DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

INTERPERSONAL COMPARISON, STATUS AND AMBITION IN ORGANISATIONS

Florian Ederer and Andrea Patacconi

Number 222

February 2005

Manor Road Building, Oxford OX1 3UQ
Interpersonal Comparison, Status and Ambition in Organisations

Florian Ederer
MIT
ederer@mit.edu

Andrea Patacconi
University of Oxford
andrea.patacconi@economics.ox.ac.uk

First Draft: December 2003
This Version: December 2004

Abstract
This paper argues that the prevalence of compensation systems which reward winners without explicitly identifying losers can be rationalized by workers’ concern for relative payoffs. If the workers’ participation constraints are binding, the firm must compensate its employees for the disutility that they may derive from low status. It follows that profit-maximizing employers may be particularly reluctant to penalize or give poor performance evaluation to employees. The theory also sheds light on many other puzzling features of incentive schemes in practice, such as small salary premia, rat races, job title proliferation, the gender wage gap, the gender/happiness paradox and the widespread use of tournaments as a sorting device.

Keywords: Reference-Dependent Preferences, Status, Ambition, Expectations, Tournaments.

JEL Classification: J31, J41.

1 Introduction
In an important paper, Baker, Jensen and Murphy (1988) discuss a number of common and important features of organisational compensation systems that are hard to reconcile with

*We are particularly indebted to Meg Meyer for her invaluable support and comments and Klaus Schmidt for his guidance at an early stage of the paper. We also would like to thank Heski Bar-Isaac, Stefano DellaVigna, Alex Edmans, Robert Gibbons, David Gill, Ian Jewitt and David Laibson for helpful discussions and remarks, as well as the Seminar for Economic Theory at the University of Munich for its hospitality. Financial support from the Economics Department at the University of Oxford, Nuffield College, ESRC and the European Commission’s Marie Curie Fellowship Programme is gratefully acknowledged. All remaining errors are ours.
the standard theory of incentives. Among these they include the prevalence of compensation systems that reward ‘winners’ without explicitly identifying ‘losers’.

“Every economist understands that a compensation scheme paying a salary of $80 plus a bonus of a $20 if a quota is met is equivalent to a compensation scheme paying a salary of $100 with a $20 penalty if the quota isn’t met. What economists don’t understand is why compensation plans almost always are of the former type instead of the latter.” (Baker et al. 1988, p. 607)

This paper provides a rationale for the asymmetric effects of rewards and punishments and for other puzzling features of actual compensation systems. We study a simple tournament model in which $N$ risk-neutral agents compete for a set of prizes. Contrary to previous contributions we make the realistic assumption that workers (and people in general) care about their relative standing, that is, their status within a group. In particular, we assume that people dislike being behind in terms of status in the organisation, but enjoy being ahead of individuals in their reference group.

In our model, individual concerns for relative standing are captured by the reference-dependent nature of workers’ preferences (see, e.g., Kahneman and Tversky 1979 and more recently Koszegi and Rabin 2004). Specifically, we posit that a worker’s utility depends not only on his wage $w$ and effort $e$, as in standard incentives models, but also on a “reference” or “typical” wage, $w^R$. Nonpecuniary gains and losses relative to $w^R$ are interpreted as measuring the magnitude of the worker’s status concerns relative to a chosen reference group.\(^1\)

A crucial element of our model is how exactly the reference wage is determined. Our answer hinges on the kind of expectations the worker has about the wages he and the other workers are likely to get in the tournament. There is considerable evidence in fact in support of the idea that recent expectations are an important determinant of reference points.\(^2\) Introspection also suggests that they are key in explaining employee satisfaction with wages. Koszegi and Rabin (2004), for instance, argue that an employee who confidently expects a 10% pay rise would probably assess a raise of only 5% as a loss. Likewise, workers may be satisfied with their current wage in stationary environments but may develop a taste for improvement in environments where workers have become accustomed to improvement.

This paper distinguishes between objective and subjective elements that contribute to the formation of expectations and hence reference points. We posit that the choice of a reference point is in part dictated by external or objective factors and in part determined by individual traits and personal aspirations. External factors refer to the objective structure of the game

\(^1\)This general framework encompasses several models of social preferences. For instance, in two-agent models of inequity aversion à la Fehr and Schmidt (2003), $w^R$ is given by the wage the other worker receives. In Bolton and Ockenfels (2000), instead, $w^R$ is the weighted average of other people’s payoffs. As will become clear in the following, however, the way in which we model $w^R$ differs significantly from these approaches.

\(^2\)Stone (2003) and Koszegi and Rabin (2004), in particular, survey the psychological evidence in support of this viewpoint.
under examination: for instance, in a tournament, the ratio of high wages to low wages, or the initial conditions/status quo (in a tournament this might be the worker’s wage before the tournament takes place). Individual aspirations instead reflect subjective (but not necessarily completely random) differences in terms of ambition (i.e., high expectations) among agents.

The theoretical economics literature has mainly focused on the first set of issues. For instance Shalev (1997), Stone (2004) and Koszegi and Rabin (2004) all derive expectations and hence reference points from the (objective) structure of the game and the players’ anticipated equilibrium strategies. By contrast, social psychologists have stressed that reference points are to some extent actively chosen (Wood and Taylor 1991; Heath, Larrick and Wu 1999). That literature also suggests that some individuals may be more ambitious than others, in the sense that they may set for themselves more demanding standards and goals. Falk and Knell (2004) present evidence in line with this hypothesis in the context of intellectual performance. In a questionnaire study Falk and Knell asked students from the University of Zurich to indicate their aspired diploma degree as a measure of their reference standard. High school grades were taken as a proxy for their abilities. Their study shows that, controlling for socioeconomic characteristics and related questions from the General Social Survey, female students have significantly lower reference levels than male students and that more able individuals set higher standards for themselves than less able agents. Moreover, people who think that ambition is important to get ahead also select higher reference standards.

Our paper incorporates insights from both the economics and psychology literature in a simple tournament model and shows that their incorporation has important consequences for the optimal provision of incentives. We assume that workers compare the wage they obtain in the tournament with a typical wage, $w^R$, and consider both external and personal factors that may influence the choice of $w^R$. In the basic model, we focus on external factors, in particular on the role of the tournament and prize structure. Specifically, we assume that $w^R$ is the wage that most workers participating in the tournament will receive. For instance, if few workers are promoted or demoted and most remain at the same intermediate level, then the ‘typical’ wage is the one paid at the intermediate level. Likewise, if the majority of workers receive a bonus of 3% of their current salary, then a bonus of 3% is the typical salary premium. Although our discussion will generally relate to promotion as the outcome of the tournament, it is important to stress that there is nothing in our framework that makes promotion inherent to winning. Indeed, the model simply assumes that workers who perform better receive a higher wage. Nevertheless, we believe that in practice the results of a tournament are more visible than higher pay, and that therefore interpersonal comparisons are more likely to matter in the former case (see also our discussion in the concluding remarks).

The above definition clearly emphasizes the role of expectations and tries to capture the idea of reference point as the customary or normal state of affairs. In our model workers categorise

---

3The conditions that identify one of the possible prizes in a tournament as the reference wage (and that we label the reference-wage constraint) are presented later when the model is described.
the possible outcomes within classes and attach a (almost) neutral connotation to one of them, namely the largest one. Thus, and this is the main motivation for our approach, workers not promoted in tournaments where there are only few winners do not perceive themselves as ‘losers’ but as being ‘normal’. This is likely to have important consequences for workers’ morale since, as Baker et al. (1988) have argued, by rewarding winners without explicitly identifying losers, “the vast majority of employees who incorrectly rank themselves near the top of their peer group can still believe, when they are passed over for promotion, that they may not be the best but are nonetheless somewhere near the top” (p. 607). By contrast, models such as those where the reference point is a weighted average of all wages cannot easily capture this effect. In those models in fact the reference wage is typically distinct from all the other prizes in the tournament and, as a result, workers are only able to classify themselves as either winners or losers (relative to the reference point).

Our framework seems also able to accommodate the well-known, paradoxical finding highlighted by Samuel Stouffer in “The American Soldier” (1949) that for each level of longevity, rank and education, the less the promotion opportunities afforded by a branch or combination of branches, the more favorable the opinion tends to be toward the promotion opportunity. So, for instance, although the Air Corps had a conspicuously high rate of promotion, Air Corps men were definitely far more critical of chances of promotion than, say, men in the Military Police, whose objective chances of promotion were about the worst in any branch of the Army. Stouffer’s explanation of this puzzling evidence revolves around the concept of “relative deprivation” and more precisely, as Merton and Kitt (1950) put it, the idea that “a generally high rate of mobility induces excessive hopes and expectations among members of the group so that each is more likely to experience a sense of frustration in his present position and disaffection with the chances of promotion.” (p. 54).

Finally, we stress that our modeling approach is cognitively very parsimonious. Indeed, although our definition of the reference point is admittedly a very coarse one, it allows the reference points to be computed with relatively little information about other employees’ wages. By contrast, alternatives such as computing a weighted average of the other workers’ wages or the equilibrium of a game requires very precise information on behalf of the workers about the general structure of the tournament, as well as the performing of complex calculations. Nevertheless, for completeness and to facilitate the comparison between our approach and more standard models of other-regarding preferences, in Appendix A we briefly consider how our results would change if the reference wage were the weighted average of all wages.

The focus of the first part of our analysis is on how a rational firm can take advantage of the interpersonal comparison of its workforce to increase profits. We study the optimal design of tournaments where workers have concerns for relative standing and compare our solutions to the standard case where workers have no such concerns as analysed by Lazear and Rosen (1981). The most important result is that, if the workers’ participation constraints are binding, the firm must compensate its employees for the disutility that they may derive from low status.
Therefore it may be particularly reluctant to penalise or give poor performance evaluation to employees. In particular, and in contrast to standard results, we show that, even with risk-neutral workers, the choice of the tournament format is not neutral, and the firm is better off choosing tournaments that reward winners instead of tournaments which penalise losers, since the former make the positive outcomes salient. When both promotion and demotions are considered, we show that under mild conditions it is optimal for the firm to minimise the number of losers and maximise the number of winners, subject to a reference-wage constraint that identifies which of the prizes in the tournament is the reference wage. The results also shed light on other puzzling features of incentive schemes in practice, such as small salary premia, rat races and job title proliferation.

We emphasise that by appropriately choosing the design of the tournament, the firm is actually trying to modify workers' expectations and therefore their reference points. In that respect, the model is similar to other papers in the behavioural economics literature that stress, for instance, how firms can offer contracts to customers to take advantage of their time-inconsistent behaviour (DellaVigna and Malmendier 2004) or can unnecessarily complicate the design of a product to exploit customers' limited attention (Gabaix and Laibson 2003).

In the second part of the paper we turn attention to the personal, idiosyncratic factors that determine reference points. In particular, we allow for individual differences in ambition, where ambition is defined as the tendency to set for oneself a high reference point. This definition is consistent with the central argument put forward by Heath et al. (1999) who argue that goals can serve as reference points. Differences in ambition are here interpreted and related to differences in reference groups: for instance, women may have lower reference wages because they compare themselves not with workers in general (male and female) but only other female workers. Analogously, high ability workers may have higher reference wages because they compare themselves only with other talented individuals.

To keep the model tractable and to focus on ambition in isolation, we consider tournaments with two agents and two jobs and set aside the role that expectations or the normal state of affairs have in determining the reference point. We may think of other female or high ability workers as belonging to other firms. Moreover, in line with psychological evidence reviewed in Section 1.1, we assume that people dislike losses (i.e. low status) relative to the reference point more than they like equivalent gains (high status). We show that if, as the empirical and experimental evidence suggests, men set for themselves higher reference points than women, then it is possible to explain the observed gender wage gap without invoking discrimination or differences in ability or in outside opportunities. We also provide an explanation for the so-called gender/happiness paradox, that is, the puzzling finding that women, although discriminated on the job in several respects, self-report higher levels of satisfaction at work than men.

Finally, we allow workers to differ also with respect to ability. In that case, it is well-known that tournament promotion systems cannot in general match employees to the jobs for which they are best suited, since performance in the current job is not necessarily the best indicator of
performance in another job. We question the practical relevance of that result in the light of the fact that empirically ambition and ability seem to be positively correlated. Our model shows that if more able workers compare themselves to equally able workers and not to low ability workers, then more able individuals will exert more effort, thereby further increasing their chances of winning the tournament and reducing the probability of an inefficient promotion decision. Our theory therefore suggests that the adverse consequences of using tournaments as a sorting device could be smaller than previously thought.

We wish to conclude our preliminary discussion with a word of caution. The aim of our paper is not to argue that status concerns are the best or the only explanation for the puzzling evidence we consider in this paper: indeed several other explanations could and have been proposed to reconcile economic theory with observed compensation practices. Nevertheless, we believe that much of the strength of the current approach derives from the fact that a small, yet very realistic modification of the orthodox economic model can explain a large variety of detailed evidence, rather than just a few stylised facts.

1.1 Evidence on Status

In his seminal book Frank (1985) argues that natural selection should favour those who are driven to seek high status. For instance, high status individuals may be more able to obtain food in adverse circumstances, compared to low status individuals. Furthermore, concern for relative standing can improve prospects in bargaining situations as it acts as a pre-commitment device. Comparing ourselves to others helps us to better understand what our talents are and to find our way in the society. Finally, status constitutes an important source of arousal and motivation. Status is associated with prestige and deference behaviour (Ridgeway and Walker 1995), and can be used as a mean to further other objectives (Lin 1990, 1994; Thye 2000). A vast body of experimental evidence shows that even an exogenous and random distribution of status among individuals has a significant impact on their relative performance (e.g., Ball, Eckel, Grossman and Zame 2001).

In the medical literature, Marmot and his collaborators have demonstrated in a series of influential studies that the social gradient in health and disease is one of the dominant features of the health situation of all industrialised countries (Marmot 2003). They showed that, even among people who are not poor, there is a social gradient in mortality that runs from the bottom to the top of society. For instance, the Whitehall II study followed British civil servants over a 10-year period and considered two possibilities: changes in health were responsible for changes in social position or changes in social position were responsible for changes in health. Results showed that, overwhelmingly, it is the effect of social position on health that dominates.4

4Interestingly, nonhuman primates exhibit a similar social gradient in the risk of cardiovascular disease and the degree of atherosclerosis. Marmot (2003) reports the following experiment involving monkeys. High status monkeys from different troops were taken and put together. In the new troop, a new hierarchy was formed. As a result, some previously dominant monkeys now had a subordinate place. Similarly, low status monkeys were
Among others, Frank (1985), Robson (1992) and Zizzo (2002) note that status might also be valued per se. Huberman, Loch and Önçüler (2004) provide experimental evidence that subjects value status independently of any monetary consequence and are willing to trade off some material gain in order to obtain it. Heath et al. (1999) argue that goals serve as reference points and alter outcomes in a manner consistent with the value function of Prospect Theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992). They present evidence that goals inherit key properties of the value function such as a reference point and loss aversion. Furthermore, there seems to be a reasonable amount of heterogeneity with respect to status-seeking behaviour, in particular marked gender and country/culture differences (Falk and Knell 2004; Huberman et al. 2004; Gneezy, Rustichini and Niederle 2003).

Organisational behaviour theorists have long been aware of the importance of status. For instance, Buchanan and Huczynski (1997) note that “status is important because it motivates people and has consequences for their behaviour. [...] Each position in a group has a value placed upon it. Within the organisation, a value is ascribed to a position by the formal organisation (e.g. Chief Controller, Vice President, Supervisor) and can be labeled formal status. Formal status is best thought of as being synonymous with rank, as in the police or in the armed forces, and reflects a person’s position on the organisational ladder” (p. 213).

1.2 Related Literature

There is a large literature on reference-dependent choice models beginning with Kahneman and Tversky (1979). Weiss and Fershtman (1998) survey the economics and sociology literature on social status and describe its role in economic analysis of saving and consumption, wages and economic growth.

Despite the vast body of empirical and experimental evidence of the importance of status and reference-dependency, until recently very little attention has been paid to the development of models of the labour market that incorporate these preferences. Fershtman and Weiss (1993) study how wages and status are determined endogenously in a simple general equilibrium model, emphasising the relationship between the distribution of wealth and occupational choice. Fershtman, Hvide and Weiss (2003) examine the effects of differences in the importance that workers give to status ranking and the reference group to which they compare themselves. They show that, for equally productive workers, wages may vary, reflecting the different incentives that firms provide to workers with different social concerns. Fershtman, Hvide and Weiss (2002) and Goel and Thakor (2004) independently show that when workers (or managers) care about relative status, the optimal provision of incentives involves some form of joint performance evaluation, where the wage of one agent is positively related to the performance and the wage of the other agents in the reference group. Meyer and Mookherjee (1987) also reach a similar conclusion. However, their result does not rely on workers having concerns for relative
standing; instead, it is the principal’s preference for ex post equality of individual welfares that introduces an equity component into the optimal reward schemes. Auriol and Renault (2003) analyse the impact of status concerns on incentives in a principal-agent model. They allow firms to provide status symbols that are independent of wages, and show that higher levels of effort can be elicited. Finally, Loch, Stout and Huberman (2000) construct a dynamic simulation model of a group in which status competition can improve or hurt team performance depending on whether competition is mainly based on merit or on political maneuvering.

By modeling status by the use of other-regarding preferences our analysis is also related to the burgeoning literature on fairness and inequity aversion surveyed by Fehr and Schmidt (2003). From an analytical point of view the contributions by Demougin and Fluet (2003) and Grund and Sliwka (forthcoming) are most closely related to the analysis of this paper. Whereas we concentrate on status concerns, they analyse the effects of inequity aversion in tournaments. Both papers focus on the simple two player case and hence, in contrast to the current model, do not explore the determination of the optimal number of winners and losers. They also do not allow for heterogeneity among agents.

Wage compression within organisations has been analysed by a number of papers. The contributions most relevant to our analysis include Frank (1984, 1985) and Lazear (1989), but constitute only small portion of this vast literature.

The remainder of the paper proceeds as follows. In Section 2 we introduce the basic model. Section 3 extends the basic model to examine demotions. Section 4 introduces worker heterogeneity in terms of ambition, and in Section 5 we also allow for ability differences. Finally, Section 6 concludes and discusses extensions to the present analysis.

2 The Basic Model

There are \(N\) identical workers, or agents, indexed \(i = 1, 2, ..., N, N \geq 2\), who compete in a tournament. The \(N_W \leq N\) agents who produce the most output receive a wage \(w_H\), while the remaining workers \(N_L \equiv N - N_W\) receive a wage \(w_L\). Individual output is given by \(x_i = e_i + \varepsilon_i\), where \(e_i\) denotes worker \(i\)'s effort and \(\varepsilon_i\) is an error term with cumulative density function \(F(\cdot)\), density \(f(\cdot)\) and zero mean. \(i\) is a ‘winner’ of the tournament, and we write \(i \in W\), if he or she is one of the \(N_W\) workers that produce the most output. Otherwise \(i\) is said to be a ‘loser’, and we write \(i \in L\). Let \(w^R \in [w_L, w_H]\) denote a reference wage. This wage represents the ‘normal’ or ‘typical’ wage that the agents expect to obtain by participating in the tournament. Agent \(i\)'s expected utility for any choice of effort is given by:

\[
Eu_i(e_i) = \Pr(i \in W)[w_H + \beta(w_H - w^R)] + \Pr(i \in L)[w_L + \alpha(w_L - w^R)] - c(e_i).
\] (1)

The function \(c(\cdot)\) denotes the cost of effort and is assumed to be strictly increasing and convex. Equation (1) captures the idea that in addition to monetary incentives, agents are also motivated by their relative standing within the group. That is, they enjoy earning a wage
that they consider high with respect to some standard, and dislike being below that threshold. The parameters $\alpha$ and $\beta$, $\alpha, \beta > 0$, measure the importance of interpersonal comparisons and, more specifically, how much workers dislike being losers and enjoy being winners, respectively.\footnote{Note that in the inequity-aversion literature instead it is assumed that $\beta \leq 0$. Although in some settings inequity aversion is clearly relevant, it is our belief that in competitive scenarios like promotion tournaments the assumption that $\beta > 0$ is an equally good, if not better, description of actual behavior. However, more empirical/experimental work is probably needed to evaluate this claim.}

In general, $\alpha$ is assumed to be greater than $\beta$ since there is very strong evidence that losses generally resonate more strongly than gains; furthermore, $\alpha > \beta$ also captures in an important way the idea of risk aversion. However, most of the results presented in the following analysis do not rely on this assumption. When they do, this will be explicitly stated.

At the optimum, the effort level of agent $i$ solves

$$\frac{\partial \Pr(i \in W)}{\partial e_i}[(w_H - w_L) + \alpha(w^R - w_L) + \beta(w_H - w^R)] = c'(e_i).$$

(2)

A key element of the analysis is to identify plausible values for the reference wage $w^R$. This is perhaps the thorniest issue in the theory of reference-dependent preferences, and several possibilities have been explored. Here we assume that the ‘typical’ wage $w^R$ is that received by the modal group of workers. To be ‘typical’, however, we also require the modal group to be considerably larger than the other group(s). We formalise this idea by assuming that the number of people belonging to the modal group exceeds the number of individuals in the other group by $D \geq 1$. Thus, for instance, in a tournament where there is only 1 winner and $N - 1$ losers and $N - 2 \geq D$, the reference wage $w^R$ equals $w_L$. Clearly this definition leaves open the possibility that no such a group exists, for instance when a group is composed by only two workers. In that case, one may assume that no group is typical and the reference wage is the mean (weighted or not) of the two groups wages. However, since the plausibility of any such assumption would be questionable, we will restrict attention to situations where a particular wage stands out as a possible reference point.

The focus of the analysis is on the profits of a firm (principal) which takes into account the interpersonal comparison component of workers’ preferences and on how the profits of this ‘behavioural’ firm compare with the profits of the ‘traditional’ firm where workers are only motivated by absolute monetary rewards. The firm’s profits are the sum of the agents’ outputs multiplied by the marginal value product $V$ which is the price of one unit of output $x$. We formulate the principal’s problem as maximising expected profits for a given level of effort (symmetric in equilibrium) subject to the incentive and participation constraints:

$$\max_{e_1, \ldots, e_N} E \Pi = V \sum_i e_i - N_W w_H - N_L w_L$$

s.t. $e_i \in \arg \max E(u_i | e_1, \ldots, e_N)$

$E(u_i | e_1, \ldots, e_N) \geq \pi$
Assuming a symmetric solution and that the participation constraints bind at the optimum,\(^6\) the participation constraint yields

\[
\frac{NW}{N}[w_H + \beta (w_H - w^R)] + \frac{N - NW}{N}[w_L + \alpha (w_L - w^R)] - c(e) = \overline{\pi}.
\]

This can be rewritten as

\[
NW w_H + (N - NW)w_L = N[\overline{\pi} + c(e)] - NW \beta (w_H - w^R) - (N - NW)\alpha (w_L - w^R). \tag{3}
\]

The right-hand side of (3) denotes the total wage bill paid at the optimum by the firm in which workers care about status. Thus expected profits can be written as

\[
E \Pi = V N e - N[\overline{\pi} + c(e)] + NW \beta (w_H - w^R) + (N - NW)\alpha (w_L - w^R). \tag{4}
\]

The firm will choose \(w_H\) and \(w_L\) to maximise (4) subject to (2). Under the same assumptions, the problem that the traditional firm faces is

\[
\max_e E \Pi^T = V N e - N[\overline{\pi} + c(e)] \tag{5a}
\]

subject to

\[
\partial \Pr(i \in W) \over \partial e_i (w_H - w_L) = c'(e_i). \tag{5b}
\]

we consider different scenarios to study the implications of interpersonal comparisons on the wage policy of a profit-maximising firm.

2.1 Tournament Scenarios

2.1.1 Promotion-based Tournaments

In promotion-based tournaments, few workers receive the high wage ("are promoted") and the majority stays at the same level and receives the low wage. Formally, \(N_L - NW \geq D\), \(NW \geq 1\), and therefore \(w^R = w_L\). Equations (4) and (2) reduce to

\[
\max_e E \Pi^{PB} = V N e - N[\overline{\pi} + c(e)] + NW \beta (w_H - w_L) \tag{6a}
\]

subject to

\[
\partial \Pr(i \in W) \over \partial e_i (1 + \beta)(w_H - w_L) = c'(e_i). \tag{6b}
\]

The problem of the organisation can therefore be expressed more concisely as

\[
\max_e V N e - N[\overline{\pi} + c(e)] + NW \frac{\beta}{1 + \beta} \left( \frac{\partial \Pr(i \in W)}{\partial e_i} \right)^{-1} c'(e). \tag{6c}
\]

\(^6\)This is always the case when limited liability is not an issue. Indeed, the firm can always set either \(w_L\) or \(w_H\) so that the participation constraint is met with equality (and profits are maximised) and set the wage gap \(\Delta w\) to provide incentives for effort.
Note that (6c) requires an additional condition on the third derivative of the cost function to guarantee that the first-order conditions actually yield a maximum (for instance, \(c'''(e) = 0\) would suffice).\(^7\) We will always assume that the second-order conditions of the programs are satisfied.

### 2.1.2 Demotion-based Tournaments

In this type of tournaments, many workers receive the high prize and only few workers the low one. We interpret this scenario as a tournament where the worst performing workers are demoted and the majority remains at the same level of the corporate hierarchy. Formally, \(N_W - N_L \geq D, N_L \geq 1\), and therefore \(w^R = w_H\). The problem of the firm in that case is

\[
\max_e E\Pi^{DB} = V Ne - N[\bar{\pi} + c(e)] - (N - N_W)\alpha(w_H - w_L) \tag{7a}
\]

subject to

\[
\frac{\partial \Pr(i \in W)}{\partial e_i} (1 + \alpha)(w_H - w_L) = c'(e_i) \tag{7b}
\]

or, more concisely stated

\[
\max_e V Ne - N[\bar{\pi} + c(e)] - (N - N_W) \frac{\alpha}{1 + \alpha} \left( \frac{\partial \Pr(i \in W)}{\partial e_i} \right)^{-1} c'(e). \tag{7c}
\]

The comparison between the above equations can help explain prevalent and salient features of organisational incentive systems.

### 2.2 Discussion

#### 2.2.1 Wage Compression

In equilibrium, for any given effort level, the optimal wage gap \(\Delta w = w_H - w_L\) is smaller when workers care about interpersonal comparisons compared to the case in which workers only care about monetary payoffs. In fact, (2) implies

\[
\Delta w = \frac{c'(e_i)}{\frac{\partial \Pr(i \in W)}{\partial e_i}} - \left[ \alpha(w^R - w_L) + \beta(w_H - w^R) \right]
\]

which, for all given effort levels, is always smaller than the wage gap for the traditional firm in (5b). Clearly, the more status-seeking workers are (i.e. the larger \(\alpha\) and \(\beta\) are), the smaller will be the optimal wage gap.

\(^7\) \(\frac{\partial \Pr(i \in W)}{\partial e_i}\) is independent of the overall level of \(e\) since all workers choose the same level of effort. This is most evident in equation (14) of Section 3 where we explicitly state \(\frac{\partial \Pr(i \in W)}{\partial e_i}\).
2.2.2 Asymmetric Effects of Rewards and Punishments

Note that (5a) and (5b) imply that if workers do not care about interpersonal comparisons, it is irrelevant to the firm whether tournaments are promotion-based or demotion-based. The introduction of interpersonal comparison however drastically changes this result. In fact, equations (5a), (6a) and (7a) make clear that for any effort level we have the following relation

\[ E_{PB} > E_{T} > E_{DB}. \]

The result offers a possible explanation for “the prevalence of compensation systems that reward ‘winners’ without explicitly identifying ‘losers’” (Baker et al. 1988, p. 607). If workers care about their relative status, compensation schemes such as demotion-based tournaments make the identity of the losers salient. This saliency generates ‘status’-costs that the firm must eventually bear since the participation constraint is binding. In contrast, promotion-based tournaments make the identity of the winners salient. This makes it easier to meet the participation constraint since the firm can also use non-monetary rewards to provide incentives.

In the promotion-based tournament case, the firm profits from a workforce that has stronger status concerns. In such a setting the potentially harmful effects of behindness-aversion, measured by the parameter \( \alpha \), on the participation constraint have been eliminated. For any given equilibrium effort level, the principal’s wage costs \( W_B(e) = N_W w_H + N_L w_L \) vary negatively with \( \beta \):

\[ \frac{dW_B}{d\beta} = -N_W \frac{\Delta w}{1 + \beta} < 0. \]

Consequently, the principal’s profits increase if individuals are more status-seeking. The opposite result obtains for the demotion-based tournaments where one obtains

\[ \frac{dW_B}{d\alpha} = (N - N_W) \frac{\Delta w}{1 + \alpha} > 0. \]

2.2.3 Optimal Effort Levels and Rat Races

The traditional firm’s maximisation problem given by (5a) and (5b) yields the familiar condition equating marginal benefit to marginal cost of effort

\[ V = c'(e^T). \]

In contrast, by maximising (6c) one obtains the condition

\[ V + \frac{N_W}{N} \frac{\beta}{1 + \beta} \left( \frac{\partial \Pr(i \in W)}{\partial e_i} \right)^{-1} c''(e^{PB}) = c'(e^{PB}). \]

The level of effort that solves the program given by (7c) is

\[ V - \frac{N - N_W}{N} \frac{\alpha}{1 + \alpha} \left( \frac{\partial \Pr(i \in W)}{\partial e_i} \right)^{-1} c''(e^{DB}) = c'(e^{DB}). \]
Since the cost function is convex, it follows that \( e^{PB} > e^T > e^{DB} \), that is, the equilibrium level of effort in the promotion-based tournament is higher than in the orthodox case without status concerns. We refer to this situation as a *rat race* for relative status. Conversely, effort is lower in the demotion-based tournament. As status concerns become more important the differences between these effort levels become more pronounced.

### 3 Demotions, Promotions and Performance Evaluations

#### 3.1 Model Outline and Solution

In this section we extend the framework studied above by introducing an additional, intermediate hierarchical level. Attention will be focused on tournaments where few workers are promoted or demoted, and the majority remains at the same hierarchical level. The expected utility of the worker is

\[
\Pr(i \in W)[w_H + \beta(w_H - w_M)] + \Pr(i \in M)w_M + \Pr(i \in L)[w_L + \alpha(w_L - w_M)] - c(e_i). \tag{8}
\]

where \( i \in M \) means that worker \( i \) belongs to the intermediate, neutral class and \( w_M \in [w_L, w_H] \) is the wage the worker receives in that case. Implicit in (8) is the assumption that staying at the same level is the ‘typical’ condition: promotion and demotions are relatively rare events. More formally, we assume that \( N_M - N_W - N_L \geq D \) where \( N_M = N - N_W - N_L \) is the number of worker who stay at the same level.\(^8\) Let \( \Delta w^+ = w_H - w_M \) and \( \Delta w^- = w_M - w_L \). In a symmetric equilibrium, each worker’s effort solves

\[
\frac{\partial \Pr(i \in W)}{\partial e_i}(1 + \beta)\Delta w^+ - \frac{\partial \Pr(i \in L)}{\partial e_i}(1 + \alpha)\Delta w^- = c'(e). \tag{9}
\]

Moreover, if the participation constraint binds, the following equation holds\(^9\)

\[
N_W w_H + N_L w_L + N_M w_M = N [\pi + c(e)] - N_W \beta(w_H - w_M) + N_L \alpha(w_M - w_L). \tag{10}
\]

The firm maximises expected profits by solving the following program

\[
\max_{\Delta w^+, \Delta w^-, N_W, N_L} E\Pi^B = VNe - N[\pi + c(e)] + N_W \beta \Delta w^+ - N_L \alpha \Delta w^- \tag{11}
\]

subject to (9). Assuming \( \frac{\partial \Pr(i \in W)}{\partial e_i} > 0 \) (which is essentially equivalent to \( N_W > 0 \)), program (11) can be rewritten as

\[
\max_{\Delta w^+, \Delta w^-, N_W, N_L} \quad VNe - N[\pi + c(e)] + \frac{N_W \beta c'(e)}{\frac{\partial \Pr(i \in W)}{\partial e_i}(1 + \beta)} + \left( \frac{\partial \Pr(i \in L)}{\partial e_i}(1 + \alpha) \right) \frac{N_W \beta - N_L \alpha}{\frac{\partial \Pr(i \in W)}{\partial e_i}(1 + \beta)} \Delta w^- \tag{12}
\]

\(^8\)An alternative assumption would be to impose the requirement that \( N_M \geq \max\{N_W + D, N_L + D\} \). The results of this section would not change under this specification.

\(^9\)In this case the firm can set \( w_M \) so that the participation constraint is met with equality, thereby maximising profits. Incentives to exert effort are provided by appropriately choosing the wage gaps \( \Delta w^+ \) and \( \Delta w^- \).
A crucial feature of this program is that, for all \( e \) and \( N_W \), it achieves its maximum when either \( \Delta w^- = 0 \) or \( N_L = 0 \), or both. Note, in fact, that the relation \( \frac{\partial \Pr(i \in L)}{\partial e_i} \leq 0 \) will hold with equality if \( N_L = 0 \). In other words, (12) implies that it is optimal to minimise the number of demotions. Thus, program (12) reduces to (6c), with the additional feature that now also \( N_W \) is a choice variable.

We now ask when it is also optimal to maximise promotions (i.e. \( N_W \), subject, of course, to the reference-wage constraints, \( N_M - N_W - N_L \geq D \). This is not trivial because, although it is clear that there is a beneficial effect in increasing the number of people who enjoy higher status, there is also the effect on the incentives to exert effort to take into account. Indeed, it is clear that there is a beneficial effect in increasing the number of people who enjoy higher incentives.

One can rewrite equation (13) as follows. Let agent \( i \)'s probability of placing \( j \)'th from the bottom be denoted by \( \Pr(i \rightarrow j) \). This probability is given by

\[
\Pr(i \rightarrow j) = \frac{(N-1)!}{(N-j)!(j-1)!} \int F(e_i - e^* + \varepsilon_i)^{j-1} [1 - F(e_i - e^* + \varepsilon_i)]^{N-j} f(\varepsilon_i) d\varepsilon_i
\]

where \( e^* \) now denotes the equilibrium effort level of the other agents and the integral is taken on the entire support of \( \varepsilon_i \). Clearly,

\[
\Pr(i \in W \mid N_W) = \sum_{j=N-N_W}^{N} \Pr(i \rightarrow j).
\]

At the symmetric equilibrium (\( e_i = e^* \)), agent \( i \) has an equal chance of ending up in any position, \( \Pr(i \rightarrow j) = \frac{1}{N} \), and therefore \( \Pr(i \in W \mid N_W) = \frac{N}{N} \). By working harder at the symmetric equilibrium there is an increased (or decreased) chance of coming in at position \( j \)

\[
\frac{\partial \Pr(i \rightarrow j)}{\partial e_i} = \frac{(N-1)!}{(N-j)!(j-1)!} \int F(\varepsilon_i)^{j-2} [1 - F(\varepsilon_i)]^{N-j-1} f(\varepsilon_i)^2 \times \{ (j-1)[1 - F(\varepsilon_i)] - (N-j)F(\varepsilon_i) \} d\varepsilon_i.
\]

It follows that

\[
\frac{\partial \Pr(i \in W \mid N_W)}{\partial e_i} = \sum_{j=N-N_W+1}^{N} \frac{(N-1)!}{(N-j)!(j-1)!} \int F(\varepsilon_i)^{j-2} [1 - F(\varepsilon_i)]^{N-j-1} f(\varepsilon_i)^2 \times \{ (j-1)[1 - F(\varepsilon_i)] - (N-j)F(\varepsilon_i) \} d\varepsilon_i. \tag{14}
\]

\(^{10}\)Note that this is only a sufficient but not a necessary condition.
In particular, note that $\frac{\partial \Pr(i \in W)}{\partial e_i}$ is independent of the level of effort. Simple algebra shows that (13) can be rewritten as

$$\frac{1}{N_W} \frac{\partial \Pr(i \in W | N_W)}{\partial e_i} > \frac{\partial \Pr(i \in W | N_W + 1)}{\partial e_i} - \frac{\partial \Pr(i \in W | N_W)}{\partial e_i}$$

or, equivalently, as

$$\frac{1}{N_W} \sum_{j=N-W+1}^{N} \frac{\partial \Pr(i \to j)}{\partial e_i} > \frac{\partial \Pr(i \to N-N_W)}{\partial e_i}. \tag{15}$$

The above result has the following interpretation. Increasing the number of prizes is optimal if the change in the probability of ending up at position $N_W + 1$ from the top ($N - N_W$ from the bottom) is less than the average change in the probability of coming in at one of the first $N_W$ positions. This is a likely outcome if the density function of the error term is 'well-behaved'. In fact, as long as more effort increases the probability of ending up in a higher position more than in a lower position, the result holds. In particular, it is possible to show (Nalebuff and Stiglitz 1983) that the result holds for the uniform density $f(\varepsilon) = a, \varepsilon \in [-A, A]$ and the exponential density $f(\varepsilon) = e^{\varepsilon-1}, \varepsilon \in [-\infty, 1]$.\footnote{These are the distributions for which it is relatively easy to compute $\frac{\partial \Pr(i \in W | N_W)}{\partial e}$. Nalebuff and Stiglitz (1983) use condition (15) to show the optimality of restricting prize structures to few distinct prizes.}

In the first case, the net increase in the probability of moving in and moving out of any interior position by working harder is zero. The increase in the probability of 'winning' is concentrated in the top position, and the corresponding decrease in the probability of not winning in the bottom one. Under the exponential distribution, increasing effort equally reduces the probability of ending up in any position but the top one. Clearly, in both cases (15) holds.

Under condition (15), the following conditions completely characterise the optimal structure of the tournament:

- $w_M$ such that (10) holds,
- either $\Delta w^-$ or $N_L = 0$, or both,
- (neglecting integer problems) $N_W = \frac{N-D}{2}$ and $N_M = \frac{N+D}{2}$,
- $w_H$ is chosen such that it solves (12),
- $w^R = w_M$.

It is worthwhile stressing that restricting attention to optimal tournaments in which $w^R = w_L$ is with no loss of generality. Obviously, a tournament in which $w^R = w_H$ would not be optimal. In Appendix B, we show that when $w^R = w_L$, tournaments with only two prizes are optimal. This is reminiscent of Lazear’s (1989) result that in multi-prize tournaments only the
highest and the lowest prize matter and all the other prizes can be chosen arbitrarily. Thus, optimal tournaments where $w^R = w_M$ and those where $w^R = w_L$ are equivalent.

The result that it is often optimal to maximise the number of winners (subject to the reference-wage constraint) and minimise the number of losers can help explain some puzzling stylised facts of compensation policies.

3.2 Discussion

3.2.1 Promotions, Demotions and the Proliferation of Job Titles

As argued before, the interpretation of tournaments as promotion contests would identify the winners and losers as workers who are promoted or demoted from a given hierarchical level, respectively. In that light, the finding that it is optimal to minimise the number of losers and maximise that of winners fits well with the empirical regularity that demotions are very rare, whereas promotions are much more common as documented by Baker, Gibbs and Holmstrom (1994a,b). In addition, if there is a ‘technologically right’ number of promotions and demotions, which may correspond to the expected number of workers needed at higher and lower levels, the result implies that demotions must be kept to the minimum (in case by hiring new workers), whereas promotions can be expanded above that limit. Thus, the result can explain the proliferation of job titles observed in some industries. Indeed, on the issue of job title proliferation, Baron and Bielby (1986) and Doeringer and Piore (1971) argue that many vertical and horizontal distinctions among jobs reflect customs and status gradations and are not simply distinctions in technical duties.

3.2.2 Performance Evaluations

According to Baker et al. (1988), even more surprising than employers’ reluctance to fire or penalise employees is the fact that “supervisors tend to assign uniform performance ratings and tend not to assign poor performance ratings”. Moreover, the human resource management literature emphasises the phenomenon of clustered performance ratings. For example, Torrington et al. (2002) report that “there is a tendency for raters to settle on the mid-point of the scale, either through lack of knowledge of the appraisee, lack of ability to discriminate, lack of confidence or through a desire not to be too hard on appraisees” (p. 301).

Table I shows both the above-mentioned central tendency as well as a clear predominance of ‘positive’ ratings. While many different explanations may play a role here, our model also predicts an asymmetry of reward and punishments around the modal classes. This asymmetry is particularly salient in the larger of the two firms, firm A. The vast majority (about three quarters) of the managers in this company are rated as “good”. That class can therefore be considered typical and used as the reference point for interpersonal comparison. Note the

\[12\] In particular, a crucial factor in our view is the fact that supervisors often lack the appropriate incentives to give poor performance ratings.
Table I
Salary Premiums Associated With Performance Ratings, and Frequency Distribution of Performance Ratings for 7,629 Managers in Two Large Manufacturing Firms

<table>
<thead>
<tr>
<th>Performance Rating</th>
<th>Salary Premium Relative to</th>
<th>Percent of Sample Receiving</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lowest Performance Rating</td>
<td>Performance Rating</td>
</tr>
<tr>
<td>Company A (4,788 managers):</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Not Acceptable</td>
<td>-0-</td>
<td>0.2</td>
</tr>
<tr>
<td>Acceptable</td>
<td>1.4</td>
<td>5.3</td>
</tr>
<tr>
<td>Good</td>
<td>5.3</td>
<td>74.3</td>
</tr>
<tr>
<td>Outstanding</td>
<td>7.8</td>
<td>20.2</td>
</tr>
<tr>
<td>Company B (2,841 managers):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unacceptable</td>
<td>-0-</td>
<td>-0-</td>
</tr>
<tr>
<td>Minimum Acceptable</td>
<td>-0-</td>
<td>-0-</td>
</tr>
<tr>
<td>Satisfactory</td>
<td>-0-</td>
<td>1.2</td>
</tr>
<tr>
<td>Good</td>
<td>1.8</td>
<td>36.6</td>
</tr>
<tr>
<td>Superior</td>
<td>3.6</td>
<td>58.4</td>
</tr>
<tr>
<td>Excellent</td>
<td>6.2</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Source: Medoff and Abraham (1980), Table I and II; Baker et al. (1988), Table I

... asymmetry between high and low ratings relative to this modal class: slightly less than 80% of the remaining managers (about 20% of the total) are ranked as “outstanding”, whereas only (approximately) 20% and 0.7% of these are ranked “acceptable” and “not acceptable”, respectively. The data of firm B is much less clear-cut, but if one takes the classes “good” and “superior” as typical (note that the difference in salary premium between these two classes is not very large, only 1.8%), one can see that there are more than three times as many workers in the top class than in the bottom three classes together.

The positive connotation of the typical class is remarkable, but difficult to explain in terms of relative comparisons. Related to this issue is the question why below average classes are present at all since according to the proposed theory these should be redundant. One possibility is that they are necessary in order to justify necessary demotions (e.g. when a prior promotion turned out to be grossly inefficient). However, it is interesting to note that firm B eventually decided to eliminate the lowest two classes.

13 Torrington et al (2002) argue that the lack of an entirely clear central tendency of performance appraisals towards a typical class, such as in the data for firm B, is often intentionally caused by the design of performance rating schemes that use an even-numbered (usually six-point) scale.
4 Ambition

In this section we introduce a possible source of heterogeneity among workers, ambition. We model ambition by assuming that more ambitious agents set more demanding standards for themselves. In a promotion-based tournament, for instance, we would assume \( w_{Amb}^R = w_H \) for ambitious workers and \( w_{Amb}^L = w_L \) for their unambitious colleagues. Note that ambitious workers need not necessarily be over-optimistic in the sense that, if ambition makes them work harder, they might have correct expectations about their relative ranking most of the time. This is true a fortiori if, as in the following section, ability and ambition are positively correlated. As mentioned in Section 1, one important way of interpreting systematic differences in ambition, on which we will elaborate further, is in terms of differences in reference groups. For instance, it is plausible that women may have lower reference wages because they compare themselves not with workers in general (male and female) but only other female workers. Analogously, high ability workers may have higher reference wages because they compare themselves only with other talented individuals.

To keep the model tractable and focus on ambition in isolation, however, we concentrate on two-workers-two-jobs tournaments and thereby neglect the role that expectations, or the normal state of affairs, have in determining the reference point. Since the firm now has no means to manipulate the reference point, the emphasis of the discussion will shift from profits to equilibrium effort levels and each worker’s probability of winning the contest. Finally, we also assume that each agent’s type is public information among the agents but is unknown to the principal and that the density \( g(\cdot) \) of the error difference is symmetric around its strict maximum, 0 (e.g. a normal density). The first-order conditions for the two types are

\[
\frac{\partial \Pr(i \in W)}{\partial e_{Amb}} (1 + \alpha) \Delta w = c'(e_{Amb})
\]

\[
\frac{\partial \Pr(i \in W)}{\partial e_{Amb}} (1 + \beta) \Delta w = c'(e_{Amb})
\]

The symmetry of the density function implies \( \frac{\partial \Pr(i \in W)}{\partial e_{Amb}} = g(e_{Amb} - e_{Amb}) = g(-e_{Amb} + e_{Amb}) = \frac{\partial \Pr(i \in W)}{\partial e_{Amb}} \). Thus, if as a large body of experimental evidence suggests losses resonate more than gains, \( \alpha > \beta \), the ambitious worker exerts more effort in equilibrium than his unambitious counterpart. In particular, Heath et al. (1999) anticipated this result in their empirical analysis: “if goals act as reference points,” they claim “then loss aversion implies that people who are below their goal by \( x \) units will perceive their current performance as a loss relative to their goal; thus they will work harder to increase their performance by a given increment than

\[14\]When wage compensation is given and identical for ex-ante asymmetric workers, in general, at least one of them will receive some surplus. Which of the agents will receive (more) surplus depends on the relative strength of the status parameters \( \alpha \) and \( \beta \): stronger status considerations make the worker more likely to win the tournament, but also hurt him more when he loses the contest. Hence, it is not clear that more ambitious types will generally end up with a higher surplus than their unambitious colleagues.
people who are above their goal by $x$ units” (p. 85). Clearly, this result also implies that for an equal treatment in a mixed tournament ambitious workers have a higher chance of winning the tournament.

Using this simple model we can rank effort levels for any given wage spread by examining the first-order conditions distinguishing between segregated and mixed tournaments for ambitious and unambitious workers. The first-order conditions for the following settings with equal prize spreads are as follows:

A. segregated ($S$) tournament, only ambitious ($Amb$) workers ($N_{Amb} = 2, N_{-Amb} = 0$)

$$g(0)(1 + \alpha)\Delta w = c'(e_{Amb}^S)$$

B. single ($S$) tournament, only unambitious workers ($-Amb$)

$$g(0)(1 + \beta)\Delta w = c'(e_{Amb}^S)$$

C. mixed ($M$) tournament, ambitious workers ($Amb$)

$$g(e_{Amb}^M - e_{-Amb}^M)(1 + \alpha)\Delta w = c'(e_{Amb}^M)$$

D. mixed ($M$) tournament, unambitious workers ($-Amb$)

$$g(e_{Amb}^M - e_{-Amb}^M)(1 + \beta)\Delta w = c'(e_{-Amb}^M)$$

As can be seen from the equations above, the ambitious workers always exert more effort in equilibrium than their unambitious colleagues. Furthermore, both types exert at least as much effort in a segregated tournament than in a mixed setting. By the properties of the distribution function, the density is largest when all agents exert the same level of effort which is only the case in the segregated tournaments. In summary, the above first order conditions yield the following relationship

$$e_{Amb}^S > e_{Amb}^M, e_{-Amb}^S > e_{-Amb}^M.$$

4.1 Discussion

4.1.1 Gender Differences

It is often argued that women earn less thanobservationally similar men because of discrimination. Furthermore, promotion rates of women are lower and women are scarcely found at high level positions (Lazear and Rosen 1990). There is ample empirical and experimental evidence of marked gender differences with respect to attitudes towards ambition and status. In particular, women are generally less ambitious and less status-seeking than men (Clark 1997; Falk and Knell 2004; Huberman et al. 2004).

Allocations across genders of high profile jobs remain largely favorable to men and are a major factor in the gender gap in earnings (Lazear 1989; Lazear and Rosen 1990). Gneezy et
al. (2003) and Antonovics, Arcidiacono and Walsh (2003) present experimental evidence that these differentials might be linked to gender differences in behaviour in competitive settings rather than due to discrimination, but do not give where this difference in behaviour stems from.

Our results can explain some of the seemingly surprising conclusions of Gneezy et al. (2003) who find marked gender differences in performance when comparing individuals’ responses to different types of incentives. Using experimental evidence they show that women may be less successful than men in mixed competitive environments (e.g. tournaments), even though they are able to perform as well as men in noncompetitive environments (e.g. piece-rate schemes) or segregated competitive environments.

4.1.2 Happiness

Despite the large and significant differences between men’s and women’s pay as well as the extensive evidence that women’s jobs are worse than men’s in terms of hiring and firing, job content, promotion opportunities and sexual harassment, women consistently report to be equally happy or even happier than men are, even controlling for all sorts of other influences on happiness. This is the so-called gender/happiness paradox. Among the most prominent explanations for this paradox is the one proposed by Clark and Oswald (1996) and Clark (1997). These authors test and confirm the hypothesis that utility depends on income relative to a ‘comparison’ or reference level. In their analysis workers’ reported satisfaction levels are shown to be inversely related to their comparison wage rates. Furthermore, they show that since women tend to compare themselves to other women and have lower reference wages and expectations, they report higher happiness scores.

To analyse this paradox in terms of the current two-workers-two-jobs model we investigate how a change in the reference wage affects the utility of workers. For this analysis, we allow $w_R$ to take any value between $w_L$ and $w_H$. As before, the utility function is given by

$$U_i(e_i, w^R_i) = \Pr(i \in W)[w_H + \beta(w_H - w^R_i)] + \Pr(i \in L)[w_L + \alpha(w_L - w^R_i)] - c(e_i)$$

The effect on equilibrium utility $U_i(e_i, w^R_i)$ of a rise in the reference wage $w^R_i$ is given by

$$\frac{dU_i(e_i, w^R_i)}{dw^R_i} = \frac{\partial U_i}{\partial e_i} \frac{de_i}{dw^R_i} + \frac{\partial U_i}{\partial e_j} \frac{de_j}{dw^R_i} + \frac{\partial U_i}{\partial w^R_i}. \tag{17}$$

At the optimum $\frac{\partial U_i}{\partial e_i} = 0$ and so we can eliminate the first term of equation (17). The third term of equation (17) is the direct effect on utility of a rise in the reference wage and is given by

$$\frac{\partial U_i}{\partial w^R_i} = -\Pr(i \in W)\beta - \Pr(i \in L)\alpha < 0.$$
This effect is negative since a rise in the reference wage reduces the utility gained if the agent wins and increases the loss in utility if the agent loses the tournament. The second term of equation (17) measures the strategic effect of a rise in ambition which Clark and Oswald (1996) do not mention. The sign of this effect is in general ambiguous. For instance, it can be easily proved that if the noise difference is assumed to be normally distributed, the strategic effect is negative if \( w_i^R > w_j^R \) and positive if \( w_i^R < w_j^R \).

All in all our results are in line with the happiness paradox documented by Clark and Oswald (1996) who find that individual happiness decreases with the reference wage. Yet, our model adds an additional complication in that the overall effect on utility of a change in the reference wage is also influenced by a strategic effect which may work in opposite directions in some cases.\(^{15}\)

5 Ambition and Ability

This section extends the model presented in Section 4 by allowing workers to differ not only in ambition but also in ability. Let the production technology be given by

\[
x_i = a_i + e_i + \varepsilon_i
\]

where \( a_i \) denotes agent \( i \)'s intrinsic ability and is public information among the agents but unknown to the principal. Note that ability enters additively in the production function so that the marginal return to effort does not vary with ability. For simplicity we assume that \( a_i \) can only take the values \( \{0, a\} \), that is agents can only be of low or high ability.

The firm runs a tournament between its two contestants. The case in which agents have symmetric ability levels has been analysed in Section 4, so here we focus on the settings in which agents have different ability levels. From the first-order conditions the following relations are obtained

A. 1 ambitious/high ability and 1 unambitious/low ability agent

\[
g(a + e_{Amb} - e_{-Amb})(1 + \alpha)\Delta w = c'(e_{Amb}) \tag{18a}
g(e_{-Amb} - e_{Amb} - a)(1 + \beta)\Delta w = c'(e_{-Amb}) \tag{18b}
\]

B. 1 ambitious/low ability and 1 unambitious/high ability agent

\[
g(e_{Amb} - e_{-Amb} - a)(1 + \alpha)\Delta w = c'(e_{Amb}) \tag{18c}
g(a + e_{-Amb} - e_{Amb})(1 + \beta)\Delta w = c'(e_{-Amb}) \tag{18d}
\]

Equations (18a) and (18b) characterise the solution when one agent is superior to his opponent in terms of both ambition and ability. Equations (18c) and (18d) instead describe the case

\(^{15}\)A third complication is that the first order condition may fail to determine the global maximum. Our analysis clearly would not apply in that case.
where the less able worker is the more ambitious than his more able opponent. Provided the
two agents are treated equally by the principal in the tournament, their respective probabilities
of winning the tournament are increasing in their level of ambition and their ability advantage.
For example, it could well be the case that \( G(e_{Amb} - e_{-Amb} - a) > \frac{1}{2} \) so that on average the
ambitious but less able worker will win the tournament more often than his unambitious but
more able colleague.

5.1 Discussion

5.1.1 Why are Tournament Promotion Systems So Common in Practice?

According to Baker et al. (1988), promotions in organisations serve two important and distinct
purposes. First, they are a way to match individuals to the job for which they are best suited.
Second, promotions “provide incentives for lower level employees who value the pay and prestige
associated with a higher rank in the organisation” (p. 599). The problem is that “tournament
promotions systems cannot in general match employees to the jobs for which they are best
suited” (p. 602), since performance in the current job might not be the best indicator of
performance in another job. For instance, law … rms promote the associates who are likely to
make the best partners, and not the highest performing associates. Many economists, including
Baker et al. (1988), therefore advocate the use of monetary incentives (bonuses) to elicit effort.
Promotion decisions, instead, should be made using all the relevant information, which usually
includes but not perfectly coincides with workers’ measurable output.

The fact that people’s utility is affected by interpersonal comparisons may help explain the
attractiveness of incentive systems based on tournaments relative to those based on bonuses.
First, since the results of a tournament are more visible than the bonus payments a worker is
able to obtain, interpersonal comparisons are easier to perform and hence are likely to matter
more in the former case. Thus, tournaments may be a cheaper way to provide incentives relative
to bonuses. Secondly when workers differ in their ability, the reference groups that they choose
might be different. In particular, more able workers may want to compare themselves to equally
able workers and not to low ability workers. Falk and Knell (2003) show that reference stan-
dards tend to increase in individuals’ abilities and that thus people tend to compare themselves
to similar others. In particular, their empirical results confirm that students with higher intel-
lectual abilities choose higher reference standards. This positive relationship between ambition
and ability makes the case given by equations (18a) and (18b) far more convincing than the
case characterised by equations (18c) and (18d), thereby mitigating the potential inefficiencies
of promoting people of lesser ability. Our findings suggest that the adverse consequences
of tournament promotion systems may be smaller than previously thought and thus the model
may help explain why tournaments are so widely used as a sorting mechanism.
5.1.2 Promotion Inefficiencies and Gender Quotas

The argument above, however, does not exclude the possibility of systematic inefficiencies in promotion decisions. In particular, more ambitious or more status-seeking individuals might end up being promoted instead of more able co-workers when ambition and talent are not perfectly correlated. As mentioned in Section 4, women, on average, are less status-seeking than men and hence despite high ability may fail to scale the corporate ladder. As a result, women are underrepresented in top management positions. For example, according to the FTSE Female Index Report (2003) only 8.6% of UK FTSE 100 directors are women.

However, at the same time the FTSE Female Index Report (2003) claims that there is “strong relationship between gender diversity, good corporate governance and high corporate performance”. These observations and guidelines are in accord with the previous argument according to which the failure of women to make it to the top of the organisation derives from a promotion inefficiency. Even if they are at least as able as their male competitors they might not be promoted due to a lack of status concerns. If the marginal return to ability is higher in more elevated positions, companies would benefit from promoting the most able worker. However, promoting individuals on the basis of ability alone blunts the sharp incentives for effort provided by a tournament described in this paper. Firms may therefore be willing to incur some degree of promotion inefficiency in favour of sharp tournament incentives.

Favorable promotion quotas for female (or other less ambitious) workers might improve the trade-off between incentive provision at the low level and efficient promotion decisions. In fact, the 2003 FTSE Female Index Report “urges companies to develop women executive directors as well as appoint more female non-executive directors”. By introducing such quotas, female workers are effectively placed into a single-sex tournament. This measure has two advantages. First, as shown in Section 4, women will compete more intensely in a segregated tournament than in a mixed tournament. This is because in such a setting women compare themselves with other women rather than with men who have higher expectations or reference wages. Second, separating workers into different types of tournaments reduces the noise stemming from differences with respect to ambition. By effectively eliminating heterogeneity with respect to status concerns the winners of the tournament will also be, on average, the most able workers. This second advantage of gender quotas means that the promotion inefficiency that arises when workers differ with respect to both ability and status concerns is effectively eliminated.

6 Concluding Remarks

There is a large body of evidence suggesting that there is a human tendency to compare outcomes with others and that individuals are often more motivated by the desire to avoid losses than they are by maximising gains. The contribution of this paper has been to introduce status concerns in an otherwise standard tournament model and show that this incorporation can shed light on several puzzling features of actual compensation systems and human resources
practices.

However, a number of interesting questions are left open. For instance, we do not analyse how the principal’s profit varies between the different forms of tournaments (mixed or segregated and ambitious or unambitious) discussed in Section 4. A fuller analysis of these issues would entail analysing which type of workers receive surplus in an equal-treatment tournament and determining the conditions under which workers choose to self-select into the tournaments designed for them.

A second issue, briefly mentioned in the Introduction, is that our discussion mainly relates to promotion as the outcome of the tournament. This was motivated by our belief that, in practice, the results of a tournament are more visible than higher pay, and that therefore interpersonal comparisons are more likely to matter in the former case. However, it is not obvious that this needs to be the case if, for instance, the firm announces all the wages publicly. How firms should choose between different degrees of disclosure and secrecy in wage setting in order to maximize workers’ incentives to work hard is, in our view, a promising avenue for future research.

We also hope that our analytical framework may inspire further empirical work. For instance, the reason for the prevalence of compensation systems that reward ‘winners’ without explicitly identifying ‘losers’ is that firms have to compensate their employees for the disutility they receive in case they end up in a low-status position. Formally, the result is a direct implication of the assumption that the participation constraint is binding. This suggests that when the participation constraint is not binding, for instance when participation is compulsory, nasty ‘reward’ schemes may be optimal. Indeed, one may argue that if the military service is compulsory, harassing the ‘weakest’ soldiers may be the most efficient way to motivate them. One empirical implication of our theory would therefore be that the progressive moves of most Central European countries from compulsory military service to a professional army should be accompanied by an increase in the soldiers’ pay and an intensification of efforts to curb harassment.

Finally, the definition of status that we adopted in this paper implicitly assumes that individual welfare is determined partly by an individual’s absolute wage and partly by the individual’s relative wage. However, range-frequency theory predicts that well-being is also gained from the individual’s ranked position of a wage within a comparison set (Brown, Gardner, Oswald and Qian 2004). The analysis therefore fails to take into account the status derived from the ordinal ranking in an organisation. Similarly, the model ignores the possibility of symbolic differentiation; status may be linked to characteristics other than relative wages. For example, the management and organisational behaviour literature stresses the concept of recognition given to workers. Despite their seemingly limited incentive effects “employee of the month” schemes that award the outstanding with nothing more than symbolic recognition, are a common feature of many organisations. By using symbolic differentiation, firms may also be able to reduce the adverse effects of inefficient promotion choices discussed in Section 5. We leave the exploration of these issues to discussion in future research.
A Weighted Average as Reference Point \( w^R = \frac{1}{N} \sum_{j=1}^{N} w_j \)

In this Appendix we show that some of the results of the analysis of this paper do not hold if the reference point is assumed to be the average of all the wages received by the agents, i.e. \( w^R = \frac{1}{N} \sum_{j=1}^{N} w_j \). Nevertheless, the specific issue we highlight in this analysis - namely status concerns always having a negative impact on profits for plausible parameter values - disappears when a specific form of non-linearity of the utility function, sometimes called the “peanuts effect” and first discussed by Markowitz (1952), is assumed.

If the expected utility of a worker is given as before by

\[
Eu_i(e_i) = \Pr(i \in W)[w_H + \beta(w_H - w^R)] + \Pr(i \in L)[w_L + \alpha(w_L - w^R)] - c(e_i),
\]

the first-order conditions yield

\[
\frac{\partial \Pr(i \in W)}{\partial e_i} [(w_H - w_L) + \alpha(w^R - w_L) + \beta(w_H - w^R)] = \ell'(e_i)
\]

which, given the new assumptions about the reference point, simplify to

\[
\frac{\partial \Pr(i \in W)}{\partial e_i} \Delta w \left(1 + \alpha \frac{N_W}{N} + \beta \frac{N - N_W}{N} \right) = \ell'(e_i).
\]

Noting that in equilibrium all agents exert the same level of effort and assuming that the participation constraint binds, allows us to rewrite the expected profits of the organisation as

\[
E\Pi = VNe - N [\pi + c(e)] - (\alpha - \beta) \Delta w \frac{(N - N_W)N_W}{N}.
\]

Since \( \alpha > \beta \), the expected profits generated under this assumption about the reference point will always be lower than those received by a traditional firm (for which \( \alpha = \beta = 0 \)) regardless of whether the firm frames the incentive scheme as a promotion-based or a demotion-based tournament. Now assume that the utility function exhibits the “peanuts effect”.

Specifically, let

\[
u_i(w_i) = \begin{cases} 
    w_i + \beta(w_i - w^R) & \text{if } w_i \geq w^R + b \\
    w_i & \text{if } w^R - b \leq w_i \leq w^R + b \\
    w_i + \alpha(w_i - w^R) & \text{if } w_i \leq w^R - b
\end{cases}
\]

where \( b > 0 \) is assumed to be small. Intuitively, worker \( i \)'s utility does not change appreciably when his realised wage \( w_i \) happens to be very close to the reference wage.\(^{16}\) The flavour of this assumption is reminiscent of results in the mental accounting and marketing literature, for instance the fact in many situations sellers and fund raisers choose to frame an annual fee as 'pennies-a-day'. (Note that this fact is inconsistent with prospect theory, which predicts

\(^{16}\) Results would not change if we allowed for a constant but sufficiently small loss whenever \( i \)'s wage is slightly below the reference point.
that losses should be integrated if the loss function is convex.) To reconcile this behaviour with prospect theory, Thaler (2000) argues that small sums may be framed “in the petty cash category, so when the expense is framed this way it tends to be compared to other items that are not booked [i.e. recorded in the accounting system]. In contrast, a $100 membership is large enough that it will surely be booked and posted [i.e. assigned to a specific expense account], possibly running into binding budget constraints in the charitable giving category” (p. 255). Like our definition of the reference wage in the present paper, Thaler’s discussion of mental accounting and the budgeting process relies heavily on categorisation or labeling.

Note that since \( w^R = \frac{N_w}{N} w_H + \frac{N-N_w}{N} w_L \), then \( w_L \in [w^R - b, w^R] \) if \( \frac{N_w}{N} \) is small enough (as in promotion-based tournaments). Clearly therefore there are promotion-based tournaments which yield greater profits than demotion-based tournaments or even tournaments in which workers are not status-seekers since only the former can generate positive ego-rents, provided of course that the reference wage is sufficiently close to the low prize (i.e., \( \frac{N_w}{N} \) is small enough).

## B Optimal Three-prize Tournament where \( w^R = w_L \)

In a three-prize tournament where \( w^R = w_L \) the expected utility of a worker is

\[
\Pr(i \in W)[w_H + \beta(w_H - w_L)] + \Pr(i \in M)[w_M + \beta(w_M - w_L)] + \Pr(i \in L)w_L - c(e_i).
\]

As in the main text, let \( \Delta w^+ = w_H - w_M \) and \( \Delta w^- = w_M - w_L \). At a symmetric equilibrium, each worker’s effort solves

\[
\frac{\partial \Pr(i \in W)}{\partial e_i} (1 + \beta) \Delta w^+ - \frac{\partial \Pr(i \in L)}{\partial e_i} (1 + \beta) \Delta w^- = c'(e).
\]

Furthermore if the participation constraint binds we have

\[
N_W w_H + N_L w_L + N_M w_M = N [\bar{u} + c(e)] - N_W \beta(w_H - w_L) - N_M \beta(w_M - w_L).
\]

The firm maximises expected profits by solving the following program

\[
\Delta w^+, \Delta w^-, N_W, N_L \max \tilde{E} = V Ne - N [\bar{u} + c(e)] + N_W \beta \Delta w^+ + (N_W + N_M) \beta \Delta w^-
\]

subject to the incentive compatibility constraint. Assuming \( \frac{\partial \Pr(i \in W)}{\partial e_i} > 0 \), the above program can be rewritten as

\[
\max_{e, \Delta w^-, N_W, N_L} V Ne - N [\bar{u} + c(e)] + \frac{\beta}{1 + \beta} \frac{N_W c'(e)}{\partial e_i} + \left( \frac{\partial \Pr(i \in L)}{\partial e_i} + 1 \right) N_W + N_M \beta \Delta w^-.
\]

Consider the expression in the square brackets and denote it by \( A \). There are three possibilities: either \( A > 0 \), \( A < 0 \) or \( A = 0 \). In the first case, it is clearly optimal to maximise \( \Delta w^- \) subject to \( w_M \in [w_L, w_H] \) and therefore \( w_M = w_H \). If \( A < 0 \), instead, it is optimal to set \( w_M = w_L \). If \( A = 0 \) any \( w_M \) would do, and in particular \( w_M = w_H \) or \( w_M = w_L \). Thus, without loss of generality, we can restrict attention to a tournament with only two prizes. Therefore, at the optimum, the firm’s profits in the scenario where \( w^R = w_M \) are the same as those in the present scenario where \( w^R = w_L \).
References


[23] *FTSE Female Index Report* (2003), Cranfield University School of Management.


