency, to make the bureaucracy accountable to the politicians (and voters). On the other hand, we envisage industries like hi-tech, with less-defined ladders and job titles, and often managers with a substantial ownership share so that accountability is less of a problem, to have a lower p. Hence Proposition 3 captures well some of the differences between Sun Hydraulics, where the degree of delegation is high and the degree of transparency is low,[21] and public bureaucracies, where the degree of delegation is lower and the degree of transparency is lower.

[21] For example, job titles being non-existent at Sun (Baron & Kreps, 1999, p. 295), it is hard for outside firms to assess the allocation of individual employees. Perhaps not surprisingly, the pay levels of individual workers is also very covert information in these firms.

Obviously, if a firm could choose its p freely and without costs, it would choose p as low as possible, to avoid cream-skimming and to obtain an informational advantage vis-a-vis the other firm. One reason for why p is not easily manipulable, and different across firms (or industries), could be that it is shaped by company culture (the degree of openness), which is slow to change. Even when p is manipulable, a potential cost is that a low visibility firm would run into problems with recruiting employees with the highest potential for learning (such learning could take the form of a productivity improvement between the first and the second period), since such employees would tend to prefer to work for firms where their learning potential later will be revealed to the market. Another cost of lowering visibility could be increased coordination costs inside the firm, such as those due to the duplication of work, since decreasing visibility from the outside would probably mean making the organization less transparent also for insiders.[22] This argument may partially explain why industries dominated by small start-ups, as segments of the software industry, seemingly have a high degree of delegation: the (incremental) coordination costs from having diffuse job descriptions are small. The reverse argument may explain why larger firms seemingly have more precise job descriptions and a lesser degree of delegation.[23]

5 Discussion

We have provided a theory of job design that in a tractable manner accommodates the delegation practices of hi-tech firms, of (government) bureaucracies, and of firms in be-

[22] Herbold (2002) gives a vivid description of the coordination problems that occured due to too much delegation at Microsoft.

[23] Osterman (1994) gives some support to this hypothesis. A related hypothesis relates delegation to ownership structure. For a publicly held firm with a dispersed ownership structure to be accountable to shareholders, the shareholders need to have access to the operations of the firm, including its personell policy. For a privately held firm there is less need for such outside visibility since the owners are either insiders to the firm, or the number of outside owners is small so that free-riding on information acquisition is a minor problem. From this, we expect publicly held firms to delegate less than privately held firms, and have a higher degree of misallocation. Among the costs of a closer ownership structure is the lesser wealth diversification by owners of privately held firms, so from a security design perspective we can envisage a trade-off between higher productivity and more diversification.

tween. This section discusses various issues related to the theory; testable implications and relation to evidence, implications for misallocation of talent within firms, some implications for which decisions are more likely to be delegated, and finally on the role of private benefits in our model.

5.1 Some testable implications and evidence

Our main message is that firms should delegate less the stronger career concerns of their workers. From this insight we can expect Japanese type of firms, with long-term employment relations and priority of job security, to delegate more than American type of firms with shorter-term contracts and higher mobility. This hypothesis is consistent with considerable evidence, as described by Aoki (1986). The same type of reasoning may also shed light on Rajan & Wulf (2003) who 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. As longer-term commitment from firms implies lesser career concerns for workers, this finding is also consistent with our line of reasoning.[24]

A different type of testable implication is that delegation is more prone to occur in firms (or levels of the organization) where the Peter's cost is larger. Comparative statics on the Peter's cost in equations (7) and (9) show that the degree of delegation and the simple wage dispersion [measured by $w_2^D - w_2^E$] increases with the Peter's cost. Intuitively, more delegation implies that more information is revealed about the ability of individual workers, and a higher degree of wage dispersion follows. This result is consistent with the empirical finding of Bauer & Brender (2003), which using a matched employer-employee dataset from Germany finds that firm level wage dispersion increases in the degree of delegation.

[24] Or maybe a firm needs to create performance incentives if it wishes to delegate, and the use of stocks/options is a response to that need, as in Prendergast (2002). This argument is consistent with the analysis in Section 6.

5.2 Misallocation of talent

Recall that misallocation of talent occurs whenever a high (low) worker is allocated to the easy (difficult) activity. We then have that,

Proposition 4 *(i) Misallocation of workers can occur in equilibrium, and is lower the higher degree of delegation. (ii) Misallocation occurs due to low ability workers performing the difficult activity.*

Proof. Follows directly from Proposition 2 and Proposition 3. ■

A natural question is what hypothesis we can derive on misallocation within firms for a cross-section of firms from different industries. Suppose that productivity parameters are stable across firms within industries but vary across firms between industries. Then the degree of delegation will be constant across firms in the same industry, while we get the following for a cross-section of firms between industries.

Proposition 5 *For a cross-section of firms, (i) The degree of misallocation and the degree of delegation are negatively related, and (ii) The wage levels and the degree of delegation are positively related.*

Proof. Follows directly from Proposition 2 and Proposition 3. ■

Increased outside observability or increased returns to talent gives less delegation and more misallocation, and for a cross-section of firms from different industries, the degree of misallocation and the degree of delegation are inversely related in equilibrium. From this result, we can expect a higher degree of misallocation of workers in industries with a high degree of outside observability than in industries with a lower outside observability. This implication is consistent with casual empiricism on the high efficiency of hi-tech firms compared to government bureaucracies.

Since more delegation is associated with a more efficient allocation of workers, we also expect wages to be higher in industries or firms with higher delegation. This hypothesis is confirmed by Bauer (2001), who finds a positive relation between wage levels and degree of delegation for workers in a panel of German firms. Rajan & Wulf (2003) on the other hand do not find conclusive evidence on the relation between pay levels and degree of delegation (decentralization) for their cross-section of firms.[25]

5.3 Which decisions should be delegated?

Theories of delegation, including this paper to this point, focus on situations where only one decision may be delegated to subordinates. However, in real-life situations principals have the option to delegate several decisions. For example, principals may delegate either task choice decisions or work method decision, or all at the same time. It is therefore of interest to ask which types of decisions are more likely to be delegated.

One question is whether more "important" decisions are more (or less) likely to be delegated.[26] Or is delegation correlated with other dimensions of a decision? Although the present model setup does not incorporate more than one activity, our results still allow us to infer that firms are less likely to delegate decisions that have high returns to talent, high transparency, or a low Peter's cost. The argument for this follows the logic of the previous sections; if a firm delegates a highly transparent decision, or a decision with a high returns to talent, then it will need to pay a high premium to low ability workers, and thereby be exposed to cream-skimming.

From this observation, we can conjecture that firms are more likely to delegate work method decisions than task choice decisions, because the latter are presumably more observable, and are also associated with a higher returns to talent.[27] In other cases, matters are more complicated because of the interaction between transparency and returns

[25] General equilibrium effects might play a role. Since wages will be higher in industries with higher degree of delegation, we would expect an inflow of workers into these industries from workers in industries with a lower degree of delegation. In the current setting, firms operate under constant returns to scale, which means that a high wage sector can absorb all the workers in the economy without wages becoming lower. More realistically, there can be demand side effects market from workers migrating into a sector, driving wages down, which can partially explain the lack of support of our hypothesis. Notice, however, that even with migration of workers, it would still be the case that the low-delegation industries would have a higher degree of delegation and a lower degree of misallocation than low-delegation industries.

[26] Rajan & Zingales (2001) argue that decisions connected to the secrecy of the firm's "critical resource" will not be delegated.

[27] Admittedly, task choice can be associated with a higher Peter's cost than work method. This effect pulls in the opposite direction.

to talent. Think for example of the firm's public relations "task". Such a function in a firm is obviously very observable, but on the other hand probably has a lower returns to talent than say decisions in strategy processes. So it is less clear from our theory to which extent delegation will occur in such a case.

5.4 Private benefits and delegation

The incentive problems we have focused on stem from the workers having private benefits from their choice of activity that is due to better career prospects from being engaged in the difficult, "prestigious", activity. The reason why the firm cannot fully counteract the private benefits in the incentive scheme, and thereby create perfect self-sorting, is that the market puts restrictions on what type of contracts can be (credibly) offered, through the cream-skimming restriction. Essentially, this restriction limits the magnitude of the premium that can be paid to low workers (that is, wage above marginal productivity) in order to induce them to self-sort into the less prestigious activity.

One could argue that there are other private benefits than career prospects that may contribute to a worker desiring the prestige activity: sense of importance, recognition by peers, friends, or parents, increased attractiveness for potential partners, and other sociological effects. Such effects will work in a similar manner to career concerns, and tend to limit the degree of delegation. One may also argue that the desire to be in the activity where you are most productive will not be foregone for a small salary increase. Naturally, such a desire makes it easier for firms to sort workers. This extra slack would work in a similar way to performance contracts, considered in the next section.

We can also think of more direct differences in private benefits between the activities. Suppose that we allow for effort costs (hours on the job) 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.[28]

[28] If the cost of effort differ for the two types of workers, this creates the possibility of the firm screening the workers, which would work in the same qualitative manner as performance contracts, analyzed in the next section.

A similar argument occurs if effort is non-observable.[29] Suppose that effort is less measurable in the difficult activity. In that case, effort will be more underprovided (relative to the first-best level) in the difficult activity, and the effective π^H and π^L decrease relative to the effective π^0. The net effect on the degree of delegation depends on whether the effective $\pi^H - \pi^0$ decreases more than the effective $\pi^L - \pi^0$. If it does, the degree of delegation will decrease, and if not it will increase.[30] So the theory does not give a definite answer on whether a less measurable effort cost in the difficult activity will increase or decrease the degree of delegation.

6 Optimal performance contracts

To amplify our points, we have made some strong assumptions. In particular, we have considered a case where firms can only distinguish between types of workers by their actions and thus can only sort workers by offering a schedule that makes both worker types indifferent between which activity to choose. What if other instruments of sorting workers than delegation were available to the firm? In this section we consider the case where contracts based on individual performance in the first period are feasible (the analysis with performance contracts being possible in both periods gives qualitatively the same results, but with more notation). The main results of the section show that our insights are *strengthened* by the introduction of (noisy) measures of individual performance, in that we obtain equilibria with the same qualitative features with respect to delegation and premium paid to low ability workers, but where the high type worker strictly prefers the difficult activity.[31]

We assume that output in the E activity is as before independent of ability, and with mean π^0. Those choosing E will therefore be offered a fixed salary denoted by F. We

[29] We also need workers to be risk-averse or to have limited liability, or other reasons for the first best levels of effort to not be implemented.

[30] The logic applied here is the same as behind equation (4), comparing the returns to ability $\pi^H - \pi^0$ with the Peter's cost $\pi^0 - \pi^L$.

[31] Performance contracts in the current setting only serves to sort workers. We can easily extend the model to encompass moral hazard problems. Such considerations would induce additional inefficiencies that are not our focus here.

furthermore assume that there are two possible output levels in the D activity, π^{low} and π^{high}, where a low (high) type worker has a probability P_L (P_H) of obtaining π^{high}. The expected output for the low (high) worker equals π^L (π^H), and therefore $P_L < P_H$. Hence, in this first period, a worker in D will be offered a wage contingent on the output π^{low} or π^{high} which we denote as B_L and B_H, respectively. In the second period, output is not contractible to either firm (as in the previous sections). While second period wages cannot be contingent on second-period output, they may indeed depend upon first-period output in addition to activity chosen (this of course would not happen in a separating equilibrium). To avoid trivial forcing contracts, we assume that workers are risk-neutral, but have limited liability, so that F, B_L, $B_H \geq 0$. In addition, to focus on the rationing mechanism, we assume that information is symmetric between the inside and the inside firm both with respect to activity ($p = 1$) and with respect to the performance of a worker.[32]

6.1 Separating equilibria

In a separating equilibrium, worker abilities are revealed to both firms before the second period and the second period wage must be π^H for high workers and π^0 for low workers. 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 worker choosing activity j in the first period for U_i^j, we then have,

$$U_H^D = P_H B + \pi^H \tag{10}$$

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, for a worker choosing the D activity (remember in a separating equilibrium only high workers choose this activity). On the other hand, the utility for a low worker for choosing the E activity equals,

$$U_L^E = F + \pi^0 \tag{11}$$

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

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,

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

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

(IC1) is the self-sorting constraint for high type workers, and (IC2) is the self-sorting constraint for the low type workers.[33]

If $F > \pi^0$, then (IC2) binds.[34] In that case, we can determine F as,

$$F = P_L B + \pi^H - \pi^0 \tag{12}$$

This implies (by $P_L < P_H$) that high-ability workers strictly prefer the D activity in a separating equilibrium, thus (IC1) holds as well.

The zero profit condition is,

$$\theta\pi^H + (1-\theta)\pi^0 = \theta B P_H + (1-\theta)F \tag{13}$$

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

$$\begin{aligned} B^* &= \frac{\pi^0 - (1-2\theta)(\pi^H - \pi^0)}{\theta(P_H - P_L) + P_L} \\ F^* &= \frac{\theta(P_H + P_L)(\pi^H - \pi^0) + P_L\pi^0}{\theta(P_H - P_L) + P_L} \end{aligned} \tag{14}$$

To have the same type of separating equilibrium as in the previous sections, where the low type is paid above marginal productivity to self-sort, we need that $F^* > \pi^0$.[35] From (14),

[33] 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.

[34] 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.

[35] 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.

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 (12), 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 is both type of workers get (expected) wage equal to marginal productivity in both periods, which is a qualitatively different separating equilibrium from that obtained previously.[36] To examine additional conditions for existence of a separating equilibrium where low workers are 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 workers have chosen which firm to work for, by raising the wage for the easy activity such that $w_1^E = F$, if the production gain exceeds the wage compensation loss. The extra compensation needed to induce a low ability worker to switch activities equals $\pi^H - \pi^0$, that is the wage loss in period 2 from being revealed as having low ability. It will pay to make this compensation only if the productivity improvement exceeds the extra compensation, that is

$$\pi^H - \pi^0 < \pi^0 - \pi^L \tag{15}$$

When this no cream-skimming condition holds, a separating equilibrium exists, which is analogous to the case without performance contracts (equation 4). 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 ability workers are compensated to self-sort, exists whenever (15) holds and $\dfrac{\pi^H}{P_H} > \dfrac{\pi^L}{P_L}$. 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 6 *When the no cream-skimming condition (15) holds and the monitoring*

[36] This solution will satisfy (IC2) if $2\pi_0 \geq P_L\dfrac{\pi_H}{P_H} + \pi_H$, 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.

technology is not too precise, a separating equilibrium exists where the low ability workers are paid above their marginal productivity. When monitoring is precise, a separating equilibrium exists where both workers are paid their marginal productivity. In both types of separating equilibria, all workers are fully delegated the job design decision, and a high ability worker strictly prefers the difficult activity.

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

6.2 Rationing equilibria

In a rationing equilibrium, a worker that chooses E in the first period will be of low ability with probability 1, and will therefore get the wage π^0 in the second period. For a worker that chooses D, the wage in the second period will depend on the fraction of low workers in D and on whether that worker obtained a bonus or not. Recall the assumption that pay can only be conditioned on performance in the first period, and hence that workers simply get their expected productivity conditional on correct sorting in the second period.

Let θ_H (θ_L) be the fraction of workers with a high (low) performance that is of high ability, and let f be the fraction of low ability workers that are assigned to D, while a fraction $1-f$ are allowed to choose freely, and hence choose E. Then,

$$\begin{aligned} \theta_H &= \frac{\theta P_H}{\theta P_H + (1-\theta) f P_L} \\ \theta_L &= \frac{\theta(1-P_H)}{\theta(1-P_H) + (1-\theta) f (1-P_L)} \end{aligned} \tag{16}$$

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

$$\begin{aligned} w_2^H &= \theta_H \pi^H + (1-\theta_H)\pi^0 \\ w_2^L &= \theta_L \pi^H + (1-\theta_L)\pi^0 \end{aligned} \tag{17}$$

$w_2^H > w_2^L$ since a high ability worker has a better chance of getting a bonus than a low ability worker. We now have the IC conditions for a rationing 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 high type workers, and (IC4) the self-sorting constraint for the low type workers in a rationing equilibrium. As with a separating equilibrium, if $F > \pi^0$ and (IC4) were not binding, a firm can improve profits by lowering F and getting a smaller fraction of low type workers. Hence we can determine F as,

$$F = P_L(B + w_2^H) + (1 - P_L)w_2^L - \pi^0 \tag{18}$$

Since (IC4) binds, (IC3) becomes redundant (by $P_L < P_H$ and $w_2^L < w_2^H$), and high ability workers must strictly prefer D also in a rationing equilibrium.

The first period zero profit condition is,

$$\theta\pi^H + (1-\theta)(1-f)\pi^0 + (1-\theta)f\pi^L = \theta BP_H + (1-\theta)(1-f)F + (1-\theta)fBP_L. \tag{19}$$

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 high ability workers, the second term is the productivity of the low ability workers that choose E, and the third term is the productivity of low ability workers that are rationed. The right hand side gives the corresponding wages for those three groups of workers. The third equilibrium condition is that firms should be indifferent between shifting low ability workers (i.e., decreasing f) on the margin, i.e., that $\pi^0 - \pi^L = F - P_L B$. Again, the productivity improvement from shifting workers 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,

$$\begin{aligned}
B^* &= \frac{\theta(\pi^H - \pi^L) + \pi^L}{\theta(P_H - P_L) + P_L} \\
F^* &= \pi^0 + \frac{\theta(P_L\pi^H - P_H\pi^L)}{\theta(P_H - P_L) + P_L} \\
f^* &= \frac{\theta P_H(P_L(\pi^H - \pi^0) + \pi^L - \pi^0)}{P_L(1-\theta)(P_L(\pi^H - \pi^0) + 2\pi^0 - \pi^H - \pi^L)}
\end{aligned} \tag{20}$$

The degree of rationing 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 low workers ($F^* > \pi^0$) whenever $\dfrac{\pi^H}{P_H} > \dfrac{\pi^L}{P_L}$, which is the same condition on monitoring as described above.[37] To see that there cannot be rationing in the case of perfect monitoring technology, that is when $P_L = 0$ and $P_H = 1$, observe that the denominator of f^* goes to 0 when P_L approaches zero. By solving for $f^* = 0$, we get that rationing 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 rationing, as shown before. We can then summarize.

Proposition 7 *If a separating equilibrium does not exist, there exists a rationing equilibrium where some workers are assigned to the D activity. In such an equilibrium, a low ability worker is paid a premium to be willing to self-sort, and a high ability worker strictly prefers the D activity to the E activity. Moreover, the degree of rationing 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 rationing, a low type worker is paid a premium to be willing to self-sort, and moreover a high type worker 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.

7 Concluding remarks

Why do firms delegate job design to workers, and what are the implications of such delegation? We have developed a private-information based explanation of delegation, where delegation provides a more efficient allocation of talent inside the firm, but at the cost that low ability workers must be compensated to self-sort. Career concerns imposes a limit to the efficiency of delegation: when the returns to ability is high, the market

[37] 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 $F^* = \pi_0$.

observability of job content is high, or the cost of misallocating low ability workers is low, firms limit delegation to avoid cream-skimming of the high-ability workers. In short, we expect firms to delegate less the lower Peter's cost and the stronger career concerns.

Two implications of the theory are that the degree of misallocation of talent inside the firm decreases in the degree of delegation, and that misallocation takes the form of too many workers undertaking activities with a high return to ability, like administration or management, and that too few perform "simple" activities, such as customer service or catalogue revision. Finally, for a cross-section of firms we expect that firms with more delegation also have a lower degree of misallocation and higher wages. These hypotheses may be useful to empirical researchers collecting data on the extent and effects of delegation and various job design practices within firms.

Let us end the paper with a speculation. A fascinating aspect of organizations is that some seem much more innovative than others. The limited evidence of Sun Hydraulics, Gore, and more prominent firms such as Microsoft, suggests that free-wheeling organizations with a high degree of delegation innovate more than more traditional, hierarchical organizations. Can there be a link between firms' degree of delegation and their innovation rates? To discuss this question, we believe one would need to take into account such factors as learning potential of employees, the ownership/financial structure of the firm, and product market conditions, in addition to factors discussed in the current paper. That is left for future work.

8 Appendix A: Separating equilibrium with asymmetrically informed firms

Proof of Remark 3. This is a first price sealed-bid auction where the inside firm bids conditional on the true productivity of the worker, and the outside firm bids conditional on its private signal. There cannot exist a pure strategy auction equilibrium, and we here derive the mixed-strategy equilibrium.[38]

Recall that if the worker is in activity D, then the outside firm's private signal is D

[38] A similar auction was solved by Wilson (1967).

with probability 1, and if the worker is in activity E, then the signal is E with probability p and D with probability $1-p$.

The inside firm uses a mixed strategy with cumulative distribution of F^L for a low worker and F^H for a high worker. Clearly, the inside firm will never bid more than π^0 for a low worker. For reasons similar to Bertrand competition (see Kaplan and Wettstein, 2003), since there is a lower bound to wages (a worker would not pay to work) neither firm will bid below π^0 for a worker. Therefore, F^L must be the distribution degenerate at π^0. Thus the inside firm can only make a profit on high workers. For F^H, the support of the distribution will be $S_{inside} = [\pi^0, \bar{\pi}]$, where $\bar{\pi} < \pi^H$. The outside firm will, conditional on the realization of the signal being i, use a cumulative bid distribution $G^i(x)$ with support $S^i_{outside}$, where $i \in \{D, E\}$. Since neither firm bids below π^0 and since the outside firm when receiving the E signal knows with certainty that the worker is a low worker, it bids π^0 for those workers. Consequently, $S^E_{outside} = \{\pi^0\}$ and $S^D_{outside} = S_{inside} \equiv S$. Since the equilibrium is separating the probability of a high worker having the signal realization D equals 1. Given that the inside firm offers x to a high worker, the expected surplus the inside firm makes on that bid equals,

$$(\pi^H - x)G^D(x),\ x \in S, \tag{A1}$$

where $G^D(x)$ is the probability that the inside firm wins the auction for a high worker, and $(\pi^H - x)$ is the surplus if it wins. Since the inside firm plays a mixed strategy when bidding for a high worker, it must be indifferent at all points in its support,

$$(\pi^H - x)G^D(x) = k_{inside},\ x \in S, \tag{A2}$$

where k_{inside} is a constant that equals the surplus the inside firm makes on a high worker. Now define the probability of a worker being high conditional on the signal realization being D as θ^D. Then,

$$\theta^D = \frac{\theta}{\theta + (1-p)(1-\theta)} \tag{A3}$$

Given that the outside firm offers y to a worker with a signal D, the outside firm gets the expected surplus per worker,

$$\theta^D F^H(y)(\pi^H - y) + (1-\theta^D)F^L(y)(\pi^0 - y) = k^D_{outside},\ y \in S. \tag{A4}$$

The first term is the expected surplus when bidding for a high worker, and the second term is the expected loss from bidding for a low worker. By the same argument as for the inside firm, the outside firm must be indifferent at all points in his support.

Here we will see that the inside firm makes a profit. First notice that the upper end of the support must be strictly less than π^H. (Note that from previous arguments the bottom of the support must be weakly greater than π^0.) If not, the outside firm at the top of the support will earn a negative expected payoff. Now since the top of the support is strictly less than π^H, the inside firm can win with certainty at a wage less than productivity. This implies the inside firm must be earning a strictly positive profit.

At the bottom of the support the inside firm must have a strictly positive chance of winning (otherwise it can deviate to make a strictly positive profit by bidding at top of S). For this to happen, the outside firm must be playing an atom at the bottom of support. Since both firms cannot be playing atoms at this point, the outside firm must have a zero chance of winning at this point. This implies the outside firm makes zero profits.

Since the outside firm makes zero profits and F^L is degenerate at π^0 we can rewrite (A4) as,

$$\theta^D F^H(y)(\pi^H - y) + (1-\theta^D)(\pi^0 - y) = 0,\ y \in S. \tag{A5}$$

We can then substitute in for θ^D in (A5) to get

$$F^H(y) = \frac{(y-\pi^0)(1-p)(1-\theta)}{\pi^H - y},\ y \in S \tag{A6}$$

This distribution must be atomless, thus the support must start from π^0. We can determine the top of the support $\bar{\pi}$ by inserting $F^H(\bar{\pi}) = 1$ into (A5) to arrive at

$$\bar{\pi} = \theta^D \pi^H + (1-\theta^D)\pi^0 \tag{A7}$$

We can also substitute $x = \bar{\pi}$ into (A2) to get $k_{inside} = \pi^H - \bar{\pi}$ and write (A2) as,

$$(\pi^H - x)G^D(x) = \pi^H - \bar{\pi},\ x \in S \tag{A8}$$

which gives,

$$G^D(x) = \frac{\pi^H - \bar{\pi}}{\pi^H - x},\ x \in S \tag{A9}$$

Notice that this cdf places an atom at $x = \pi^0$, where the magnitude of the atom equals $\frac{\pi^H - \bar{\pi}}{\pi^H - \pi^0}$. We can observe that the induced density function increases in x, since the second derivative of G^D is positive. Furthermore, we can note that when $p < 1$, the inside firm makes positive information rents in the second period (on the high workers). These rents must be offset by negative profits in the first period.

The equilibrium (expected) wage for an agent of type j in the second period equals the expected maximum offer in the bidding before that period. For a low worker, the outside firm determines the expected wage,

$$w_2^E = p\pi^0 + (1-p)\int_{\pi^0}^{\bar{\pi}} z dG^D(z) \tag{A10}$$

The expected wage for a high worker equals,

$$w_2^D = \int_{\pi^0}^{\bar{\pi}} z dG^D(z)F^H(z) \tag{A11}$$

That $w_2^E > \pi^0$ and $w_2^D < \pi^H$ follows directly from $\pi^0 < \bar{\pi} < \pi^H$. Moreover, since $H(\cdot) \equiv G^D(\cdot)F^H(\cdot)$ first order stochastically dominates $G^D(.)$, it follows that $w_2^D > w_2^E$.

We now show that $w_2^D - w_2^E$ is monotonically increasing in p. Since in the second period the outside firm makes zero profits and there is full efficiency, we have

$$\theta\pi^H + (1-\theta)\pi^0 = \theta w_2^D + (1-\theta)w_2^E + \theta k_{inside} \tag{A12}$$

The left hand side is total production in the second period and the right hand side is total wages plus profits made by the inside firm. Using the derived expression for k_{inside}, the right hand side of (A12) equals $\theta(w_2^D - w_2^E) + \theta(\pi^H - \bar{\pi}) + w_2^E$, which implies

$$w_2^D - w_2^E = \frac{\theta\pi^H + (1-\theta)\pi^0 - w_2^E}{\theta} - (\pi^H - \bar{\pi}) \tag{A13}$$

By integrating (A10), w_2^E can be expressed as,

$$w_2^E = p\pi^0 + (1-p)[\bar{\pi} + (\pi^H - \bar{\pi})\ln(G^D(\pi^0)] \tag{A14}$$

Substituting the right hand side of (A14) into (A13),

$$w_2^D - w_2^E = \frac{\theta\pi^H + (1-\theta)\pi^0 - \{p\pi^0 + (1-p)[\bar{\pi} + (\pi^H - \bar{\pi})\ln(G^D(\pi^0)]\}}{\theta} - (\pi^H - \bar{\pi}) \tag{A15}$$

Notice that the only exogenous variables in this expression are p, θ, π^0, and π^H. Without loss of generality, we can normalize by setting $\pi^0 = 0$ and $\pi^H = 1$ and get,

$$w_2^D - w_2^E = \frac{\theta - (1-p)[\bar{\pi} + (1-\bar{\pi})\ln(G^D(0)]}{\theta} - (1-\bar{\pi}) \tag{A16}$$

Define $z = \dfrac{(1-p)(1-\theta)}{1-p+p\theta}$. Since $\bar{\pi} = 1 - z$ and $(1-\bar{\pi})\ln(G^D(0)) = z\ln(z)$, we have

$$w_2^D - w_2^E = 1 - z - \frac{(1-p)[(1-z)+z\ln(z)]}{\theta} \tag{A17}$$

Since $\dfrac{dz}{dp} = -\dfrac{(1-\theta)\theta}{(1-p+p\theta)^2} < 0$, the inverse function $p(z)$ exists and equals,

$$p(z) = \frac{1-z-\theta}{(1-\theta)(1-z)} \tag{A18}$$

Therefore, we can substitute in for $1 - p = \dfrac{\theta z}{(1-\theta)(1-z)}$ into (A17) to get,

$$w_2^D - w_2^E = 1 - z - \frac{z[1-z+z\ln(z)]}{(1-z)(1-\theta)} \tag{A19}$$

It is then sufficient to show that $w_2^D - w_2^E$ decreases in z for $z \in [0, 1-\theta]$. By differentiating (A19), we find that this is the case if,

$$(2-z)z\ln(z) + (1-z)(1+(1-\theta)(1-z)) > 0,\ z \in [0, 1-\theta] \tag{A20}$$

We can see this holds by taking the Taylor's series expansion around $z = 1$.

$$(1-\theta-\frac{1}{2})(z-1)^2 - \frac{2}{3}(z-1)^3 + \frac{1}{4}(z-1)^4....$$

All odd terms have negative coefficients and all even terms (starting with 4) have positive coefficients. The sum of the first two terms is positive if $(1-\theta-\frac{1}{2}) - \frac{2}{3}(z-1) > 0$ or if $\frac{3}{2}(1-\theta) + \frac{1}{4} > z$. This holds since $z < (1-\theta)$. Thus (A20) is always positive and we have our desired result that $w_2^D - w_2^E$ is increasing in p. ■

Proof of Proposition 3, part (i). We show that cream-skimming is more prone to occur the higher p. By an analogous argument with symmetrically informed firms, a separating equilibrium exists if it is sufficiently cheap to revise a cream-skimming attempt (a low wage for the easy activity) upwards, or if

$$w_2^D - w_2^E < \pi^0 - \pi^L. \tag{A21}$$

On the other hand, when (A21) does not hold, a firm can profit by deviating through (credible) cream-skimming, and a separating equilibrium cannot exist. Hence a separating equilibrium is more likely to exist the lower the wage difference $w_2^D - w_2^E$. Since this difference is increasing in p by Remark 3, a separating equilibrium is more likely to exist for lower p. The second part of Proposition 3 (i) is proved under Example 1 in Appendix B. (ii) and (iii) of Proposition 3 are proved in Appendix B. ■

9 Appendix B: Rationing Equilibrium

In this appendix, we characterize the rationing equilibrium that occurs when a separating equilibrium does not exist.

Proof of Proposition 3, part (ii) and (iii). The proof follows along the same lines as under symmetric information, with added complexity due to the auction equilibrium. By the same argument as under symmetric information, a rationing equilibrium must involve the E activity slots being rationed and the slots in the D activity being freely available.

Before the second period, the two firms bid for workers conditional on their information, where the inside firm knows the activity a worker was engaged in and the outside firm bids conditional on its signal X. Since the auction equilibrium under rationing is very similar to the auction equilibrium without rationing (derived in Remark 3), we sketch the former here.

Let f be the fraction of low workers that are forced into D in the first period. Let $\hat{\pi}$ be the average productivity in D, i.e., $\hat{\pi} = \dfrac{\theta\pi^H + f(1-\theta)\pi^0}{f(1-\theta)+\theta}$. In the auction before the second period, the outside firm makes zero profits and the inside firm makes the profit Δ_2 (we also denote Δ_1 as the profit of the inside firm in the first period), where

$$\Delta_2 = (\hat{\pi} - \bar{\pi})(\theta + f(1-\theta)) \tag{B1}$$

where $\bar{\pi}$ is the top of the support for the bidding used by the inside firm for a worker in the D activity (we will still call the cumulative distribution for this F^H) and the outside firm for a worker receiving a D signal (using cumulative distribution G^D). As before, this amount is determined by the condition that an inside firm can deviate to bidding on top of the support and winning with certainty. The profit for each worker in the D activity is then $\hat{\pi} - \bar{\pi}$ and the fraction of workers in this activity is $\theta + f(1-\theta)$ for a total profit of

Δ_2. The top of the support is determined by at what point an outside firm would make zero profits by bidding at this point, thus $\bar{\pi} = \phi\hat{\pi} + (1-\phi)\pi^0$ and ϕ is the probability of a given worker been occupied in D conditional on $X = D$, i.e.,

$$\phi = \frac{\theta + f(1-\theta)}{f(1-\theta) + \theta + (1-p)(1-f)(1-\theta)} \tag{B2}$$

Notice that $\hat{\pi} - \bar{\pi} = (1-\phi)(\hat{\pi} - \pi^0)$ by the definition of $\bar{\pi}$. Therefore,

$$\Delta_2 = (1-\phi)(\hat{\pi} - \pi^0)(\theta + f(1-\theta)) = (1-\phi)\theta(\pi^H - \pi^0). \tag{B3}$$

The distribution functions that support this solution are,

$$\begin{aligned} F^H(y) &= \frac{(y-\pi^0)(1-p)(1-\theta)(1-f)}{(\theta + (1-\theta)f)(\bar{\pi} - y)}, \ y \in S \\ G^D(x) &= \frac{\hat{\pi} - \bar{\pi}}{\hat{\pi} - x}, \ x \in S, \end{aligned} \tag{B4}$$

As before, the inside firm bids π^0 for a low worker in E, and the outside firm bids π^0 for a worker with $X = E$.

The second period auction equilibrium determines w_2^D and w_2^E as functions of f. The wage difference of the first period, $w_1^E - w_1^D$, can then be determined by the self-sorting constraint, i.e.,

$$w_1^D + w_2^D = w_1^E + w_2^E. \tag{B5}$$

This condition has the same interpretation as in a separating equilibrium. The total wage levels are determined by the overall zero profit constraint $\Delta_1 + \Delta_2 = 0$. For a given degree of rationing f, we have then determined the equilibrium wages $\{w_1^D, w_2^D, w_1^E, w_2^E\}$. We now determine the equilibrium rationing f^*.

With symmetrically informed firms, a firm will choose the minimal degree of rationing consistent with the no cream-skimming constraint. Due to the knowledge gain the inside firm makes from decreasing rationing, this constraint becomes more complex than with symmetrically informed firms.

The first period the profit of the inside firm equals Δ_1, where

$$\Delta_1 = \theta(\pi^H - w_1^D) + (1-\theta)[(\pi^L - w_1^D)f + (\pi^0 - w_1^E)(1-f)]. \tag{B6}$$

Suppose that a firm decreases the degree of rationing (and pays workers to switch) at the interim stage. The effect on first period profits from a marginal change in f equals,

$$\frac{d\Delta_1}{df} = -(1-\theta)[\pi^0 - \pi^L - (w_2^D - w_2^E)], \tag{B7}$$

The first term is the productivity gain and the second term is the added wage bill from changing the degree of rationing in the interim. The effect on the second period profits from a marginal change in the degree of rationing equals the gain a firm makes in the second period auction by knowing more about their workers $K(p, f)$, i.e.,

$$\frac{d\Delta_2}{df} = \frac{d[(1-\phi)\theta(\pi^H - \pi^0)]}{df} = -K(p, f) = -\frac{d\phi}{df}\theta(\pi^H - \pi^0) < 0 \tag{B8}$$

Hence increasing the degree of rationing leads to lower profits in the second period for the inside firm. The no cream-skimming constraint then becomes,

$$\frac{d\Delta}{df} = \frac{d\Delta_1}{df} + \frac{d\Delta_2}{df} = -(1-\theta)[\pi^0 - \pi^L - (w_2^D - w_2^E)] - \frac{d\phi}{df}\theta(\pi^H - \pi^0) = 0 \tag{B9}$$

This equation determines our candidate f^*. Let us now simplify this expression.

In the second period, the outside firm makes zero profits and there is full efficiency. Therefore,

$$\theta\pi^H + (1-\theta)\pi^0 = (\theta + (1-\theta)f)w_2^D + (1-f)(1-\theta)w_2^E + (\theta + f(1-\theta))(\hat{\pi} - \bar{\pi}) \tag{B10}$$

On the left hand side is total production in the second period, and on the right hand side are total wages plus profits made by the inside firm. The right hand side of (B10) equals $(\theta + (1-\theta)f)(w_2^D - w_2^E) + (\theta + f(1-\theta))(\hat{\pi} - \bar{\pi}) + w_2^E$, which implies that

$$w_2^D - w_2^E = \frac{\theta\pi^H + (1-\theta)\pi^0 - w_2^E}{\theta + (1-\theta)f} - (\hat{\pi} - \bar{\pi}) \tag{B11}$$

By integration of (B11), w_2^E can be expressed as,

$$w_2^E = p\pi^0 + (1-p)[\bar{\pi} + (\hat{\pi} - \bar{\pi})\ln(G^D(\pi^0))] \tag{B12}$$

Substituting the right hand side of (B12) into (B11),

$$w_2^D - w_2^E = \frac{\theta\pi^H + (1-\theta)\pi^0 - \{p\pi^0 + (1-p)[\bar{\pi} + (\hat{\pi} - \bar{\pi})\ln(G^D(\pi^0))]\}}{\theta + (1-\theta)f} - (\hat{\pi} - \bar{\pi}) \tag{B13}$$

Normalizing by setting $\pi^0 = 0$ and $\pi^H = 1$ (notice that π^L must be negative after the normalization), we get that,

$$w_2^D - w_2^E = \frac{\theta - \{(1-p)[\phi\hat{\pi} + (1-\phi)\hat{\pi}\ln(G^D(0))]\}}{\theta + (1-\theta)f} - (1-\phi)\hat{\pi} \tag{B14}$$

and substituting for $w_1^E - w_1^D = w_2^D - w_2^E$, we can write the first order condition as,

$$\frac{d\Delta}{df} = (1-\theta)[\pi^L + \frac{\theta - \{(1-p)[\phi\hat{\pi} + (1-\phi)\hat{\pi}\ln(G^D(0))]\}}{\theta + (1-\theta)f} - (1-\phi)\hat{\pi}] - \frac{d\phi}{df}\theta = 0 \tag{B15}$$

We now move to considering whether the f^* determined above satisfies the renegotiation constraint, i.e., that a firm cannot increase profits by increasing delegation in the interim, and whether the induced relation between f^* and p is positive.

Assuming that there exists a unique $f^* \in (0,1)$ for a given p, then the condition $\frac{d\Delta}{df} = 0$ implicitly defines a function $f^*(p)$. By the implicit differentiation rule,

$$\frac{df^*}{dp} = -\frac{\frac{d^2\Delta}{df\,dp}}{\frac{d^2\Delta}{df^2}} \tag{B16}$$

For a candidate f^* to be an equilibrium, it needs to be renegotiation-proof, i.e., $\frac{d^2\Delta}{df^2} < 0$. Hence it is necessary to show that $\frac{d^2\Delta}{df^2} < 0$ implies $\frac{d^2\Delta}{df\,dp} > 0$. Unfortunately, the algebraic complexity of the derivatives makes us only able to numerically verify that this condition holds. Numerical analysis confirmed that there exists a unique f^* that satisfies $\frac{d^2\Delta}{df^2} < 0$, and moreover that the function $f^*(p)$ implicitly defined is increasing.[39] ■

To illustrate the solution, we consider the polar case $p = 0$.

Example 1 (Proof of Proposition 3(i), second part) $p = 0$

For $p = 0$, we have $K(0, f) = (\pi^H - \pi^0)\theta > 0$, which is independent of f. Furthermore, the highest possible wage offered in the support is $\overline{\pi} = \theta\pi^H + (1-\theta)\pi^0$. The inside firm can offer this when the worker is high and make profit $\pi^H - \overline{\pi}$, making the inside

[39] We sampled a million different combinations of (θ, f, p) and was not able to find a counterexample. Furthermore, numerical analysis showed that for intermediate values of p, there are two solutions for f^*, defined by (B9), one which satisfies the renegotiation constraint $\frac{d^2\Delta}{df^2} < 0$, and one that does not.

firm's profit equal to $(\pi^H - \overline{\pi})\theta$. We know that there is full efficiency in the second period and the outside firm makes zero profits so $\theta w_2^D + (1-\theta)w_2^E + (\pi^H - \overline{\pi})\theta = \overline{\pi}$. Rearranging yields $\theta(w_2^D - w_2^E) = (1+\theta)\overline{\pi} - \theta\pi^H - w_2^E$. By substituting in for $\overline{\pi}$ we have $w_2^D - w_2^E = \theta(\pi^H - \pi^0) + (\pi^0 - w_2^E)/\theta < \theta(\pi^H - \pi^0) = K(0, f)$. Thus, the wage difference is less than the knowledge gained and we always have incentive to get workers to sort for $p = 0$ and there cannot be cream-skimming. Therefore, a separating equilibrium always exists for $p = 0$.

10 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., 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. (2001). Flexible Workplace Practices and Labor Productivity, IZA working paper #700, University of Bonn.

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.

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.

Fairburn, J. A. & J. M. Malcomson (2001). Performance, Promotion, and the Peter Principle. *Review of Economic Studies*, 68, 45-66.

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.

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

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

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

Lazear, E. P. (2001). The Peter Principle: A Theory of Decline. NBER Working paper no. 8094. Forthcoming, *Journal of Political Economy.*

Marschak, T. & S. Reichelstein (1998). Network Mechanisms, Informational Efficiency, and Hierarchies. *Journal of Economic Theory*, 79, 106-41.

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

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

Osterman, P. (1994). Supervision, Discretion, and Work Organization. *American Economic Review, Papers and Proceedings*, 84, 380-84.

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 & Zingales (2001). The Firm as a Dedicated Hierarchy: A Theory of the Origin and Growth of Firms. *Quarterly Journal of Economics*, 116, 805-51.

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.

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

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

Wilson, R. (1967). Competitive Bidding with Asymmetric Information. *Management Science*, 13, 816-820.

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