Incentive Hierarchies

Emmanuelle AURIOL*and Régis RENAULT†

April 2000

Abstract

Because much of work incentives are provided through promotions, their effectiveness depends to a large extent on the structure of the organization's hierarchy. Here we investigate the impact of the incentive motive on the optimal hierarchy using the Auriol-Renault (1999) framework which highlights the role of recognition in the work place. The latter framework provides a rationale for using promotions as an incentive device which relies on a complementarity between recognition and income: those who earn more should also earn more recognition. We identify factors which affect the hierarchy in terms of number of ranks, population size at each rank and the extent of the differentiation between ranks. We show that the harder it is for an employee to improve performance through effort the more pyramid-like is the hierarchy, with a small group of successful individuals at the top earning high income and recognition. If a high performance may be easily achieved, a seniority based promotion system may be optimal.

*ARQADE and IDEI, University of Toulouse I

†, GEMMA, University of Caen

1 Introduction

Labor earnings represent roughly two thirds of the developed economies total national income. These economies rely to a large extent on their human-resources. The management of these "resources" is one of the most important task of formal organizations such as firms, administrations or teaching institutions. The ability to choose the right people, to favor the development of their knowledge and skills, to create incentives for them to work, is at the heart of success in organizations. Organizations make an extensive use of promotions as a means of providing incentive to their members. A promotion system is a system where individuals are ranked in a hierarchy and paid according to their rank (monetary rewards are attached to positions). The positions are revised regularly so that high past performances are rewarded by a promotion (the successful agent is moved upward in the hierarchy, thus benefiting in general from an increase both in his rank and income). This paper uses the framework of Auriol-Renault (1999) to study the design of optimal hierarchical structure as part of a promotion system.

To understand the use of promotions for incentive purposes it is necessary to take into account the potentially long term relationship involved in a labor contract.[1] Then, the appropriate theoretical framework is that of repeated moral hazard. An important result in this literature is that the optimal long term incentive contract should involve some memory: the types of incentives which are currently given to an agent depends on her past performance (see for instance Rogerson (1985) and Chiapori et al. (1992)). The idea is that, if agents are risk averse, it is optimal to spread over time the effect of income shocks resulting from good or bad performances; this is the need for consumption smoothing emphasized by Malcomson and Spinnewyn (1988). Then the agent's wage in a given period should not vary too much as a function of his current performance. Rather, current performance should affect the way the agent will be treated in future periods. A typical car insurance contract, where the history of each driver influences the amount currently paid, is an example of such long term incentive contracts.

Although a promotion system may be a way to implement an intertemporal incentive scheme it is presumably neither the only one, and more importantly, nor the most efficient. Indeed in a promotion system the benefits of work tend to be delayed to later stage of the career, and awarded only to a lucky few. An extreme example of this is provided by "up or out" policies such as tenure track contract. Since the insurance motive calls on the contrary for consumption smoothing, these contracts are very difficult to explain in the classical principal-agent framework.

[1] "In the early 1980s the typical U.S. worker was in a jobs that would last about eight years, with a quarter of the work force being in jobs that will last 20 years or more." (Milgrom-Robert (1992) page 358).

Auriol-Renault (1999) show that to better understand promotion policies it is useful to acknowledge the social motive to individual action, namely that economic agents care both about private –monetary– and social –social status– retributions.[2] If social recognition matters for individuals' performances, the organization will try to manipulate it. An important result in Auriol-Renault (1999) is that such manipulation is costly. In a short term relationship (i.e., in a static context) introducing symbolic differentiation among identical workers is harmful to the organization. Indeed social status being relative in essence, providing *a priori* some individuals with a higher social status is done at the expanse of status degradation for others. This is reinforced by the fact that at the equilibrium social status and monetary rewards being complement, individuals with higher status also earn higher monetary rewards. As individual marginal performance is decreasing in aggregate rewards, the benefit on the high status/income workers is not compensating for the loss on the low status/income ones.

Once dynamic considerations come into play Auriol-Renault (1999) show that it is optimal for the organization to introduce differentiation among the organization members. While young the undifferentiated workers work hard for minimal status and wage in the hope of obtaining a promotion in case of success. In case of failure they get a lower position with lower reward. Though it distorts the profit on the senior workers, the differentiation between previously successful and unsuccessful agents is optimal; it lowers the total wage bill. This result emerges because the elasticity of substitution between social and monetary rewards is positive. Then it is less costly to achieve a certain level of expected utility (in term of monetary reward) to concentrate both resources in a particular state of nature. For incentives purpose they are concentrated in the late part of the working life, and awarded to previously successful agents. This result which holds when agents are risk-neutral with respect to income, is robust to the introduction of risk-aversion. This general framework is used here to analyze hierarchies as an incentive device in a promotion system.

Practitioners put a lot of emphasis on issues such as the number of distinct hierarchies (for instance one for executives and one for engineers), the number of levels in each (number of rank, salary grades), pay differential between ranks (15% is the rule of thumb), time it takes an employee to progress through the hierarchy: "Managers and union officials devote considerable resources to developing and administering pay structure and behave as if the number of distinct structures, number of levels, size of the differentials, and rate of progress matter." (Gerhart-Milkovich 1991) Yet theoretical attempts at analyzing hierarchies remain surprisingly sparse, and systematic data are rarely reported

[2]Individuals care about social rewards in so far as they affect their behavior as well as the behavior of others towards them. An explanation, in addition to the obvious one that human is social, can be found in Fershtman and Weiss (1996) who characterize the circumstances under which evolution leads to the survival of socially minded individual even though relative fitness is determined only by economic payoff.

in the research literature. A notable exception on the theoretical ground are the works by Radner (1993) and Bolton-Dewatripont (1994). These authors focus on the impact of the hierarchical structure on an organization's ability to process information and to minimize reaction lags. It is modeled as a grid to channel information flows. Their analysis highlights the technical aspect of hierarchy. By contrast, the literature on repeated moral hazard derives optimal contracts that specifies an intertemporal structure for monetary incentives. It says nothing about agents taking different positions in a hierarchy with pay attached to jobs. Still the hierarchical structure is a key feature of any organization. There are always several ways of organizing an activity and the actual way chosen matters for the incentives of the organization members. For instance in the university there is a distinction between assistant, associate and full professor. Such a hierarchy does not exist for technical reason since everybody does the same work. It exits for incentive purpose. A formal hierarchy puts a clear order among the organization members who are ranked from top to bottom according to the transitive relation "superior to". Assuming people care about social rewards, this ranking directly affects their preferences. Under the assumption that agents with a higher "rank" are more responsive to monetary incentives,[3] our goal is to analyze the optimal hierarchical structure associated with a promotion system. To the best of our knowledge, this problem have not yet been addressed in the economic literature.

Our approach builds on the results in the dynamic framework of Auriol-Renault (1999). While these results clearly establish that it is optimal to provide incentives through promotions, they provide limited insight as to what should be the shape of a hierarchy induced by a promotion system. In any organization there are several distinct hierarchies (e.g., a different one for executives, clerical staff, or blue collar workers). As part of a promotion system these hierarchies are a management variable designed to provide work incentives. We thus refer to them as "incentive hierarchies" to distinguish them from the technical hierarchies in Radner (1993) and Bolton-Dewatripont (1994). Here we investigate what should be the number of ranks, the population size at each rank and the extent of the differentiation between ranks in an optimal incentive hierarchy. To this end, we specify parameterized functional forms for tastes and technology. We introduce a parameter measuring an agent's ability and/or willingness to improve his performance through effort. We show that there may be two types of hierarchies which we label "Seniority Hierarchies" and "Merit Hierarchies" respectively. Seniority hierarchies are optimal when an agent's willingness/ability to improve his performance is high enough so that it is optimal to induce him to exert an effort level that guarantees a high performance. Because all young workers perform well, they are all promoted. There are only two ranks in the hierarchy and an agent's rank is solely determined by his seniority. When a high performance is

[3]For a discussion of this assumption see Auriol-Renault (1999).

not so easily achieved, then promotions are based on merit as well as seniority. More specifically, the merit hierarchy has 3 ranks with the young at the bottom, the old who were unsuccessful when young in the middle and the old who were successful when young at the top. As the willingness/ability to improve performance decreases, the top group becomes smaller and is granted a very high status along with a high wage while the middle group is comprised of most of the old population.

Section 2 sets up an overlapping generations model of work. Section 3 studies the optimal steady state hierarchy and pay structure in a dynamic problem within an infinite lived organization. Our findings are summarized in Section 4.

2 The model

We consider an overlapping generations model with infinite horizon based on Auriol-Renault (1999). A risk-neutral principal wants to maximize the "profit" of the unit (bureau, subdivision, firm,...) she supervises. At each date, the organization is comprised of two "generations": the "young" (the junior) who enter the organization in the current period and the "old" (the senior) who joined the organization in the previous period and will not be around in the next one. Hence each cohort only lives two periods. It is assumed that the population of agents is large so that it may be represented by a continuum. It is composed at each period of an equal number of young and old normalized without loss of generality to 1. It has thus a constant population of agents of 2. The principal's objective function, which is assumed to be intertemporally separable with a discount factor $\delta \leq 1$, is strictly increasing in total output and strictly decreasing in the total wage bill. The marginal rate of substitution between the two is assumed to be constant and may thus be normalized to 1 without loss of generality. Such preferences are represented by:

$$\sum_{t=0}^{+\infty} \delta^t \pi_t = \sum_{t=0}^{+\infty} \delta^t \Big(Q_t - W_t\Big). \tag{1}$$

where Q_t is total output and W_t is the wage bill in period t. This specification is consistent with the objectives of a profit maximizing firm.

2.1 The workers

In order to focus on the incentive implications of the choice of an optimal hierarchical structure, workers are assumed to be *ex-ante* identical individuals, hired to do the same type of work. That is, there is no *a priori* legitimate hierarchy among them, and, since they are intended to do the same work, there is no need for one. To be more specific each worker living at date t, either junior or senior, contributes to the collective outcome by exerting an effort $e_t^i \geq 0$. The harder agent i works (the higher e_t^i is), the larger is the probability of a high output for the unit. Individual i working at date t contributes to the total output for an amount q_t^i which may be either high $q_t^i = \overline{q}$, with probability $\mu(e_t^i)$ or low $q_t^i = \underline{q}$, with probability $1 - \mu(e_t^i)$ ($\overline{q} > \underline{q}$), so that the total output of the unit at date t, $Q_t = \int_0^2 q_t^i di$, is random. The individual output q_t^i is verifiable and its realization is independent across time. The probability of a high performance for agent i, $\mu(e_t^i)$, increases with e_t^i. More specifically we assume the following.

A1 $\qquad \mu(e) = \min\{e, 1\} \qquad \forall e \geq 0$

While the output q_t^i is *ex-post* verifiable, the effort level e_t^i is not. The principal is confronted with a moral hazard problem since exerting an effort is costly to worker i. The cost to work is captured by a disutility function $\psi(e_t^i)$, where

A2 $\qquad \psi(e) = A\dfrac{e^2}{2} \qquad A > 0.$

When A increases it becomes more costly to provide an effort. That is, A reflects the difficulty of the task at hand. An equivalent formulation is to set $\mu(e) = min\{\frac{e}{A}, 1\}$ and $\psi(e) = \frac{e^2}{2}$. It is a matter of convention to put A on the probability side or on the disutility side. The two formulations are equivalent in term of the results. The important point is that the larger A is, the more difficult is it for an agent to achieve a high output. In equilibrium the probability of success is inversely related to A.

We now turn to the description of the workers' preferences. Standard principal/agent theory assumes that individuals preferences in the work place are fully characterized by a unanimous dislike of work and a liking of money. Yet, even if it implies costly effort, work is often perceived by people as a source of achievement. It does not mean that all workers find their job beaming. The management literature, and human resources practitioners stress that the motives to work, apart from material and security needs, being social, satisfaction in work is related to the perception of its social importance and usefulness, that is to say to its social recognition. Hence the satisfaction differs depending on whether a task is perceived as crucial or, on the contrary, as subordinate. To capture the idea that the utility attached to work depends also on the social recognition it yields we assume that the marginal rate of substitution between effort and income varies with the social status of the workers. That is, instantaneous preferences of agent i over status $s_t^i \geq 0$, income $w_t^i \geq 0$, and effort level $e_t^i \geq 0$, are represented by the following utility function:

$$U_i(s_t^i, w_t^i, e_t^i) = s_t^i w_t^i - \psi(e_t^i). \qquad (2)$$

The implication of the utility function is as follows. The marginal rate of substitution between effort and income is decreasing in status while the marginal rate of substitution between effort and status is decreasing in income. This means that a higher status induces a greater willingness to exert effort in exchange for some additional income, and a higher income induces a greater willingness to exert effort in exchange for an improved status. The formalization allows for comparisons with the classical repeated moral hazard literature. Justification and discussion of this utility function can be found in Auriol-Renault (1999).

2.2 Social status and hierarchy

Since social status increases the effectiveness of the monetary reward, whoever is in charge of the organization will try to manipulate the hierarchy to provide the cheapest incentives to work. Acknowledging the influence of social status on individual motivation to work leads to an obvious question: through what means may organizations control social recognition of their members? It is often hard to disentangle pure social recognition considerations from the cheer enjoyment of better work conditions. For instance a pleasant office and expensive furniture are likely to make work more enjoyable, but they are also perceived as the sign of a person's high status. The introduction of a formal hierarchy in which positions are labeled differently and ordered vertically represents an unambiguous way of allocating social recognition in the workplace. The present paper looks at hierarchies as a means of allocating status. Hierarchy is used here in a very broad sense. It is just a system which establishes a complete order among organization members. It can be for instance a collection of numbers (ranks) to be assigned to each member of the organization together with the order relation "superior to" among them.

The organization designer chooses a hierarchical structure for the social status allocation it yields. By manipulating the hierarchy, she changes the set of attainable status levels. However she is constrained in this manipulation. For instance if there is a large number of individual –a continuum each of whom recieves a rank in a closed interval $[\underline{r}, \overline{r}]$, the choice of a hierarchy by a principal is equivalent to the choice of a density function $f(r)$ over the rank set. Then a natural social status index for an individual situated at rank r in the hierarchy $f(.)$ is $s_r = \int_{\underline{r}}^{r} f(\tilde{r})d\tilde{r} = F(r)$. Indeed this index provides the exact position of individual r on the social scale (it is a sufficient statistic). It is easy to check that the sum of such a status index is constant independent of the hierarchy considered (i.e. no matter what the density f is). That is, $\int_{\underline{r}}^{\overline{r}} F(r)f(r)dr = \frac{1}{2}\ \forall f(.)$. This status index is used in most of the papers which aim at studying the impact of status motive at the macro-economic level (see for instance Frank (1985), Robson (1992)). A discrete version of this index in a hierarchy with N individuals attributed ranks r_i

($i = 1, ..., N$) such that $r_i > r_j$ if and only if i is the strict superiors of j, is the following: $s_i = \frac{N-1+\sum_{k=1}^{N}\left(\mathbb{1}_{r_k<r_i}-\mathbb{1}_{r_k>r_i}\right)}{2(N-1)}$. Such a status index provides the position of individual i on a social scale defined on $[0,1]$. For instance if an individual is alone at the top of the hierarchy (the other $N-1$ agents are his inferiors), he gets a social status index of 1. Symmetrically if an individual is alone at the bottom of the hierarchy he gets 0. If everybody gets the same rank (no hierarchy at all) the social status index is $\frac{1}{2}$ for everybody. It is easy to check that whatever the particular hierarchy -number of layers, number of agents by layers– chosen by the principal for her N agents, the sum of social status index remains constant: $\sum_{i=1}^{N} s_i = \frac{N}{2}$.

In these two examples, the social status allocation is a "zero-sum game" in the organization (the increase in someone's status is exactly compensated by the decrease in someone else's status). From an economic point of view this corresponds to a constant return to scale assumption in the social status "production" function. This is appealing, since social status is a substitute for income, increasing return to scale in status allocation, would tend to introduce a bias in favor of hierarchy where decreasing return to scale a bias against it. This property is not an artifact of the particular status indexes chosen. It still holds in any set of attainable social status if one postulate the following.

a) Scarcity: *For a given number of people in the organization, everybody prefers to have more subordinates than less, and less superiors than more.*

b) Anonymity: *Since individuals are ex ante identical, the set of attainable social status in any hierarchy is symmetric among agents.*

c) Convexity: *The set of attainable social status allocation is convex.*

Under assumption a) status in hierarchy is a scarce resource: the increase of an individual status is always made at the expanse of someone else's status. That is, assume worker i is moved to a higher rank while the other workers' rank remain unchanged, a) implies that the status of all the new subordinates of i has decreased, as well as the status of the workers who now share the rank of i and who previously were his superiors. In his efforts to allocate status optimally, the principal is faced with the constraint that more social status for someone necessarily involves less social status for somebody else. To this we add the intuitive requirement that since individuals are ex ante identical the set of attainable social status is not affected by who gets which position in the hierarchy (assumption b). Finally for technical simplicity in the optimization problem we consider a convex set of status allocations (Assumption c). This is a restrictive assumption but it can be relaxed somewhat (see Auriol-Renault 1999). Since there are only a finite number of possible hierarchies when ranking a finite number of people, in practice the set of feasible status allocations is finite. Maximizing on such a set leads to technical difficulties without really improving the economic insight. We consider an organization with a large number of workers (i.e., a continuum). Then even if at the steady state solution there are at most three ranks in the hierarchy, which obviously is finite, the principal still choose how many people (which percentage of a cohort of junior workers) will be entitled to a promotion. Then according to assumption a) the value of being in the first rank is endogenously determined and varies continuously with the percentage of peers who get the promotion.

Auriol-Renault (1999) show that under assumptions a), b) and c), the set of feasible social status allocation is characterized as follow (the equality to 2 is a normalization).

$$\text{(F)} \qquad \int_0^2 s_t^i di = 2 \qquad s_t^i \geq 0 \forall i, t$$

Rather than optimizing on the set of underlying hierarchical structure, we optimize on the set of feasible social status allocation defined by (F).

2.3 The contracts

Since social ranking matters for individuals, the principal is going to rely both on monetary and social rewards to provide them incentive to work. For each agent he chooses a social status allocation s_t^i in (F) (a position in the hierarchy), a fixed wage $\underline{w}_t^i$, and a bonus Δw_t^i in case of a high performance. Status is allocated before the worker exerts any effort. That is, when an agent joins the organization he is assigned to a rank somewhere in the hierarchy. The position is revised at the end of the first period. To be more specific, the timing for a cohort joining the organization at date t is as follows.

date t: The new cohort of workers is offered contracts that include a beginning status level, a monetary scheme and a promotion system (future status and monetary scheme in case they succeed and in case they fail).

date $t+0,5$: Workers choose an effort level according to the current monetary incentive they face, their current status and their promotion perspective in the organization.

date $t+1$: Outputs are observed, transfers and promotions occur.

date $t+1,5$: Workers choose an effort level according to their current monetary incentive and status (which may depend whether they have been successful in the first period or not).

date $t+2$: Outputs are observed, transfers occur, workers retire.

The outcome q_t^i which depends on effort at date t being random, some agents will be successful and others not. This randomness allows for ex-ante identical agents living through different histories. In other words, individuals become differentiated through their performance over time. Then an agent can be characterized by the current period t he is living in, and by the fact that he is either a junior worker, denoted by 1, or a senior one with a history of high past performance, denoted h, or a history of low past performance, denoted l.

Each worker's intertemporal utility is assumed to be additively separable with a discount factor of $\delta \leq 1$. The expected utility of an old worker whose past performance has been $p \in \{l, h\}$ and exerting an effort e_{pt} is:

$$EU^i_{pt} = [\mu(e^i_{pt})\Delta w^i_{pt} + \underline{w}^i_{pt}]s^i_{pt} - \psi(e^i_{pt}). \qquad \forall p \in \{l, h\}.$$

Let $\Delta U^i_t = EU^i_{ht} - EU^i_{lt}$. A young worker's expected utility if his effort level is e^i_{1t} is

$$EU^i_{1t} = s^i_{1t}[\mu(e^i_{1t})\Delta w^i_{1t} + \underline{w}^i_{1t}] - \psi(e^i_{1t}) + \delta[\mu(e^i_{1t})\Delta U^i_{t+1} + EU^i_{l(t+1)}].$$

We assume that workers can sell their work force outside the firm. This yields at any point of time an instantaneous reservation utility of $\underline{U}$. The individual rationality constraints are:

(IR) $\qquad EU^i_{pt} \geq \underline{U} \qquad$ and $\qquad EU^i_{1t} \geq (1+\delta)\underline{U}$

We now turn to a characterization of the optimal hierarchical structure.

3 Optimal hierarchy

Contrary to a monetary bonus, the social status is awarded ex ante; that is, before the worker exerts an effort. Still it influences his effort because it affects the responsiveness to monetary incentive. Auriol-Renault (1999) show that in a static context there is a cost for an organization to introduce a hierarchy among identical workers because the benefit on the high status/income workers will not compensate for the loss on the low status/income ones. To maximize instantaneous profit, it is optimal to give identical agent identical contracts (same status, same reward). This result points to the cost of relying on symbolic differentiation to provide incentives in organizations. This suggest to restrict the analysis in the dynamic framework to symmetric contracts. That is, we postulate the following.

d) Equity: *Individuals who are identical (same seniority, same past performance) are treated identically (same status, same monetary scheme, same promotion perspective).*

Assumption d) implies that all young workers have the same contract. Similarly old workers with identical past performances must have the same contract. We can thus drop the index of i. Individual are fully characterized by their seniority and past performance within the firm at date t. Under assumption d) there are at most three ranks in the optimal hierarchy.

Since the population is represented by a continuum of measure 2, the proportion of old who have been successful when young, denoted γ_{ht}, is equal to the probability that a young worker in the previous period had a high performance h. Symmetrically the proportion of old who have been unsuccessful when young, denoted γ_{lt}, is equal to the probability that a young worker in the previous period had a low performance l. That is, under assumption d), $\gamma_{ht} = \mu(e_{1(t-1)})$ and $\gamma_{lt} = 1 - \mu(e_{1(t-1)})$ for $t > 0$. Then the feasibility constraint on status allocation is:

(F) $$s_{1t} + \gamma_{ht}s_{ht} + \gamma_{lt}s_{lt} = 2.$$

Let us denote $c_{1t} = (s_{1t}, \underline{w}_{1t}, \Delta w_{1t})$ The contract of a young worker at date t, and $c_{pt} = (s_{pt}, \underline{w}_{pt}, \Delta w_{pt})$ the contract of an old at date t with performance $p \in \{h, l\}$ at date $t-1$. The principal must pick a sequence of contract combinations $< (c_{1t}, c_{ht}, c_{lt}) >$ that maximizes intertemporal profit. He solves:

$$(3)\ \text{Max} \sum_{t=0}^{+\infty} \delta^t E\Pi_t = \sum_{t=0}^{+\infty} \delta^t \Big\{ \mu(e_{1t})(\Delta q - \Delta w_{1t}) - \underline{w}_{1t} + \gamma_{ht}\Big[\mu(e_{ht})(\Delta q - \Delta w_{ht}) - \underline{w}_{ht}\Big] + \gamma_{lt}\Big[\mu(e_{lt})(\Delta q - \Delta w_{lt}) - \underline{w}_{lt}\Big] + 2\underline{q} \Big\}.$$

subject to the following constraints.

F (feasibility): the array of status levels must be feasible: $s_{1t} + \gamma_{ht}s_{ht} + \gamma_{lt}s_{lt} = 2$.

IR (individual rationality): each agent's utility must be above some reservation level $\underline{U}$: $EU_{1t} \geq (1+\delta)\underline{U}$ and $EU_{pt} \geq \underline{U}$.

LL (limited liability): the lowest wage cannot be under a given (legal) threshold $\underline{w}$ set to 0 without loss of generality: $\underline{w}_{tk} \geq 0$ $(k \in \{1, h, l\})$.[4]

IC (incentive compatibility): each agent chooses his effort optimally given his status, the incentive scheme he is offered and his promotion perspective: $e_{1t} = Arg\ \text{Max}\ EU_{1t}$ and $e_{pt} = Arg\ \text{Max}\ EU_{pt}\ \ p \in \{h, l\}$.

We consider first the problem of a senior worker. Let $e_{pt}(s^{pt}, \underline{w}_{pt}, \Delta w_{pt})$ denote the optimal effort level of a senior worker with status s_{pt} and monetary scheme $(\underline{w}_{pt}, \Delta w_{pt})$. The agent chooses his effort to solve:

(4) $$\text{Max}\, EU_{pt} = (\mu(e)\Delta w_{pt} + \underline{w}_{pt})s_{pt} - \psi(e_{pt}).$$

Under assumption A1 and A2, the agent's problem is strictly concave in effort; it has a unique solution. Agent pt's effort, $e_{pt}(s_{pt}, \underline{w}_{pt}, \Delta w_{pt}) = e^*(s_{pt}\Delta w_{pt})$, where $e^*(.)$ is defined by:

(5) $$e^*(x) = \min\{\frac{x}{A}, 1\}$$

[4] As long as we allow $\underline{U}$ to vary, we can set $\underline{w} = 0$; they play symmetric role.

Then from equation (??) e_{pt} is increasing in $s_{pt}\Delta w_{pt}$ and is decreasing with A. It is independent of $\underline{w}_{pt}$ due to the risk neutrality assumption. We consider next the problem of a junior worker. Since workers are expected utility maximizer, it is easy to check that the incentive compatibility constraints for the young may be written as follows $e_{1t} = e^*(s_{1t}\Delta w_{1t} + \delta \Delta U_{t+1})$. The incentive compatibility constraints are written as follows.

(IC) $$e_{1t} = e^*(s_{1t}\Delta w_{1t} + \delta \Delta U_{t+1}) \quad \text{and} \quad e_{pt} = e^*(s_{pt}\Delta w_{pt}) \qquad p \in \{h, l\}.$$

Taking into account the different constraints (F), (IR), (LL) and (IC), the principal maximizes her profit with respect to status and monetary rewards.

Initial populations γ_{h0} and γ_{l0} are exogenously given. Our aim now is to characterize a solution to the principal's dynamic problem. To do this we restrict the analysis to a steady state of such a solution. In a steady state, γ_{ht} and γ_{lt} are constant over time. It is also natural to restrict the solution to be such that s_{kt}, w_{kt} ($k \in \{1, l, h\}$) do not change across periods. In turn, effort levels e_{1t}, e_{ht}, e_{lt} are constant over time. There is no loss of generality in dropping the time subscripts. From now on they are thus omitted. The next proposition is proved in Auriol-Renault (1999).

Proposition 1 *(Auriol-Renault 1999) In a steady state of a profit maximizing solution we have:*

(6) $$\Delta w_1 = \underline{w}_1 = s_1 = 0$$

(7) $$s_h > s_l$$

(8) $$\underline{w}_l \leq \underline{w}_h \text{ and } \Delta w_l \leq \Delta w_h$$

where at least one of the inequalities in (??) is strict.

This proposition makes two important points. First, older employees should be treated differently, depending on their past performance. This is coherent with results in a standard repeated moral hazard setting. However, the second point made in Proposition (??) is that the young should receive minimal monetary incentives. They should instead be induced to exert effort through a promotion system. Rewards are thus concentrated late in the career. Concentrating the rewards on a particular state of nature helps to reduce the total wage bill because the rewards are mutually reinforcing. We show in Auriol-Renault (1999) that this result remains valid if agents are risk averse regarding income in spite of the consumption smoothing motive highlighted by the repeated moral hazard literature. By differentiating their status and incentive scheme the principal loses on the olds (the optimal contract for old workers is the egalitarian solution). This loss is fully balanced by the benefit made on the young workers. This model thus provides a rationale for the use of a promotion system as an incentive device. It also explains the observed long term character of work relationships. Workers have an incentive to stay in the same company in order to garner the benefits of their efforts early in their career and it is optimal for employers to commit to delayed rewards.

The use of promotions as an incentive device clearly has important consequences for the shape of the hierarchy in an organization. Proposition 1 provides little insight in this respect. In particular, it does not say how many ranks there should be, what fraction of the agents' population should be at each rank or how much differentiation there should be between the various ranks in terms of status and income. The specific functional forms used here for the technology, μ, and the disutility of effort ψ allow us to answer these questions. In particular, we show that, in equilibrium, there may be only two layers in the hierarchy.

Let $\delta = 1$. Under assumptions A1 and A2, and in virtue of proposition **??**, the program **??** becomes:

$$(9)\quad \text{Max} \sum_{t=0}^{+\infty} \left\{ e_1\Big(\Delta q + e_h\big(\Delta q - \Delta w_h\big) - \underline{w}_h\Big) + (1-e_1)\Big(e_l\big(\Delta q - \Delta w_l\big) - \underline{w}_l\Big) + 2\underline{q} \right\}.$$

subject to

F $e_1 s_h + (1-e_1)s_l = 2.$

IR $EU_1 = e_1\Delta U + EU_l - A\frac{e_1^2}{2} \geq 2\underline{U}$ and $EU_p = s_p\Big(e_p\Delta w_p + \underline{w}_p\Big) - A\frac{e_p^2}{2} \geq \underline{U}.$

LL $\underline{w}_p \geq 0 \quad p \in \{h, l\}.$

IC $e_1 = \min\{\frac{\Delta U}{A}, 1\}$ and $e_p = \min\{\frac{\Delta w_p s_p}{A}, 1\} \quad p \in \{h, l\}.$

We next show that depending on the value of $\frac{A}{\underline{U}}$ different regimes might appear.

3.1 Promotion by seniority

When A is small enough it is easy to provide incentives to the workers (in equilibrium the probability of success is high). This corresponds to the case of an easy task to perform and to monitor. Then the optimal management policy consists in promoting all the workers based on their seniority. There are only two ranks in the optimal hierarchy. In equilibrium all the workers go through them. The incentive hierarchy has an equal number of people at the top and at the bottom. It is then tubular. The next proposition establishes this result. It is proven in appendix 1.

Proposition 2 *When $\frac{A}{\underline{U}} \leq 2$ in equilibrium $e_1 = 1$. The optimal hierarchy has two ranks, the "junior": $s_1 = 0$ with $\Delta w_1 = \underline{w}_1 = 0$, and the "senior": $s_h = 2$ and if*

- $\Delta q \leq \frac{A}{2}$ *then* $\Delta w_h = \Delta q$, $\underline{w}_h = \underline{U} + \frac{A}{4} - \frac{\Delta q^2}{A} > 0$, $EU_h = 2\underline{U} + \frac{A}{2}$, *and* $e_h = \frac{2\Delta q}{A}$.
- $\Delta q > \frac{A}{2}$ *then* $\Delta w_h = \frac{A}{2}$, $\underline{w}_h = \underline{U}$, $EU_h = 2\underline{U} + \frac{A}{2}$, *and* $e_h = 1$.

In this context of an easy task to perform, everybody gets a promotion based on seniority. Then according to the management literature the hierarchical structure is *flat* in the sense that there are few ranks. The pay differential between the two ranks is relatively small: $Ew_h - Ew_1 = \underline{U} + \frac{A}{2} \leq 2\underline{U}$, and the social reward associated with the promotion also $(s_h - s_1 = 2)$. Since everybody gets it, its social value is low. On the other hand, if a worker fails it means that he has shirked. This results into firing. Promotion by seniority can thus be seen as an extreme example of an up or out policy where everybody is promoted in equilibrium. We should expect this type of incentive hierarchies for workers in standardized industries, on assembly lines, or clerical positions for instance. This is indeed the case in Japan, but also in the US, with a set of seniority rules at the factory floor such as "last to come first to go". Now if the task is difficult to fulfill and the outcome of work is random the optimal hierarchy becomes, in the words of the management literature, *tall.*

3.2 Promotion by merit

When A is large, it is costly to provide work incentives to the agents simply because it is difficult for them to achieve a high output. The optimal hierarchy then is comprised of three ranks. One at the top for the high performers, an intermediate one for the low performers, and one at the bottom for the beginners. The structure of the hierarchy varies with $\frac{A}{\underline{U}}$. The higher A, the higher is the difficulty of the task, and the lower the probability of success. Then fewer are promoted at the top, and the more pyramidal the incentive hierarchy looks. The next result is proven in appendix 2.

Proposition 3 *When $\frac{A}{\underline{U}} \geq 2$ in equilibrium $e_1 < 1$. The optimal hierarchy has tree ranks. There is the "junior" rank: $s_1 = 0$ with $\Delta w_1 = \underline{w}_1 = 0$. For the senior there are two ranks: $s_l < s_h$. Under the assumption that $\frac{\Delta q}{A}$ is small enough*[5] *we have:*

$$s_l = \frac{2}{1 + \left(\frac{A}{2\underline{U}}\right)^{0.5}\left(\left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5} + 1\right)} \qquad (10)$$

$$s_h = \frac{2\left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5}}{1 + \left(\frac{A}{2\underline{U}}\right)^{0.5}\left(\left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5} + 1\right)} \qquad (11)$$

Associated to these we have $\Delta w_h = \Delta w_l = \Delta q$, $\underline{w}_h = \left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5} w_l$, $e_h = \frac{s_h \Delta q}{A}$, $e_l = \frac{s_l \Delta q}{A}$ which implies $EU_l = \underline{U}$, $\Delta U = (2A\underline{U})^{0.5}$ and $e_1 = \left(\frac{2\underline{U}}{A}\right)^{0.5}$.

Proposition **??** implies that as $\frac{A}{\underline{U}}$ increases, the probability of success and the number of people being promoted diminish. It is then easy to check by comparing equations

[5] The exact condition is given in appendix 2.

(??) and (??) that the status associated with the promotion rises. This illustrates the fact that social status being relative in essence, it is not the same to get a promotion when everybody gets it, or when only an exclusive group of people gets it. Proposition **??** implies that the more difficult a task is, the larger the difference in rewards between successful and unsuccessful agents should be. In particular, when A goes to infinity, both s_h and w_h go to infinity, s_l converges to 2 and w_l to $\frac{U}{2}$. In this case the probability of success e_1 goes to zero. This phenomenon is well illustrated by professions where achieving a high performance is very difficult, such as managers of large publicly owned companies or artists (e.g. musicians, painters, movie actors). As predicted by the model the pay structure in these professions is of a "winner-take-all" type. There are very few slots to be filled at the top but the rewards are high. For instance Rosen (1981) estimates that there are approximately 200 full-time comedians in the United States. Similarly the number of full-time soloists on any given instrument is also on the order of a few hundred. The number of CEOs of large publicly owned companies is in the same range. The lucky few who manage to reach these top positions receive huge sums of money for what they do. In the case of CEOs of large companies, this has been widely criticized. The average total compensation for the 365 CEOs of the largest publicly-held corporations in the United-States was \$1.95 million according to a survey in *Business Week* covering 1990 earnings. The critics point out that these payments are largely unrelated to the performance of the firms on the stock market, and that they create a feeling of unfairness among the workers that undermines their commitment to the organization.

Propositions **??** and **??** imply that the pay differential should be larger in hierarchies for executives or managers than in hierarchies for office staffs or blue collars. This result seems to be supported by the little evidences that exist on the pay structure in hierarchies. In particular differentials between the maximum and minimum paid within a rank is on average 50% for office and clerical employees and at least 120% for professionals and managerial staff (Gerhart-Milkovich 1991). Similarly, in professions involving a great deal of uncertainty and requiring a lot of creativity, initiative, such as the ones in high-technology industries (computer, aerospace products and biotechnology), the incentive hierarchy should be taller, with wider differential in pay than in industries where production is standardized. For instance Rabin (1987) found that the pharmaceuticals industry which traditionally invests heavily in research and development has greater pay differentials than the conservative insurance industry. Undoubtedly, more systematic studies are called to understand the determinants of pay structure in incentive hierarchies across industries.

Finally Propositions **??** and **??** suggest that the more "meritiocratic" an incentive system is, the taller is the hierarchy and the wider the pay differentials. This result seems to be supported by evidences from international comparisons. Incentive hierarchies are flatter in Japan than in the US. For instance regarding the hierarchy for management, Ford has seventeen layers of management between its CEO and its employees on the factory floor, and GM has as many as twenty-two. By contrast Toyota has seven. Regarding wage differences: "The average CEO of a very large Japanese firm (the equivalent of $ 30 billion in sales) earns 17 times what the average Japanese worker does. For comparable firms in France and Germany, the figure is about 24 times. In the United States, it is 109" (Milgrom-Robert 1992 pages 425-426). Now it is well established that promotions within large Japanese corporations are mostly based on seniority: "A new white-collar recruit can expect to spend at least a decade with the firm before being considered (along with the others in the same cohort, who entered the firm together) for a promotion. Then members of any experienced cohort are evaluated for further promotions only after more senior people have had their chances. Pay is tied to seniority as well, with individual merit or performance pay being rare." (Milgrom-Robert 1992, page 351).

According to our model, these differences in the pay structure should arise because of differences in the cost of providing incentives. Why should these costs be lower in Japan than in the US? One possible explanation is that the Japanese practice of life-time employment favors a more effective human ressource management, so that there is a better adequacy between an employee's skills and the tasks he has to perform. As a result, it is easier for a Japanese worker to achieve a high performance. This may be the case for the two following reasons. First, life time employment encourages Japanese firms to invest in their employees, in particular by training them, and it also encourages employees to acquire firm specific skills. It is not uncommon that Japanese companies pay employees to earn a graduate degree in business or law. Second, because the whole career unfolds within the organization, the personnel department can track employees training and performance throughout their working life and select for each an appropriate position. Indeed, the personel department is one of the most important departments in large Japanese corporations. Contrary to what happens in the US, it attracts the best employees and is highly respected (see Milgrom and Roberts 1992 p 352).

4 Conclusion

This paper studies optimal hierarchies viewed as a pure incentive variable, and the associated optimal pay structure. The analysis is based on an overlapping generation model developed in Auriol-Renault (1999). The extent to which members of an organization are willing to exert some effort to improve their performance depends to a large extent on how much their contribution is valued relative to that of other members. Agents that get more social recognition for their achievements will display a lower marginal disutility of effort. They are more responsive to monetary incentives than those with a lower social status. Here the problem of awarding recognition to members of the organization for their achievements consists in allocating ranks in a formal hierarchy. Moving some people up the hierarchy necessarily leads to degrading the relative position of others. There typically are several incentive hierarchies within an organization (e.g. one for executives, one for blue collars).

We identify the determinants of the number of ranks inside any particular incentive hierarchy, and the extent of pay differentials between the ranks. The more predictable and easy to fulfill a task (involving routine procedure, standardized production,...), the flatter the optimal incentive hierarchy is. It has few ranks and a tubular shape. Everybody goes through the different ranks at a predictable pace. Promotion is based on seniority. Then the promotion system is egalitarian in the sense that everybody is treated alike throughout his working life. On the other hand, when the work requires creativity, initiative and involves uncertainty, the incentive hierarchy should be tall. If the task to achieve is very difficult (i.e. the probability of success is very low), then the optimal hierarchy has a winner-take-all structure.

Appendix 1: Promotion by seniority ($\frac{A}{\underline{U}} \leq 2$)

Let first consider the case where in equilibrium $e_1^* = 1$ (i.e., $\Delta U^* \geq A$). In case of failure we assume that the worker is fired in which case he gets his reservation utility: $EU_l = \underline{U}$. The program **??** becomes:

$$\text{Max}\, E\Pi = \Delta q + e_h(\Delta q - \Delta w_h) - \underline{w}_h + 2\underline{q} \tag{12}$$

s.t. $EU_h = s_h\Big(e_h \Delta w_h + \underline{w}_h\Big) - A\frac{e_h^2}{2} \geq \underline{U}$ (RIh)

$EU_1 = EU_h - EU_l - \frac{A}{2} \geq 2\underline{U}$ and $EU_l = \underline{U}$ (RI1)

$s_h = 2;\quad \underline{w}_h \geq 0;\quad e_h = min\{\frac{s_h \Delta w_h}{A}; 1\}$

We establish the following preliminary result.

Lemma 1 *Without loss of generality at the optimum:* $\Delta w_h \leq \frac{A}{2}$.

Proof: Assume that Δw_h is such that $\Delta w_h > \frac{A}{2}$ then $e_h = 1$. Let $\epsilon = \Delta w_h - \frac{A}{2}$ and $\Delta w_h' = \Delta w_h - \epsilon$, and $\underline{w}_h' = \underline{w}_h + \epsilon$. This implies $e_h' = e_h = 1$, $EU_h = EU_h'$, and $E\Pi = E\Pi'$. QED

The constraint (RI1) is equivalent to $EU_h \geq \underline{U} + \frac{A}{2}$. Since $\underline{U} + \frac{A}{2} > \underline{U}$, (RI1) is stronger that (RIh). The program **??** becomes:

$$\text{Max}\, E\Pi = \Delta q + \frac{2\Delta w_h}{A}(\Delta q - \Delta w_h) - \underline{w}_h + 2\underline{q} \tag{13}$$

s.t. $EU_h = 2\Big(\frac{\Delta w_h^2}{A} + \underline{W}_h\Big) \geq \underline{U} + \frac{A}{2}$

$\underline{w}_h \geq 0 \quad \Delta w_h \leq \frac{A}{2}$.

The Lagragian is:

$$L = \Delta q + \frac{2\Delta w_h}{A}(\Delta q - \Delta w_h) - \underline{w}_h + 2\underline{q} + \lambda\Big(2\frac{\Delta w_h^2}{A} + 2\underline{W}_h - \underline{U} - \frac{A}{2}\Big) + \gamma\Big(\frac{A}{2} - \Delta w_h\Big).$$

We derive the following conditions:

$\lambda \geq 0;\quad \gamma \geq 0;$

$\lambda\Big(2\frac{\Delta w_h^2}{A} + 2\underline{W}_h - \underline{U} - \frac{A}{2}\Big) = 0;\quad 2\frac{\Delta w_h^2}{A} + 2\underline{W}_h \geq \underline{U} + \frac{A}{2};$

$\gamma\Big(\frac{A}{2} - \Delta w_h\Big) = 0;\quad \frac{A}{2} \geq \Delta w_h.$

$\frac{\partial L}{\partial \Delta w_h} = \frac{2}{A}\Big(\Delta q - 2\Delta w_h\Big) + 4\lambda\frac{\Delta w_h}{A} - \gamma = 0$

$\frac{\partial L}{\partial \underline{w}_h} = -1 + 2\lambda = 0.$

>From $\frac{\partial L}{\partial \underline{w}_h} = 0$ we deduce that $\lambda = 0.5 > 0$. This implies that (RI1) is binding, $\underline{w}_h = \underline{U} + \frac{A}{4} - \frac{\Delta w_h^2}{A}$, and in $\frac{\partial L}{\partial \Delta w_h} = 0$ that $\Delta w_h = \Delta q - \frac{A}{2}\gamma$. We deduce from $\gamma\Big(\frac{A}{2} - \Delta w_h\Big) = 0$ that if $\Delta q \leq \frac{A}{2}$ then $\gamma = 0$. If $\Delta q > \frac{A}{2}$ then $\Delta w_h = \frac{A}{2}$. A necessary condition to have promotion by seniority is $\Delta U \geq A$. Since (RI1) is binding $\Delta U = \underline{U} + \frac{A}{2}$, which implies condition $\frac{A}{\underline{U}} \leq 2$. We deduce proposition **??**.

Appendix 2: Promotion by merit ($\frac{A}{\underline{U}} > 2$)

Let now consider the case where in equilibrium $e_1 < 1$ (i.e., $\Delta U = EU_h - EU_l < A$). We establish the following preliminary result.

Lemma 2 *Without loss of generality at the optimum:* $s_p \Delta w_p \leq A$ *($p = h, l$).*

Proof: Assume that Δw_p is such that $\Delta w_p > \frac{A}{s_p}$ then $e_p = 1$. Let $\epsilon = \Delta w_p - \frac{A}{s_p}$ and $\Delta w_p' = \Delta w_p - \epsilon$, and $\underline{w}_p' = \underline{w}_p + \epsilon$. This implies $e_p' = e_p = 1$, $EU_p = EU_p'$, and $E\Pi = E\Pi'$ $(p = h, l)$. QED

By virtue of Lemma **??** $e_p = \frac{\Delta w_p s_p}{A}$, and by assumption $e_1 = \frac{\Delta U}{A} < 1$. This implies that $EU_p = s_p\Big(e_p \Delta w_p + \underline{w}_p\Big) - A\frac{e_p^2}{2} = \frac{(s_p \Delta w_p)^2}{2A} + s_p \underline{w}_p$ $(p = l, h)$, and $EU_1 = e_1(\Delta U) + EU_l - A\frac{e_1^2}{2} = \frac{\Delta U^2}{2A} + EU_l - 2\underline{U}$. Let $\Delta U = EU_h - EU_l$. The program **??** becomes:

(14) $\text{Max}\, E\Pi = e_1\Big[\Delta q + e_h(\Delta q - \Delta w_h) - \underline{w}_h\Big] + (1 - e_1)\Big[e_l(\Delta q - \Delta w_l) - \underline{w}_l\Big] + 2\underline{q}$

s.t. $EU_l \geq \underline{U}$ (RIl)

$\frac{\Delta U^2}{2A} + EU_l - 2\underline{U} \geq 2\underline{U}$ (RI1)

$e_1 s_h + (1 - e_1)s_l = 2;\quad \underline{w}_h \geq 0;\quad \underline{w}_l \geq 0;\quad s_h \Delta w_h \leq A;\;\; s_l \Delta w_l \leq A.$

The Lagragian is:

$$L = \frac{\Delta U}{A}\Big[\Delta q + \frac{s_h \Delta w_h}{A}(\Delta q - \Delta w_h) - \underline{w}_h\Big] + \Big(1 - \frac{\Delta U}{A}\Big)\Big[\frac{s_l \Delta w_l}{A}(\Delta q - \Delta w_l) - \underline{w}_l\Big] + 2\underline{q} + \alpha(EU_l - \underline{U})$$
$$+\beta\Big(\frac{\Delta U^2}{2A} + EU_l - 2\underline{U}\Big) + \gamma\Big(2 - \frac{\Delta U}{A}s_h - \Big(1 - \frac{\Delta U}{A}\Big)s_l\Big) + \lambda_l \underline{w}_l + \lambda_h \underline{w}_h + \epsilon_h(A - \Delta w_h s_h) + \epsilon_l(A - \Delta w_l s_l)$$

Let $B = \Delta q + \frac{s_h \Delta w_h}{A}(\Delta q - \Delta w_h) - \underline{w}_h - \frac{s_l \Delta w_l}{A}(\Delta q - \Delta w_l) + \underline{w}_l + \beta(\Delta U) - \gamma(s_h - s_l)$.

We derive the following conditions:

$\alpha \geq 0;\quad \beta \geq 0;\quad \gamma \geq 0;\quad \lambda_l \geq 0;\quad \lambda_h \geq 0;\quad \epsilon_l \geq 0;\quad \epsilon_h \geq 0.$

$\alpha(EU_l - \underline{U}) = 0$

$\beta\Big(\frac{\Delta U^2}{2A} + EU_l - 2\underline{U}\Big) = 0$

$\gamma\Big(2 - \frac{\Delta U}{A}s_h - (1 - \frac{\Delta U}{A})s_l\Big) = 0$

$\lambda_l \underline{w}_l = 0; \quad \lambda_h \underline{w}_h = 0;$

$\epsilon_h(A - \Delta w_h s_h) = 0; \;\; \epsilon_l(A - \Delta w_l s_l) = 0$

$\frac{\partial L}{\partial \Delta w_h} = 0 \Leftrightarrow \frac{s_h \Delta w_h}{A^2}B + \frac{\Delta U}{A^2}(\Delta q - 2\Delta w_h) - \epsilon_h = 0 \quad (1)$

$\frac{\partial L}{\partial s_h} = 0 \Leftrightarrow (\frac{s_h \Delta w_h}{A^2} + \frac{\underline{w}_h}{A\Delta w_h})B + \frac{\Delta U}{A^2}(\Delta q - \Delta w_h) - \gamma\frac{\Delta U}{A\Delta w_h} - \epsilon_h = 0 \quad (2)$

$\frac{\partial L}{\partial \Delta w_l} = 0 \Leftrightarrow \frac{s_l \Delta w_l}{A^2}(-\frac{B}{A} + \alpha + \beta) + (1 - \frac{\Delta U}{A^2})(\Delta q - 2\Delta w_l) - \epsilon_l = 0 \quad (3)$

$\frac{\partial L}{\partial s_l} = 0 \Leftrightarrow (\frac{s_l \Delta w_l}{A^2} + \frac{\underline{w}_l}{\Delta w_l})(-\frac{B}{A} + \alpha + \beta) + (1 - \frac{\Delta U}{A})\frac{\Delta q - \Delta w_l}{A} - \frac{\gamma}{\Delta w_h}(1 - \frac{\Delta U}{A}) - \epsilon_l = 0 \quad (4)$

$\frac{\partial L}{\partial \underline{w}_h} = 0 \Leftrightarrow \frac{s_h}{A}B - \frac{\Delta U}{A} + \lambda_h = 0 \quad (5)$

$\frac{\partial L}{\partial \underline{w}_l} = 0 \Leftrightarrow s_l(-\frac{B}{A} + \alpha + \beta) - (1 - \frac{\Delta U}{A}) + \lambda_l = 0 \quad (6)$

Depending on the value of the parameters different solutions occur. For the sake of simplicity we concentrate on explicit solutions.[6] To be more specific we focus on the case where at the equilibrium:

i) $\underline{w}_h > 0$ and $\underline{w}_l > 0$ which implies $\lambda_h = \lambda_l = 0$.

ii) $\alpha > 0$ which implies $EU_l = \underline{U}$

iii) $\beta > 0$ which implies $\Delta U = (2A\underline{U})^{0.5}$ and $e_1 = (\frac{2\underline{U}}{A})^{0.5}$.

>From (1), (5) and (i) we get: $\frac{\Delta U}{A^2}(\Delta q - \Delta w_h) = \epsilon_h$. Similarly from (3) and (6) we get: $(1 - \frac{\Delta U}{A})\frac{1}{A}(\Delta q - \Delta w_l) = \epsilon_l$. We deduce from these two equations that if

iv) $\dfrac{\Delta q}{A} \leq \dfrac{1}{s_h}$

then $\epsilon_h = \epsilon_l = 0$ so that $\Delta w_h = \Delta w_l = \Delta q$.

>From (1) and (2) we get $\gamma = \frac{\underline{w}_h B}{\Delta U} + \frac{\Delta w_h^2}{A}$. Applying (i) to (5) we get $B = \frac{\Delta U}{s_h}$. We deduce that $\gamma = \frac{\underline{w}_h}{s_h} + \frac{\Delta w_h^2}{A}$. >From (5), (6) and (i) we deduce that $\alpha + \beta = \frac{1-e_1}{s_l} + \frac{e_1}{s_h}$. Substituting in (3)and (4), we get $\gamma = \frac{\underline{w}_l}{s_l} + \frac{\Delta w_l^2}{A}$. Equating the two γ and using $\Delta w_p = \Delta q$, we deduce that (s_h, s_l) is such that: $\frac{\underline{w}_h}{s_h} = \frac{\underline{w}_l}{s_l}$. Moreover we have: $\frac{\Delta U}{A}s_h + (1 - \frac{\delta U}{A})s_l = 2$ $(\gamma > 0)$. Combining these two equations and using (iii) we find:

$$s_l = \frac{2}{1 + \left(\frac{A}{2\underline{U}}\right)^{0.5}\left(\left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5} + 1\right)}$$

$$s_h = \frac{2\left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5}}{1 + \left(\frac{A}{2\underline{U}}\right)^{0.5}\left(\left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5} + 1\right)}.$$

[6] It exists other solutions where the optimal contracts are only defined implicitly.

We deduce from $\frac{\underline{w}_h}{s_h} = \frac{\underline{w}_l}{s_l}$ that $\underline{w}_h = \left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5} w_l$,
and using (ii) we get $\frac{(s_l \Delta q)^2}{2A} + s_l \underline{w}_l = \underline{U}$. That is: $\underline{w}_l = \frac{\underline{U}}{s_l} - \frac{s_l \Delta q^2}{2A}$.

We need now to check that this solution meet the assumptions of the problem. Using the fact that $B = \frac{\Delta U}{s_h}$ we deduce that $\beta = \frac{3}{s_h} - \frac{\Delta q}{\Delta U} - 2\frac{EU_l}{\Delta U}\left(\frac{1}{s_l} - \frac{1}{s_h}\right)$ which implies $\alpha = \frac{\Delta q}{\Delta U} - \frac{2}{s_h} + \left(\frac{1}{s_l} - \frac{1}{s_h}\right)\left(1 - \frac{\Delta U}{A} + 2\frac{EU_l}{\Delta U}\right)$. Then one can check that conditions (i) to (iv) are satisfied under the following condition:

$$(15) \quad \frac{\Delta q}{A} < \min\left\{\left(\frac{A}{2\underline{U}}\right)^{0.5}; \left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5}; 3 - \left(\frac{2\underline{U}}{A}\right)^{0.5}\left[\left(\left(\frac{A}{2\underline{U}}\right)^{0.5} + 1\right)^{0.5} - 1\right]\right\} \frac{\left(\frac{2\underline{U}}{A}\right)^{0.5}}{s_h}$$

Under condition **??** we get: $e_h = \frac{s_h \Delta q}{A} < 1$ and $e_l = \frac{s_l \Delta q}{A} < 1$. Finally $e_1 = (\frac{2\underline{U}}{A})^{0.5} < 1$ is equivalent to $\frac{A}{\underline{U}} > 2$. QED

References

[1] Auriol, E. and R. Renault (1999): Costs and Benefits of Symbolic Differentiation in the Work Place, mimeo Toulouse I.

[2] Ball, S. and C. Eckel (1996): Buying Status: Experimental Evidence on Status in Negotiation. forthcoming in *Psychology and Marketing.*

[3] Ball, S., Eckel, C., Grossman, P. J. and W. Zame (1996): Status in Markets. Mimeo.

[4] Bolton, P. and M. Dewatripont (1994): The Firm as a Communication Network. *Quarterly Journal of Economics.* 104, 809-839.

[5] Chiappori, P. A., Macho, I., Rey, P. and B. Salanié (1994): Repeated Moral Hazard: The Role of Memory, Commitment and the Access to Credit Markets. *European Economic Review.* 38.8, 1527-1555.

[6] Frank, R. (1985): The Demand for Unobservable and other Nonpositional Goods, *American Economic Review*, vol. 75, No 1, 101-116.

[7] Gerhart, B. and G. Milkovich (1991): Employee Compensation: Research and Practice. *Handbook of Industrial and Organizational Psychology* 2nd edition Editors M. D. Dunette, and L. M. Hough. Palo Alto CA, Consulting Psychologists Press, pages 481-569.

[8] Ghemawat, P. (1995): Competitive Advantage and Internal Organization: Nucor Revisited. *Journal of Economics and Management Strategy.* 3.4, 685-717.

[9] Maslow, A. (1954): Motivation and Personality. 2nd edition 1970. Harper & Row.

[10] Malcomson, J. and F. Spinnewey (1988): The Multiperiod Principal-Agent Problem. *Review of Economic Studies.* 55, 391- 408.

[11] Milgrom, P. and J. Roberts (1992): Economic Organization and Management. Prentice Hall.

[12] Radner, R. (1993): The organization of Decentralized Information Processing. *Econometrica.* 61, 1109-1146.

[13] Robson, A. (1992): Status, the distribution of wealth, private and social attitudes to risk, *Econometrica*, vol. 60, No 4, 837-857.

[14] Rogerson, W. (1985): Repeated Moral Hazard. *Econometrica.* 53, 69-76.

[15] Rosen, S. (1981): The Economics of Superstars, *American Economic Review*, vol. 71, No 5, 845-858.