Optimal Incentive Schemes when Only the Agents’ "Best" Output Matters to the Principal

Steven D. Levitt


Your use of the JSTOR database indicates your acceptance of JSTOR’s Terms and Conditions of Use. A copy of JSTOR’s Terms and Conditions of Use is available at http://www.jstor.org/about/terms.html, by contacting JSTOR at jstor-info@umich.edu, or by calling JSTOR at (888)388-3574, (734)998-9101 or (FAX) (734)998-9113. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*The Rand Journal of Economics* is published by The RAND Corporation. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/rand.html.

*The Rand Journal of Economics*

©1995 The RAND Corporation

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2001 JSTOR
Optimal incentive schemes when only the agents’ “best” output matters to the principal

Steven D. Levitt*

Standard principal-agent models assume that the principal’s payoff is a function of the total output of all agents. In many real-world situations, however, the principal’s payoff is based solely on the “best” of the agents’ outputs (e.g., the first agent to make an innovation, the most creative advertising campaign, or the cheapest product design). The results obtained from such a model differ from the standard results in a number of respects. For instance, even when identical agents perform identical tasks, the optimal incentive scheme will often differ across agents. Also, the principal may want to increase the variance of the agents’ output or reduce the correlation of output across agents, even when the agents are risk averse.

1. Introduction

In the standard formulation of the multiagent principal-agent problem, the principal’s payoff is taken to be a function of the total combined output of the agents employed minus the total wage bill. For certain situations, such as production by workers on an assembly line, management of business units within the firm, or the output of salespeople, formulating the principal’s payoff as the sum of the agents’ output is natural and appropriate.

There are, however, many situations in which the principal’s payoff is determined not by the total quantity of output produced, but rather by the “best” of the agents’ outputs. The definition of best, of course, depends upon the specific context. In the case of product design, the best proposal may be the one that meets a certain set of specifications at the lowest cost. For an AIDS vaccine, best may simply correspond to the first viable solution developed. With an advertising campaign, best might correspond to most creative.

In the case where only one agent’s output is eventually used by the principal, the cost-minimizing approach for the principal would be to assign only one agent to the task, relying on that agent’s solution. It may, however, be in the interests of the principal to assign multiple agents to a problem, even though ex post the principal derives no benefit from any but the highest of the agents’ output. Two benefits accrue to the

---


I would like to thank Drew Fudenberg, Klaus Schmidt, Jean Tirole, an anonymous referee, and MIT seminar participants for many helpful comments and suggestions. All remaining errors are solely my own. Financial support of the National Science Foundation is gratefully acknowledged.
principal from employing multiple agents on the same task. The first benefit is the well-known "insurance" effect of relative performance evaluation. If there exists an information-laden signal (in this case the output of a second agent), the sufficient statistic result (Holmström, 1979) tells us that such information should be incorporated into an optimal incentive scheme.\footnote{This logic has seen application in numerous settings, including rank-order tournaments (Nalebuff and Stiglitz, 1983; Lazear and Rosen, 1981; Mookherjee, 1984), yardstick competition (Shleifer, 1985), dual sourcing (Demski, Sappington, and Spiller, 1987; Auriol and Laffont, 1992), and vertical restraints (Rey and Tirole, 1986).} The second benefit of hiring multiple agents is what I shall call the "sampling" effect. If there is ex ante uncertainty about each agent's output (and the distribution of the agents' outputs overlap), increasing the number of agents increases the expected value of the best output. Put another way, many draws from a distribution are better for the principal than a single draw.

In the real world, multiple agents are often assigned to the same task even though ex post only one of the outputs is used. In the development of a new advertising campaign, as many as ten teams within an advertising agency will generate three potential campaigns each. Of those thirty ideas, only one or two will be presented to the client. In public funding of medical research, the National Institutes of Health will often support multiple teams attacking the same problem. It is standard in procurement to solicit multiple design concepts, even though only one design will eventually be put into production. Bendor (1985) details numerous case studies of redundancy in government-commissioned design projects. Kamien and Schwartz (1982) report that Upjohn simultaneously exploited six research teams following different approaches to the synthesis of cortisone.

This article considers the design of optimal incentive schemes when the principal's payoff depends on the agents' best output rather than the sum of the outputs.\footnote{Olsen (1994) also considers the case where the principal's payoff depends only on the agents' best output. Whereas this article is concerned with the design of incentive schemes, Olsen's focus is on whether the allocation of resources to R&D is socially optimal in the presence of asymmetric information.} Using the tractable model of Holmström and Milgrom (1987), a number of important differences emerge vis-à-vis the results of the standard principal-agent problem.

(1) Even when identical agents are assigned to the same task, the optimal incentive scheme will often differ across agents. One agent may be given a high-powered incentive scheme to encourage a high level of effort, while another agent is given low-powered incentives (or even a flat wage). Although the second agent's output will rarely be used, his presence makes the insurance aspect of relative performance effect possible.

(2) In general, incentive schemes will be lower powered when multiple agents are employed at a task than if one agent is assigned to perform the same job alone (unlike the standard multiagent principal-agent, where relative performance measures lead to higher-powered incentives). The intuition for that result is as follows: on the margin, only an increase in the effort of the best performer increases the principal's payoff ex post. To entice an extra unit of output ex ante, however, the principal must provide incentives for all agents to work harder. In designing an optimal incentive scheme, the principal takes into account that the marginal cost of an agent's extra effort is borne with certainty, but the associated marginal benefit is realized only probabilistically.

(3) The value of multiple agents in this context is generally increasing in both the variability of output across agents and the value of output to the principal relative to the reservation wage of the agents. Multiple agents performing redundant tasks
are therefore more likely to be observed in environments where output is highly variable and value added is large relative to costs, such as research and development, product design, and advertising. Substantial anecdotal evidence supports that prediction.

(4) It may be in the principal’s interest to either increase the variance of the agents’ output or reduce the correlation of output across agents, even though the principal is risk neutral and the agents are assumed risk averse. The intuition is that high variance and low correlation, while increasing the principal’s wage bill, also make the sampling effect more valuable. As a consequence, the principal may want to encourage risk-taking and innovative approaches on the part of the agents, and perhaps may want to be intentionally vague in defining a task or even restrict information flows between agents tackling the same problem.

The article is organized as follows. Section 2 presents the general setup of the problem and outlines the solution approach to be taken. Section 3 presents the optimal incentive scheme when the principal hires only a single agent. The results of Section 3 are simply a restatement of results first presented in Holmström and Milgrom (1987); the unique properties of the model developed here arise only in the presence of multiple agents. Section 4 characterizes the optimal incentive scheme when two agents are assigned to the same task. (For simplicity, the analysis is restricted to a comparison of the one-agent and two-agent cases, although the logic generalizes to a multiagent problem.) Section 5 examines the optimal incentive scheme in four special cases in which results can be stated more precisely than in the general setting. Section 6 considers the implications of the model for job design and organizational structure. Section 7 offers a brief conclusion.

2. The problem

A profit-maximizing, risk-neutral principal has unit demand for a good. The principal’s payoff depends only on the highest quality of the agents’ outputs and the total wage bill. Quality is measured such that each unit of quality increases the principal’s payoff by one unit.

Each agent employed works alone to produce exactly one indivisible unit of output. The assumption that agents work alone eliminates consideration of the free-rider problem arising in teams (Holmström, 1982). The assumption that output is indivisible implies that the principal must choose a single agent’s output in satisfying her unit demand for the good (as opposed to combining the attractive elements of multiple agents’ output). Indivisibility is by no means a necessary assumption for the results that follow; it is assumed partly to simplify the algebra and partly to focus attention on the aspects of this analysis that are new. The conclusions of the article continue to hold with only minor modifications when this assumption is relaxed.

The quality of an agent’s output is jointly determined by the agent’s effort level (not observable by the principal) and a stochastic component. Specifically, the quality \( x_i \) of agent \( i \)'s output (observed by the principal and all of the agents) is determined by \( x_i = a_i + \epsilon_i + \eta \), where \( a_i \in [0, \infty) \) is the level of effort exerted by agent \( i \), \( \epsilon_i \) is an agent-specific shock, and \( \eta \) is a common shock across all agents. The agent-specific shock is uncorrelated with the common shock.\(^3\) Shocks are assumed

\(^3\)The formulation of the error terms in the model, while very convenient for finding a solution, comes at some cost to realism. In particular, the assumptions imply that all agents, regardless of the effort exerted, are affected in exactly the same way by the common shock. In the model, the principal ascertains as much information about the common shock from one agent working hard and another agent not working at all as she does from two agents both working hard.
individually and independently distributed normal across agents with mean zero and variance $\sigma_i^2$ and $\sigma_n^2$ respectively. The total variance of an agent’s output, denoted by $\sigma^2$, is equal to $\sigma_i^2 + \sigma_n^2$. The agents observe the sum of the shocks, but they cannot distinguish between the agent-specific and common components of the shock. Agents are unable to coordinate their efforts, ruling out considerations of collusion discussed in Tirole (1986).

Agents are assumed to be identical in every respect. Therefore, adverse selection is not a consideration in the model. Agents exhibit constant absolute risk aversion (with identical coefficients of absolute risk aversion $r$). Agents have identical reservation utilities, the certainty equivalent of which is denoted $v$. All agents have the same quadratic cost of effort, $C_i(a_i) = a_i^2/2$.

Agent $i$’s choice of action $a_i$ is properly viewed as the reduced form of a more complex problem in which the agent controls the drift rate (but not the covariance matrix) of a stochastic Brownian motion process (Holmström and Milgrom, 1987). The principal observes only the path of output. Holmström and Milgrom demonstrate that control of such a process can be represented by a once-and-for-all choice of action by the agent. Moreover, optimal incentive schemes are linear in the principal’s observables (in this case, the observed quality level of each agent’s output).

In addition to the linearity result, two other features of the model simplify the problem. First, utility is transferable. As a consequence, maximization of the principal’s payoff is equivalent to maximizing joint surplus. Secondly, the assumption of constant absolute risk aversion (CARA) utility functions implies that expected utility can be expressed in terms of the certainty equivalent.

The principal’s problem can be broken down into two stages. First, she determines the optimal incentive scheme taking the number of agents as given. Then, she selects the number of agents to maximize her payoff. The following section examines the one-agent case; Section 4 addresses the two-agent case (the logic of the two-agent case readily extends to the $N$-agent case, but the algebra becomes burdensome).

3. The one-agent case

The results of this section are simply a restatement of results first presented in Holmström and Milgrom (1987). The unique properties of the model arise only when multiple agents are present.

With only one agent, the optimal incentive scheme can be shown to be linear in the agent’s output, i.e.,

$$W(x) = \alpha + \beta x.$$  

(1)

The principal’s problem is as follows:

$$\max_{a, \beta} \mathbb{E}[x(a^*(\beta))] - \frac{(a^*(\beta))^2}{2} - \frac{r\sigma_n^2(\beta)}{2}$$

such that

---

4 The assumptions adopted here implicitly rule out cases where organization and individual performances are related in a highly nonlinear way. For a discussion of organizational design under such circumstances, see Jacobs (1981).

5 If the reader is uncomfortable with the characterization of R&D and other creative activities as a stationary Brownian motion process, then the incentive schemes in this article are nonetheless applicable if the principal is restricted to linear schemes for exogenous reasons.
(IC) \[ a^*(\beta) \in \argmax a\beta - \frac{a^2}{2} \]

(2)

(IR) \[ a^*(\beta) \cdot \beta + \alpha - \frac{r\sigma^2_\omega(\beta)}{2} - \frac{(a^*(\beta))^2}{2} \geq \bar{v}, \]

where \( \sigma^2_\omega = \beta^2 \sigma^2_\epsilon \) is the wage variance of the agent.

The principal chooses the parameters of the incentive scheme to maximize joint surplus. Joint surplus is given by the principal’s gross (i.e., before-wage) payoff minus the agent’s costs of effort and risk bearing. The IC constraint dictates that the agent choose the action \( a^* \) to maximize expected utility. (The risk-bearing costs and the fixed portion of the wage do not vary with the agent’s choice of actions and therefore are omitted from the IC constraint.) The first term after the argmax in the IC represents the variable wage utility of the agent; the second term is the disutility of effort. The agent’s IR constraint requires that the certainty equivalent utility of participating is at least as great as the certainty equivalent of his outside option.

The assumption of constant absolute risk aversion on the part of the agent allows the principal to deal separately with the IC and IR constraints. Taking first-order conditions of the IC constraint, which are necessary and sufficient due to concavity of the agent’s problem, the agent’s optimal choice of action is given by \( a^* = \beta \). The principal then sets \( \alpha^* \) to satisfy the IR constraint with equality, i.e.,

\[ \alpha^* = \bar{v} - \beta^2 + \frac{\beta^2}{2} + \frac{r\sigma^2_\epsilon \beta^2}{2}. \] (4)

The first term is the agent’s certainty equivalent reservation utility, the second term is the variable portion of the wage, the third term reflects effort costs, and the fourth term is the cost to the agent of bearing risk.

Substituting for \( \alpha^* \) and \( a^* \) in (4) and maximizing with respect to \( \beta \) yields the following solution to the principal’s problem with one agent:

\[ a^* = \beta^* = \frac{1}{1 + r\sigma^2_\epsilon}, \quad \alpha^* = \bar{v} + \frac{(r\sigma^2_\epsilon - 1)}{2(1 + r\sigma^2_\epsilon)^2} \]

\[ E\Pi_1(r, \bar{v}, \sigma^2_\epsilon) = \frac{1}{2(1 + r\sigma^2_\epsilon)} - \bar{v}, \] (5)

where \( E\Pi_1 \) is the principal’s expected net payoff when one agent is employed.

The principal’s net payoff is decreasing in the agent’s level of risk aversion, the variance of total output, and the agent’s reservation utility. As is standard in moral hazard problems, effort in the second-best case is below that of the first-best case (the first-best effort is given by \( a^\text{FB} = 1 \)) because of the tradeoff between optimal risk sharing and incentives. Only in the special cases where the agent is risk neutral or there is no noise in the production function will the first best be achieved.

4. The solution to the two-agent problem

- The introduction of a second agent changes the problem in two ways. First, relative performance evaluation becomes feasible. Second, the principal’s payoff depends on the maximum of the two outputs.
It follows from Theorem 7 of Holmström and Milgrom (1987) that the optimal incentive scheme with two agents takes the form

$$W_i(x_i, x_j) = \alpha_i + \beta_i x_i + \gamma_i x_j,$$

where $\gamma_i < 0$ if shocks to the agents’ output are positively correlated and $\gamma_i > 0$ if shocks are negatively correlated. For a given slope $\beta_i$ of the incentive scheme, the introduction of a second agent reduces the risk borne by the first agent (the insurance effect) but does not affect the first agent’s optimal choice of action ($\alpha_i^* = \beta_i$). Therefore, writing the agents’ actions directly in terms of the slopes of the incentive schemes, the principal’s optimization problem is as follows:

$$\max_{\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2} \quad E[\max(\beta_1 + \epsilon_1, \beta_2 + \epsilon_2)] - \frac{1}{2} \sum_{i=1}^{2} \beta_i^2 - \frac{1}{2} \sum_{i=1}^{2} r \frac{\sigma^2_{\epsilon_i}(\beta_i, \gamma_i, \sigma^2_\epsilon)}{2}$$

such that

$$\beta_i^2 + \alpha_i - \frac{\beta_i^2}{2} - \frac{r \sigma^2_{\epsilon_i}(\beta_i, \gamma_i, \sigma^2_\epsilon)}{2} \geq \bar{v} \quad \text{for } i = 1, 2.$$  

Having already solved for the agents’ optimal actions, the IC constraint is no longer relevant.

Finding the solution to the principal’s problem can be greatly simplified via a number of transformations. Without loss of generality (since the agents are identical), assume that $\beta_1 \geq \beta_2$. The following identity always holds:

$$E[\max(\beta_1 + \epsilon_1, \beta_2 + \epsilon_2)] = \beta_1 + E[\max(\epsilon_1, \beta_2 - \beta_1 + \epsilon_2)].$$

Equation (8) simply says that the principal’s expected gross (i.e., before wages) payoff can be decomposed into two parts: the expected level of output from the hardest-working agent and a sampling effect that arises because of the idiosyncratic shock $\epsilon_i$. The following lemma characterizes the sampling effect. No formal proof of the lemma is provided; all results follow directly from the definition of the sampling effect and the characteristics of extreme value functions.\(^6\)

Lemma 1 (the sampling effect). Let $S(\beta_2 - \beta_1; \sigma_\epsilon) = E[\max(\epsilon_1, \beta_2 - \beta_1 + \epsilon_2)]$. There is no closed-form solution for $S(\beta_2 - \beta_1; \sigma_\epsilon)$. Nonetheless, the behavior of the sampling function can be characterized by the following results:

\[
\begin{align*}
\frac{\partial S(\beta_2 - \beta_1)}{\partial \beta_1} &= -\frac{\partial S(\beta_2 - \beta_1)}{\partial \beta_2} \leq 0 \\
\frac{\partial S(0)}{\partial \beta_2} &= \frac{1}{2} \quad \text{for } \rho \neq 1 \\
\frac{\partial^2 S(\beta_2 - \beta_1)}{(\partial \beta_2)^2} &= \frac{\partial^2 S(\beta_2 - \beta_1)}{(\partial \beta_1)^2} \geq 0
\end{align*}
\]

\(^6\) The density function of the maximum of multiple draws from a normal distribution is distributed as a Type III extreme order function. For further details on the characteristics of such functions, see Galambos (1978) or Gumbel (1958).
\[ S(0) = .564 \sigma_e \quad \text{for } \rho \neq 1 \]
\[ \text{if } \rho = 1, \text{ then } S(\beta_2 - \beta_1) = 0 \quad \forall \beta_1, \beta_2 \]
\[ \frac{\delta S(\beta_2 - \beta_1)}{\delta \sigma_e} \geq 0 \]
\[ \frac{\delta^2 S(\beta_2 - \beta_1)}{(\delta \beta_2)^2 \delta \sigma_e} \leq 0, \]

where \( \rho \) is the correlation of output shocks across agents 1 and 2 (i.e., \( \rho = \frac{\sigma^2_\eta}{\sigma^2_\epsilon + \sigma^2_\eta} \)).

The sign in S1 is intuitive: the further apart the effort levels of the agents, the less likely it is that the lower-effort agent (agent 2, by assumption) will have a large enough positive error term to overcome the gap, thereby reducing the value of the sampling effect. The equality in S1 follows directly from the fact that the sampling effect depends upon \( \beta_1 \) and \( \beta_2 \) only through their difference. Result S2 simply says that when the agents are working equally hard, the probability that a given agent has a higher output is one-half. The sign of the second partial derivatives in S3 implies that the sampling function is convex, i.e., the rate of change of the sampling effect decreases as the two agents’ effort levels get farther apart. The equality of the second partials in S3 is due to the sampling effect’s depending on \( \beta_1 \) and \( \beta_2 \) only through their difference. Result S4 provides a numeric approximation of the size of the sampling effect when agents work equal amounts; its magnitude increases linearly with the standard error of the idiosyncratic shocks, \( \epsilon_e \). Result S5 notes the fact that there is no sampling effect when the agents’ output shocks are perfectly correlated. In such cases, the agent who works harder is certain to have the higher output. Result S6 states that the size of the sampling effect increases as the agent-specific shocks increase. Result S7 says that the convexity of the sampling function with respect to \( \beta \) decreases as the agent-specific output shocks become large. Intuitively, when shocks are large relative to the effort levels, the sampling effect remains important even as the effort levels of the agents move far apart. By contrast, if shocks are very small, the sampling effect drops very quickly to zero: even with a small difference in effort levels, the high-effort agent is very likely to have the higher output.

The following lemma, which states that the variance of the wage does not depend directly on the slope of the incentive scheme with respect to the other agents’ output, is also useful in simplifying the solution to the principal’s problem.

**Lemma 2.** In the two-agent case, \( \sigma^2_{\omega_i} = \sigma^2_\delta(1 - \rho^2) \beta_i^2. \)

**Proof.** With two agents, the optimal wage scheme is of the form

\[ W_i = \alpha_i + \beta_i(a_i + \epsilon_i + \eta) + \gamma_i(a_j + \epsilon_j + \eta). \]

The wage variance can therefore be expressed as

\[ \sigma^2_{\omega_i} = \beta_i^2(\sigma^2_\epsilon + \sigma^2_\eta) + \gamma_i^2(\sigma^2_\epsilon + \sigma^2_\eta) + 2\beta_i\gamma_i^2(\sigma^2_\eta + \sigma^2_\epsilon) + 4\beta_i\gamma_i^2|\beta_i - \beta_j| \]

with the second equality following from the two identities \( \sigma^2_\epsilon + \sigma^2_\eta = \sigma^2_\delta \) and \( \rho = \sigma^2_\eta/\sigma^2_\epsilon \). Substituting the above formula into (7) and taking first-order conditions with respect to \( \gamma_i \) (ignoring the IR constraint) yields \( \gamma^2_i = -\rho \beta_i \). Substituting for the optimal \( \gamma_i \) leads to Lemma 2. **Q.E.D.**
Using the results of Lemmas 1 and 2, and the earlier observation that the choice of the optimal \( \alpha_i \)'s (the fixed portion of the incentive scheme) can be made after \( \beta_i^* \) is chosen, the principal’s determination of the optimal \( \beta_i \)'s simplifies to

\[
\max_{\beta_1, \beta_2} \beta_1 + S(\beta_2 - \beta_1) - \frac{\sum_{i=1}^{2} \beta_i^2}{2} - \frac{\sum_{i=1}^{2} r \sigma_i^2(1 - \rho^2)\beta_i^2}{2}.
\]  

(9)

Note, however, that because of the convexity of \( S(\beta_2 - \beta_1) \), (9) is not necessarily a concave problem.

Taking first-order conditions,

\[
\frac{\partial \text{(equation 9)}}{\partial \beta_1} \rightarrow 1 + \frac{\partial S(\beta_2 - \beta_1)}{\partial \beta_1} - \beta_1 - r \sigma_1^2(1 - \rho^2)\beta_1 = 0
\]

(10)

\[
\frac{\partial \text{(equation 9)}}{\partial \beta_2} \rightarrow \frac{\partial S(\beta_2 - \beta_1)}{\partial \beta_2} - \beta_2 - r \sigma_2^2(1 - \rho^2)\beta_2 = 0.
\]

These first-order conditions are necessary but not sufficient to ensure a global maximum, since the program is not necessarily concave.\(^7\)

Using result S1 and rearranging, the optimal choice of \( \beta_1 \) and \( \beta_2 \) must satisfy

\[
\beta_1^* = \frac{1 - \frac{\partial S(\beta_2^* - \beta_1^*)}{\partial \beta_2}}{1 + r \sigma_1^2(1 - \rho^2)} \quad \beta_2^* = \frac{\frac{\partial S(\beta_2^* - \beta_1^*)}{\partial \beta_2}}{1 + r \sigma_2^2(1 - \rho^2)}.
\]

(11)

The conditions of (11) are exactly equivalent to the following conditions, which provide a more direct interpretation:

\[
\beta_1^* + \beta_2^* = \frac{1}{1 + r \sigma_i^2(1 - \rho^2)} \quad \beta_1^* - \beta_2^* = \frac{1 - \frac{\partial S(\beta_2^* - \beta_1^*)}{\partial \beta_2}}{1 + r \sigma_i^2(1 - \rho^2)}.
\]

(12)

Equation (12) implies the following proposition.

**Proposition 1.** The optimal solution to the two-agent problem has the following properties: (i) The total effort exerted by the agents (represented by \( \beta_1 + \beta_2 \)) is identical to the second-best effort in the standard one-agent equilibrium-agent problem, where the principal observes an exogenous signal that is correlated with the shock to the single agent’s output (with correlation coefficient \( \rho \)).

(ii) The optimal incentive schemes for two identical agents can be asymmetric \( (\beta_1 - \beta_2 \neq 0) \). The optimal asymmetric wage structure will always take the form of \( \{ \beta_1^* = \beta_{\text{sym}}^* + \Delta, \beta_2^* = \beta_{\text{sym}}^* - \Delta \} \), where \( \beta_{\text{sym}}^* \) is the optimal slope of the incentive scheme within the class of symmetric incentive schemes.

**Proof.** Proposition 1 follows directly from (12) and the solution to the one-agent problem with an exogenous signal derived in Holmström and Milgrom (1987). \( Q.E.D. \)

---

\(^7\) The claim that the first-order conditions in (10) are necessary for a global maximum depends on the assumption made earlier that effort levels are nonnegative. If agents can engage in “destructive acts” (i.e., negative effort levels), then the principal’s objective function is discontinuous at \( \beta = 0 \). Given that discontinuity, the optimal incentive scheme could involve \( \beta_i = 0 \), and (10) would not need to be satisfied.
The precise results of Proposition 1 depend heavily on the assumption of quadratic effort costs. As a consequence, Proposition 1 is best viewed as a convenient simplification of the task at hand rather than an important result in and of itself.

The second result of Proposition 1 proves very useful in determining the conditions under which the principal prefers an asymmetric wage scheme to one in which both agents face the same incentives, by dramatically reducing the class of asymmetric schemes that must be considered. Rearranging the second equality in (12) yields

\[
(\beta_1^* - \beta_2^*)[1 + r\sigma_2^2(1 - \rho^2)] = 1 - 2\frac{\partial S(\beta_2 - \beta_1)}{\partial \beta_2}.
\]  

(13)

The left-hand side of (13) corresponds to the marginal cost of increasing the level of asymmetry in the wage bill, holding the total effort of the agents constant. The wage bill is increasing in the amount of asymmetry for a given level of effort because of the convexity of effort costs. It costs more to entice a marginal unit of effort from an agent who is already working hard. The right-hand side of (13) represents the marginal benefit (not including wage payments) to the principal of increasing the asymmetry while holding total effort constant. The first term on the right-hand side represents the direct benefit to the principal from the increased effort of agent 1; the second term reflects the reduction in the sampling effect as a result of the greater gap between the two agents’ effort levels. Result S1 ensures that the right-hand side is increasing because the sampling effect shrinks as asymmetry increases.

Equation (13) suggests a graphical means of characterizing the globally optimal two-agent incentive scheme. Figures 1 through 3 plot the three possible relationships between the left-hand and right-hand sides of (13). The horizontal axes measure the gap between \(\beta_1\) and \(\beta_2\); the vertical axes represent the marginal costs or benefits respectively of increasing the asymmetry. Potential optima are represented by the intersections of the curves. At those points, the marginal costs of increased asymmetry equal the marginal benefit.

It is clear from inspection of (13) that the left-hand side goes through the origin and is linear in the gap between \(\beta_1\) and \(\beta_2\). Result S2 ensures that the right-hand side also equals zero when \(\beta_1 = \beta_2 = 0\). The right-hand side is strictly concave in \(\beta_1 - \beta_2\) by result S3 (for \(\rho \neq 1\)).

Figure 1 reflects the situation where the optimal two-agent incentive scheme is symmetric. The unique equilibrium occurs at the origin. For all asymmetric wage schemes, the costs outweigh the benefits. In contrast, Figures 2 and 3 represent cases where the optimal two-agent incentive scheme is asymmetric. In Figure 2 there are two possible equilibria, one symmetric and one asymmetric. The asymmetric wage scheme is clearly preferred by the principal, however, since the benefits exceed the costs at all points between the equilibria. The total benefit to the principal of adopting the asymmetric scheme over the symmetric scheme is given by the area between the two curves. Figure 3 reflects the special case where the outputs of the two agents are perfectly positively correlated (\(\rho = 1\)). In that case (to be discussed in greater depth in Section 5), the sampling effect is nonexistent, so the right-hand side of (13) is always equal to one. Under such a circumstance, a symmetric scheme will never be optimal.

---

8 If effort costs are more convex, total effort will be higher in the two-agent case than in the parallel one-agent case with an exogenous signal; also, the optimal asymmetric wage will involve lowering one agent’s effort (vis-à-vis the symmetric wage) more than the other agent’s effort is increased. The opposite conclusions would follow in the case where effort costs are less convex than the quadratic.

9 There is also, of course, a second asymmetric equilibrium in which the positions of agents 1 and 2 are reversed. For our purposes, however, the two asymmetric equilibria are equivalent.
The graphical analysis of Figures 1 to 3 implies the following proposition.

**Proposition 2.** The necessary and sufficient conditions for the principal to prefer a symmetric wage scheme to an asymmetric wage scheme in the two-agent case are

(i) \[ \rho < 1 \]

and

(ii) \[ \frac{\partial^2 S(0)}{(\partial \beta_2)^2} < \frac{1 + r\sigma_2^2(1 - \rho^2)}{2} \] (14)

**Proof.** (i) Proof of this condition is deferred to the more lengthy treatment of the case \( \rho = 1 \) in Section 5.

(ii) Because the left-hand side of (13) is linear in \((\beta_2 - \beta_1)\) and the right-hand side is concave, the necessary and sufficient condition for the best symmetric wage scheme to be preferred to all asymmetric schemes is that the partial derivative of the left-hand side of (13) evaluated at \( \beta_2 - \beta_1 = 0 \) be greater than the equivalent partial derivative of the right-hand side. The result in (14) follows directly, since

\[ \frac{\partial S(\beta_2 - \beta_1)}{\partial (\beta_2 - \beta_1)} = \partial S(\beta_2 - \beta_1)/\partial \beta_2. \]

**Q.E.D.**

The second condition of Proposition 2 requires some interpretation. It says that the more convex is the sampling function with respect to \( \beta_2 \) (or equivalently, the smaller is the sampling effect relative to the effort levels, by result S7), the more attractive is an asymmetric scheme. The intuition is as follows. Proposition 1 showed that the optimal asymmetric scheme involves increasing one agent’s effort by some amount
while decreasing the other agent’s effort by the same amount. If the sampling effect is very small, the lower-effort agent in the asymmetric scheme will only rarely provide the highest output. Therefore, the fact that his effort has declined does not affect the principal substantially. In the limiting case—perfectly correlated output shocks—the symmetric scheme is never optimal, as reflected in condition (i) of Proposition 2. When the sampling effect is very large, on the other hand, the low-effort agent will still produce the highest output with nonnegligible probability. Lowering his effort level (as dictated by the optimal asymmetric scheme) proves costly to the principal whenever the low-effort agent has the highest output. In the limiting case of $\sigma_e = \infty$, the asymmetric scheme is never optimal.

To conclude the general two-agent case, simple manipulations show that in the symmetric case, the principal’s expected net payoff is
\[ E\Pi_2^{ym} = S(0) + \frac{1}{4(1 + r\sigma_2^4(1 - \rho^2))} - 2\bar{\nu}. \]  

(15)

No simple equation for the principal’s payoff emerges for asymmetric schemes.

5. The principal’s choice of one versus two agents

- Sections 3 and 4 characterized the optimal one-agent and two-agent incentive schemes. Given that the choice of agents is endogenous in this problem, this section considers the conditions under which the principal prefers two agents to a single agent.

A fully general treatment of the question is ruled out because one cannot explicitly solve for the optimal asymmetric incentive scheme. Consequently, I limit the analysis of this section to the examination of four special cases: (1) risk neutrality \((r = 0)\), (2) symmetric two-agent schemes preferred to asymmetric two-agent schemes, (3) perfect correlation of outputs \((\rho = 1)\), and (4) no correlation of outputs \((\rho = 0)\). These four cases highlight the substantial differences between the results of this model and the standard multiagent principal-agent literature. In what follows, the discussion is limited to the case where the principal’s payoff is nonnegative when one agent is employed. Although other situations can be examined with little alteration of the results, that case is inherently the most interesting.

□ Case 1: risk-neutral agents (optimal sampling). The assumption of risk-neutral agents eliminates the need to consider optimal risk sharing in the design of incentive schemes, making the principal’s problem an exercise in optimal sampling.

Proposition 3. (i) If both the principal and the agents are risk neutral, the first best is attainable.

(ii) A necessary (but not sufficient) condition for one agent to be optimal is

\[ \bar{\nu} + \frac{1}{4} > S(0) = .564\sigma. \]  

(16)

If one agent is optimal, \( \beta^* = 1 \) (the first-best effort level).

(iii) If the inequality in (16) does not hold, a symmetric scheme is optimal in the two-agent case if and only if

\[ \frac{\partial^2 S(0)}{(\partial \beta)^2} > \frac{1}{2}, \]  

(17)

with \( \beta_1^* = \beta_2^* = \frac{1}{2} \), the first-best level.

(iv) If both (16) and (17) fail to hold, an asymmetric scheme is optimal, with the gap between the two agents’ efforts given by

\[ \beta_1^* - \beta_2^* = 1 - 2 \frac{\partial S(\beta_2 - \beta_1)}{\partial \beta_2}. \]  

(18)

Proof: (i) Follows directly once (ii)–(iv) of this proposition are derived.

(ii) A necessary (but not sufficient) condition for one agent to be optimal is that \( E\Pi_1 \geq E\Pi_2^{ym} \). Comparing (5) and (15), the respective payoffs for the principal, yields (16) when \( r = 0 \). The approximate equality in (16) follows from result S4. \( \beta^* \) is trivially computed from (5).
(iii) The desired result, (17), is just a special case of (14), with the optimal incentive schemes coming from (11) and S2.
(iv) A special case of (12). Q.E.D.

The principal chooses one agent as long as the increased fixed costs associated with the second agent (the first term on the left-hand side) and the decline in output due to the lower effort level of each agent (second term on the left-hand side) outweigh the sampling effect. One agent looks relatively more attractive when reservation utilities are high relative to the variance of the agent-specific shocks.

In contrast to the first best in the standard incentive problem, an increase in the variance of output raises both the likelihood of adopting multiple agents and the principal’s expected profit if two agents are hired. Increasing output variance does not affect the wage payment to a risk-neutral agent, but it does increase the magnitude of the sampling effect.

Case 2: symmetric two-agent schemes vs. one-agent schemes. Assuming that the conditions of Proposition 2 are satisfied (i.e., in the two-agent case the principal prefers a symmetric incentive scheme), the necessary and sufficient conditions for the principal’s payoff to be greater employing one agent rather than two agents are given by Proposition 4.

Proposition 4. Assuming the conditions of Proposition 2 are satisfied, the principal will hire one agent to perform the task if and only if

$$\bar{\nu} + \frac{1}{2(1 + r\sigma^2)} - \frac{1}{4[1 + r\sigma^2(1 - \rho^2)]} > S(0) \approx .564\sigma_e. \quad (19)$$

Proof. If the conditions of Proposition 2 are satisfied, the necessary and sufficient condition for the principal to hire one agent is $$\Pi_1 > \Pi_2^{\text{nm}}$$. Equation (19) therefore follows directly from (5) and (15). Q.E.D.

Proposition 4 simply presents the conditions under which expected profits are higher with one agent as opposed to two symmetrically paid agents. Expected profits are decomposed into three pieces; the change in the fixed costs (i.e., the reservation utility of the second agent hired), the change in expected variable profits excluding the sampling effect (the second and third terms on the left-hand side), and the value of the sampling effect (the right-hand side). The combined effort exerted in the two-agent case ($$\beta_1^e + \beta_2^e$$) is strictly greater than the effort of an agent working alone, as long as $$\rho \neq 0$$—a consequence of the insurance value of relative performance evaluation.

It is also useful to note the conditions required for the first agent to increase his level of effort when a second agent is added. Algebraic manipulation of the optimal incentive schemes implied by Proposition 4 shows that the first agent faces higher-powered incentives with a second agent is introduced if and only if

$$\rho^2 \geq \frac{1}{2} + \frac{1}{2r\sigma^2_e}. \quad (20)$$

Intuitively, there are two effects to consider when determining the slope of the incentive scheme. The insurance effects of relative performance evaluation make it cheaper to entice an extra unit of effort from the agents. The principal, however, must compensate both agents for their extra effort but receives the benefit of that effort from only one of the agents. Unless the correlation between the agents’ output is relatively high, the first agent will be given lower-powered incentives in the two-agent case.
The comparative statics of Proposition 4 hinge upon whether the first agent’s optimal incentive scheme is higher-powered in the two-agent case or the one-agent case. If incentives are higher powered in the two-agent case, then an increase in either the variance of output (holding the correlation across agents constant) or risk aversion makes the two-agent solution more attractive because the insurance value of the second agent outweighs the other costs. If the incentive scheme for the first agent is lower-powered in the two-agent case than in the one-agent case (the more likely scenario), increasing risk aversion works against a second agent. The effect of increasing the variance of output is uncertain in this case. Sampling becomes more valuable, but the costs of enticing effort outweigh the insurance benefits. Finally, in contrast to the standard relative performance evaluation literature, increasing the correlation between agents has an ambiguous effect: it increases the insurance value but decreases the sampling effect. I devote further attention to that observation in Section 6 in the discussion of job design.

Case 3: perfect correlation of output across agents. The case of perfect correlation, while not particularly realistic, leads to a strong set of results. Perfect correlation has two important implications. First, all agent-borne risk can be eliminated by using multiple agents. Second, the sampling effect disappears because there are no individual-specific shocks.

Proposition 5. If output shocks are perfectly correlated across agents, then:

(i) If two agents are optimal, one agent exerts an effort level corresponding to the first best in the single-agent case, while the other exerts no effort at all. The first-best output is obtained, but at a greater cost than if an agent’s action were observable.

(ii) More than two agents will never be optimal.

(iii) If one agent is optimal, less than first-best effort is observed.

(iv) The necessary and sufficient condition for the principal to adopt two agents is

\[
\frac{r\sigma^2}{2(1 + r\sigma^2)} \geq \bar{\nu}.
\]

Proof. With perfectly correlated output shocks, two agents provide perfect insurance, just as in the standard multiagent problem. Also, with perfect correlation, the sampling effect disappears. The agent who exerts the highest effort will always have the highest output.

(i) Because there is no sampling effect, (12) leads directly to the desired result. The effort level of agent 1 corresponds to the first-best effort level observed in Proposition 3. If effort were observed directly, however, the principal would not have to pay \( \nu > 0 \) to the second agent to identify the shock.

(ii) Since \( \nu > 0 \), the addition of a third agent is costly. The third agent provides no benefit (neither sampling nor insurance) and so will not be hired.

(iii) This is an immediate result of the observation that a single agent will bear risk under the optimal one-agent incentive scheme.

(iv) This follows from (5) and (15). Q.E.D.

Proposition 5 says that it may pay the principal to have an agent around simply to provide a signal about another agent’s output, even though the low-effort agent’s output will never be used. A similar result is presented in Mookherjee (1984).

The specific results in Proposition 5 depend upon the assumption that all agents, regardless of their effort, are affected identically by the common shock. Although in the real world such a result is unlikely to hold exactly, the underlying intuition remains
basically valid. As long as there is some correlation between the shocks to the two agents, the hiring of a low-effort agent may be warranted based primarily upon the information content rather than the direct value of the agent’s output.

- **Case 4: no correlation of shocks across agents.** Although relatively few concrete conclusions can be drawn in this case, there are two points of interest. First, when the agents’ output shocks are uncorrelated, multiple agents may nonetheless benefit the principal due to the presence of sampling. Second, incentive schemes with multiple agents are unambiguously lower powered than when a single agent is employed, due to the absence of insurance effects.

The latter result highlights a difference between my model and the standard multiagent principal-agent problem. In the standard model, if output shocks are uncorrelated, the presence of other agents will not affect the optimal incentive scheme for a given agent. In my model, however, the existence of a second agent imposes an externality at the margin on the first agent, since that agent’s output is less likely to be of value to the principal. As a consequence, the optimal incentive scheme is lower powered because the marginal unit of a given agent’s output is less valuable to the principal.

6. **Implications for job/organizational design**

- The model has a number of implications for both job design and the structure and boundaries of firms, many of which are supported by the anecdotal evidence available. The most overarching of these predictions is that tasks with high value added and substantial variability in outcomes (such as R&D, advertising, or product design) should be organized in a very different manner than are low-value-added tasks (e.g., manufacturing, quality control, clerical) or jobs with little uncertainty over time or across agents (e.g., accounting). For tasks that have high value added and variability in outcomes, one would expect to observe substantial job overlap (numerous agents assigned to the same task) and compensation schemes based on performance relative to other agents working on similar tasks. Casual observation supports that dichotomy. In the advertising industry, creative activities are organized along precisely those lines, whereas account-management tasks at advertising agencies exhibit the standard hierarchy and controls. Similarly, contrast the structure of two government organizations, the Social Security Administration and the National Science Foundation. The tasks performed at the former have little uncertainty and low value added. As a result, the jobs are performed by civil servants with low-powered incentives and, at least in theory, little redundancy in tasks. The National Science Foundation, the goal of which is to promote basic research, uses a dispersed group of independent researchers to achieve its aims. Funding of multiple independent teams studying the same question is common. Career concerns ensure that the agents (researchers) have relatively high-powered incentives.

A second implication of the model is that it may be in the principal’s interest to take actions that reduce the correlation across agents’ output and/or increase the overall variance of output. (Such actions are never profitable in the standard principal-agent framework with risk-averse agents.) Increased variability in outcomes or reduced correlation across agents will increase the principal’s profits as long as the increased value of sampling outweighs the rise in the wage bill due to higher risk-bearing costs. Consistent with this prediction, IBM has a formal policy encouraging multiple approaches to problems (Peters and Waterman, 1982). Similarly, ICI’s development of a commercially viable process for manufacturing its first systemic fungicide was the result of a competition between two divisions with contrasting scientific experiences and approaches (Loasby, 1986). Finally, this result also explains the move by the Department
of Defense away from procurement contracts with very strict product specifications toward broader requests based upon performance attributes.

My model also implies differential treatment of identical agents performing the same tasks.¹⁰ Leading examples might be the use of both in-house and external lawyers and consultants. Although in-house teams are generally less costly to the principal (due partly to lower-powered incentives but perhaps also to less human capital, a feature not included in the model), they are often less effective than outside experts who receive extremely high-powered incentives. The in-house team may provide a useful benchmark against which to measure the performance of the outsiders.

Finally, my model provides an explanation for the existence of specialized firms such as advertising agencies, consulting firms, and research houses. If multiple agents attempt to solve the particular problem at hand, there may be a number of viable solutions, only one of which is actually implemented. To the extent that the insights gained in pursuit of the answers to a specific problem are transferable to other related problems, a firm that specializes in solving a particular type of problem will have an advantage over a firm that faces such problems infrequently, providing a rationale for large firms to contract out such activities.

7. Concluding remarks

- This article has examined the principal-agent problem when only the best agent’s best output is used by the principal. A number of important differences emerge vis-à-vis the standard principal-agent problem. Even when identical agents perform identical tasks, the optimal incentive scheme will often differ across agents. As multiple agents are introduced in this model, the power of incentive schemes generally falls; in the standard problem, relative performance evaluation leads to higher-powered incentive schemes. Finally, the principal may want to increase the variance of the agents’ output or reduce the correlation of output across agents—a practice that is never optimal with risk-averse agents in the standard framework.

Multiple agents with redundant outputs are most likely to be observed in those activities characterized by high value added and substantial uncertainty. Casual empiricism suggests that such redundancy is quite common in external contracting and procurement settings. Within the firm, however, behavior of this kind is very much the exception rather than the rule. While part of that difference may be traced to the nature of activities undertaken internally versus externally, it may also be related to the same factors responsible for the use of low-powered incentives within the firm more generally.

References


¹⁰ This result differs from Holmström and Milgrom (1991), where identical agents are treated differentially, but only because they are assigned to activities in which the levels of monitoring differ.


