Positive self-image and incentives in organizations

Santos-Pinto, Luís

Universidade Nova de Lisboa

3 September 2003

Positive Self-Image and Incentives in Organizations†

Luís Santos-Pinto

Universidade Nova de Lisboa, Departamento de Economia
Campus de Campolide, PT-1099-032, Lisboa, Portugal

February 14, 2007

Abstract

This paper investigates the implications of individuals' mistaken beliefs of their abilities on incentives in organizations using the principal-agent model of moral hazard. The paper shows that if effort is observable, then an agent's mistaken beliefs about own ability are always favorable to the principal. However, if effort is unobservable, then an agent's mistaken beliefs about own ability can be either favorable or unfavorable to the principal. The paper provides conditions under which an agent's over estimation about own ability is favorable to the principal when effort is unobservable. Finally, the paper shows that workers' mistaken beliefs about their coworkers' abilities make interdependent incentive schemes more attractive to firms than individualistic incentive schemes.

Journal of Economic Literature Classification Numbers: A12; J41.
Keywords: Positive Self-Image; Incentives; Moral Hazard; Behavioral Economics.

†This paper is part of my Ph.D. dissertation in University of California, San Diego. I am indebted to my Ph.D. supervisor, Joel Sobel, for extremely helpful suggestions. I would also like to thank seminar participants at University of Montreal, Universidade Nova de Lisboa, University of California, San Diego, University of East Anglia, University of Innsbruck, and two anonymous referees for many insightful comments and suggestions. I gratefully acknowledge financial support from Praxis XXI and Fundação Calouste Gulbenkian.

1 Introduction

Most individuals have a tendency to make overly positive evaluations of their absolute and relative abilities. Positive self-image is a staple finding in psychology and has been shown to be present in individuals' self-assessments of performance in their jobs.[1]

Baker et al. (1988) cite a survey of General Electric Company employees according to which 81 percent of a sample of white-collar clerical and technical workers rated their own performance as falling within the top 20 percent of their peers in similar jobs.[2] Myers (1996) cites a study according to which, in Australia, 86 percent of people rate their job performance as above average, 1 percent as below average. Brozynski et al. (2003) find that fund managers' hold overly positive views of their relative performance. Oberlechner and Osler (2004) find that 75 percent of currency traders in foreign exchange markets think they are better than average.

The evidence suggests that in settings where individuals' assessments of their abilities is a factor that matters for making decisions, then positive self-image may alter substantially the standard equilibrium predictions of economic theory. This raises a host of theoretical and empirical questions that so far have been left unanswered. For example, does the existence of positive self-image change the firm's optimal incentive scheme? Does heterogeneity in workers' self images have interesting implications for the composition of the workforce? This paper addresses these questions by focusing on the impact of mistaken beliefs of ability on incentives in organizations.

The paper starts by studying the impact of an agent's mistaken beliefs about own ability on the principal's welfare using the standard principal-agent model of moral hazard. An agent with mistaken beliefs of own ability misunderstands his productivity of effort. An agent with a positive (negative) self-image of own ability over(under)estimates his productivity of effort.[3] The firm is assumed to know about the agent's mistaken beliefs.[4]

When effort is observable, the agent's mistaken beliefs only intervene in the agent's willingness to accept the contract offered by the principal.

[1] According to Myers (1996), a textbook in social psychology, "(...) on nearly any dimension that is both *subjective* and *socially desirable*, most people see themselves as better than average."

[2] In another study based on the same survey conducted on managerial and professional employees it was found that 83 percent rated their performance in the top 10 percent.

[3] Most individuals display positive self-image but some display negative self-image.

[4] The assumption that the agent holds mistaken beliefs together with the assumption that the principal is perfectly informed about the agent's mistaken beliefs are the main difference between this model and conventional moral-hazard models.

Because the agent believes that his effort will result in a different level of output than it actually will (i) the optimal contract should not fully insure the agent, and (ii) it is cheaper for the principal to implement the intended action, than in the standard model.

The intuition behind these results is as follows. If effort is observable the principal can benefit from the agent's mistaken beliefs by modifying the contract that would be optimal for an accurate worker to include a bonus in states that the worker believes are more likely to occur than is actually the case and a penalty in states that he believes are less likely to occur than is actually the case. The worker mistakenly views this side-bet as generating a positive expected payoff and is willing to accept the resulting contract so long as side-bet is sufficiently small that it does not expose him to too much risk. In other words, for a small deviation from full insurance the agent's mistaken beliefs generate a first-order gain to the principal, whereas risk aversion generates a second-order loss.

Matters are not so straightforward when effort is unobservable. In this case, the agent's mistaken beliefs are no longer necessarily favorable to the principal. This happens because the agent's mistaken beliefs now affect his incentives to exert effort as well as his willingness to accept the contract. However, the paper shows that it is possible to find conditions under which the impact of positive (negative) self-image beliefs is always (un)favorable to the principal: these must ensure that (1) the incentive scheme is nondecreasing with output, (2) the only binding constraints are the local downward incentive compatibility constraints, and (3) positive self-image beliefs and effort are complements.

If the incentive scheme is nondecreasing with output, then a positive (negative) self-image agent thinks that he is more (less) likely to earn high wages than he actually is. Thus, a positive (negative) self-image agent over(under)estimates the value of the contract when the incentive scheme is nondecreasing with output. If positive self-image and effort are complements, then positive self-image makes higher effort relatively more attractive, slackening the downward incentive constraints which is sufficient for an overall favorable (unfavorable) impact on incentive constraints given (2).

The paper also argues that workers' mistaken beliefs about their coworkers' abilities make interdependent incentive schemes more attractive to organizations than individualistic incentive schemes.[5] To make this point the

[5] Under an individualistic incentive scheme a worker's payment schedule is only contingent on the outcome of his own task. Under an interdependent incentive scheme a worker's payment schedule is contingent on the outcome of his own task as well as the outcome of others' tasks.

possibility that workers hold mistaken beliefs about their coworkers' abilities is incorporated in Mookherjee's (1984) principal multiple-agent model of moral hazard. A worker with mistaken beliefs of his coworkers' abilities is assumed to have an accurate assessment of the productivity of his own effort but to misperceive the productivity of his coworkers. The firm is assumed to know about the workers' mistaken beliefs and workers are assumed to know about their coworkers' mistaken beliefs.

The paper finds that if workers have mistaken beliefs about their coworkers' productivity, then a change of the optimal individualistic incentive scheme in the direction of an interdependent incentive scheme increases the firm's welfare. The intuition behind this result is straightforward. When workers have mistaken beliefs about their coworkers' productivity and the firm knows this, the firm can make an advantageous "side bet" by making workers' compensation dependent on their coworkers' output. The firm does this by offering workers interdependent contracts that include a bonus in output states which workers mistakenly believe are more likely to occur than it is the case, and a penalty in output states which workers mistakenly believe are less likely to occur than is actually the case. However, the bonus and penalties can not be too large otherwise this would expose workers to too much risk.

The remainder of the paper is organized as follows. Section 2 studies the implications of an agent's mistaken beliefs about own ability in a principal agent model of moral hazard. Section 3 shows that workers' mistaken beliefs about their coworkers' abilities make it more likely that firms want to use interdependent rather than individualistic contracts. Section 4 discusses the findings. Section 5 reviews related literature. Section 6 concludes the paper. The Appendix contains the proofs of all results.

2 Mistaken Beliefs about Own Ability in a Principal Agent Model

In this section we study the impact of a worker's mistaken beliefs about own ability on the firm's welfare using the canonical principal-agent model of moral hazard. According to this model a principal hires an agent to produce output. Output is a stochastic function of effort. When the agent's effort is not directly observable by the principal, the principal can only offer a contract based on the observable and verifiable output levels.

There exists a finite set of possible levels of output, $Q = \{q_1, \ldots, q_M\}$, a finite set of possible effort levels, $\mathcal{A} = \{a_1, \ldots, a_N\}$, and output is stochastic, given effort. Let $p_m(a_n)$ denote the probability of output level q_m resulting if the effort level is a_n. Every output level is possible under any effort level, that is, $p_m(a_n) > 0$ for all n and m. An increase in effort results in a higher probability of higher levels of output, that is, for all pairs n and n' such that $n > n'$, $p(a_n)$ first-order stochastically dominates $p(a_{n'})$, where $p(a_n) = (p_1(a_n), \dots, p_M(a_n))$.

The agent is allowed to hold mistaken beliefs of his productivity. That is, $\tilde{p}(a_n) \neq p(a_n)$, for at least one n, where $\tilde{p}(a_n)$ denotes the agent's perception of the marginal distribution of output induced by effort level n. An agent with mistaken beliefs about own ability is assumed to always perceive high effort to be more productive than low effort, that is, for all pairs n and n' such that $n' > n$, $\tilde{p}(a_{n'})$ first-order stochastically dominates $\tilde{p}(a_n)$.

The agent exhibits positive self-image of own ability if $\tilde{p}(a_n) \neq p(a_n)$, for at least one n, and $\tilde{p}(a_n)$ first-order stochastically dominates $p(a_n)$, for all n. The agent exhibits negative self-image of own ability if $\tilde{p}(a_n) \neq p(a_n)$, for at least one n, and $p(a_n)$ first-order stochastically dominates $\tilde{p}(a_n)$, for all n. The agent exhibits accurate self-image of own ability if $\tilde{p}(a_n) \equiv p(a_n)$, for all n.[6]

The agent is assumed to be risk averse with a utility function that is separable in income and effort, $U(y,a) = u(y) - c(a)$, with u strictly increasing and concave and c increasing. The agent's ex ante utility from accepting the contract must be at least his reservation utility level, $\bar{U}$.

The principal is risk neutral and cares only about maximizing profits, the difference between expected benefits and expected wages. The principal is assumed to be perfectly informed about the agent's true productivity and also about the agent's mistaken beliefs.

The assumption that the principal knows about the agent's mistaken beliefs should be viewed as a reasonable simplification. The psychological evidence that individuals are prone to overestimate their abilities is widespread and publicly available. If the bias is systematic, the principal could at least be expected to make a good guess about the mistaken beliefs of a particular agent by assuming that the agent overestimates his ability.

[6] Note that according to this definition an agent with positive self-image of own ability might not overestimate his productivity at all effort levels. However, it is certain that he never underestimates it.

2.1 Observable Effort

We begin our analysis by studying the impact of mistaken beliefs about own ability on the principal's welfare when effort is observable. In this case the contract that the principal offers to the agent specifies the agent's effort and his wage as a function of observed output. It is a standard result in the agency literature that when effort is observable and the agent is risk averse the optimal solution to the principal's effort implementation problem is to fully insure the agent against risk. Hence, given the contract's specification of effort, the principal offers a fixed wage such that the agent receives exactly his reservation utility. Proposition 1 shows that this result is no longer valid when the agent misperceives his productivity.

Proposition 1 *In the principal-agent model with observable effort, risk neutral principal, and risk averse agent, if the agent holds mistaken beliefs about his productivity, the principal knows about the agent's beliefs as well his true productivity, then (i) the optimal contract should not fully insure the agent, and (ii) the cost to the principal of implementing an arbitrary action is lower or equal than in the standard model.*

When the agent holds mistaken beliefs about his productivity and effort is observable full insurance is still a feasible solution to the principal's effort implementation problem so the principal can't do worse. However, the difference in beliefs generates a beneficial side-bet to the principal. The principal can do better by offering the agent a contract that includes a side bet that pays a bonus in states that the worker believes are more likely to occur than is actually the case and a penalty in states that he believes are less likely to occur than is actually the case.

The agent's risk aversion implies that the shift from a contract that fully insures the agent to a contract with a side bet has a cost. The side bet makes the agent's payment contingent on output and therefore increases risk. Proposition 1 shows that for a small enough side bet the reduction in compensation cost due to the impact of the agent's mistaken beliefs on the participation constraint dominates the increase in compensation cost associated with the agent's risk aversion. This is because, for a small side bet, the impact of the agent's mistaken beliefs is of first-order whereas the impact of risk aversion is of second-order.[7]

[7] If an agent is risk neutral and has mistaken beliefs the principal's problem does not have a solution. In this case the principal can always increase profits by raising the stakes of the side bet. However, if the agent is risk neutral and is protected by limited liability this situation does not arise since at a certain poin the limited liability constraint becomes biding and this prevents the principal from increasing the bet anyfurther.

2.2 Unobservable Effort

When effort is unobservable and the principal wants the agent to choose effort level a_n the principal's effort implementation problem is given by

$$\begin{aligned} \min_{\{v_m\}} \quad & \sum_m p_m(a_n)\, h(v_m) \\ \text{s.t.} \quad & \sum_m \tilde{p}_m(a_n)\, v_m - c(a_n) \geq \bar{U} \qquad (1) \\ & \sum_m \left[\tilde{p}_m(a_n) - \tilde{p}_m(a_{n'})\right] v_m \geq c(a_n) - c(a_{n'}), \ \forall\, a_{n'} \neq a_n, \qquad (2) \end{aligned}$$

where $h \equiv u^{-1}$, (1) is the participation constraint and (2) represents the set of incentive compatibility constraints.

When effort is unobservable the impact of mistaken beliefs about own ability on the principals' welfare becomes more interesting. Now, besides their impact on the participation constraint, mistaken beliefs about own ability also have an impact on the agent's incentives to exert effort.

When effort is unobservable, the paper distinguishes between two effects that an agent's mistaken beliefs about own ability have on the principal's welfare. The *participation effect* concerns the impact of mistaken beliefs about own ability on the agent's willingness to accept the contract offered by the principal for a fixed wage incentive scheme. The *incentive effect* concerns the impact of mistaken beliefs about own ability on the set of incentive compatibility constraints for a fixed wage incentive scheme.

To explain the impact of these two effects of mistaken beliefs on the principal's effort implementation problem suppose that there are only two output and two possible effort levels. In this case, if the principal prefers to implement low effort the results of Proposition 1 apply: the principal is better off with an agent with mistaken beliefs than with an accurate agent. However, if the principal prefers to implement high effort then the agent's mistaken beliefs about own ability have an impact on the agent's incentives to exert effort. If the agent perceives the productivity of high and low effort to be more similar than they actually are, then there is an unfavorable impact of mistaken beliefs on incentives since the returns to high effort become less attractive to the agent by comparison with the returns to low effort. Of course, if the opposite happens there is a favorable impact of mistaken beliefs on incentives.

When there are only two output and two effort levels and the principal prefers to implement high effort the impact of mistaken beliefs on the participation constraint can be either favorable or unfavorable to the principal.

To see this let us start by recalling the features of the optimal solution to the principal's high effort implementation problem when effort is unobservable and the agent has an accurate self-image. In this case high effort can only be implemented if the principal offers the agent a high wage for high output and a low wage for low output. Now suppose that the principal faces a positive self-image agent who overestimates the productivity of high effort and that the principal offers this agent the optimal contract designed for an accurate agent. The fact that the optimal contract for the accurate agent has a wage-incentive schedule that is strictly increasing in output implies that the positive self-image agent values it more than the accurate agent. This happens because the positive self-image agent thinks that he is more likely to produce high output and hence to earn a high wage than an accurate agent. In this case the impact of mistaken beliefs on the participation constraint is favorable to the principal. The opposite happens if the principal faces a negative self-image agent who underestimates the productivity of high effort.

The result that positive self-image has a favorable impact on the participation constraint while negative self-image has an unfavorable impact is valid for two output levels and two effort levels case but does not extend to cases where there are more than two output levels. For example, it is well known in the agency literature that if there are two effort levels and more than two output levels the wage-incentive scheme that implements high effort need not be strictly increasing in output. If the wage-incentive scheme is non-monotonic in output then it is easy to construct examples where positive self-image has an unfavorable impact on the participation constraint. This observation leads us to our next result.

Lemma 1 *In the principal-agent model with unobservable effort, if the principal prefers to implement effort level a_n and the optimal wage-incentive scheme for an accurate agent is nondecreasing in output, then for an agent with positive self-image*

$$E_{\tilde{p}(a_n)}\left[U\left(\left(y_m^*\right), a_n\right)\right] \geq E_{p(a_n)}\left[U\left(\left(y_m^*\right), a_n\right)\right],$$

while for an agent with negative self-image

$$E_{\tilde{p}(a_n)}\left[U\left(\left(y_m^*\right), a_n\right)\right] \leq E_{p(a_n)}\left[U\left(\left(y_m^*\right), a_n\right)\right],$$

where (y_m^*) is the optimal wage-incentive scheme for an accurate agent.

Lemma 1 follows because integrating an increasing function against ordered measures preserves the order and so we omit its proof.[8] This result states that if the optimal contract is nondecreasing in output, then a positive (negative) self-image agent over(under)estimates its value since he thinks that he is more (less) likely to earn high wages than he actually is.

Note that a wage-incentive scheme that is nondecreasing with output can be the result of characteristics of the technology and preferences of the problem.[9] For example, if the monotone-likelihood ratio condition (MLRC) and the concavity of the distribution function condition (CDFC) are satisfied then the optimal wage-incentive scheme is nondecreasing with output.[10]

We will make use of MLRC, CDFC, and Lemma 1 in our next result. Before that we need to introduce one definition. Let $f(a_n) \equiv \tilde{p}(a_n) - p(a_n)$. We say that self-image and effort are *complements* if and only if $\sum_m^k f_m(a_n) \leq \sum_m^k f_m(a_{n'})$ for all $k = 1, \ldots, M$, and for all pairs n and n' such that $n > n'$. We say that self-image and effort are *substitutes* if and only if $\sum_m^k f_m(a_n) \geq \sum_m^k f_m(a_{n'})$ for all $k = 1, \ldots, M$, and for all pairs n and n' such that $n > n'$. These two definitions capture the idea that an increase in effort can be associated with an increase or a decrease in the increment of positive self-image beliefs. We are now ready to state our next result.

Proposition 2 *In the principal-agent model with unobservable effort, risk neutral principal, and risk averse agent, if MRLC and CDFC hold, self-image and effort are complements, and the principal knows about the agent's beliefs, then the cost to the principal of implementing an arbitrary action is lower (higher) or equal with an agent with positive (negative) self-image beliefs than with an agent with accurate beliefs.*

Proposition 2 identifies conditions under which the impact of positive self-image is always favorable to the principal and the impact of negative self-image is always unfavorable. The intuition behind this result is as follows. If MRLC and CDFC hold, then the optimal incentive scheme is nondecreasing with output. If the optimal incentive scheme is nondecreasing with output, then we know from Lemma 1 that, for a fixed wage incentive scheme, positive self-image slackens the participation constraint. Addition-

[8] We have that y_m increases in m, and U increases in y. When the agent has positive self-image $\tilde{p}(a_n)$ stochastically dominates $p(a_n)$.

[9] A wage-incentive scheme that is nondecreasing with output can also be the result of constraints on the set of feasible incentive schemes. For an example see Grossman and Hart (1983).

[10] The monotone-likelihood ratio condition says that a higher effort level increases the likelihood of a high output level more than the likelihood of a low production level. The concavity of the distribution function condition says that increases in effort have decreasing marginal impact on the probability of higher output levels being attained.

ally, if MRLC and CDFC hold the only incentive constraints that matter are the local downward incentive constraints. This and the assumption that self-image and effort are complements imply that, for a fixed wage incentive scheme, positive self-image also slackens the incentive constraints since it makes high effort relatively more attractive to the agent than low effort.[11]

3 Mistaken Beliefs about Coworkers' Abilities in a Principal Multiple-Agent Model

We are interested in finding out if workers' mistaken beliefs about their relative abilities have an impact on the firm's choice of optimal incentive scheme. More specifically, we want to know if workers' mistaken beliefs about their coworkers' abilities make interdependent contracts more attractive to firms than individualistic contracts. To find an answer to this question we introduce worker mistaken beliefs about their coworkers' abilities in Mookherjee's (1984) principal-multiple-agent model of moral hazard.

In this model a principal (firm) hires two agents (workers). After signing a contract with the firm each worker takes an unobservable action. The level of output, which depends on the worker's action and the realization of a random variable, is publicly observed and workers are compensated on the basis of the output realizations.

There exists for each worker k a finite set of possible levels of output, $Q^k = \{q_1, \ldots, q_n\}$, a finite set of possible actions, $\mathcal{A}^k$, a random variable θ^k (with finite range $\mathcal{H}^k$) and a production function $f^k\left(a^1, a^2, \theta^k\right) : \mathcal{A}^1 \times \mathcal{A}^2 \times \mathcal{H}^k \rightarrow Q^k$. The production function determines for any pair of actions $\left(a^1, a^2\right)$ and any realization of the random variable θ^k, a unique output for worker k. The random variables θ^1 and θ^2 have a joint probability distribution represented by $g\left(\theta^1, \theta^2\right)$. The production functions f^1 and f^2 together with the joint distribution g of θ^1 and θ^2 induce a probability distribution over output pairs $\left(q^1, q^2\right)$ for any given pair of actions $\left(a^1, a^2\right)$. Let $p_{ij}\left(a^1, a^2\right)$ denote the probability of output pair $\left(q_i^1, q_j^2\right)$ resulting if the actions are $\left(a^1, a^2\right)$.

Workers are assumed to be risk averse with a von Neumann-Morgenstern utility function U^k which is additively separable in monetary reward, y^k, and

[11] If self-image and effort are substitutes, then the impact of positive self-image on the participation constraint would still be favorable to the principal but the impact on the incentive constraints would be unfavorable. In this case the impact of positive self-image beliefs on the principal's welfare is ambiguous.

action chosen. That is,

$$U^k\left(y^k, a^k\right) = u^k\left(y^k\right) - c^k\left(a^k\right),$$

where possible rewards by the firm range over a closed interval $\left[y_L^k, y_H^k\right]$. Worker k is assumed to have a reservation utility of $\bar{U}^k$ which represents the minimum payment he must be given to sign the contract with the firm. It is further assumed that:

(i) u^k is continuous, strictly increasing and concave over $\left[y_L^k, y_H^k\right]$;
(ii) if $a_L^k \in \mathcal{A}^k$ minimizes $c^k\left(a^k\right)$ over $\mathcal{A}^k$ then $u^k\left(y_L^k\right) - c^k\left(a_L^k\right) < \bar{U}^k$;
(iii) if $a_H^k \in \mathcal{A}^k$ maximizes $c^k\left(a^k\right)$ over $\mathcal{A}^k$ then $u^k\left(y_H^k\right) - c^k\left(a_H^k\right) > \bar{U}^k$.

The firm is assumed to be risk neutral and to be concerned exclusively with the maximization of profits, the difference between expected benefits and expected compensation costs. Hence if $\hat{B}\left(q^1, q^2\right)$ is the benefit function of the firm then expected benefits when actions chosen by the workers are $\left(a^1, a^2\right)$ are equal to

$$B\left(a^1, a^2\right) = \sum_i \sum_j p_{ij}\left(a^1, a^2\right) \hat{B}\left(q_i^1, q_j^2\right).$$

If the firm is unable to observe workers' actions, then payment can only be made contingent on workers' outputs. In this case an incentive scheme for worker k is an $n_1 \times n_2$ dimensional vector $\left(y_{ij}^k\right) \in \left[y_L^k, y_H^k\right]^{n \times n}$ where y_{ij}^k is the compensation given to worker k if the output pair $\left(q_i^1, q_j^2\right)$ results. Given a pair $\left(\left(y_{ij}^1\right), \left(y_{ij}^2\right)\right)$ of incentive schemes and an action pair $\left(a^1, a^2\right)$ the firm incurs an expected compensation cost of

$$C\left(a^1, a^2, \left(y_{ij}^1\right), \left(y_{ij}^2\right)\right) = \sum_i \sum_j p_{ij}\left(a^1, a^2\right)\left(y_{ij}^1 + y_{ij}^2\right).$$

The firm will choose a pair of incentive schemes $\left(\left(y_{ij}^1\right), \left(y_{ij}^2\right)\right)$ and an action pair $\left(a^1, a^2\right)$ which maximize expected profits subject to the constraints that the action pair $\left(a^1, a^2\right)$ is a Nash equilibrium for the workers under $\left(\left(y_{ij}^1\right), \left(y_{ij}^2\right)\right)$, and that they attain an expected utility of at least $\bar{U}^k$ in equilibrium.

Recall that our goal is find out if workers mistaken beliefs about their coworkers' abilities make interdependent contracts more desirable to the firm than individualistic contracts. This implies that we need to rule out alternative explanations for interdependent contracts to be more attractive to the firm than individualistic contracts.

We know from Mookherjee (1984) that, in the absence of workers' mistaken beliefs, a sufficient condition for optimal payments to worker 1 to be independent of worker 2's output is that worker 1's output is a sufficient statistic for a^1.[12] A corollary of this result is that if the production functions are separable in actions (an absence of production externalities), the random variables θ^1 and θ^2 are independent (an absence of common uncertainty), and workers have accurate beliefs, then individualistic contracts are optimal.[13]

Thus, we assume that workers' production functions are separable in actions and that the random variables θ^1 and θ^2 are independent. If this is the case we have that

$$p_{ij}\left(a^1, a^2\right) = p_i\left(a^1\right) p_j\left(a^2\right),$$

where $p_i\left(a^1\right)$ denotes the probability of output q_i^1 resulting from action a^1 and $p_j\left(a^2\right)$ denotes the probability of output q_j^2 resulting from action a^2.

We now extend Mookherjee's model to include the possibility that workers' hold mistaken beliefs about their coworker's abilities. Let $p^k\left(a^k\right)$ denote worker k's beliefs regarding his own marginal distribution of q^k given a^k and let $p^k\left(a^l\right)$ denote worker k's beliefs regarding worker l's marginal distribution of q^l given a^l. To allow for workers to have mistaken beliefs about their coworkers' abilities we assume that each worker $k = 1, 2$ has a correct belief of his own ability but has an incorrect belief of worker l's ability, $l \neq k$. That is, we assume that worker k's beliefs regarding his own marginal distribution of q^k given a^k match his actual marginal distribution, $p^k\left(a^k\right) = p\left(a^k\right)$ $k = 1, 2$, and that worker k's beliefs regarding worker l's marginal distribution of q^l given a^l differ from worker l's actual marginal distribution of q^l given a^l, $p^k\left(a^l\right) \neq p\left(a^l\right)$, $l \neq k$, $k = 1, 2$.

We also assume that if a worker has mistaken beliefs about his coworker's relative ability, then he thinks that his coworker's perception of relative ability is mistaken, while he thinks that his own perception of relative ability is

[12] That is, worker 2's output is uninformative about worker 1's action and therefore worker 1's compensation should not depend on worker 2's output (if it did this would only increase the randomness in worker 1's compensation without any gains in terms of inference about worker 1's action).

[13] This is because if θ^1 and θ^2 are independent the joint distribution of $\left(q^1, q^2\right)$ given $\left(a^1, a^2\right)$ can be decomposed into the product of the marginal distributions of q^1 given a^1 and q^2 given a^2.

correct. Finally, we assume that workers "agree to disagree" about their perceptions of relative ability. Squintani (2005) shows that this last assumption rules out a conflict between mistaken beliefs and equilibrium in the workers' simultaneous action choice subgame.[14] The firm is assumed to know about the workers mistaken beliefs.

Analogous to Grossman and Hart's (1983) analysis of the single-agent problem we know that the firm's problem can be decomposed into two parts: the effort implementation problem and the effort selection problem.[15] The analysis will focus on the implementation problem since we know from Mookherjee (1984) that the qualitative properties of the optimal incentive scheme can be deduced from it. The goal is to show that if workers have mistaken beliefs of their relative abilities, interdependent contracts are better for the firm than individualistic contracts. The method of proof consists in showing that for any actions that can be implemented by an individualistic incentive scheme there exists a feasible interdependent incentive scheme that implements the same actions at a lower cost.

Proposition 3 *If production functions are separable in actions, random variables θ^1 and θ^2 are independent, workers are risk averse and have mistaken beliefs of their coworkers' productivities, the firm is risk neutral and knows about workers' beliefs, then the cost to the firm of implementing an arbitrary action pair is lower under an interdependent incentive scheme than under an individualistic one.*

Proposition 3 shows that when the firm knows that workers have mistaken beliefs of their coworkers' abilities then the firm can make an advantageous "side bet." The firm does this by offering workers interdependent contracts that increase compensation for outcomes that the workers' mistakenly think are more frequent and decrease compensation for outcomes that the workers' mistakenly think are less frequent. If workers have positive self-image of their relative abilities the firm should increase compensation in output pairs where a worker's opponent has a low output and reduce

[14] Kruger (1998) presents evidence that suggests that people expect others to have positive self images. If a worker is not aware of his opponent's mistaken beliefs of relative ability then his opponent may choose an action that is not expected by him, that is, there is no equilibrium.

[15] In the implementation problem the firm, for any arbitrary action pair $\left(a^1, a^2\right)$, chooses the incentive schemes that minimize the firm's expected compensation cost subject to the Nash incentive compatibility constraint and the participation constraint. In the effort selection problem the firm, given the solution to the implementation problem, chooses the action pair $\left(a^1, a^2\right)$ that maximizes the difference between expected benefits and implementation costs.

compensation in output pairs where a worker's opponent has a high output by comparison with the optimal individualistic incentive scheme. If workers have negative self-image the opposite procedure should be taken.

Worker risk aversion implies that the shift from an individualistic incentive scheme to an interdependent incentive scheme has a cost. Making each worker's payment contingent on his own and on his opponent's output increases risk by comparison with the case when a worker's payment is only contingent on his own output. Proposition 3 shows that for small enough changes in the direction of an interdependent incentive scheme the reduction in implementation cost due to the workers' mistaken beliefs dominates the increase in implementation cost due to the workers' risk aversion. The reason behind this result is that the impact of workers' mistaken beliefs on implementation cost is of first-order (influences the first derivative of the objective function) whereas the impact of workers' risk aversion is of second order (influences only the second derivative).[16]

The method of proof of Proposition 3 uses the assumption that workers have accurate beliefs about their own productivity. The result also holds with the same method of proof if workers have mistaken beliefs about their own productivity and the firm is aware of that. However, the result does not hold if workers have mistaken beliefs about their own productivity and the firm is not aware of that. If workers are mistaken about their own abilities unbeknownst to the firm, then the firm might offer them a contract which they mistakenly perceive to give them a lower payoff than their reservation payoff.

The assumption that there is no common uncertainty allows the firm to design a side bet that reduces implementation cost without having any impact on the workers' incentives. We cannot show that relaxing this assumption allows the firm to design a side bet that reduces implementation cost.

4 Discussion

This section analyzes the impact of worker mistaken beliefs on worker welfare, discusses alternative definitions of positive self-image, and explores the implications of the findings for firm and worker behavior.

[16] If workers are risk neutral the interdependent implementation problem has no solution: the firm and the workers are both willing to make a side bet and scale it to infinity. However, if workers are risk neutral and are protected by limited liability this situation does not arise since the limited liability constraints become binding and prevent the firm from increasing the bet.

4.1 Worker Welfare

In this paper, workers' ex-ante perceived payoffs are equal to their reservation utilities but their ex-ante actual payoffs may be different from their reservation utilities. This result is not surprising given that workers are making their decisions based on biased perceptions of their abilities, that is, they do not maximize the correct payoff function.

To study the impact of worker mistaken beliefs on worker welfare we take the point of view of an outside observer, who knows the worker's actual productivity of effort.[17] If effort is observable, then the worker is worse off by holding mistaken beliefs about own ability since these lead him to accept unfavorable gambles. If effort is unobservable and the conditions of Proposition 2 hold, then worker positive (negative) self-image of own ability makes the worker worse (better) off relative to an accurate worker.

4.2 Alternative Definitions

In this paper workers overestimate their absolute or relative productivity in the firm but have an accurate assessment of their cost of effort and of their outside option. There could be other ways of thinking about workers views of their abilities.

A worker with a positive self-image might overestimate the value of his outside option while having a correct assessment of his productivity in the firm and of his cost of effort. In this case, whether effort is observable or unobservable, positive self-image is always unfavorable to the firm but favorable to the worker. This happens because a worker who overestimates his outside option must be paid more to accept the contract than a worker who assesses his outside option correctly. This definition of worker positive self-image has opposite welfare implications for the firm (and for the worker) as the ones obtained in this paper.[18]

Another possibility might be that a worker with a positive self-image underestimates his cost of effort but assesses correctly his productivity in the firm and his outside option. In this case, if worker underestimation of cost

[17] The worker's welfare could also be assessed as he perceives it ex ante, that is, using the worker's perception of his productivity of effort. However, the worker will not experience this welfare on average ex post.

[18] A referee suggested that a positive self-image agent might simultaneously overestimate his productivity in the firm and the value of his outside option. The impact of this type of positive self-image on the agent's willingness to accept the contract is ambiguous, both in the observable as well as in the unobservable effort case. So, this alternative definition has no clear cut welfare implications for the firm and for the worker.

of effort is increasing with effort, then there is a favorable incentive effect for the firm.[19] A worker who underestimates the cost of effort can also be paid less to participate in the tournament than a worker who assesses his cost of effort correctly. So, according to this definition, if worker underestimation of cost of effort is increasing with effort, then positive self-image is favorable to the firm.[20]

4.3 Implications of Findings

The paper shows that, everything else equal, if effort is observable, then the firm should have a preference for hiring workers who have mistaken beliefs of their abilities rather than workers who hold correct beliefs of their abilities. When effort is unobservable this need no longer be the case.

Proposition 3 shows that if a firm is aware that workers have mistaken assessments of the ability of their coworkers, effort is unobservable, and there are no production externalities, then the firm prefers to offer interdependent incentive schemes to workers than individualistic ones. Under these circumstances, holding everything else equal, the firm would prefer to hire workers who are mistaken about their coworkers' skills rather than workers who are not.

In settings where performance depends on relative ability positive self-image leads individuals to overestimate the probability of favorable outcomes. If this is the case individuals should, on average, prefer incentive schemes featuring payments contingent on relative performance (e.g., rank-order tournaments or incentive schemes composed partly by fixed pay and partly by variable pay dependent on the magnitude of relative performance) to individualistic incentive schemes (e.g., fixed salary plans or piece rates).

5 Related Literature

This paper is related to a recent strand of literature in behavioral economics that studies the welfare consequences of biases in judgment. The paper focuses on the welfare implications of worker positive self-image and is closely related to papers that study the impact of biased beliefs on the employment relationship.[21]

[19] The opposite happens if underestimation of cost of effort is decreasing with effort.

[20] If overestimation of cost of effort is increasing with effort, then negative self-image is unfavorable to the firm.

[21] Some papers–e.g. Keiber (2003)–study the impact of overestimation of the precision of one's knowledge on the employment relationship. Other papers–e.g. Fairchild (2005)–

This literature can be divided into two approaches. The first approach assumes that the firm is better informed about the worker's skill than the worker. The worker and the firm may have common or different prior beliefs about the worker's skill. The firm uses the contract offered to the worker to given him incentives but also to give him a signal about his skill. Under this approach a worker's beliefs about skill are endogenous to the model. Bénabou and Tirole (2003) and Moscarini and Fang (2005) are examples of papers that follow this approach.

The second approach also assumes that the firm is better informed about a worker's skill than the worker. The firm and the worker have different prior beliefs about the worker's ability. The firm knows about the worker's skill and may also know the worker's beliefs about skill. The firm takes advantage of being better informed to design the incentive scheme. Under this approach there is no signalling by the firm and worker's beliefs about skill are exogenous to the model. The current paper follows this approach. Other examples are Hvide (2002), Gervais and Goldstein (2004) and De la Rosa (2005).

Hvide (2002) considers a principal-agent model in which a worker chooses beliefs about his own ability, and provides conditions under which there is a strategic advantage for selecting beliefs that overestimate true ability. The advantage arises through commitment effects. A worker can gain from overestimating his skill if that improves his bargaining position against the firm (outside option). The firm is made worse off by the workers biased beliefs.

Gervais and Goldstein (2004) study the impact of worker overestimation of own productivity in teams. In that context it is shown that if there are complementarities between workers' efforts, then the team is better off with a workforce that overestimates absolute productivity. The intuition for this result is simple, a worker who overestimates her own marginal productivity works harder, thereby increasing the marginal productivity of her teammates who then work harder as well. Thus, positive self-image can help alleviating the free-riding and coordination problems in teams.[22]

De la Rosa (2005) considers the welfare effects of worker overestimation

look at the implications of managerial optimism for financing decisions by firms (debt or equity). Here, I focus on papers that study the implications of worker overestimation of skill on the agency relationship.

[22] This finding is valid as long as worker overestimation of relative productivity comes mainly from overestimation of own productivity. In fact, the welfare results obtained in their paper are reversed if overestimation of relative productivity comes mainly from underestimation of the productivity of the other team members.

of own productivity when different firms compete to contract with a risk averse worker. He considers a set-up where the worker has two actions (effort or no effort) and there are two possible outcomes: success or failure. The firm knows about the worker's true productivity and knows that the worker overestimates the probability of success given that he exerts effort. This is the dual approach to the one used here since competition between firms drives expected profits to zero and the worker enjoys all the surplus generated by the employment relationship. De la Rosa finds that a risk averse worker benefits from moderate positive self-image but may be hurt from high positive self-image.

The main contribution of this paper to this literature is to show, using an optimal contracting framework, how firms can structure incentives to take advantage of individuals' inflated self-perceptions of skill. This result stands in contrast to Hvide (2002) and does not rely on the assumption of complementarity between workers' efforts present in Gervais and Goldstein (2004).

6 Conclusion

This paper shows that workers' mistaken beliefs of their abilities have interesting implications for the design of incentives in organizations. If effort is observable and the worker has mistaken beliefs of his productivity then, (i) the optimal contract should not fully insure the worker, and (ii) it is cheaper for the firm to implement the intended action, than in the standard model. By contrast, if effort is unobservable worker's mistaken beliefs are no longer necessarily favorable to the firm. The paper also shows that if workers misperceive the productivity of their coworkers, then this makes interdependent incentive schemes more attractive to firms than individualistic ones.

7 Appendix

Proof of Proposition 1 If effort is observable, the principal faces an agent with an accurate self-image, and the principal prefers to induce effort level a_n, then it is a standard result that if the agent is risk averse, the optimal compensation scheme is a fixed wage payment such that the agent receives exactly his reservation utility level. That is $v_m^* = \bar{U} + c(a_n)$, for all m. So, when the agent has an accurate self-image the principal's implementation costs are given by

$$C(a_n, (v_m^*)) = h(\bar{U} + c(a_n)).$$

If effort is observable, the principal faces an agent with mistaken beliefs about own ability, and the principal prefers to induce effort level a_n, then the principal's implementation problem is given by

$$\begin{aligned} \min_{\{v_m\}} \quad & \sum p_m(a_n) h(v_m) \\ \text{s.t.} \quad & \sum \tilde{p}_m(a_n) v_m - c(a_n) \geq \bar{U}. \end{aligned}$$

If $\tilde{p}(a_n) = p(a_n)$ then the agent does not hold mistaken beliefs about effort level a_n and the principal's implementation cost is the same as with an accurate agent. If $\tilde{p}(a_n) \neq p(a_n)$ then the agent holds mistaken beliefs about effort level a_n. If $\tilde{p}(a_n) \neq p(a_n)$ then there exist at least two output levels q_i and q_j such that $\tilde{p}_i(a_n) > p_i(a_n)$ and $\tilde{p}_j(a_n) < p_j(a_n)$, with $i \neq j$. If $z_i = 1$, $z_j = -p_i(a_n)/\tilde{p}_j(a_n)$, and $z_k = 0$, for all $k \neq i, j$, then

$$\sum_m p_m(a_n) z_m < 0 < \sum_m \tilde{p}_m(a_n) z_m,$$

for $p(a_n) \neq \tilde{p}(a_n)$. Let $(\hat{v}_m)$ denote the incentive scheme that adds εz_m to the agent's output contingent utility payments when the agent's output is q_m. That is, $\hat{v}_m = v_m^* + \varepsilon z_m$, with $\varepsilon > 0$. We will show that: (1) $(\hat{v}_m)$ is feasible in the principal's implementation problem with an agent with mistaken beliefs and (2) $(\hat{v}_m)$ has a lower implementation cost than (v_m^*) for small ε. Let us consider (1) first. It is obvious that $(\hat{v}_m)$ satisfies the agent's participation constraint. Let us now consider (2). Taking a second-order Taylor series expansion of h around the individualistic output contingent utility payments we have that

$$h(v) \approx h(v^*) + h'(v^*)(v - v^*) + \frac{1}{2} h''(v^*)(v - v^*)^2. \tag{3}$$

Making use of (3), the implementation cost of $(\hat{v}_m)$ is approximately equal to

$$C\left(a_n, (\hat{v}_m)\right) \approx \sum p_m\left(a_n\right)\left[h\left(v_m^*\right) + h'\left(v_m^*\right)\varepsilon z_m + \frac{1}{2}h''\left(v_m^*\right)\left(\varepsilon z_m\right)^2\right],$$

or, after taking into account that $v_m^* = \bar{U} + c\left(a_n\right)$, for all m, and simplifying terms

$$\begin{aligned} C\left(a_n, (\hat{v}_m)\right) &\approx h\left(\bar{U} + c\left(a_n\right)\right) + \varepsilon h'\left(\bar{U} + c_H\right)\sum p_m\left(a_n\right) z_m \\ &\quad + \frac{1}{2}\varepsilon^2 h''\left(\bar{U} + c_H\right)\sum p_m(a_n) z_m^2. \end{aligned} \tag{4}$$

The second term on the right hand side of (4) is negative given that h' is positive and $\sum p_m\left(a_n\right) z_m$ is negative. The third term on the right hand side of (4) is positive since agent risk aversion implies that h'' is positive. However, for a small enough ε the second term on the right hand side of (4) is larger than the third. This implies that $C\left(a_n, (\hat{v}_m)\right) < C\left(a_n, (v_m^*)\right)$. *Q.E.D.*

Proof of Proposition 2 If MRLC and CDFC hold, then we know from Proposition 8 in Grossman and Hart (1983) that the optimal incentive scheme is nondecreasing with output. If the optimal incentive scheme is nondecreasing with output, then it follows from Lemma 1 that positive self-image slackens the participation constraint. If MRLC and CDFC hold, then we also know from Grossman and Hart (1983) that the only binding incentive constraints are the local downward incentive constraints. If that is the case the principal's problem is reduced to

$$\begin{aligned} \min_{\{v_m\}} \quad & \sum_m p_m\left(a_n\right) h\left(v_m\right) \\ \text{s.t.} \quad & \sum_m \tilde{p}_m\left(a_n\right) v_m - c\left(a_n\right) \geq \bar{U} \\ & \sum_m \left[\tilde{p}_m\left(a_n\right) - \tilde{p}_m\left(a_{n'}\right)\right] v_m \geq c\left(a_n\right) - c\left(a_{n'}\right), \quad a_n \neq a_{n'},\ n > n'. \end{aligned}$$

The definition of complementarity between positive self-image and effort implies that we can rewrite the incentive constraints as

$$\sum_m \left[p_m\left(a_n\right) - p_m\left(a_{n'}\right)\right] v_m + \sum_m \left[f_m\left(a_n\right) - f_m\left(a_{n'}\right)\right] v_m \geq c\left(a_n\right) - c\left(a_{n'}\right),$$

$a_n \neq a_{n'}$, $n > n'$. Since $v_1 \leq \ldots \leq v_M$, $\sum_m f_m(a_n) = 0$, and $\sum_m^k f_m(a_n) \leq \sum_m^k f_m(a_{n'})$ $\forall k = 1, \ldots, M$ it follows that $\sum_m [f_m(a_n) - f_m(a_{n'})] v_m \geq 0$, which implies that the incentive constraint is slackened. If, for a fixed wage incentive scheme, this type of positive self-image beliefs slackens both the participation and the incentive constraints, then the principal's set of feasible incentive schemes is larger with an agent with this type of positive self-image beliefs than with an agent with accurate beliefs. If that is the case then the cost of implementing an arbitrary action pair is lower or equal with an agent with this type of positive self-image beliefs than with an agent with accurate beliefs. Similarly, since under the condition assumed negative self-image worsens both the participation and incentive constraints, the cost of implementing an arbitrary action pair is higher or equal with an agent with negative self-image beliefs than with an agent with accurate beliefs. *Q.E.D.*

Proof of Proposition 3 To make the argument simple and clear we focus on a single agent. Suppose the firm wishes to implement action pair $\left(a^k, a^l\right)$ through an individualistic incentive scheme. Let the individualistic incentive scheme (v_i^*), where v_i^* is the utility payment to worker k if he produces q_i, be the solution to

$$\begin{aligned}
\min_{\{v_i\}} \quad & \sum_i p_i\left(a^k\right) h\left(v_i\right) \\
\text{s.t.} \quad & \sum_i p_i\left(a^k\right) v_i - c\left(a^k\right) \geq \bar{U} \\
& \sum_i \left[p_i\left(a^k\right) - p_i(a)\right] v_i \geq c\left(a^k\right) - c(a), \quad a \in \mathcal{A}^k,\ a \neq a^k,
\end{aligned}$$

where $h \equiv u^{-1}$. We call this problem the individualistic implementation problem. The implementation cost to the firm of using the optimal individualistic incentive scheme is equal to

$$C\left(a^k, (v_i^*)\right) = \sum_i p_i\left(a^k\right) h\left(v_i^*\right).$$

Now consider the interdependent implementation problem. Since we have assumed that the firm knows that worker k has mistaken beliefs regarding worker l's productivity, that is $p^k\left(a^l\right) \neq p\left(a^l\right)$, the firm should take this into account when choosing the optimal interdependent scheme. Making use of this assumption, the assumption that workers' production functions are separable in actions, and the assumption that the random variables θ^1 and θ^2 are independent, the firm's interdependent implementation problem can be written as

$$\begin{aligned}
\min_{\{v_{ij}\}} \quad & \sum_i \sum_j p_i\left(a^k\right) p_j\left(a^l\right) h\left(v_{ij}\right) \\
\text{s.t.} \quad & \sum_i \sum_j p_i\left(a^k\right) p_j^k\left(a^l\right) v_{ij} - c\left(a^k\right) \geq \bar{U}
\end{aligned} \tag{IR}$$

$$\sum_i \sum_j \left[p_i\left(a^k\right) - p_i\left(a\right)\right] p_j^k\left(a^l\right) v_{ij} \geq c\left(a^k\right) - c\left(a\right), \quad a \neq a^k. \tag{IC}$$

If $p^k\left(a^l\right) \neq p\left(a^l\right)$ then there exist at least two output levels q_m and q_n such that $p_m^k\left(a^l\right) > p_m\left(a^l\right)$ and $p_n^k\left(a^l\right) < p_n\left(a^l\right)$, with $m \neq n$. If $z_m = 1$, $z_n = -p_m^k\left(a^l\right)/p_n^k\left(a^l\right)$, and $z_i = 0$, for all $i \neq m, n$, then

$$\sum_j p_j\left(a^l\right) z_j < 0 = \sum_j p_j^k\left(a^l\right) z_j,$$

for $p\left(a^l\right) \neq p^k\left(a^l\right)$. Let $(\hat{v}_{ij})$ denote the interdependent incentive scheme that adds εz_j to worker k's individualistic output contingent utility payments when worker l's output is q_j. That is, $\hat{v}_{ij} = v_i^* + \varepsilon z_j$, with $\varepsilon > 0$. We will show that: (1) $(\hat{v}_{ij})$ is feasible in the interdependent implementation problem and (2) $(\hat{v}_{ij})$ has a lower implementation cost to the firm than (v_i^*) for small ε. Let us consider (1) first. It is obvious that $(\hat{v}_{ij})$ satisfies the worker's participation constraint, (IR). Replacing v_{ij} by $v_i^* + \varepsilon z_j$ in the left hand side of the worker's incentive compatibility constraint, (IC), yields

$$\sum_j p_j^k\left(a^l\right) \sum_i \left[p_i\left(a^k\right) - p_i\left(a\right)\right] v_i^* + \varepsilon \sum_i \left[p_i\left(a^k\right) - p_i\left(a\right)\right] \sum_j p_j^k\left(a^l\right) z_j,$$

or, after simplifying terms,

$$\sum_i \left[p_i\left(a^k\right) - p_i\left(a\right)\right] v_i^*.$$

Since, by assumption (v_i^*) is a solution to the individualistic implementation problem we have that

$$\begin{aligned}
&\sum_i \sum_j \left[p_i\left(a^k\right) - p_i\left(a\right)\right] p_j^k\left(a^l\right) \hat{v}_{ij} = \\
&\qquad \sum_i \left[p_i\left(a^k\right) - p_i\left(a\right)\right] v_i^* \geq c\left(a^k\right) - c\left(a\right), \; a \neq a^k.
\end{aligned}$$

So, $(\hat{v}_{ij})$ is feasible in the interdependent implementation problem. Let us now consider (2). Taking a second-order Taylor series expansion of h around the individualistic output contingent utility payments we have that

$$h\left(v_{ij}\right) \approx h\left(v_i^*\right) + h'\left(v_i^*\right)\left(v_{ij} - v_i^*\right) + \frac{1}{2}h''\left(v_i^*\right)\left(v_{ij} - v_i^*\right)^2. \tag{5}$$

Making use of (5), the implementation cost of $(\hat{v}_{ij})$ is approximately equal to

$$C\left(a^k, (\hat{v}_{ij})\right) \approx \sum_i \sum_j p_i\left(a^k\right) p_j\left(a^l\right)\left[h\left(v_i^*\right) + h'\left(v_i^*\right)\varepsilon z_j + \frac{h''\left(v_i^*\right)\left(\varepsilon z_j\right)^2}{2}\right]$$

or, after simplifying terms

$$\begin{aligned} C\left(a^k, (\hat{v}_{ij})\right) \quad \approx \quad & \sum_i p_i\left(a^k\right) h\left(v_i^*\right) + \varepsilon \sum_i p_i\left(a^k\right) h'\left(v_i^*\right) \sum_j p_j\left(a^l\right) z_j \\ & + \frac{1}{2}\varepsilon^2 \sum_i p_i\left(a^k\right) h''\left(v_i^*\right) \sum_j p_j\left(a^l\right) z_j^2. \end{aligned} \tag{6}$$

The second term on the right hand side of (6) is negative given that h' is positive and $\sum p_j\left(a^l\right) z_j$ is negative. The third term on the right hand side of (6) is positive given that $\sum p_j\left(a^l\right) z_j^2$ is positive and worker risk aversion implies that h'' is positive. However, for a small enough ε the second term on the right hand side of (6) is larger than the third. This implies that $C\left(a^k, (\hat{v}_{ij})\right) < C\left(a^k, (v_i^*)\right)$. *Q.E.D.*

References

Baker, G., Jensen, M., and K. Murphy (1988). 'Compensation and Incentives: Practice vs. Theory,' *Journal of Finance*, Vol. XLIII, No. 3, 593-616.

Bénabou, R. and J. Tirole (2003). 'Intrinsic and Extrinsic Motivation,' *Review of Economic Studies*, Vol. 70, No. 3, 489-520.

Brozynski, T., Menkhoff, L., and U. Schmidt (2004). 'The Impact of Experience on Risk Taking, Overconfidence, and Herding of Fund Managers: Complementary Survey Evidence,' mimeo, University of Hannover.

De la Rosa, L. (2005). 'Overconfidence and Moral Hazard,' University of California, Berkeley, mimeo.

Fairchild, R. (2005).'The Effect of Managerial Overconfidence, Asymmetric Information, and Moral Hazard on Capital Structure Decisions,' *The ICFAI Journal of Behavioral Finance* vol II, no 4 . December 2005.

Gervais, S. and I. Goldstein (2004). 'Overconfidence and Team Coordination,' Duke University mimeo.

Grossman, J. and O. Hart (1983). 'An Analysis of the Principal Agent Problem,' *Econometrica* 51, 7-45.

Hvide, H. K. (2002). 'Pragmatic Beliefs and Overconfidence,' *Journal of Economic Behavior and Organization*, vol. 48, no. 1, 15–28.

Keiber, K. (2003). 'Overconfidence in the Continuous-Time Principal-Agent Problem,' Available at SSRN: http://ssrn.com/abstract=315299.

Krueger, J. (1998). 'Enhancement Bias in the Description of Self and Others,' *Personality and Social Psychology Bulletin*, vol. 24, 505–516.

Moscarini, G. and H. Fang (2005). 'Morale Hazard,' *Journal of Monetary Economics*, Vol. 52, 749-777.

Mookherjee, D. (1984). 'Optimal Incentive Schemes with Many Agents,' *Review of Economic Studies*, LI, 433-446.

Myers, D. (1996). *Social Psychology.* New York: McGraw-Hill.

Oberlechner, T. and C. Osler, 2004, 'Overconfidence in Currency Markets,' mimeo, University of Vienna and Brandeis University.

Rabin, M. (1998). 'Psychology and Economics,' *Journal of Economic Literature*, Vol. XXXVI, 11-46.

Squintani, F. (2006). "Equilibrium and Mistaken Self-Perception." *Economic Theory*, Vol. 27, No. 3, 615 - 641.