Optimal Contract for Ambitious Team Workers

Kong-Pin Chen
Research Center for Humanities and Social Sciences
Academia Sinica
and Department of Economics
National Taiwan University

Tsung-Sheng Tsai *
Department of Economics
National Tsing Hua University

Keywords: Relative performance, Absolute performance, Promotion tournament, Wage contract

JEL classification: J3

* Correspondence: Tsung-Sheng Tsai, Department of Economics, National Tsing Hua University, Hsinchu 300, Taiwan. Tel: (03) 573-1197; Fax: (03) 572-2746; E-mail: tstsai@mx.nthu.edu.tw. We thank Chun-Lei Yang for useful comments. Special thanks go to Yoram Weiss for very useful discussion and detailed comments. We are also very grateful to two anonymous referees for their helpful comments.
ABSTRACT

This paper investigates the optimal compensation scheme for workers in a team who value not only absolute but also relative incomes. A worker is said to be more ambitious if his utility places more weight on relative income. In this case, the firm can exploit the worker’s preference for relative comparison to design a compensation scheme which induces the same effort level with lower cost. Under the optimal compensation scheme, the workers’ wages are shown to depend on relative performance, and exhibit wage compression. More importantly, even if the production technology calls for absolute performance evaluation in the traditional principal-agent model, the optimal wage structure still relies on relative performance. Finally, in contrast to past literature, worker heterogeneity is shown to reduce the firm’s profit.
1. INTRODUCTION

Theoretical works have long postulated that the utility of an economic agent might depend not only on individual income, but also on relative income.\(^1\) This theoretical postulation is also supported by empirical findings. For example, using Dutch cross-section and panel data, respectively, Kapteyn et al. (1980) and van de Stadt et al. (1985) show that individual utility functions depend on relative, rather than absolute, incomes. Easterlin (1974, 1995) shows that self-reported happiness has a positive correlation with income across individuals within a country, but this correlation does not increase when the country grows richer over time. This is interpreted as evidence that individual well-being depends on relative, rather than absolute, level of income. Clark and Oswald (1996) uses British House Panel Survey data to show that workers’ reported levels of well-being are at best weakly correlated with their absolute income alone, but are significantly negatively correlated with comparison incomes.\(^2\) A recently paper by Luttmer (2004) uses National Survey of Families and Households (NSFH) panel data matched to Public Use Microdata Areas (PUMA) to test whether individuals feel worse when others around them earn more. With careful specifications that filter out individual fixed effects and local factors, he convincingly argues that individual utility should depend on relative in addition to absolute earning. This tendency to compare one’s income with others is also confirmed by recent experimental data. For example, Zizzo and Oswald (2001) design an experiment in which the subjects can reduce others’ money at their own costs. They find that about 65% of the subjects are on average willing to sacrifice 25 cents in order to reduce their peer’s incomes by one dollar.\(^3\)

In the context of a firm, this implies that the utility of a worker depends not only on his own wage, but also on the wages of his co-workers as well. Specifically, a worker derives an additional utility when he earns a higher wage than his co-workers, and suffers a disutility when earning a lower one. Since this represents the desire to do better than his co-workers, we will call it the worker’s “ambition.”

---

\(^1\) See, for example, Boskin and Sheshinski (1978) on redistributive taxation, Frank (1985) on status seeking and Abel (2005) on taxation to achieve optimal balanced growth path.

\(^2\) Two measures of comparison income are used in that paper: Standard mincer earnings and peers’ wages.

\(^3\) For extensive discussion on the literature, see Luttmer (2004) and Fershtman et al. (2003b).
If a worker is ambitious, the marginal utility for income will be greater, regardless of whether he earns more or less than his co-workers. This means that, given a fixed wage schedule, the utility gap between high and low wages will be greater, which in turn implies that in order to implement the same effort level, the firm only needs to set a smaller wage gap. In particular, the firm can lower the wage for a worker who produces higher output (and is paid more than this co-worker) and at the same time retains the same incentives. There are three consequences for this. First, the wage cost of the firm is lower in order to provide the same incentives to the workers. Consequently, the profit of the firm will increase. In other words, the firm can increase its profit by hiring ambitious workers. Second, there is wage compression in the sense that the wage gap between workers with high and low outputs will be narrowed. Third and most importantly, even when the information structure does not require relative compensation in the traditional principal-agent model, the optimal wage structure here is still strictly a relative performance evaluation. This is of great interest because it has been a puzzle in the personnel economics literature; given the drawbacks of relative performance evaluation, why is it still so widely used? There can be many explanations for it, and this paper offers an answer from a new perspective. The reason is not to infer a worker’s effort level more precisely by observing other workers’ output (as is the rationale in the standard principal-agent model), but to exploit the workers’ concern for relative income and to save wage cost.

Recently there are many (mostly experimental) papers which show, or take as assumption, that people are strongly influenced by other-regarding preferences. In the context of the principal-agent relation, the basic ingredient of this literature is to view the relation between a worker (or an agent) and a firm (or a principal) as a gift-giving game. It then considers how the workers’ concern for fairness and the motive of reciprocity between the firm and the workers affect the firm’s wage policy and the workers’ effort decisions. (For a survey see Fehr and Schmidt, 2003.) Therefore, the main focus of this literature is the one-principal-one-agent relationship without formal contractual obligations, and it departs from the traditional principal-agent model in two aspects. First, the agent’s utility depends not only on how much he is paid, but also on how he is paid. Second, the level of wage directly influences the effort level of the agent, instead of through the incentive compatibility constraints.

In contrast to the gift-giving game model, our model focuses on the formal con-

---

4 See, for example, Lazear (1989) and Chen (2003).
5 DeVaro (2002) finds empirical evidence that promotion is strongly related to relative performance.
tractual relationship. We depart from the traditional principal-agent model only in that the worker’s utility is assumed to depend on other workers’ income as well. Moreover, unlike the gift-giving literature, the workers’ preferences depend on their co-workers’ wages, instead of on their concerns of fairness with the principal. Specifically, we consider a one-principal-many-agent model in which the psychological utility arises from the relation between the workers, instead of between the principal and the agent. The only papers we are aware of that also use the one-principal-many-agent framework in this regard are Charness and Kuhn (2004) and Fershtman et al. (2003b). In the former, however, a worker’s effort level is assumed to be an increasing function of wage, so that the basic relation between the principal and the agent is still a gift-giving game. In particular, in their model the incentive compatibility constraint is assumed away, so that the only way to increase effort is to raise wage instead of redesigning the output-contingent wage structure (as in our model). Moreover, in their model the workers differ in productivity, while in ours the workers differ in the degree to which they care about other workers’ incomes.

Fershtman et al. (2003b) is perhaps the closest to our paper in the literature. If restricted to the context of a single firm, theirs is essentially a one-principal-many-agent model in which a worker is one of two types. He either cares about relative income, or does not care at all. Thus they are concerned not with how the degree to which the workers care about relative income affects the wage policy, but with the form of optimal contract, given that a certain proportion of workers care about relative income. They derive the optimal effort decision of the workers and the equilibrium wage structure of the firm. There are three main differences between their model and ours. First, in their model the worker’s wage is assumed to be linear in his own output. This implicitly assumes away relative performance evaluation. In particular, a worker’s wage is independent of other workers’ wages or outputs. Our model, in contrast, investigates how a worker’s concern for relative income affects the firm’s decision to link his wage to other workers’ performances. Put differently, our model endogenizes the choice between absolute and relative performance evaluations. As we will show, even if production technologies are independent (so that relative performance evaluation is never needed on informational grounds), the optimal wage structure is still a relative performance evaluation. Second, they assume the firm is able to offer different contracts

---

6 This model also differs from the literature that assumes altruistic utility function (see, e.g., Rotemberg, 1994, and the survey of Fehr and Schmidt, 2003, for details) in that this literature assumes a positive relation between utility and other agents’ wages, while in our model, there is a negative relation.
to workers with different characteristics, and the first-best result can thus be obtained. That means there is essentially neither moral hazard nor adverse selection concern in their model, and the firm’s objective is virtually to maximize joint surplus. We assume that the firm is not allowed to offer different contracts to different workers, and the only way the workers are paid by different compensation schemes is through self-selection. Our model thus combines the moral hazard and adverse selection considerations of contracts. Third, their model is much more general regarding the workers’ effort choices. They assume a continuous effort level and an explicit relation between effort and output so that the optimal effort level can be endogenously derived. Our model, on the other hand, allows only binary choice of effort and binary output level.

2. THE MODEL

Consider a model that has the same informational structure as in Che and Yoo (2001). A principal (firm) hires two workers, each to perform a project. The worker has a binary choice of either high \((k = 1)\) or low \((k = 0)\) effort level, at a cost of \(ke\), for \(e > 0\).\(^7\) Effort levels are not observable to the firm.

After effort is exerted, the principal learns of each worker \(i\)’s performance (output) \(y_i\), which can be either high \((y_i = 1)\) or low \((y_i = 0)\). The probability distribution of \(y_i\) depends on worker \(i\)’s effort level and a common shock. A favorable shock occurs with probability \(\sigma \in (0, 1)\), under which both workers will have high performance for sure, regardless of their effort levels. On the other hand, an unfavorable shock occurs with probability \(1 - \sigma\), under which the probability for a worker to obtain a high output is \(q_1\) when he supplies high effort level and \(q_0\) if low, where \(1 > q_1 > q_0 \geq 0\). Since with probability \(\sigma\) both workers get high output, their performance is correlated. It follows that each worker obtains high output with probability \(\sigma + (1 - \sigma)q_k\) and the low output with probability \((1 - \sigma)(1 - q_k)\), given his effort decision \(k\). We assume that high output is so valuable to the firm that it intends to implement high effort for both workers.\(^8\)

Wages can only be contingent on each agent’s output. Let \(w^i_{mn}\) denote the wage for worker \(i\) when his output is \(y_i = m\) and his co-worker’s output is \(y_j = n\). The

---

\(^7\) This setup is for simplification. We believe our analysis can still be valid if there is a continuum of choices of effort.

\(^8\) This means that high output is large in comparison to low output and effort cost \(e\).
wage structure of the firm is thus \((w^1, w^2)\), where \(w^i = (w^i_{11}, w^i_{10}, w^i_{01}, w^i_{00})\) is the contract for worker \(i\). Wages are subject to limit liability (LL) so that the principal cannot offer negative wages to the workers, i.e., \(w^i_{mn} \geq 0\) for all \(i, m\) and \(n\).

All parties are risk neutral. The main assumption of our model is that a worker derives utility (suffers disutility) from receiving a higher (lower) wage than his co-worker. There are many possible ways to model this. The simplest specification to capture this idea is to assume that worker \(i\) has a utility function 

\[
u_i(x_i) = x_i + \alpha_i(x_i - x_j) = (1 + \alpha_i)x_i - \alpha_ix_j.\tag{1}
\]

Here, \(\alpha_i\) is a measure of how strongly worker \(i\)’s utility depends on his income relative to his co-worker. It can be seen from (1) that the marginal utility for income is greater if the worker cares about relative income. Moreover, since the value of \(\alpha_i\) measures how strong is worker \(i\)’s desire to do better than his co-worker, we will use the value of \(\alpha_i\) to measure how “ambitious” worker \(i\) is. Note that if \(\alpha_i > 1\), then worker \(i\) cares even more about relative income than own income. Given contracts \(w^1\) and \(w^2\) and the effort decisions \(k \in \{0, 1\}\) by worker \(i\) and \(l \in \{0, 1\}\) by worker \(j\), worker \(i\)’s expected utility from wage is:

\[
u_i(k, l; w^1, w^2) = [\sigma + (1 - \sigma)g_kq_l][(1 + \alpha_i)w^i_{11} - \alpha_iw^i_{11}] \\
+ (1 - \sigma)g_k(1 - q_l)[(1 + \alpha_i)w^i_{10} - \alpha_iw^i_{01}] \\
+ (1 - \sigma)(1 - g_k)q_l[(1 + \alpha_i)w^i_{01} - \alpha_iw^i_{10}] \\
+ (1 - \sigma)(1 - g_k)(1 - q_l)[(1 + \alpha_i)w^i_{00} - \alpha_iw^i_{00}].\tag{2}
\]

Assume that a worker’s reservation utility is 0. Contracts \(w^1\) and \(w^2\) induce both agents to supply high effort level in a Nash equilibrium if and only if

\[
(\text{IC}_1) \quad u_1(1, 1; w^1, w^2) - e \geq u_1(0, 1; w^1, w^2),
\]

and

---

9 This is exactly the specification used in Fershtman et al. (2003a, 2003b). Another possible specification is the ERC model (Bolton and Ockenfels, 2000) in which worker \(i\)’s utility can be written as \(u_i(w_i, w_i/w_j)\). This utility is inconvenient to use in our context since, as can be seen in our later discussion, the wage of a worker is often zero.
An important issue concerning the wage policy is that in the traditional principal-agent model, (LL) constraint is enough to ensure that a worker’s expected utility is non-negative. As a result, no individually rational (IR) constraint is needed, as (LL) implies (IR). In our setup, however, this is not true. The reason is that a worker’s utility is decreasing in his co-worker’s wage. Even if he is always paid a positive wage, if his co-worker is paid a wage far greater than his own, his utility can be negative. Consequently, even if (LL) is imposed, we still need the (IR) to ensure that every worker’s expected utility attains its reservation value. This means that

\[(\text{IR}_i) \quad u_i(1, 1; w^1, w^2) - e \geq 0 \quad \text{for all } i.\]

The firm’s objective is thus to design a wage contract to minimize the expected cost of inducing both workers to supply high effort level, subject to limited liability and individually rational constraints. That is,

\[
\min_{w^1, w^2} \quad c(1, 1; w^1, w^2) \equiv [\sigma + (1 - \sigma)q_1^2](w_{11}^1 + w_{11}^2) + (1 - \sigma)q_1(1 - q_1)(w_{10}^1 + w_{01}^2) + (1 - \sigma)q_1(1 - q_1)(w_{01}^1 + w_{10}^2) + (1 - \sigma)(1 - q_1)^2(w_{00}^1 + w_{00}^2),
\]

subject to \[(\text{IC}_1), (\text{IC}_2), (\text{IR}_1), (\text{IR}_2), w^1 \geq 0 \quad \text{and} \quad w^2 \geq 0.\] (3)

In the case when the workers do not care about relative incomes, i.e., when \(\alpha_1 = \alpha_2 = 0\), Che and Yoo (2001) have shown that the optimal wage scheme is a relative performance evaluation (thereafter RPE), in which \(w^1 = w^2 = w^S = (0, w_{10}^S, 0, 0)\), where \(w_{10}^S = [e/(1 - \sigma)(q_1 - q_0)(1 - q_1)]\). The worker’s expected utility is thus \(u^S(1, 1, w^S, w^S) = q_1e/(q_1 - q_0)\). Under this scheme, a worker is paid a positive wage only if his performance is strictly better than his co-worker. This wage structure is actually equivalent to a promotion tournament. The essence of a promotion tournament is to provide incentives by setting up a prize (generally a promotion) that is awarded to a worker only if he out-performs all others. The workers are motivated to supply effort simply because of the prospect of becoming the winner.\(^{10}\) Under the RPE wage

\(^{10}\) See Lazear and Rosen (1981) and Green and Stokey (1983) for details on promotion tournaments.
structure derived above, a worker is paid a positive wage only if his output is high and his co-worker’s output is low. This is exactly the way a promotion tournament motivates workers, in which \( w^i_{10} \) is a worker’s prize in wage when he is promoted. Because of this, in the case when one worker has high output and the other low, we will also call the former the “winner,” and the latter the “loser.”

3. THE CASE OF HOMOGENEOUS WORKERS

This section considers the homogeneous case in which the workers are equally ambitious, that is, \( \alpha_1 = \alpha_2 = \alpha > 0 \). From (1) we can see that when the workers are ambitious, there is stronger incentive for workers to provide high effort level, since the marginal utility of wage is higher. Because of this, in order to induce high effort level, the winner does not need to be paid as high as when he is not ambitious. We formally prove this in the following proposition.

**Proposition 1** When workers are equally ambitious, then:

1. If \( \alpha \leq q_0(1 - q_1)/(q_1 - q_0) \), the optimal wage contract is \( w^1 = w^2 = w^R = (0, w^R_{10}, 0, 0) \), where \( w^R_{10} = e/[(1 - \sigma)(q_1 - q_0)(1 + \alpha - q_1)] \).
2. If \( \alpha > q_0(1 - q_1)/(q_1 - q_0) \), there exists an optimal contract \( w^1 = w^2 = w^R = (0, w^{RR}_{10}, 0, 0) \), where \( w^{RR}_{10} = e/[(1 - \sigma)q_1(1 - q_1)] \).

**Proof:** The incentive constraints for the workers, \((IC_1)\) and \((IC_2)\), can be re-written as

\[
(IC_1) \quad q_1 [(1 + \alpha)w^2_{11} - \alpha w^2_{11}] + (1 - q_1) [(1 + \alpha)w^1_{10} - \alpha w^2_{01}] - q_1 [(1 + \alpha)w^1_{01} - \alpha w^2_{01}] - (1 - q_1) [(1 + \alpha)w^1_{00} - \alpha w^2_{00}] \geq \frac{e}{(1 - \sigma)(q_1 - q_0)}, \tag{4}
\]

and

\[
(IC_2) \quad q_1 [(1 + \alpha)w^2_{11} - \alpha w^2_{11}] - q_1 [(1 + \alpha)w^2_{01} - \alpha w^1_{10}] + (1 - q_1) [(1 + \alpha)w^2_{01} - \alpha w^1_{10}] - (1 - q_1) [(1 + \alpha)w^0_{00} - \alpha w^0_{00}] \geq \frac{e}{(1 - \sigma)(q_1 - q_0)}. \tag{5}
\]
In a symmetric equilibrium, \( w_{1n}^1 = w_{2n}^2 \equiv w_{mn} \),\(^{11}\) and the firm’s minimization problem becomes (where superscripts denoting the identity of workers are omitted):

\[
\min_w 2[(\sigma + (1-\sigma)q_1^2)w_{11} + (1-\sigma)q_1(1-q_1)(w_{10} + w_{01}) + (1-\sigma)(1-q_1)^2w_{00}],
\]

s.t. \( (IC) \quad q_1 w_{11} + (1+\alpha-q_1)w_{10} - (\alpha+q_1)w_{01} - (1-q_1)w_{00} \geq \frac{e}{(1-\sigma)(q_1-q_0)}, \)

\( (IR) \quad [\sigma + (1-\sigma)q_1^2]w_{11} + (1-\sigma)q_1(1-q_1)(w_{10} + w_{01}) + (1-\sigma)(1-q_1)^2w_{00} \geq e, \)

and \( w \geq 0. \) (6)

Note that the firm’s target function is simply two times the left-hand-side of (IR). We discuss two cases separately. First, suppose (IR) is not binding. In this case, because the coefficients for \( w_{01} \) and \( w_{00} \) are positive in the objective function but negative in the constraint (IC), and because of the limit liability constraints, it is optimal for both \( w_{01} \) and \( w_{00} \) to be zero. Moreover, since

\[
\frac{q_1}{1+\alpha-q_1} < \frac{\sigma + (1-\sigma)q_1^2}{(1-\sigma)q_1(1-q_1)},
\]

we can decrease the value of the objective function by increasing (decreasing) the value of \( w_{10} \) (\( w_{11} \)) while keeping (IC) intact. This means that it is also optimal to set \( w_{11} = 0 \). Therefore, the optimal solution is \( w^R = (0, w_{10}^R, 0, 0) \), where \( w_{10}^R \) is set to satisfy constraint (IC). As a result,

\[
w_{10}^R = \frac{e}{(1-\sigma)(q_1-q_0)(1+\alpha-q_1)}.
\]

Under \( w^R \) a worker’s utility can be easily computed to be \([q_1(1-q_1)e]/[(q_1-q_0)(1+\alpha-q_1)]\). That (IR) is not binding requires \( \alpha < q_0(1-q_1)/(q_1-q_0) \).

It remains to show that the wage scheme \( w^R \) indeed induces both workers to provide high effort level in a Nash equilibrium. It can be easily seen that

\[
u_i(1,k; w^R) - e = u_i(0,k; w^R), \quad k = \{0, 1\}.
\]

\(^{11}\) By using the simplex algorithm method, we can show that this solution is indeed the unique one when workers are equally ambitious.
This means that a worker prefers to provide high effort regardless of the other worker’s effort level. Consequently, the unique Nash equilibrium is for both workers to provide high effort.

In the second case when (IR) is binding, we know from previous discussion it must be when \( \alpha > q_0(1 - q_1)/(q_1 - q_0) \). In this case since the target function is only two times the left-hand-side of (IR), the firm’s profit is always \( 2e \). Consequently, any \( w^R = (w_{11}, w_{10}, w_{01}, w_{00}) \) satisfying

\[
\begin{align*}
    w_{mn} &\geq 0, \\
    q_1w_{11} + (1 + \alpha - q_1)w_{10} - (\alpha + q_1)w_{01} - (1 - q_1)w_{00} &\geq \frac{e}{(1 - \alpha)(q_1 - q_0)}, \\
    [\sigma + (1 - \sigma)q_1^2]w_{11} + (1 - \alpha)q_1(1 - q_1)(w_{10} + w_{01}) + (1 - \sigma)(1 - q_1)^2w_{00} &= e,
\end{align*}
\]

can be a solution. In particular, let \( w^R = (0, w^{RR}_{10}, 0, 0) \), where \( w^{RR}_{10} = \frac{e}{[(1 - \sigma)q_1(1 - q_1)]} \) is a solution.

Proposition 1 shows that when workers are equally ambitious, then as in Che and Yoo (2001), the optimal contract is a tournament: the firm pays a worker positive wage if and only if he out-performs his co-worker. However, we now have \( w^R_{10} < w^S_{10} \) and \( w^{RR}_{10} < w^{SS}_{10} \). That is, although the nature of the compensation scheme is still a tournament, the winner is now paid less. In the language of promotion tournament literature, this means that there is a wage compression. The intuition is clear: since the winner derives additional utility by being paid more than his co-worker, the firm can provide the same incentive by paying the winner less. This result is consistent with several models that imply wage compression from different perspectives.\(^{12}\)

The following proposition shows how the presence of ambition affects the firm’s profit and the workers’ utilities.

**Proposition 2** As long as \( \alpha \leq q_0(1 - q_1)/(q_1 - q_0) \), the profit for the firm is increasing, and the expected utility for each worker is decreasing, in the degree of workers’ ambition. If \( \alpha > q_0(1 - q_1)/(q_1 - q_0) \), the profit of the firm and the utility of the worker are fixed.

---

Proof: The firm’s wage cost is

\[
2(1 - \sigma)q_1(1 - q_1)w_{10}^R = \begin{cases} 
2q_1(1 - q_1)e \left( \frac{1}{1 + \alpha - q_1} \right) & \text{if } \alpha \leq \frac{q_0(1 - q_1)}{(q_1 - q_0)}, \\
2e, & \text{if } \alpha > \frac{q_0(1 - q_1)}{(q_1 - q_0)}. 
\end{cases}
\]

This means that the profit of the firm is increasing as \(\alpha\) increases if \(\alpha \leq q_0(1 - q_1)/(q_1 - q_0)\), and fixed otherwise. The expected utility for an ambitious worker is

\[
u_i(1, 1; w^R) = (1 - \sigma)q_1(1 - q_1)(1 + \alpha)w_{10}^R - \alpha w_{10}^R = (1 - \sigma)q_1(1 - q_1)w_{10}^R
\]

\[
= \frac{q_1(1 - q_1)e}{(q_1 - q_0)(1 + \alpha - q_1)} < \frac{q_1 e}{(q_1 - q_0)} = u^S,
\]

if \(\alpha \leq q_0(1 - q_1)/(q_1 - q_0)\), and is \(e < q_1 e/(q_1 - q_0) = u^S\) if \(\alpha > q_0(1 - q_1)/(q_1 - q_0)\). Obviously, utility is a decreasing function of \(\alpha\) if \(\alpha \leq q_0(1 - q_1)/(q_1 - q_0)\), and fixed otherwise.

The intuition for Propositions 1 and 2 is as follows. Since the workers derive utility from making more money than their peer, the firm does not need to pay the winner as much in order to induce the same effort level. Put differently, recall that the optimal contract is essentially a promotion tournament, and in that case what matters for the workers’ incentives is the utility gap between the winner and the loser, rather than the absolute values of utility. When the workers are ambitious, since both the gain in utility (when they are paid more than their co-workers) and the loss in utility (when they are paid less) are magnified, the utility gap between the winner and the loser is widened. The firm can therefore exploit this fact and lower the value of \(w_{10}\) (while keeping other components of wage at 0) to retain incentive for the workers to exert high effort. As a result, the firm induces the same effort level with lower cost. The workers have lower utilities because although they are still paid a zero wage as a loser, the utility of being a loser is now negative.

The proof of Proposition 2 also shows that, since both \(w_{10}^R\) and \(u_i(1, 1; w^R)\) are decreasing in \(\alpha\), the profit of the firm is increasing in \(\alpha\) and the utility of the worker is decreasing in \(\alpha\). That is, the firm can make more profit by hiring workers who are more ambitious. The advantage of hiring ambitious workers, however, has a limit. When \(\alpha\) is sufficiently large, the firm exploits the workers’ ambition so much so that
their utilities hit the reservation value. Beyond that, ambition will not help the firm. The worker’s utility is always 0, and the firm’s profit always $2e$. This can be seen clearly from the fact that as $\alpha$ increases, the wage paid to winner, $w_{10}^R$, is decreasing, until $\alpha$ equals \(q_0(1 - q_1)/(q_1 - q_0)\), beyond which the winner’s wage, $w_{10}^{RR}$, remains a constant.

The fact that the firm’s profit increases with $\alpha$ is in contrast to Charness and Kuhn (2004), who show that the firm’s profit stays the same when the workers care about relative incomes. The reason for the difference in result is that in their model the workers have different productivities, and that effort is assumed to be a linear function of own wage. Since (by their assumption) the workers are concerned equally about relative income, the firm needs to pay more (less) to worker with higher (lower) productivity in order to equalize the workers’ marginal productivities. This is only a reshuffling of wage bills, and has no effect on the total wage cost.

Of special interest is the extreme case when $\sigma = 0$; that is, when the realization of the workers’ performances are independently distributed. In that case, it is well known in the principal-agent literature that the optimal wage for a worker is independent of the performance of others, since other workers’ performance offers no information in inferring his effort level. That is, the optimal wage for each worker is an absolute performance evaluation (hereafter APE). If this is also the case in our model, then since a worker’s wage depends only on his own performance, the contract for worker $i$ reduces to $w^i = (w^1_i, w^0_i)$, where $w^j_i$ is the wage for $i$ when his output is $j$.

Due to symmetry, the firm’s minimization problem becomes:

$$\min_{w_1, w_0} 2[q_1 w_1 + (1 - q_1)w_0],$$

s.t. \((1 + \alpha)(q_1 - q_0)(w_1 - w_0) \geq e, \quad \text{and} \quad w_1 \geq 0, \quad w_0 \geq 0. \tag{7}\) $$

The following result is immediate.

**Lemma 1** If the realizations of the workers’ performances are independently distributed, then the APE optimal contract for each worker is $w^A = (w^A_1, 0)$, where

---

13 In their model, if the worker’s effort level responds to wage incentives only when his wage is lower than that of his co-worker (i.e. $\alpha > 0$ only when a worker’s wage is lower than his co-worker’s), then in even more contrast to our result, the firm is worse off.

14 See, for example, Mookherjee (1984).
The following proposition shows that the APE contract derived in Lemma 1 is actually not optimal, even if the information structure calls for an APE in the traditional principal-agent model.

**Proposition 3** In the case when workers are homogeneous and there is no correlation in performances, the profit for the firm is higher under the optimal RPE than under the optimal APE.

**Proof:** According to Lemma 1, the firm’s expected wage cost under the optimal APE is

\[ c(1, 1; w^A) = 2q_1w_1 = \frac{2q_1e}{(q_1 - q_0)(1 + \alpha)}. \]

From Proposition 2, the firm’s cost when \( \sigma = 0 \) is

\[
c(1, 1; w^R) = 2q_1(1 - q_1)w^R_{10} = \frac{2q_1(1 - q_1)e}{(q_1 - q_0)(1 + \alpha - q_1)} < \frac{2q_1e}{(q_1 - q_0)(1 + \alpha)}
\]

That is, the firm can induce both workers to exert high effort at a lower cost under an RPE.

Proposition 3 is of great interest, since it offers an explanation for why the firm will still use relative performance evaluation, when a worker’s performance provides no information in inferring the effort level of any other worker. As is well known in the principal-agent literature, the main reason why a worker’s wage will depend on other workers’ performances is that if there is a common shock that affects the performance of all workers in the same way, then an RPE can filter out this common shock, both to infer the workers’ effort more precisely, and to reduce the workers’ wage risks caused by the shock. The optimal contract is thus a relative performance evaluation.\(^\text{15}\) Absent this consideration, an RPE only brings in disturbance to a worker’s wage, and should not be used, as is shown in Lemma 1. However, relative-performance-based promotion is so commonly used that it seems to go beyond the consideration of screening

\(^{15}\) See, for example, Mookherjee (1984).
out risks. Here we offer an explanation which does not depend on consideration of reducing risks. It comes from the workers’ desire to do better than their co-workers. If an ambitious worker’s wage is higher than his co-worker, he is willing to accept a lower wage in order to exert the same effort. The firm can thus save wage cost by artificially creating a wage structure that depends on RPE. This is achieved by setting the worker’s wage to be zero if he is a loser (this comes from the limited liability constraint), and pays a winner only enough to induce high effort. Since the worker derives additional utility when he is a winner, a lower wage for high performers suffices to provide enough utility gap in order to induce high effort.

4. THE CASE OF HETEROGENEOUS WORKERS

This section considers the case when the workers do not have the same degree of ambition. Without loss of generality, we assume that worker 1 is more ambitious than worker 2, that is, $\alpha_1 > \alpha_2$. Although the workers are heterogeneous in ambition, the firm might have difficulty in offering different wage contracts to different workers. There are two possible reasons for this. First, the firm might not know which worker is more ambitious, and therefore cannot offer wage contracts based on the worker’s characteristics. Second, and perhaps more importantly, the firm might be legally forbidden from discriminating between the workers. This is particularly so in our model because all workers have the same productivity, and are induced to exert high effort. If, say, worker 1 is paid less than 2 when both have exerted high efforts and produce high outputs, then the firm risks being sued by the workers. As a result, there are two possible types of contracts that the firm can offer. The first is to offer a uniform contract to both types of worker. That is, although the workers have different degrees of ambition, the firm does not try to distinguish between them. We will call this a pooling contract. In the second kind of contract, the firm offers a menu of two contracts to discriminate between two workers. However, following the adverse selection literature, the workers choose the contracts they prefer, rather than being forced by the firm. Since the workers will reveal their own types by choosing different contracts, we call this a separating contract. The advantage of a separating contract is that it can exploit the difference in the worker’s degree of ambition. However, the disadvantage is that additional self-selection constraints have to be satisfied. The following proposition

---

16 See DeVaro (2002).
Proposition 4

(1) If $\alpha_2 \leq q_0(1 - q_1)/(q_1 - q_0)$, then
   (a) Under the optimal pooling contract, $w_1 = w_2 \equiv w^H = (0, w_{10}^H, 0, 0)$, where $w_{10}^H = e/[(1 - \sigma)(q_1 - q_0)(1 + \alpha_2 - q_1)]$.
   (b) Under the optimal separating contract, $w_1 = (0, \overline{w}_{10}^1, 0, \overline{w}_{00}^1)$, where $\overline{w}_{10}^1 = [(1 + \alpha_2)e]/[(1 - \sigma)(q_1 - q_0)(1 + \alpha_2 - q_1)(1 + \alpha_1)]$, $\overline{w}_{00}^1 = [q_1(\alpha_1 - \alpha_2)e]/[(1 - \sigma)(q_1 - q_0)(1 - q_1)(1 + \alpha_2 - q_1)(1 + \alpha_1)]$; and $w_2 = (0, \overline{w}_{10}^2, 0, 0)$, where $\overline{w}_{10}^2 = e/[(1 - \sigma)(q_1 - q_0)(1 + \alpha_2 - q_1)]$.

(2) If $\alpha_2 > q_0(1 - q_1)/(q_1 - q_0)$, then
   (a) There exists an optimal pooling contract with $w_1 = w_2 \equiv w^{HH} = (0, w_{10}^{HH}, 0, 0)$, where $w_{10}^{HH} = e/[(1 - \sigma)q_1(1 - q_1)]$.
   (b) There exists an optimal separating contract with $w_1 = (0, \overline{w}_{10}^1, 0, \overline{w}_{00}^1)$, where $\overline{w}_{10}^1 = [(1 - q_0)e]/[(1 - \sigma)(q_1 - q_0)(1 - q_1)(1 + \alpha_1)]$, $\overline{w}_{00}^1 = [(\alpha_1 q_1 - q_0 - q_0(1 - q_1))e]/[(1 - \sigma)(q_1 - q_0)(1 - q_1)^2(1 + \alpha_1)]$; and $w_2 = (0, \overline{w}_{10}^2, 0, 0)$, where $\overline{w}_{10}^2 = e/[(1 - \sigma)q_1(1 - q_1)]$.

Proof: The proof is available on the webpage http://mx.nthu.edu.tw/~tstsi.

The optimal pooling contract, as in the case of homogeneous workers, is an RPE. Note that although the workers have different degrees of ambition, they have exactly the same utilities. This can be seen by plugging $w^H$ into the utility functions in (2). What needs to be explained is then why $w^H$ depends only on $\alpha_2$, and not on $\alpha_1$. The reason for this is as follows. Since worker 1 is more ambitious, he derives more utility for being a winner, and thus only requires a lower wage when he is a winner than worker 2 to derive the same utility. So the value of wage paid to the winner, $w_{10}^i$, is binding only for worker 2 (and is thus a function of $\alpha_2$), who requires a higher wage as a winner to attain the same utility. By the same argument, the lower bound for the loser’s wage, $w_{01}^i$, should be binding for worker 1, who suffers more as a loser. However, the limited liability requirement limits its bound as 0, which is independent of $\alpha_1$. Consequently, the value of $w^H$ is a function of $\alpha_2$ only, and wage structure is thus independent of $\alpha_1$.

A direct comparison between $w^H$ and $w^R$ shows that as long as $\alpha \geq \alpha_2$, the

---

17 This is a reasonable restriction because worker 2 is the least ambitious worker. In general, an adequate comparison between homogeneous and heterogeneous cases might be to assume that $\alpha$ is the mean value of $\alpha_1$ and $\alpha_2$, where the mean is taken over the relative composition of type 1 and type 2 workers. In this
Total wage cost for the firm will be higher, implying that the workers receive higher utilities in the heterogeneous case. The reason for this is that under a pooling contract, the firm is paying the workers as if both are the less ambitious worker, which requires higher wage as a winner. This means that although the firm still profits from hiring ambitious workers, its profit is lower than in the homogeneous case. In other words, the firm is worse off when it has a mixture of workforce with different degrees of ambition.\(^{18}\)

In the separating contract, the firm offers a menu of two wage contracts, each voluntarily selected by one type of worker. The benefit of a separating contract is that the firm can discriminate against the more ambitious worker by paying him a lower wage as a winner than worker 2. This can be seen from Proposition 4 that \(\overline{w_1^2} > \overline{w_1^1}\).

In order to prevent worker 1 from choosing \(w_2^2\), he must be compensated in some other contingency when he is not the winner. It turns out that he is paid a positive wage when both workers have low outputs, i.e., \(w_{10}^1 > 0\). The reason why it is \(w_{10}^1\), rather than \(w_{11}^1\) or \(w_{11}^1\) which positive is as follows. If \(w_{01}^1 > 0\), then worker 1 is given positive wage as a loser. This weakens his incentive in effort, and is thus less desirable than setting \(w_{11}^1 > 0\) or \(w_{10}^0 > 0\). Although raising the value of either \(w_{11}^1\) or \(w_{10}^1\) will increase the utility of worker 1, it will at the same time reduce worker 2’s utility, but by a different degree. A unit raise in \(w_{11}^1\) (resp. \(w_{10}^1\)) will increase the utility of worker 1 by \([\sigma + (1 - \sigma)q_1^2](1 + \alpha_1)\) (resp. \((1 - \sigma)(1 - q_1)^2(1 + \alpha_1)\)) but reduce the utility of worker 2 by \([\sigma + (1 - \sigma)q_1^2]\alpha_2\) (resp. \((1 - \sigma)(1 - q_1)^2\alpha_2\)). Simple comparative advantage comparison implies that it is less costly for the firm to raise the value of \(w_{10}^1\).\(^{19}\)

Although worker 1 is paid a positive wage when both have low outputs, the wage structure still resembles a relative performance evaluation. In particular, worker 2 is paid a positive wage only when he is a winner. Moreover, both \(\overline{w_1^1}\) and \(\overline{w_2^2}\) are increasing in \(\alpha_2\), and \(\overline{w_{10}^0}\) is decreasing in \(\alpha_2\), meaning that as the worker becomes more ambitious, the wage structure will more resemble a tournament. Similar to the case \(\alpha_1 > \alpha > \alpha_2\).

\(^{18}\) Fershtman et al. (2003b) also shows that wage cost will be higher under the heterogeneous case. However, since the workers exert more effort, the firm actually has higher output and, in contrast to our result, higher profit. We will discuss their results in more detail shortly.

\(^{19}\) In reality, it is not very unusual that an individual manager is still paid highly when the whole company (or team) performs bad. For example, excerpted from Buyupsid.com: information for stock investor, “In 2004, Blockbuster, a movie rental chain, increased chairman and chief executive John F. Antioco’s salary by 17 percent to $2.05 million despite a $1.25 billion loss for the company. In addition, Antioco received a $5 million bonus and millions of dollars worth of stock options.”
pooling contract case, unless \( \alpha < \alpha_2 \), \( \bar{w}_{10}^1 (\bar{w}_{10}^1) > w_{10}^R \) and \( \bar{w}_{10}^2 (\bar{w}_{10}^2) > w_{10}^R \), which in turn implies that the firm has lower profit. That means unless the worker in the homogeneous case is even less ambitious than the least ambitious in the heterogeneous case, both workers are paid higher wage in the heterogeneous case.

Also note that \( w_{10}^H \) is decreasing in \( \alpha_2 \), and that \( w_{10}^{HH} > w_{10}^H \) for all \( \alpha_2 \leq q_0 (1 - q_1) / (q_1 - q_0) \). This means that, in the pooling contract case, the firm can pay the worker a lower wage when the workers are more ambitious. The workers’ utility thus decreases with the degree of their ambition. This is also true for separating contract case. But there is a limit on the firm’s ability to manipulate the wage policy. As the value of \( \alpha_2 \) passes a threshold, \( q_0 (1 - q_1) / (q_1 - q_0) \), the utility of the worker hits the reservation value, and the winner’s wage will be fixed no matter how large \( \alpha_2 \) becomes. This is true for both separating and pooling contracts.

Although the optimal wage structures are different for pooling and separating contracts, a common feature emerges: the firm is worse off with a mixed workforce. That is, worker heterogeneity encroaches on the ability of the firm to manipulate wage policy, and thus reduces its profit. This is in contrast to the result in either Fershtman et al. (2003b) or Charness and Kuhn (2004); both accentuate the merit of a mixed workforce in improving the profit of the firm. There are two reasons for this difference in result. First, in our model the firm cannot discriminate against the workers by giving different contracts to workers with different degrees of ambition. The only case in which the workers can be paid differently is through self-selection. As a result the wage contracts have to satisfy additional incentive compatibility constraints, which reduces the firm’s profit. Alternatively, the firm can offer a uniform contract to both types of workers, but in that case all workers have to be treated as low-ambition type. This, similarly, reduces the firm’s profit. Fershtman et al. (2003b) and Charness and Kuhn (2004) both assume that the firm is allowed to offer different contracts to different types of workers, so there is no cost in treating different workers differently.\(^20\) Second, in their models the workers have a continuum of choices of effort. In that case workers are not always motivated to exert the highest effort possible. Different contracts tailored to different types of workers (without needing to satisfy the incentive compatibility constraints by assumption) enables the firm to motivate the workers to exert the optimal level of effort. Consequently, although workers are paid more in the heterogeneous

\(^{20}\) In term of our discussion in the beginning of this section, in their models the firm has the advantage of directly discriminating between the workers without the disadvantage of having to satisfy the self-selection constraints.
case (which is the same as our result), their effort level also increases, which also increases the total profit of the firm.

It turns out that pooling and separating contracts yield exactly the same profits for the firm and the same utilities for the workers. That is, in terms of the result, they are exactly the same for both the firm and the workers. This can be seen by comparing the expected wage for both types of workers under pooling and separating contracts. However, although the two contracts result in the same expected utility for both the principal and the agent, they are different in nature. The pooling contract is strictly a relative performance evaluation in that only the winner receives positive wage. The separating contract, while setting an RPE for worker 2, also pays a positive wage to worker 1 when both have low outputs. Note that if we measure wage risk of a worker by the variance of his wage distributions, then both workers have lower wage risks under a separating contract.\(^21\) In that case the separating contract will Pareto dominate the pooling contract, and is a better prediction of the wage structure that the firm will offer.

5. **CONCLUSION**

This paper investigates the property of a firm’s compensation scheme when the workers’ utilities depend not only on their own incomes, but also on the incomes of their co-workers. The optimal contract is shown to exhibit wage compression, and relies on relative performance evaluation. Moreover, even if the information structure calls for absolute performance evaluation in the traditional principal-agent model, the optimal contract will still be a relative performance evaluation. Thus our model offers a strong rationale for why, despite many of its drawbacks, relative performance evaluation is so commonly used. The firm’s profit is greater than when the workers only care about own-incomes. This is because given a fixed wage schedule, the utility gap for a worker between different wages levels is widened. The firm can thus lower the wage paid to the high-performer while still maintaining the same incentives. We also show that, in contrast to past literature, worker-heterogeneity encroaches on the ability of the firm to manipulate wage policy.

A natural extension of the present model is the case when the workers differ not only in ambition, but also in productivity. For example, different workers might have

\(^{21}\) Note that this is true even for worker 2.
different probabilities of producing high output even when they all exert high efforts. This, however, requires very careful specification of the worker’s utility function. For example, in our setup both workers are motivated to exert high effort. As a result, there is stronger reason that a worker will suffer psychological loss if he is paid a less. If, however, the worker who is being paid a lower wage is aware that his co-worker has higher productivity, then he might feel this difference in wage is justified. Or, on the contrary, a low-ability worker might believe that he should be paid more if he has the same output as a high-ability worker, since this signals that he has exerted more effort. Obviously, more fundamental investigation on how and in what way the workers’ utilities depend on others’ incomes, when they are heterogeneous in abilities, is much needed in order to have a thorough understanding of team compensation.

REFERENCES


給野心員工的最適契約

陳恭平
中央研究院人文社會科學研究中心
暨國立台灣大學經濟學系

蔡崇聖 *
國立清華大學經濟學系

關鍵詞: 相對績效評估，絕對績效評估，升遷，勞動契約

JEL 分類代號: J3

* 聯繫作者: 蔡崇聖, 國立清華大學經濟學系, 新竹市 300 光復路二段 101 號。電話: (03) 573-1197; 傳真: (03) 572-2746; E-mail: tstsai@mx.nthu.edu.tw。作者感謝楊春雷教授、Yoram Weiss 教授及兩位匿名審稿者所提供的寶貴意見。
摘 要

我們考慮在一個公司團隊當中，員工不只在乎他的絕對所得，也在乎與其他成員比較之下的相對所得。若員工越在意相對所得，我們就說他越有野心。當員工有要勝過團隊中其他成員的野心時，公司實際上可以利用員工這種心理，以較低的工資成本，激勵員工投入相同程度的努力。我們證明此時員工的最適契約符合相對績效評估的性質。更重要的是，即使員工之間的生產技術是獨立的，當員工有野心時，最適契約仍然應該採用相對績效評估。這個結果不同於傳統模型中認為在这种情況下，應該採用絕對績效評估的見解。另一個與文獻不同的結果是，當員工間有不同程度的野心時，相對於同質員工的狀況，公司有較低的利潤。