The Costs and Benefits of Symbolic Differentiation in the Work Place *

Emmanuelle AURIOL† and Régis RENAULT‡

November 2002

Abstract

We introduce, in a multiple agents moral hazard setting, a status variable which reflects an agent’s claim to social recognition in her work. Status is a scarce resource so that increasing an agent’s status requires that another agent’s status is decreased. High status agents are more willing to exert effort in exchange for monetary compensations while well-paid agents care more about recognition so that they would exert a higher effort in exchange for a higher status. We obtain results coherent with actual management practices and management experts recommendations such as: (i) status and income should be complements; (ii) egalitarianism is desirable in a static context; (iii) in a long-term work relationship, promotions are more effective than direct monetary incentives.

Keywords: repeated moral hazard, internal labor markets, social status.
JEL classification: D82, L23, M12, J33.

*We should like to thank seminar participants at Ecole des Ponts et Chaussées (Paris), Institut d’Analisi Economica (Barcelona), Université d’Aix-Marseille 2, Université Catholique de Louvain, Université de Caen, Université des Sciences Sociales de Toulouse, University of Virginia, Stockholm School of Economics and the participants to the workshop Social Interaction and Economic Behavior in Paris December 1999 for stimulating criticism and comments on an early version of the paper. Finally we would like to thank Thomas Mariotti and Marie-Christine Henninger for their help and suggestion. All remaining errors are ours.

†ARQADE and IDEI, Université des Sciences Sociales de Toulouse.
‡Université de Cergy-Pontoise, ThEMA.
1 Introduction

Although economists have put out a substantial amount of research on work incentives, their approach, remains at odds with much of the management and organization literature on the subject. The logic of using money to induce effort which is the main focus of economic analysis is definitely a key feature of actual incentive packages. Yet, a mere description of monetary incentive schemes falls short of providing a full account of management practices. Even in cases where direct monetary incentives are used extensively, as for instance for sales force employees, they are associated with other types of benefits ranking from travel or merchandise to symbolic rewards like trophies or medals. It is often argued that merchandise, although a poor substitute for money according to standard economic theory, is an effective means of providing incentives because of its trophy value: it reminds the winner and others of her/his high past performance (see Nelson, 1994, Wood, 1998). Wood (1998) reports that Will Haffer vice-president of sales with Bowne-publishing, while reminiscing about his winning a large screen TV said “Actually the main reason I wanted it was that it was the top prize. I could afford to buy a big screen but it was not the same as winning it.”

Whereas the above examples suggest that there are some benefits in stressing differences among employees, the opposite point is often made that it is appropriate to adopt an egalitarian approach by de-emphasizing symbolic differences (see Pfeffer, 1994). Human resource management in companies known for their outstanding performances provides striking illustrations such as the manager of the contract manufacturer Selectron giving up his/her private office (see section 4 below for more examples). In the present paper we propose a simple framework in which the desirability of using symbols to stress differences among organization members can be assessed.

The need for recognition is typically viewed as immaterial and nonetheless a vital one. In 1805 John Adams said “The desire of the esteem of others is as real a want as hunger - and the neglect and contempt of the world as severe a pain as the gout or stone…”

1Quoted in Fussell (1983).
The importance of social recognition in professional life is well illustrated by the following remark by a steelworker reported by Terkel (1972): “I would like to see a building, say the Empire State, I would like to see on one side of it a foot-wide strip from top to bottom with the name of every bricklayer, the name of every electrician, with all the names. ... Picasso can point to a painting. What can I point to? A writer can point to a book. Everybody should have something to point to.”

Typically, sociologists use social status to capture the need for social recognition. As defined by Weber (1922), social status is “an effective claim to social esteem in terms of negative or positive privileges”. He insists that a status ranking is not directly related to wealth or income though it may be affected by them. Empirically, there is obviously a strong correlation between social status and material well-being. There is for instance a clear positive correlation between the ranking of occupations in term of social status by respondents in surveys and the average income in these occupations. However, the status ranking of occupations may be much better explained if education is added along with income as an explanatory variable (see Perrot, 1999, for details). Hence the theory of Veblen (1899), according to which status stems mostly from relative income or wealth, is somewhat restrictive.\(^2\) An opposite argument could actually be made for the reverse causality: a higher status is the basis for earning a higher income. There is some experimental evidence, both from psychologists (Jemmott and Gonzalez, 1989) and economists (Ball and Eckel, 1996, Ball and Eckel, 1998, and Ball, Eckel, Grossman and Zame, 2001) that an exogenous and random distribution of status among individuals has a significant impact on their relative performance.\(^3\) Belliveau et ali. (1996) study how CEO compensations are affected by the CEO’s status relative to that of the compensation committee.

\(^2\)See Weiss and Feshtman (1998) survey for references on the implications of Veblen’s theory in economic models. In this tradition, and closer to our preoccupation, Fershtman and Hvide and Weiss (2001) paper studies how status concern (i.e., ranking based on workers comparison of their wage to the wage of other workers within the same firm) affects wage distribution. Their main result is that status concern increase wage inequality.

\(^3\)Ball et al. (2001) created status by arbitrarily awarding a “gold star” (a pin) through a “ceremony” to half of the subjects. Next the subjects were asked to play a standard buyers/sellers game (oral double auction). They found that status was a significantly positive (and unconscious –the gold star was never mentioned in the strategy the players reported to follow) determinant of a subject’s earnings. The result holds whether it was clear or not to the participants that the gold star was awarded on an arbitrary basis.
Symbols in organizations are a means of commanding recognition by others. Even when monetary incentives are widely used, companies devise ways of providing their successful employees with conspicuous awards. For instance it is a common practice to grant top sales people medals, rings, sculptures, plaques and so on, handed out during lavish ceremonies (see Nelson 1994). Oftentimes, recognition for deserving employees is institutionalized through exclusive clubs such as the “100 Percent Club” at Norwest Corporation Financial Services or the “Top Élite Club” at AGF Insurance which honors the top one hundred sales agents each year. Recognition is therefore a key feature of work incentives even when output is easily measured. Yet it would be misleading to view recognition as a cheap substitute for money. A common theme in organization theory is that there is some cost associated with a differential treatment of people performing similar tasks. There has been a substantial body of research, in the wake of Adams (1965), concerned with the impact of “unequal” or “unfair” treatment on work motivation. Indeed according to Adams’ equity theory, people react to inequity by making up for it. For instance they lower their input if they feel that what they obtain in return is insufficient, relative to what others around them obtain. While symbolic differentiation is enjoyed by those with high status, it is disliked by those with low status who, as a result, lose some motivation. Recognition has a cost because it is valued in relative terms: what matters is earning more recognition than others. The framework presented here captures this cost of granting recognition in a context where there is a benefit to treating workers differently for incentive purposes.

We consider a simple multi-agents moral hazard problem and allow for an agent’s preferences to depend on her status as well as income and effort. Our specification of preferences has the property that high status agents are willing to exert more effort in exchange for additional income while better paid agents are willing to exert more effort.

\footnote{For economic arguments against large pay differences see Milgrom (1988) and Lazear (1989).}
in exchange for an improved status. As sociologists would put it, agents exhibit a taste for status congruence: those with high status are more eager to earn a large income while those with a high income strive to obtain a status in line with their income ranking. In Section 2, we argue that this is coherent with results in the management and organization literature.

Our results differ depending on whether the work relationship is long term or short term. In a short term (once and for all) relationship, status in the organisation, which can be thought of as a rank in a formal hierarchy, may not be handed out as a reward for good past performances. Then the relevant question is whether an employer would \textit{ex ante} choose to differentiate status among \textit{a-priori} identical workers. In line with Adams’ equity principle we find that this is not a good management practice. We first show that individuals with a higher status should receive a higher expected wage. This is coherent with the observed complementarity between status and income. The egalitarian result may then be understood as follows. For a fixed status allocation, the principal benefits from taking advantage of the substitutability between status and income by paying a lower expected wage to those agents with higher status. This, however, contradicts the requirement that status and income should be complements. Hence only a flat hierarchy may be optimal. We conclude that introducing symbolic differentiation among workers performing similar tasks is costly for an organization. Although agents with high status are more responsive to monetary incentives, the resulting benefits are outweighed by the impact of a lower motivation to work of those with lower status. This short term result emphasizes the cost of status differentiation often stigmatised in the management and human resource management literature.

In order to bring in the benefits of differentiation, we adopt a long term perspective and consider an organization comprised of overlapping generations of agents. We show that the optimal solution for the organization involves giving young agents a status as low as possible along with no monetary incentives. Their motivation to work stems solely from the prospect of being promoted. Old agents are promoted and are granted a higher status and a higher expected income. For incentive purposes promotions are more substantial
for those who have been successful in the past; they end up with prestigious positions paid above their marginal productivity. Because individuals’ preferences exhibit positive cross effects between status and money, symbolic and material rewards reinforce each other. By concentrating both types of compensations in the same time period and in the same state of nature, the organisation exploits their complementarity to reduce the total wage bill. Although this differential treatment of older employees reduces instantaneous profit, the loss is more than compensated by the benefit resulting from the added incentives for junior employees. We show that these results are robust to the introduction of income risk aversion, a case where a standard repeated moral hazard model would prescribe to smooth consumption over time (see for instance Rogerson, 1985, and Chiappori et al., 1994).\footnote{Becker, Murphy and Werning (2000) also find that when status and income are complements risk-averse people may be willing to take gambles (for more on this point see section 3.2 and section 6).}

Our results show that career profiles differ greatly depending on whether or not the employer may commit to long term incentive schemes. Without commitment the employer chooses to introduce limited symbolic differentiation which usually translates into a relatively flat hierarchy. Monetary compensations are performance based so that wages should reflect productivity differences and should not depend too much on the employee’s tenure with the organization. In contrast, an employer that is able to commit organizes an internal labor market where pay is attached to jobs, rewards are delayed in time and a larger income is associated with more recognition (e.g. a higher rank in the hierarchy). Differentiation among junior workers is limited and there are substantial differences in rank and pay for more senior employees. Whereas wage differences are small early in the career they become very substantial and reflect productivity differences as tenure increases. A comparison of job tenures and career profiles in Japan and the US illustrates the empirical relevance of these results.

We describe an agent’s preferences in Section 2, with particular attention to the sources of cross effects between status and income. Section 3 presents the organization and the way status may be allocated among agents. The egalitarian solution in a short-term
setting and the optimality of incentives based on promotions when the relationship is long-term are derived in Section 4 and Sections 5 and 6 respectively. Section 7 illustrates the empirical relevance of our results through a comparison of job tenures and career profiles in Japan and the US and Section 8 concludes.

2 Preferences

We consider the provision of work incentives to agents whose effort level is unobservable. If, as is usually assumed, an agent’s preferences are fully characterized by a taste for money and a distaste for effort, incentives may be provided through monetary rewards and penalties. As we argued in the introduction, actual incentive procedures typically involve many non-monetary attributes. Although some of these attributes clearly provide material benefits (more independence, more influence, better work conditions) many others are symbolic and their value to employees stems mostly from the social or psychological benefit they entail (self esteem or social recognition). In this section we describe how this symbolic dimension may enter into an agent’s preferences. We use the concept of status to summarize the overall access to those psychological or social benefits that an employee may secure through her position in the organization. Following the work by Frank (1984) we assume that status is rooted in tastes.\footnote{For a discussion of this hypothesis see Frank (1985). In Frank’s (1984) pioneering work status stems from relative income ranking.}

We postulate the following utility function:

$$u(w, s, e) = sw - \psi(e), \quad s \geq 0, \ w \geq 0, \ e \geq 0.$$ 

where $s$ is status, $w$ is wage income and $e$ is effort. The disutility of work, $\psi$, is taken to be a strictly increasing, strictly convex and twice continuously differentiable function.

\[ A2 \quad \psi'(e) > 0 \quad \psi''(e) > 0 \quad \text{for} \quad e \geq 0. \]

This specification reflects in a simple manner the agents’ taste for money and status and their distaste for effort. Setting status equal to 1 yields as a special case the standard
quasi-linear utility of a risk neutral agent, so that our results may be readily compared with predictions of standard moral hazard framework. Linearity with respect to wage indicates that agents are risk neutral regarding income. In subsequent sections, we discuss how our results may be affected if this assumption is relaxed.\textsuperscript{7} The requirement that status and wage should be positive is a normalization. Utility could easily be rewritten to allow for non zero lower bounds. The important point is that there are such lower bounds. It is not essential for our results that at the minimal income an agent derives no added satisfaction from an increase in status or that at the minimal status level an agent does not benefit from a larger income.

Since income and status are both positively valued, indifference curves relating these two variables for a given effort level are strictly decreasing. This reflects the substitution between status and income: an agent is willing to accept a lower pay if she is given a more prestigious position in the organization, and conversely. However preferences over status and income are strictly convex so that there is not a perfect substitution between these two variables: a prestigious title does not compensate for the absence of wages, nor does a good wage make up for the contempt of others.

Utility also exhibits positive cross effects between status and income. These cross effects have important implications for the income-effort and status-effort tradeoffs.\textsuperscript{8} Formally, 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. Here marginal rates of substitution are respectively the slopes of indifference curves in the effort/income and effort/status spaces. In other words, 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. Figures 1 depicts the indifference curves in the effort/income space when status varies (the arrows represent the marginal rates of substitution).

\textsuperscript{7}The interpretation of linearity with respect to status is provided in section 7.

\textsuperscript{8}These cross effects matter because preferences are also defined over effort levels. If preferences were defined on status and income alone, then taking the log of the Cobb-Douglas utility would wipe out cross effects with no change in preferences. They also matter because we consider situations involving risk and intertemporal substitution.
\[ \bar{w} < 1 \]
\[ \bar{w} > 1 \]

Figure 1
These cross effects may be best interpreted by relating them to the sociological theory of status, the psychological analysis of work motivation and the conventional wisdom prevailing among management practitioners.

We first consider the impact of a change in status on the income/effort tradeoff. One obvious prediction of our specification of the agent’s preferences is that, for a given level of monetary incentives, an agent should be all the more productive that she has a high status. The literature on job satisfaction suggests that a higher status enhances work commitment. On the one hand, status is closely related to the need for recognition which has been found to be a key factor in job satisfaction (e.g. Dunette, Campbell and Hakel, 1967). On the other hand, many studies have shown that a low job satisfaction tends to result in high turnover and absenteeism rates. Empirically Tahlin (1999) found in a study on job mobility in Sweden that everything else being equal people with low status (i.e., Treiman’s, 1977 prestige score) are more likely to make a voluntary job shift than people with high status. It should be expected that a low satisfaction also results in shirking which, contrary to absence and resignation, is not observable. A positive effect of status on productivity has been found by Greenberg (1988) in a study on office reallocation: this author found that a group of employees temporarily moved to lower-status offices lowered their performance while lower-status employees assigned to better offices increased their output. Finally, Tsui et al. (1992) found that workforce heterogeneity (i.e. mixing different sex groups or ethnic groups) had a negative impact on work commitment for white men whereas this impact was small for women and minorities. To explain this difference they argue that white men feel that their status is deteriorated if they work with women or minorities while the latter may feel that their status is enhanced when they hold positions comparable to those held by white males.

---

9 See for instance Day and Hamblin (1964), Baum and Youngblood (1975).
10 Many studies have shown that there is a positive correlation between job satisfaction and quality of services (see Varma et al. (1999). In contrast researchers like Vroom (1964) failed to establish a clear relationship between job satisfaction and productivity. This may be due in part to the fact that a high job satisfaction is sometimes associated with groups known as the “happiness for lunch bunch” who mostly enjoy the social life associated with the work place (see McCroskey, Larson and Knapp, 1971).
11 They actually found that women working mostly with men had a high degree of attachment to their
We now examine how the trade-off between effort and status is affected by a person’s income. According to our specification of preferences, richer agents care more about their status in the sense that they are willing to exert more effort in order to improve it. The hierarchy of needs proposed by Maslow (1954) provides a nice interpretation of this phenomenon. Maslow argues that there is a five-level hierarchy of human needs ranking from bottom to top: physiological needs, safety needs, social needs, esteem needs, and self-actualization needs. Higher level needs correspond to less material (more psychological) preoccupations. A person develops a taste for higher level needs only after fulfilling those at lower levels. In the present context, income is the means of fulfilling material satisfaction while status is the means of fulfilling psychological satisfaction. Then, individuals with low income are mostly preoccupied with material needs and care little about status while those with higher income having satisfied their material needs are mostly concerned about increasing their status. Various observations, either in the work place or in broader social contexts, illustrate the relevance of Maslow’s construction. Certers and Bugertal (1966) find evidence that factors at the top of Maslow’s hierarchy play a more important role for employees earning higher wages. This is consistent with the logic applied by practitioners when they use non-monetary compensations. A human resource management guide indicates that using merchandise to reward employees is inappropriate for those earning low wages while such prizes are highly valued by those who are paid sufficiently well (see Nelson, 1994). Similarly, rich people seeking social recognition through the funding of charity, fine arts or higher education reflects such a shift in tastes caused by a higher income.12

In the next section we present the organization and in particular we describe how status may be allocated among agents.

12For instance, it has been observed that children with high-income parents typically select high-status positions (see Treiman and Ganzeboom, 1990 and Lillard and Reville, 1997). On a more anecdotal note, Cornelius Vanderbilt Whitney earning a Ph.D. for the sheer pleasure of being referred to as Doctor Whitney illustrates this appetite for status among rich people (see Fussell, 1983).
3 The Organization.

3.1 Production.

We consider an organization (bureau, subdivision, firm,...) supervised by a risk-neutral principal. There are \(n \geq 2\) workers identified by an index \(i = 1, \ldots, n\). They are assumed to be \textit{ex-ante} identical individuals, hired to do the same type of work. That is, there is no \textit{a priori} legitimate motive for treating them differently. The principal aims at maximizing expected profit, with profit \(\pi\) defined by

\[
\pi(Q, w_1, \ldots, w_n) = Q - \sum_{i=1}^{n} w_i.
\]

(1)

where \(Q = \sum_{i=1}^{n} q_i\) is total output (its price is normalized to 1) and \(w_i\) is the wage paid to agent \(i\).

Each worker contributes to the collective outcome by exerting an effort \(e_i \geq 0\). The harder agent \(i\) works (the higher \(e_i\) is), the larger is the probability of a high output for the unit. Formally, individual \(i\) contributes to the total output for an amount \(q_i\), which may be either high \(q_i = \overline{q}\), with probability \(\mu(e_i)\) or low \(q_i = \underline{q}\), with probability \(1 - \mu(e_i)\) \((\overline{q} > \underline{q} > 0)\). Individual output, \(q_i\), is verifiable. That is, contrary to Lazear and Rosen (1981) where only \textit{relative} performance is observable, here \textit{absolute} performance is observable. This is a case where direct individual monetary incentives are particularly appropriate. Under our assumptions any departure from these direct incentive contracts may be ascribed to the introduction of the social status motive. The probability of a high performance for agent \(i\), \(\mu(e_i)\), increases with \(e_i\) at a decreasing rate (the impact of effort on performance declines for larger effort levels). The function \(\mu(.)\) is assumed to be strictly increasing, strictly concave and twice continuously differentiable.

\[
\begin{align*}
A1 & \quad \mu'(e) > 0, \quad \mu''(e) < 0 \quad \text{for} \quad e \geq 0.
\end{align*}
\]
3.2 Status in the Organization

Organizations may grant recognition to their members through various formal sources of status: wage distribution, distribution of scarce non monetary resources (such as offices, furniture, computers, locker rooms, dining facilities...), conspicuous awards or, most commonly, positions in the organization’s hierarchy. In the case of hierarchies, the choice of a status allocation is constrained by the production process (i.e., the technology). Yet there are many instances of firms in the same industry resorting to different hierarchies despite similar production technologies. In the auto industry Toyota has seven layers of management between its CEO and employees on the factory floor, whereas Ford has seventeen and GM has as many as twenty-two (see Milgrom and Roberts, 1992). At Nucor Corporation, a steel producer, the number of layers in the executive hierarchy has been restricted to 4 against a dozen on average for the rest of the industry (see Ghemawat, 1995). From the above examples it is apparent that organizations are somewhat able to manipulate hierarchical differentiation to provide work incentives. In order to emphasize the relationship between status and work incentives we abstract from the technical role played by the hierarchy and leave much latitude to the principal in allocating status among organization members.\footnote{In the paper we consider departure from the ‘minimal’ hierarchy (i.e., from the technical hierarchy). We study whether the principal has interest in adding ranks for incentives purposes.}

In spite of her ability to act as a social engineer, the principal still faces a constraint because status is enjoyed through interpersonal comparisons. Independently of the method used to grant social recognition, its value is perceived in relative terms. Then the increase in an individual’s status is always achieved at the expanse of someone else’s status. For instance, if status is derived from a person’s position in a formal hierarchy, in order to increase one agent’s status, it is necessary to improve her position in the hierarchy relative to some other members who, inevitably, suffer some loss. In other words, status in organizations is a scarce resource. Formally, let \( s = (s_1, \ldots, s_n) \) denote a status allocation in a feasibility set \( S \subseteq \mathbb{R}^n_+ \), the \( i \)th component measuring the status of agent \( i \). The scarcity assumption may be stated as follows.
**A3** It is not possible to improve an agent’s status without degrading some other agent’s status: \( \forall s, s' \in S, \text{if } s_i > s'_i \text{ for some } i, \text{ then } s'_j > s_j \text{ for some } j. \)

Individuals being *ex ante* identical, the feasibility set should also satisfy the following anonymity condition.

**A4** If a status allocation \( s \) is feasible then any permutation of this allocation, denoted \( s^p \), is also feasible: \( s \in S \Rightarrow s^p \in S. \)

Finally, in order for the principal to have as much freedom as possible in fine tuning the status allocation, the feasibility set should be connected. We impose the more stringent restriction that it is convex.

**A5** The feasibility set \( S \) is convex.

By assumption A3, the feasibility set is akin to a Pareto frontier and assumption A4 requires that this frontier be symmetric across individuals. Assumption A5 implies that this frontier is linear which is somewhat restrictive and is meant to ease the exposition of the results (especially in the optimization problem). Some discussion of the robustness of our results to more general functional forms is provided in Section 6. It is straightforward to establish that under assumptions A3-A5 a status allocations is feasible if and only if it satisfies a linear constraint that may be written as

\[
(F) \quad \sum_{i=1}^{n} s_i - n = 0, \ s \in \mathbb{R}^n_+.
\]

Overall status summing up to \( n \) is a normalization. Any other strictly positive constant would lead to the same results. However \( n \) has the convenient property that, when no status disparity is introduced, all agents have a status of 1 so that our results may easily be contrasted with those of the classical moral hazard literature with quasi-linear agent preferences.

Finally when an agent enters the organization, she must have a position in the hierarchy so that, contrary to wages, status is awarded *before* the agent exerts effort. Thus,
in a static context, status may not be used as a reward (see section 4). It may however depend on past performance in a long term relationship (see section 5).

If employees are treated differently as a result of their past performance, this differentiation seems legitimate. However if the principal chooses to treat new employees in different ways the differentiation will be deemed arbitrary. In our framework, there is no basis for \textit{a-priori} legitimacy. Agents are \textit{ex-ante} identical and perform identical and independent tasks. This is clearly a simplification. In real world situations there is always some heterogeneity among workers: sex, race, age, personal history, and so on. This heterogeneity may be the basis for a discriminatory treatment.\footnote{For instance Cawley (2000) finds evidence that in the US weight lowers wages for white women. In his study a difference in weight of two standard deviations (roughly sixty-five pounds) is associated with a difference in wages of 7\%.} In section 4, we show that such an \textit{a-priori} unequal treatment is actually undesirable for the organization. This however is due to the employee’s inability to commit in actual work relations. To see why this is the case, we now briefly discuss what would be the optimal incentive scheme in the first-best situation where there is no moral hazard problem.

3.3 First Best Allocation

We first consider the benchmark case where each agent may fully commit to a contractible effort level as well as to an unconditional participation in the organization. It is then optimal for the principal to offer each agent to participate in a lottery where there is only one winner who receives all of the status and who is the only employee being paid. The main argument in the proof is that, instead of having two agents with positive status, the joint status could be given to only one, where each of them would receive this total status with some probability. The added status for each agent when she is paid exactly compensates her for a lower probability of being paid. This allows for paying each agent less often, thus lowering the expected wage bill by a factor $n$.\footnote{For the sake of simplicity consider the case where the individual probability to win the lottery is $\frac{1}{n}$. The prize is $s^{\text{win}} = \bar{s}$, where $\bar{s}$ is a strictly positive constant (e.g. $n$ in (F)) and $w^{\text{win}} = \frac{\bar{s}}{n}(U + \psi(e^*))$ where $e^*$ is the first best effort level (i.e., it is solution to $\psi'(e) = \mu'(e)\Delta q$). With such a lottery the individual expected utility is $\bar{U}_i$, each agent commit to effort level $e^*$ and the total wage bill is $\frac{n}{\bar{s}}(U + \psi(e^*))$. In comparison when $s_i = \frac{\bar{s}}{n} \forall i$ (i.e., the classical principal-agent model), the individual}
cross-effects between status and income, it is optimal to concentrate status and monetary rewards on one individual.

One might think that the optimality of a lottery depends on income risk neutrality or on the linearity of the feasibility constraint. It turns out that the result is quite robust. A lottery is still optimal if, utility is linear in one argument and either the agent is risk averse regarding income, or utility is strictly concave in status (which is equivalent to introducing an additively separable quasi-convex feasibility constraint for status).\textsuperscript{16} Becker, Murphy and Werning (2000) obtain related results while studying the evolution of inequalities when individuals care both about income and status and the two are complements. They show that individuals are willing to participate in lotteries even though they have a diminishing marginal utility of income. As is the case here, complementarity between consumption and status implies that individuals are willing to take gambles in which winners get to consume more, along with a higher status.\textsuperscript{17}

Actual work relations allow for much less commitment on the part of the agent than what was postulated here. Subsequent sections investigate the implications of our model in more realistic settings.

4 The Cost of Status Manipulation

Real world work relations typically involve a moral hazard problem since effort levels are not perfectly verifiable. Furthermore, the ability of an agent to commit is limited by work legislation which usually forbids clauses that would prevent her from quitting at any time. For both these reasons, the first-best solutions may not be implemented: losing agents would either not participate or, if they did participate, they would exert no effort. In this section we reconsider the static problem and show that moral hazard and the lack of

\[ w_i = \frac{\overline{U}}{\Psi} \left( U + \psi(e^*) \right) \]

The total wage bill is

\[ w^2 = \frac{\overline{U}}{\Psi} \left( U + \psi(e^*) \right) \]

\textsuperscript{16}See section 6 for related arguments.

\textsuperscript{17}Focusing on the evolution of inequality they do not consider moral hazard problem. There is no effort in their setting. They get the nice result that starting from different distribution of wealth, the society ends up with an unique unequal distribution. They conclude that there is an incompressible level of inequality if people care about social status.
of commitment on the part of agents force the principal into choosing a solution which drastically differs from the first-best.

The moral hazard problem and the agent’s lack of commitment translate into incentive compatibility constraints and interim participation constraints respectively. The information structure of a static relationship is as follows:

\textit{stage 1}: the principal offers contracts which specify for each agent status and monetary rewards;

\textit{stage 2}: agents choose whether or not to participate;

\textit{stage 3}: interim information (the draw of a lottery, if any) is revealed and agents choose whether to quit or not;

\textit{stage 4}: agents chose their effort levels;

\textit{stage 5}: outputs are observed and payments are made.

The new constraints are a consequence of stages 3 and 4. The interim stage 3 may seem unnatural in this context and is solely introduced for the sake of comparability with the first-best solution. As we already pointed out, the lottery involved in the first-best clearly violates both the interim participation constraints of stage 3 and the incentive compatibility constraints of stage 4.\footnote{It is \textit{a-priori} less apparent whether the added constraints rule out lotteries all together. Proposition 2 shows that they do.}

At stage 5, payments may depend on output. Let \( w_i \) be agent i’s fixed salary and \( \Delta w_i \) be agent i’s bonus in case of a high performance (i.e., \( w_i + \Delta w_i \) and \( w_i \) are agent i’s wages associated to outputs \( q \) and \( q \) respectively). Worker i chooses her effort so as to maximize:

\[
EU_i = \left( \mu(e_i)\Delta w_i + w_i \right)s_i - \psi(e_i).
\]

Under assumptions A1 and A2, the agent’s utility is strictly concave in effort and therefore has a unique maximum point. Agent i’s optimal effort, \( e^*(s_i\Delta w_i) \), solves the following
first order condition,
\[ \frac{\psi'(e^*(s_i \Delta w_i))}{\mu'(e^*(s_i \Delta w_i))} = s_i \Delta w_i. \]

(3) Standard comparative statics shows that, from the concavity of \( \mu \) and the convexity of \( \psi \), \( e^* \) is increasing in \( s_i \Delta w_i \). It is independent of \( w_i \) due to income risk neutrality.

Taking into account additional constraints, the principal’s program may be written as
\[
\max E \sum_{i=1}^{n} \mu(e_i)(\Delta q - \Delta w_i) - w_i + q
\]
subject to
\[
\sum_{i=1}^{n} s_i = n, \text{ with probability 1,}
\]
\[
s_i[\mu(e_i)\Delta w_i + w_i] - \psi(e_i) \geq U \quad \forall i = 1, ..., n, \text{ with probability 1,}
\]
\[
e_i = e^*(s_i \Delta w_i) \quad \forall i = 1, ..., n \text{ with probability 1.}
\]

Note that the \textit{ex-ante} participation constraints are dropped since they are implied by the interim participation constraint.

The following proposition states three conditions that should hold in an optimal allocation and which, in short, say that a higher status goes hand-in-hand with a higher income.

**Proposition 1 (status/income complementarity)** In an optimal solution, the three following conditions must hold with probability 1.

(i) \( \Delta w_i \leq \Delta q \quad \forall i = 1, ..., n. \)

(ii) \( \Delta w_i = \Delta q \text{ or } w_i = 0 \quad \forall i = 1, ..., n. \)

(iii) Suppose \( e^* \) is concave. Then, \( s_i < s_j \text{ if and only if } w_i = w_j = 0 \text{ and } \Delta w_i < \Delta w_j, \) or \( w_i < w_j \).

The proof is in the appendix. Part (i) is the standard result that there is no point for the principal in giving more than full incentives. Part (ii) is also quite standard;

\footnote{Sufficient conditions for \( e^* \) to be concave are that \( \mu''' < 0 \) and \( \psi''' > 0. \)}
given that the agent is risk neutral, the principal abstains from giving full incentives only when she is restricted in the choice of the low performance wage. The novel insight from Proposition 1 is in part (iii) which states that agents with differing status, either receive different low performance wages (the higher status agent being better paid) or receive different incentives (the larger high performance reward going to the higher status agent). That is, different status levels imply an unequal treatment in monetary as well as symbolic rewards. This logic is exploited fully in the first best solution, where the whole status and money is concentrated on one agent. However, as the next proposition shows, the lack of commitment on the agents’ part, makes unequal treatment among agents suboptimal.

**Proposition 2 (symbolic egalitarianism)** To maximize instantaneous profit it is optimal to give identical agents identical contracts (same status, same monetary scheme).

Appendix 2 provides a proof of Proposition 2 under some technical conditions on $\psi$ and $\mu$.\(^{20}\) To give some intuition for this result, consider the case where at least one agent, $i$, receives a strictly positive low performance wage. Then it is easy to show, using Proposition 1, that if some other agent’s status differs from that of agent $i$, profit may be increased. To see this, note that (iii) in Proposition 1 implies that the agent with the larger status necessarily has a strictly larger expected utility (which is therefore strictly above $U$) and a weakly larger low performance wage. The larger low performance wage must be strictly positive since it is at least as large as that of agent $i$. Hence the low performance wage of the agent with a larger status may be decreased without violating her individual rationality constraint so that profit would increase.\(^{21}\)

The situation where the principal chooses to give strictly positive low performance wages occurs when $\underline{U}$ is large enough, namely when

$$\underline{U} > \mu(e^*(\Delta q))\Delta q - \psi(e^*(\Delta q)).$$

This lower bound is obtained as follows. First note that the status of the agent getting the worst treatment may not exceed 1. Since, from (i) in Proposition 1, it is not optimal to

\(^{20}\)From the argument that follows, those technical conditions are only needed in the case where $\underline{U}$ is low enough so that the lower bound constraint for the wage is binding.

\(^{21}\)Note that because of risk neutrality, effort is unaffected by a change in the low performance wage.
give monetary incentives exceeding $\Delta q$, if (4) holds, her individual rationality constraint requires that she receives a strictly positive low performance wage. All agents must therefore have an equal status of 1. Then (ii) in Proposition 1 prescribes that all agents should receive full incentive, thus being rewarded $\Delta q$ for a high performance. This is not surprising since agents are risk neutral regarding money and the principal is, in this case, not constrained by a lower bound on monetary transfers. The case where $U$ is low which is more complex, is analyzed in the appendix.

Proposition 2 is the first formalization with the tools of economists of the equity theory in social psychology according to which it is harmful to introduce differences among workers performing identical tasks (see Adams, 1965). A major theme in human resource management is that symbolic differentiation among workers is an obstacle to communication, cooperation, and commitment for those who are in lower positions. Pfeffer (1994) argues that “symbolic egalitarianism” is a key feature of human resource management in successful companies. He points to such examples as the car manufacturer NUMMI, where the executive dining room has been eliminated, or the contract manufacturer Selectron whose CEO has no private office. The well documented success story of Nucor Corporation is another striking illustration (see Ghemawat, 1995). At Nucor, external signs of hierarchical differences are systematically de-emphasized (no personal secretary, common parking lot, everybody flying economy class, and so on). Nucor relies on direct monetary rewards to provide work incentive. The average Nucor salary is comparable to competitors’ average salary, but the structure differs (it has more bonus and less fix). Yet the turn-over rate is low. Obviously Nucor has succeeded in stimulating individual involvement in the firm’s operation by its original work management.

By virtue of Proposition 2 symbolic differentiation is sub-optimal. The benefit made on those with improved status does not compensated for the loss made on those whose status is deteriorated. According to this egalitarian logic, in a short term relationship the principal chooses to rely on direct monetary incentives and avoids hierarchical differences. Workers’ incentives stem solely from the variable part of their pay,
as recommended by standard micro-economics. This result is consistent with empirical evidences on short term contracts (e.g., for temporary workers). In a once and for all relationship only technological constraints motivate the introduction of a hierarchy. The next sections show that this conclusion is no more valid if the work relation is long term.

5 Overlapping Generations in the Organization

Work relationships between individuals and organizations are in general medium to long term. As workers stay longer than one period within the organization, the principal has more instruments to provide them with work incentives than in the previous section. Indeed she can replicate the static contract, but can also propose an intertemporal incentive scheme that links future rewards to past performances.

We study this problem within an overlapping generations setup with an infinite horizon. At each date, the organization is comprised of two “generations”: the “young” (juniors) who enter the organization in the current period and the “old” (seniors) who joined the organization in the previous period and will not be around in the next one. Hence each cohort only stays two periods. Lotteries are ruled out and we assume that the principal is able to commit. Finally we restrict the analysis to equitable contracts: all young agents at period $t$ are offered the same two period contract. Thus identical agents (i.e. with identical résumés) receive an identical treatment. Proposition 2 suggests that this restriction is reasonable. Moreover

The timing for a cohort joining the organization at date $t$ is as follows.

date $t$:

stage 1, the new cohort of workers is offered contracts that include a beginning status level, a monetary incentive scheme and a promotion system (future status and monetary scheme depending on past performance);

stage 2, agents choose whether or not to participate;

\footnote{For more on this see Milgrom and Roberts (1992).}
stage 3, workers choose an effort level based on their current monetary incentives and status as well as their promotion prospects;

stage 4, outputs are observed, transfers and promotions occur;

date \( t + 1 \):

stage 5, agents choose whether to stay or to leave;

stage 6, workers choose an effort level according to their current monetary incentive and status (which may depend on whether they have been successful in the first period or not);

stage 7, outputs are observed, transfers occur, workers retire.

Note that stage 5 implies that, as is the case in actual work contracts, an agent may not commit for two periods. An individual rationality constraint for old workers must therefore be included.

Each worker’s intertemporal utility is assumed to be additively separable with a discount factor of \( \delta < 1 \). The expected utility of an old worker exerting an effort \( e_{pt} \) and whose past performance has been \( p \in \{ l, h \} \) (\( l \) is for “low” and \( h \) is for “high”) is that of the static model (see equation 2):

\[
EU_{pt} = [\mu(e_{pt}) \Delta w_{pt} + w_{pt}]s_{pt} - \psi(e_{pt}).
\]

Let \( \Delta U_t = EU_{ht} - EU_{lt} \). A young worker’s expected intertemporal utility if her effort level is \( e_{1t} \) is

\[
EU_{1t} = s_{1t}[\mu(e_{1t}) \Delta w_{1t} + w_{1t}] - \psi(e_{1t}) + \delta[\mu(e_{1t}) \Delta U_{t+1} + EU_{l(t+1)}].
\]

The individual rationality constraints are:

\[
(\text{IR'}) \quad EU_{pt} \geq U, \ p \in \{ h, l \} \quad \text{and} \quad EU_{1t} \geq (1 + \delta)U.
\]

Let \( e^* \) be implicitly defined by equation (3). It is easy to check that the incentive compatibility constraints for the young and for the old may be written as follows.
The population is assumed to be large so that it may be represented by a continuum with a measure normalized to 2. Then, at each period, the proportion of old who have been successful when young, denoted $\gamma_t$, is equal to the probability $\mu(e_{1,t-1})$ that, in the previous period, a young agent had a high performance. The feasibility constraint on status allocation is:

$F'\quad s_{1t} + \gamma_t s_{ht} + (1 - \gamma_t) s_{lt} = 2$ \quad with $\gamma_t = \mu(e_{1,t-1})$.

Let us denote $c_{1t} = (s_{1t}, w_{1t}, \Delta w_{1t})$ the contract of a young worker at date $t$, and $c_{pt} = (s_{pt}, w_{pt}, \Delta w_{pt})$ the contract of an old at date $t$ with performance $p \in \{h, l\}$ at date $t - 1$. As in the static model the principal faces three types of constraints at each period: $(F')$, $(IR')$, $(IC')$. She must pick a sequence of contract combinations $<(c_{1t}, c_{ht}, c_{lt})>$ that maximizes intertemporal profit subject to those constraints. The principal has the same discount factor as workers, $\delta < 1$, so that there is no exogenous bias against, or in favor, of delayed monetary rewards. Her intertemporal profit may be written as:

$\mathbb{E}\Pi_t = \sum_{t=0}^{\infty} \delta^t \{ \mu(e_{1t})(\Delta q - \Delta w_{1t}) - w_{1t} + \gamma_t [\mu(e_{ht})(\Delta q - \Delta w_{ht}) - w_{ht}]$

$+ (1 - \gamma_t) [\mu(e_{lt})(\Delta q - \Delta w_{lt}) - w_{lt}] + 2q \}$.

Initial conditions, $\gamma_0$, $c_{h0}$ and $c_{l0}$, are exogenously given. Finally we define a steady state to be a situation in which $(c_{1t}, c_{ht}, c_{lt})$ is independent of time (i.e. all generations are offered the same intertemporal contract).

6 Incentives and Promotions

One well known puzzle in the literature on internal labor markets (ILM) is the extensive use of promotions as a means of providing incentives.\textsuperscript{23} It is widely accepted that, although

\textsuperscript{23} According to Doeringer and Piore (1971) the main features of an internal labor market are: long term employment relationships, limited port of entry for hiring, career paths within the firm and promotion.
promotions have an obvious role as a screening device, they are, to a large extent, used as a reward for good past performances. As Milgrom and Roberts (1992, Chapter 11) note, it is not necessarily the case that the best person in a top position is the one who performed the best at lower levels. This may result in a conflict between the two roles of promotions: incentives and screening. To solve this conflict, companies such as 3M or IBM have created separate career ladders for scientists and engineers so that they may be promoted without having to go into management. Similarly, faculty members in universities, or physicians in hospital are generally promoted without changing job.

Even when there is no conflict between incentive goals and screening goals, the extensive use of promotions for incentive purposes remains puzzling. Direct monetary transfers allow for a better fine tuning of the incentive scheme contrary to promotions which are discrete and rare. One explanation for the use of discrete incentive schemes is that it is not always possible to assess absolute performance whereas relative performances are somewhat easier to establish. Then promotions may be viewed as a prize in a tournament between employees as in Lazear and Rosen (1981). Our analysis suggests that because of the trophée value of winning a prize, tournaments are appropriate even if absolute performance may be observed. Management practices directed at sales force employees, whose output is easily observed, are consistent with this result. Then viewing promotions and tournaments as an instance of differentiation among workers yields valuable insights into their role in work incentives; They help fulfilling individuals’ need for recognition.

While the static approach emphasized the costs of symbolic differentiation, the present dynamic setting highlights its benefits as part of an intertemporal incentive scheme. In view of the various constraints which pertain to the dynamic profit optimization problem, one would expect that the exact nature of the solution depends very much on which of these constraints are binding. Although this is true to some extent, the results in the next from within.

24Along with the “Top Élite Club”, AGF has created the “Club Pro” designed to encourage steady high performances. In order to join, a sales person must meet some prespecified goals (they are absolute, not relative). It takes three years of high performance to join, and only one year to join back for a former member who had been dropped after a bad year. The “Club Pro” thus represents a persistent change in status. A high past performance is formally acknowledged as in a promotion system.
proposition are quite general.

**Proposition 3 (incentives through promotion)** In a steady state of a profit maximizing solution we have

\[ s_1 = w_1 = \Delta w_1 = 0, \]
\[ s_h > s_l, \]
\[ w_h \geq w_l \text{ and } \Delta w_h \geq \Delta w_l, \]

where at least one of the inequalities in (8) is strict.

The above Proposition, which is proved in the appendix, provides a crisp characterization of the optimal intertemporal incentive scheme. It is optimal to endow young agents with the lowest possible status level while providing them with no direct monetary incentive.\(^{25}\)

Junior workers earn the same salary independently of their performances. They are induced to exert effort by the prospect of a future promotion. That is, pay is attached to job and earnings profiles become individual specific only as careers unfold. When old, an agent’s status and monetary incentive scheme depend on her past performance. As in the static context, it is optimal to associate a higher wage to a higher status. However, in contrast with the egalitarian solution of Proposition 2, it is optimal to introduce some differentiation between generations and among old agents. A higher past performance induces a higher status as well as higher monetary compensations. This solution allows for taking advantage of positive cross effects between status and income by concentrating benefits in both dimensions on one state of nature. This is reminiscent of the first-best solution in the static problem where the whole status and wage are concentrated on one individual.

An important result in the literature on repeated moral hazard is that the optimal long term incentive contract should involve some memory: the type of incentives currently

\(^{25}\)As explained earlier utility could easily be rewritten to allow for non zero lower bounds (e.g., \(u(w, s, e) = (w + 1)(s + 1) - \psi(e))\). The important point is that there are such lower bounds.
given to an agent depends on her past performance (see for instance Rogerson, 1985, and Chiappori et al., 1994). 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). The need for consumption smoothing implies that it is not optimal to delay all rewards and penalties as is done in actual promotion systems and in accordance with the prescriptions of Proposition 3. An obvious difference between the model of this paper and the standard repeated moral hazard framework is the agents’ attitude towards risk on income. Here agents are taken to be risk neutral while they are usually assumed to be risk averse. With risk neutral agents in the standard model (where all have equal status), the principal is indifferent as to when the young should receive monetary rewards as long as all parties discount the future at the same rate and their limited liability constraint is not binding. Then postponing all rewards as is done in Proposition 3 is one solution, but not the only one.

We now briefly explore the robustness of our results to the introduction of some income risk aversion in the agents’ preferences. It turns out that in order to investigate this point, it is fruitful to concurrently discuss the robustness of the results to a change in the status technology. In the model considered here, the status constraint is linear and utility is linear in status. This may loosely be interpreted in saying that there are constant returns to concentrating status on one group of individuals. It might be expected that, if those returns were sufficiently decreasing, the result that the young should have a minimal status would be upset. There are two possible options for making returns to concentrating status decreasing: either the

left-hand side of the status feasibility constraint could be made strictly quasiconvex or utility could be made strictly concave in status. The second route is followed in the argument below.

Let us rewrite instantaneous utility as

\[ u(w, s, e) = g(s) h(w) - \psi(e), \quad s \geq 0, \quad w \geq 0, \quad e \geq 0, \]

24
where $h$ and $g$ are concave functions satisfying $h(0) = g(0) = 0$.

In Step 1 of the proof of Proposition 3, it is shown that if $h$ and $g$ are linear, then a steady state in which the young have positive status cannot be optimal since it is possible to obtain the same effort levels and the same intertemporal utility level while increasing the utility of the old and paying a lower wage bill. The following proposition states that this argument may be generalized to situations where either $h$ or $g$ is strictly concave.

**Proposition 4** Suppose that the agent’s instantaneous utility is linear in income ($h$ linear) or linear in status ($g$ linear) while the discount rate is small enough. Then in any steady state of an optimal solution we have $s_1 = w_1 = \Delta w_1 = 0$.

The proposition is proved in the appendix. It shows that the result that the young should receive a minimal status is upheld when either income risk aversion is introduced or utility is strictly concave in status. Because earnings and status are complement, individuals are willing to take gambles (in exchange for their effort while young) in which winners receive both a higher income and a higher status (for related results with income risk aversion see Becker, Murphy and Werning, 2000).

### 7 Job tenure and career profiles

Combining Propositions 2 and 3 our results indicate that an organization will choose to resort to symbolic differentiation for incentive purpose only when it can set up an internal labor market. More specifically, the predictions of the model are as follows. In a long term relationship rewards for a high performance are delayed in time and a pay increase is associated with a change in status which usually is achieved by a move up in the hierarchy. Furthermore, the use of direct monetary incentives is limited and wages are to a large extent attached to jobs, so that promotions are the main source of incentives. Differences in productivity are then reflected in wages only for senior employees. In contrast, if commitment is not possible, no symbolic differentiation is introduced. Incentives are provided through direct monetary rewards. The hierarchy is flatter and pay is no longer
attached to jobs. Employees with different productivities are paid different wages early on. The feasibility of an internal labor market hinges on the employees’ expected tenure within the organization which conditions the organization’s ability to commit. A comparison of work relations in the United States and in Japan illustrates the two situations of strong and weak commitment.

According to the US Bureau of Labor Statistics the average person in the US holds 9.2 jobs from age 18 to age 34. More than half of these jobs are held between the ages of 18 and 24 (Department of Labor 2000). This does not mean that there is no internal labor market in the US. However they tend to begin late in the career (i.e., after age 35). As Brown et al. (1997) note “retention rates are bifurcated for male workers in the US. American junior workers have only a 25 percent to 35 percent chance of being with the same employers five years hence, while workers with five to ten years’ experience have 90 percent retention rates in the utilities, transportation, and service industries and 60 percent to 70 percent retention rates in government, finance, manufacturing, and agriculture (Nakata, 1990).” As Farber (1999) shows, most new jobs in the US end early, and the probability of a job ending declines with tenure. In contrast in Japan labor mobility is low for young core workers. For instance 3/4 of Japanese engineers will have only one employer during their entire career (Jacobs and Herbig, 1998). Hashimoto and Raisian (1985), using data from the 1960s and 1970s, indicate that among male workers holding a job for at least 5 years when they are 20-24 years old, 65% retain that job 15 years later in Japan against 30% in the United-States. These differences have been remarkably stable since the early 1970s.

The present analysis then predicts that, while young, Japanese workers receive relatively low wages, independently of their education level. Differentiation comes later in the career so that the earnings profile is increasing with seniority at an increasing rate. That is, the earnings profile is roughly upwards sloping and convex. In contrast in

---

26 Internal labor markets do exist in the US, and, contrary to what is often believed, they are quite stable (see Groshen and Levine, 1998).
27 For updated data see Brown et al 1997 pp 31.
the US young workers who are very mobile do not accept delayed rewards. Their earnings profiles should be relatively steeper at young ages (i.e., under 35), and flatter later (not necessarily upwards sloping). Earnings, which better reflect workers’ productivity, should also be more differentiated across education levels. This implies that the disparity in earnings is higher for young workers in the US than in Japan.

According to the Bureau of Labor Statistics real earnings of individuals in the US increase more rapidly at young ages than at older ages. From the ages of 18 to 24, real hourly earnings grow on average by 6.6 percent per year. This growth rate falls to 4 percent between age 25 and age 29 and then to 2.4 percent between age 30 and age 34 (US Department of Labor 2000). In contrast in Japan young core workers receive undifferentiated treatment. “White-collar and blue-collar pay tables are integrated into a single table that erases distinctions between the two categories. There is also no major gap between production workers and craft workers. New workers are placed at the bottom of the ability rank table and given simple assignments.” (Brown et al., 1997 pp 105). Differentiation appears with seniority so that a pay raise is coupled with a change in status. "University graduates may reach management in 10 years, typically by the time they reach ages 35 to 40. High school graduates may reach management in twenty-two years, and most have reached a management grade by age 50. Much of the career-based pay increases take place only when, and if, workers are promoted to managerial positions that are not in the union, generally after age 35." (Brown et al., 1997 pp 111).

This implies that for young workers (i.e., below age 30-35) the level of earnings and the variance of earning are low. Figures 2 and 3 which are borrowed from Brown et al. (1997) pp. 117 and 118 provide a striking illustration of the results discussed above.

[Figures 2 and 3]

Figure 2 depicts earnings by age and education in the automobile industry and in the electrical industry in Japan and in the US; Figure 3 gives earnings profiles by age and

---

28 Young American workers who face flat tenure-earnings profile change job to increase their earnings. For instance Topel and Ward (1992) found considerable returns to between-job mobility in a study of white male high school graduates.
Earnings by Age and Education, 1990

Sources: A: Computed by the authors from Basic Survey of Wage Structure, Japan Ministry of Labor. B: Computed by the authors from the Current Population Survey, March 1990.

Note: A: Regular male workers, scheduled earnings. B: Male private wage and salary workers, annual earnings.

Figure 4.9 – p 118
From Brown et al (1997)
education on a national-level basis. In Japan differentiation in earnings comes after age 35 and the earnings gap between different type of workers is widening with age.\textsuperscript{29} In contrast in the US earnings which increase (sharply for educated workers) in the junior years but not necessarily afterwards, are differentiated early. The earnings gap between educated and non educated workers widens until age 35-39 and then stabilizes. In the US the earnings profile is roughly increasing and concave; in Japan it is increasing and convex. Since the industries studied are standardized the difference in earnings profile cannot be explained by technological differences. They reflect different management practices.

Our theoretical analysis predicts that internal labor markets are a superior mode of work organization. The theoretical prediction is somewhat corroborated by a comparison of per worker growth rates of GDP in the US and in Japan. Over the period from 1979 to 1990, its annual average has been 2.9 percent in Japan against 1 percent in the US (OECD 1992). It is generally accepted that the distinctive human resource management of Japanese firms provided an important source of increased efficiency and growth. If this is the case one may wonder why firms in the US do not resort to them more systematically. There is no obvious answer to this question. There are arguably cultural differences in work related value orientations that affect the labor market organisation.\textsuperscript{30} For instance US workers are supposedly more individualistic and more concerned with annual income than with long-term employment since money income, rather than company affiliation, is important in marking success. Furthermore, dismissal regulation is more stringent in Japan and firms use alternatives to layoffs for employment adjustment of core workers. Houseman (1997) shows that countries with a strict dismissal regulation have longer job tenures. Finally to organize an internal labor market, firms need not only to commit to keep employees, but also to be large enough or growing to propose stimulating career paths. In recession phases this might not be possible. Recently in Japan there has been a growing number of middle-aged

\textsuperscript{29}At an aggregate level, earnings profile in Japan increases with age at an increasing rate until 55. After that age companies encourage workers to retire.

\textsuperscript{30}Hofstede (1980) identified four dimensions along which dominant patterns of culture can be ordered: power distance, uncertainty avoidance, individualism, masculinity. Later he added long-term orientation. Japan scores higher than the US on all dimensions except for individualism.
employees, referred to as "marginal employees" (genkai shain), who are on payroll but are unemployed. This situation is not stable because firms in recessions cannot afford to pay high wages to redundant employees. Moreover junior employees will not work very hard to later become the “bulge generation”, the “overpaid”, or the “window gazers”. This is opposed to the logic of grouping rewards in status and income later in the career.

The Japanese system of life time employment, and more generally promotion systems, are not appropriate for unstable economic environment. When flexibility matters so that commitment is not possible there is no benefit in creating hierarchical structure for incentives purpose. Firms should rely on direct monetary incentives. In recent years there has been a significant move towards delayering in industrial countries. For instance the study by Bauer and Bender (2001) on a representative German employer-employee data set reveals that between 1993 and 1995 50.73% of the 251 firms sampled went through a reduction in hierarchy levels. Similarly it has been noted that during the 1980s, many US companies have engaged in reducing the number of employment levels. For instance, General Electric (chemical division) cut the number of pay grades from 22 to 5.31 According to our analysis, this evolution may be the result of a weakening employer’s commitment which could be explained by an anticipated increase in the job loss rate. Indeed there have been evidences of such an increase during the 1990s (see for instance Farber, 1997).

8 Conclusion.

We have argued that social recognition has a major role in the work place and we have shown that taking this role into account may enhance our understanding of actual management practices. Social aspects are all the more significant that much of labor relations take place outside the market. Our analysis relies on the following two premises: recognition and income are complements and recognition is scarce because it is valued in relative terms. Our main findings are that, while it is costly to introduce differentiation between

identical coworkers in a static environment, such a differentiation may prove to be quite a powerful incentive device in a dynamic setting. Stylised facts are consistent with our theory. In particular the grouping of rewards implies that those who are promoted are paid above their marginal productivity. This helps to explain why the wage increase associated with a promotion is oftentimes out of proportion with any reasonable guess on the increase in marginal productivity resulting from switching to a job higher in the firm’s hierarchy (Lazear 1991).\footnote{In Auriol and Renault (2001) we investigate the implications of Proposition 3 for the specific shape of the optimal incentive hierarchy assuming that $\mu(e) = \min\{e, 1\}$ and $\psi(e) = A^2 e^2$. We find that the harder it is for an employee to improve performance through effort (i.e., the larger $A$), the more pyramid-like is the incentive hierarchy. Indeed when $A$ is very large success is rare; it is extremely prestigious to get promoted and the associated pay raise is huge (it diverges in the limit). On the other hand if a high performance is easily achieved a seniority based promotion system may be optimal (i.e., everybody is successful and is promoted).}

This paper focuses on one particular aspect of the incentive role of promotions. There are of course other theoretical explanations. In particular a stylised fact that received much attention in the economic literature is the positive relationship between tenure and pay. The use of large prizes only attributed at specific times in a career is often interpreted as an attempt by firms to improve employee attachment (see for instance Becker, 1962, Salop and Salop, 1976 or Lazear, 1979). Lazear (1979) proposes an appealing explanation which is based on firm-specific capital. He shows that firms that want to invest in specific capital in their workers offer a back-loaded compensation structure to retain them. Farber (1999) tries to explain the high return to tenure by testing firm-specific human capital theories. He concludes that “the capital that accrues with tenure has a strong industry-specific rather than firm-specific component. To the extent that this is the case, it is harder to argue that the accrual of firm-specific capital is what drives the decline in the probability of job change with tenure”. This empirical result shows that

firm-specific investment in human capital cannot be the sole determinant of the design of promotion systems and hierarchies. It is also affected by the need to honor deserving employees and to meet functional goals such as production efficiency or screening.

Taking a broader perspective, the present paper shows that an analysis of incentive
packages that would only focus on monetary incentives would overlook many important 
issues in human resources management. Yet, taking a full account of all relevant aspects 
and achieving a proper assessment of the costs and benefits of differentiation may prove to 
be a difficult task. For instance some authors have stressed that ill defined recognition 
systems lead to mental health disorders for employees. Hood (2002) notes “In North 
America, mental health problems cost US$80 billion/year, while stress and mental health 
disorders increased about 200% over the five years up to and including the year 2000.”

Accounting for such considerations is presumably not a simple challenge.

33According to the author the top 5 sources of workplace stress are: 1) the feeling of not contributing and 
lacking control, 2) lack of two way communication up and down, 3) being unappreciated, 4) inconsistent 
performance management processes, 5) career and job ambiguity.
Appendix 1: Proof of Proposition 1.

The proof of the first two conditions is straightforward. First note that for a given status level $s$, total surplus is a strictly concave function of effort which reaches a maximum at $e^*(s_i \Delta q)$. Thus if $\Delta w_i > \Delta q$, total surplus may be increased by decreasing $\Delta w_i$, while keeping the agents’ utility unchanged by increasing $w_i$. This clearly increases profit. By a symmetric argument, if $\Delta w_i < \Delta q$ and $w_i > 0$, profit could be increased by increasing $\Delta w_i$ and decreasing $w_i$.

To prove that condition (iii) must hold first note that $(w_j, \Delta w_j) = (0, 0) \iff s_j = 0$. Thus if $s_j = 0$ the result clearly holds. The remainder of the proof deals with the case where $s_i > s_j > 0$.

It is useful to define

$$g(w, \Delta w, s) = \frac{\mu(e^*(s \Delta w))\Delta w + w}{s}.$$ 

Step 1: If $s_i > s_j > 0$, and $g(w_i, \Delta w_i, s_i) = g(w_j, \Delta w_j, s_j)$ then condition (iii) must hold.

Applying the implicit function theorem to the requirement that $g$ be constant ensures the existence of a differentiable function $(w(s), \Delta w(s))$ with partial derivatives

$$\frac{\partial w}{\partial s}(s) = \frac{- (\partial g/\partial s)(w, \Delta w, s)}{(\partial g/\partial w)(w, \Delta w, s)},$$

$$\frac{\partial \Delta w}{\partial s}(s) = \frac{- (\partial g/\partial s)(w, \Delta w, s)}{(\partial g/\partial \Delta w)(w, \Delta w, s)}.$$

Since the partial derivatives of $g$ with respect to $w$ and $\Delta w$ are clearly strictly positive, the above partials will be strictly positive as long as $(\partial g/\partial s)(w, \Delta w, s) < 0$, or $[-\Delta w \mu(e^*(\Delta ws)) - w + \Delta w^2 e''(\Delta ws)\mu'(e^*(\Delta ws)) < 0.$ Since $(w, \Delta w) \neq (0, 0)$, if $\Delta w = 0$ the above condition clearly holds. If $\Delta w > 0$, it holds under the sufficient condition that $[\mu(e^*(\Delta ws)) - \Delta we''(\Delta ws)\mu'(e^*(\Delta ws))] > 0$. Since $\mu(e^*(0)) = 0$ the above condition holds as long as $\mu \circ e^*$ is concave, which in turn is true if $e^*$ is concave.
Step 2. If $s_i > s_j > 0$, $g(w_i, \Delta w_i, s_i) = g(w_j, \Delta w_j, s_j)$.

Since all agent have strictly positive status, it is always possible, for any pair of agents, to increase the status of one while reducing that of the other one by the same amount. Consider now what would happen if this is done while keeping both agent’s effort level and utility constant. This amounts to keeping the sum $s_i + s_j$ and the products, $s_k \Delta w_k$, and $s_k w_k$, $k = i,j$, constant. These requirements may be written out as $\Delta w_k = f_1(s_k)$, $w_k = f_2(s_k)$ and $s_j = f_3(s_i)$, where $f_1$, $f_2$ and $f_3$ are three differentiable functions with derivatives respectively defined by $f_1'(s_i) = -\frac{\Delta w_i}{s_i}$, $f_2'(s_i) = -\frac{w_i}{s_i}$ and $f_3'(s_i) = -1$. Then the derivative of profit with respect to $s_i$, given that $s_j$ changes by the same amount with an opposite sign and that efforts and utilities remain unchanged is given by

$$-\mu(e^*(s_i \Delta w_i))f_1'(s_i) - f_2'(s_i) - f_3'(s_i)\mu(e^*(s_j \Delta w_j))f_1'(s_j) + f_2'(s_j).$$

If this derivative was non zero, it would be profitable to transfer status from one agent to the other. Now the requirement that this derivative is zero yields the result.

Appendix 2: Proof of Proposition 2.

First we show that if status differs across agents then $w_i = 0$ for all $i$.

Suppose not. Then for some agent $i$, $w_i > 0$. Then from Lemma 2(ii), $\Delta w_i = \Delta q$. Now suppose that for some agent $j$, $s_j > s_i$. Then by Lemma 2(ii) and 2(iii) $\Delta w_j = \Delta q$ and $w_j > w_i$. Then agent $j$’s expected utility $U_j$ is larger than agent $i$’s. Since individual rationality must hold for agent $i$, $U_j > 0$. Then profit could be increased by decreasing $w_j$. If on the other hand, $s_j < s_i$ for some $j$, a similar argument shows that it is possible to increase profit by decreasing $w_i$.

Now if $w_i = 0$ and (IR) is binding, then $\Delta w_i$ must be larger if status is lower. Then, from Lemma 2(iii), all agents in that situation must have the same status.

Now consider agents for whom $w_i = 0$ and (IR) does not bind. Setting the first derivative of expected profit with respect to $\Delta w_i$ to 0 we find that the optimal solution
must satisfy
\[ s_i e^s(s_i \Delta w^*(s_i)) \mu'(e^*(s_i \Delta w^*(s_i)))(\Delta q - \Delta w^*(s_i)) - (e^*(s_i \Delta w^*(s_i))) = 0. \]

Standard comparative statics shows that
\[ s_i \Delta w^*(s_i) = -\Delta w^* + \frac{\Delta q e^s \mu'(e^*)}{-s_i(\Delta q - \Delta w)[e^{s''}\mu'(e^*) + [e^s]^2 \mu''(e^*)] + 2e^s \mu'(e^*)} \]

Some computations show that \( \Delta w^*(s_i) \leq 0 \), if and only if
\[ -\Delta w^* s_i[e^{s''}\mu'(e^*) + [e^s]^2 \mu''(e^*)] \geq e^s \mu'(e^*) \frac{\Delta q - 2\Delta w^*}{\Delta q - \Delta w^*}. \]

A sufficient condition for the above inequality is that the derivative of \( xe^s(x)\mu'(e^*(x)) \) with respect to \( x \) should be non positive. Using the first order conditions for optimal effort, this derivative should be equal to that of \( e^s(x)\psi'(e^*(x)) \) which is non positive if \( \mu'' < 0, \psi''' > 0 \) and
\[ \frac{\psi''(e)}{\psi'(e)} \leq -\frac{2\mu''(e)}{\mu'(e)}. \]

for all \( e \).

Then, \( \Delta w^* \) is decreasing in \( s_i \) and, from Lemma 2(iii), all agents with a zero low performance wage with (IR) not binding must have identical status.

Finally suppose that there are two agents \( i \) and \( j \) with \( w_i = w_j = 0 \) and such that (IR) is binding for \( i \) only. Then, from Lemma 2(iii), this is possible only if \( s_i < s_j \). Thus \( \Delta w^*(s_i) > \Delta w^*(s_j) \). Yet, since (IR) is binding for agent \( i \), she must receive a wage differential even higher than \( \Delta w^*(s_i) \) and this contradicts Lemma 2(iii)

**Appendix 3: Proof of Proposition 3.**

Consider a steady state. Then there exists \( (c_1, c_l, c_h) \) such that \( (c_{1t}, c_{lt}, c_{ht}) = (c_1, c_l, c_h) \) for all \( t \). The proof proceeds in three steps.

**Step 1.** \( c_1 = (0, 0, 0) \).
If $s_1 = 0$, then it is clearly optimal to set $w_1 = \Delta w_1 = 0$. Thus proving the result amounts to showing that $s_1 = 0$. Suppose to the contrary that $s_1 > 0$. At some date $t$ the principal may switch to

$$c'_1 = (0, 0, 0), \quad c'_h = (s_h + s_1, \frac{\delta s_h w_h + s_1 (w_1 + \Delta w_1)}{\delta (s_h + s_1)}, \frac{s_h \Delta w_h}{s_h + s_1})$$

$$c'_l = (s_l + s_1, \frac{\delta s_l w_l + s_1 w_1}{\delta (s_l + s_1)}, \frac{s_l \Delta w_l}{s_l + s_1})$$

It is readily verified that if each generation from $t$ on is offered these contracts, the young’s expected intertemporal utility is maintained while the old’s intertemporal utility is increased. Furthermore, all effort levels are maintained. On the other hand, the intertemporal wage bill from each generation is clearly lower. Hence intertemporal profit has increased. It follows that a steady state with $s_1 > 0$ cannot be part of an optimal solution.

**Step 2:** If $U_h > U_l$, then (7) and (8) must hold.

First note that the arguments used to prove Proposition 1 may be applied to the old population at each period so that the two propositions, (7) and (8), are equivalent. Furthermore, it is obvious that if $U_h > U_l$, we cannot have $s_l \geq s_h$, since it would imply that high and low performance wages for type $l$ old workers should be at least as high as those of type $h$ old workers.

**Step 3:** $U_h > U_l$.

Let us first show that a steady state in which the young’s effort is zero cannot be part of an optimal solution. In such a steady state, at each date, only the old exert effort. Now suppose that at some date $t$, the principal commits to giving only half of the status to the old at date $t+1$. Then she is in a position to implement the egalitarian solution of Proposition 1 which is optimal in the static problem. Since the solution in which only the old exert effort is also feasible in the static problem, it yields a strictly lower per period profit. Thus the young’s effort must be strictly positive in the steady state of an optimal solution.

Since the young exert effort in spite of a zero status, we must have $U_h > U_l$. 35
Appendix 4: Proof of proposition 4.

Consider a steady state. An agent may face four possible states of nature.

For one such state of nature, \( \omega \), let \( s_1(\omega) \) and \( w_1(\omega) \) denote the agent’s status and wage when young and, \( s_2(\omega) \) and \( w_2(\omega) \) denote the agent’s status and wage when old, Let \( v(\omega) = g(s_1(\omega))h(w_1(\omega)) + \delta g(s_2(\omega))h(w_2(\omega)) \). In order to ease notation, the reference to the state \( \omega \) is dropped in the remainder of the proof. Now suppose that \( s_1 > 0 \). If the principal switches to a solution \((s'_1, w'_1, s'_2, w'_2)\), with \( s'_1 = w'_1 = 0 \) and \( s'_2 = s_1 + s_2 \), \( v \) is unchanged as long as

\[
(9) \quad h(w'_2) = \frac{g(s_1)h(w_1) + \delta g(s_2)h(w_2)}{\delta g(s_1 + s_2)}.
\]

It is readily verified that if this is done for all states of nature, effort levels and intertemporal expected utility are unchanged while the agent’s utility when old increases.

First suppose that the agent is risk neutral regarding income, so that \( h(w) = w \). Then (9) becomes \( [w'_2 = \frac{g(s_1)w_1 + \delta g(s_2)w_2}{\delta g(s_1 + s_2)}] \). Since \( g \) is strictly increasing, the discounted wage bill \( \delta w'_2 \) is lower than \( w_1 + \delta w_2 \). Thus the principal may obtain the same effort levels while paying the agents less.

Now suppose that agents are risk averse but utility is linear in status (i.e. \( h \) is linear and \( g \) is strictly concave). Then (9) reads

\[
(9) \quad h(w'_2) = \frac{s_1 h(w_1) + \delta s_2 h(w_2)}{\delta(s_1 + s_2)}.
\]

Strict concavity of \( h \) implies

\[
h \left( \frac{s_1w_1 + s_2w_2}{s_1 + s_2} \right) > \frac{s_1 h(w_1) + s_2 h(w_2)}{(s_1 + s_2)}.
\]

Thus, with sufficiently little discounting, if \( w'_2 \) satisfies (9), then \( \delta w'_2 < w_1 + \delta w_2 \).
References


