Competition for Agency Contracts

R. Preston McAfee; John McMillan

Competition for agency contracts

R. Preston McAfee*

and

John McMillan**

This article introduces a market for the services of agents into a principal-agent model. The principal and the potential agents are risk neutral. The contract trades off adverse selection against moral hazard. In a broad range of circumstances the optimal contract is linear in the outcome. In an incentive-compatible contract the more able is an agent, the larger is his contractual share of his marginal output; thus, a more able agent is induced to work at a rate closer to the first-best.

1. Introduction

■ When a principal contracts with an agent for the provision of some good or service, the contract often does not specify a fixed price. Instead, the contract contains provisions for price adjustment. For example, when a firm subcontracts some of its input requirements to another firm, or when a government contracts for the production of a public good, the contract often provides for the price to be adjusted in response to unpredictable changes in the supplier's costs. A salesman's remuneration is usually dependent, *via* commissions, on the amount he sells; but often he also receives a fixed sum regardless of the level of his sales. The contracts between publisher and author, franchisor and franchisee, landlord and share-cropper, and patent holder and licensee similarly often have both fixed and variable components.[1] As Arrow (1985, p. 44) notes, "a fee function is a significant departure from the arms-length fixed-price relation among economic agents usually postulated in economic theory." Typically, the fee functions used in practice are simple: contracts involving royalties or commissions make the payment linear in the output.

At first glance, it might seem that contractual provisions that lessen the agent's responsibility for his own actions, such as provisions for price adjustment in response to cost fluctuations, or fixed payments to salesmen, cannot be in the principal's interest. To some extent the costs incurred or the sales achieved depend upon the effort made by the agent:

[1] A share contract of some historical consequence was the agreement between the Spanish Crown and Christopher Columbus, under which Columbus was entitled to 10% of the output of the territories he discovered (Elliot, 1970, p. 61).

in other words, there is moral hazard. Such contractual provisions weaken the agent's incentives to act as the principal would want.

The principal-agent literature (Harris and Raviv, 1979; Holmström, 1979; Shavell, 1979; MacDonald, 1984; Holmström and Milgrom, 1987) provides one possible explanation for price-adjustment provisions. If the agent is more risk averse than is the principal, then it is in their mutual interest to share the burden of the risk. The principal, in designing the contract that is best for him, trades off risk sharing against moral hazard.

Often, however, the parties to a contract are large, and the risks associated with any one contract are small relative to their overall operations. In such cases contractual provisions for price adjustment cannot be explained as risk-sharing devices. This article offers an alternative principal-agent model, applicable when both principal and agent are risk neutral. The model provides an explanation of contractual provisions for price adjustment without appealing to risk aversion.

We introduce a market for the services of agents. Potential agents compete with each other for the contract with the principal. The potential agents have different types (for example, innate ability) that the principal cannot observe: in other words, there is adverse selection in the market for agents' services. The principal designs a contract that exploits the competition among the potential agents and induces them to reveal their types. Instead of trading off risk sharing against moral hazard, as in the usual principal-agent model, in this setting the contract trades off adverse selection against moral hazard.[2]

Arrow (1985, p. 48) in evaluating principal-agent theory notes that the principal-agent relationships observed in the real world differ from those predicted by principal-agent theory:

> Most importantly, the theory tends to lead to very complex fee functions. It turns out to be difficult to establish even what would appear to be common-sense properties of monotonicity and the like. We do not find such complex relations in reality.

Holmström and Milgrom (1987) make a similar point. In contrast, we shall show that the contract that, for the principal, optimally trades off adverse selection against moral hazard can be relatively simple.

In a broad set of circumstances, the predicted contract is linear in the observed outcome. The contract varies with the selected agent's type: in an incentive-compatible contract, the more efficient is an agent, the larger is his contractual share of his marginal output. Thus, the contract screens the agents according to their abilities, and a more efficient agent is induced to work at a rate closer to the first-best. But except when the selected agent has the highest possible efficiency, the contract elicits less than the ideal amount of effort. Despite the principal's ability to commit himself to his mechanism, the asymmetry of information leaves the selected agent with some of the gains from trade.

2. Adverse selection and moral hazard

■ A risk-neutral principal wishes to contract with a single agent. There are n risk-neutral potential agents who compete for the contract. The potential agents have different characteristics, denoted z (measuring, for example, efficiency, ability, or expected production cost). Each potential agent knows only his own type z; the other potential agents and the principal perceive types as being independently drawn from a distribution $G(z)$, with density $g(z)$. Thus, differences among the potential agents are assumed to reflect inherent differences in their productivity, so that, after learning that he has an unusually high or low z, a potential

---

[2] The tradeoff between adverse selection and moral hazard was also modelled in our earlier article (McAfee and McMillan, 1986). That model was more general than the present one in allowing the potential agents to be risk averse. But it considered only a restricted class of contracts, whereas the present analysis solves for the contract that is fully optimal given the informational constraints.

agent has no reason to revise his estimates of the others' z values.[3] Let the support of g be an interval $[z_0, z_1]$.

The principal, having announced the contractual payment terms, asks the potential agents to report their types. Denote by $\hat{z}$ an agent's reported type. On the basis of these reports, the principal selects one agent.

The agent who wins the contract chooses a level of effort (possibly multidimensional), which is not directly observable by the principal. Output x, which is observed by the principal, depends on the agent's effort and ability, as well as on a random variable. We leave implicit in the notation both the effort and the random shock, and depict the agent as choosing an expected level of output μ. We assume $C(\mu, z)$ is the cost, in monetary terms, to an agent of obtaining a given expected output μ when his type is z. We assume that higher types have a lower cost and a lower marginal cost: $C_z(\mu, z) \leq 0$, $C_{\mu z}(\mu, z) < 0$, where subscripts denote partial derivatives. We assume also that the marginal cost of output is positive and increasing: $C_\mu(\mu, z) > 0$, $C_{\mu\mu}(\mu, z) > 0$.[4]

The assumptions of risk neutrality and independent types imply that the principal would gain nothing by making the payment to the selected agent depend on the other potential agents' reports, or by requiring payments from the other potential agents. This is so since under these two assumptions any payment scheme that depends on the reports of the other potential agents can be replaced by one with the same expected payment, given the selected agent's type, without affecting the incentives.[5] Thus, we assume that the principal makes his payment to the selected agent depend only upon the agent's reported type and the observed outcome. Let $S(\hat{z}, x)$ denote the payment function to which the principal commits himself. Thus, the principal's return is $x - S(\hat{z}, x)$, and the selected agent's return is $S(\hat{z}, x) - C(\mu, z)$. The net (social) expected return is $\mu - C(\mu, z)$, which we presume is positive for some z, μ.

This problem involves both adverse selection (since the principal cannot observe the potential agents' types) and moral hazard (since the principal cannot observe the selected agent's action). The purpose of this article is to derive the principal's optimal payment function $S(\hat{z}, x)$.

3. The agent's optimization

■ In this section we shall examine the contract offered by the principal in terms of the action μ that an agent of type z chooses, that is, in terms of the function $\mu(z)$. We show that

---

[3] This is the independent-private-values assumption: for a thorough discussion of the meaning and the limitations of this assumption, see Milgrom and Weber (1982, pp. 1090–1098).

[4] We thank Paul Milgrom for the observation that depicting the agent as choosing expected output simplifies the analysis. This can be derived from a more familiar moral-hazard formulation as follows. The agent who wins the contract chooses an action or effort $a \in R^n$. Assume that there is a function $h: R^{n+1} \to R$ such that the distribution of output x has the form $\hat{F}(x|h(a, z))$; that is, effort and type combine to form univariate measure of efficiency units of input $h(a, z)$. (An assumption of this sort was first used by Laffont and Tirole (1986).) Let $\bar{\mu}(h(a, z))$ be the expected value of x, given $h(a, z)$. Let $\phi(a, z)$ be the monetary cost to the agent of his action a. Now define the cost of μ by

$$C(\mu, z) = \min_a \phi(a, z)$$

subject to

$$\bar{\mu}(h(a, z)) = \mu.$$

If we assume that $\bar{\mu}(h)$ is increasing in h, we may write the distribution of outputs as a function of μ, $F(x|\mu) = \hat{F}(x|\bar{\mu}^{-1}(\mu))$. Thus, the efforts are entirely suppressed in the notation. The assumption that there is a univariate measure of efficiency units of inputs h effectively means that the problem is isomorphic to one where agents choose only the mean, as stated in the text.

[5] Compare with Harris and Raviv (1981), Milgrom and Weber (1982), Myerson (1981), and Riley and Samuelson (1981).

the function $\mu(z)$ can be used to characterize how much the agent is paid, without specifying the contract explicitly. This corresponds to the generalized principal-agent formulation of Myerson (1982), in which the principal recommends a decision $\mu(z)$ to the agent of type z, and the agent finds it in his interest to obey this recommendation.

Consider first the optimizing actions of the potential agents. As Myerson (1982) proved, by the revelation principle we can, without loss of optimality, presume that the potential agents are given an incentive to reveal their types honestly to the principal.[6]

Let $p_i(z_1, \ldots, z_n)$ be the probability that agent i is selected. Clearly, $0 \le p_i$ and $\sum_{i=1}^{n} p_i \le 1$. Then, given a report z_i, agent i is selected with probability:

$$P_i(z_i) = \int_{z_0}^{z_1} p_i(z_{-i}, z_i) \prod_{j \ne i} g(z_j) dz_1, \ldots, dz_{i-1} dz_{i+1}, \ldots, dz_n. \tag{1}$$

We shall show in the proof of Theorem 1 that $P_i(z_i) = G(z_i)^{n-1}$, that is, the principal invariably chooses the agent with the highest type. The *ex ante* expected utility of a potential agent of type z who reports his type to be $\hat{z}$ is therefore

$$\pi(\hat{z}, \mu, z) = [ES(\hat{z}, x) - C(\mu, z)] P_i(\hat{z}). \tag{2}$$

The incentive-compatibility condition, to be further examined in the next section, is

$$\pi(\hat{z}, \mu, z) \le \pi(z, \mu', z), \qquad \text{for all } (\hat{z}, \mu), \tag{3}$$

where μ' is the agent's optimal action, given that he reports z. That is, μ' satisfies $\pi_\mu = 0$ or[7]

$$\frac{\partial}{\partial \mu} ES(z, x) = C_\mu(\mu, z). \tag{4}$$

We assume that each potential agent has freedom not to participate, so that expected profits must be nonnegative.

Define a function A by

$$A(\mu, z) = \mu - C(\mu, z) + \frac{1 - G(z)}{g(z)} C_z(\mu, z), \tag{5}$$

and assume that the functions C and G are such that A is twice continuously differentiable. Define

$$\mu^*(z) = \arg \max_\mu A(\mu, z). \tag{6}$$

If $A(\mu^*(z_0), z_0) < 0$, let $\underline{z} = z_0$; otherwise let $\underline{z}$ be the smallest solution to $A(\mu^*(z), z) = 0$.

To understand the significance of the functions A and μ^*, note that since the principal keeps the total gain from trade minus the selected agent's profit, the principal's expected return from potential agent i is

$$\int_{\underline{z}}^{z_1} \{[\mu - C(\mu, z_i)] P_i(z_i) - \pi(z_i)\} g(z_i) dz_i. \tag{7}$$

From this and the potential agent's optimization, we obtain the Hamiltonian

$$\mathcal{H} = \{([\mu - C(\mu, z_i)] P_i(z_i) - \pi_i(z_i)\} g(z_i) - \lambda C_z(\mu, z_i) P_i(z_i). \tag{8}$$

---

[6] Although Myerson's proof applies to the case of a finite set of types, the same method of proof extends to the continuum case.

[7] Note that (4) presumes that the contract chosen by the principal satisfies $\partial^2 ES/\partial \mu^2 \le C_{\mu\mu}$. We shall see in the next section that the optimal linear contract satisfies this condition.

Now, $d\lambda/dz_i = -\partial H/\partial \pi = g(z_i)$. Integration of this, with the boundary condition $\lambda(z_1) = 0$, yields $\lambda = G(z_i) - 1$. Thus, from (5), we may write the Hamiltonian as

$$\mathscr{H} = A(\mu, z_i)P_i(z_i)g(z_i). \tag{9}$$

Thus, $A(\mu, z_i)g(z_i)$ is the coefficient of the control variable $P_i(z_i)$ in the Hamiltonian for the principal's problem; it therefore represents the marginal return to the principal from increasing the probability that type z_i wins, if it is present, when the incentive constraints are taken into account. This marginal return consists of the agent's output minus his cost, minus any rent he earns, and minus the increase in expected rents earned by the other types of agents. The latter effect arises because, by increasing the probability that a report of z_i will win the contract, the principal makes it more attractive that higher types pretend to be z_i. To prevent this the principal must improve the terms for all types higher than z_i, and thus raises their rents.

If potential agent i chooses his report to maximize his expected rents, the envelope theorem applied to (2) yields $d\pi/dz_i = -C_z(\mu, z_i)P_i(z_i)$. Thus, each unit increase in $P_i(z_i)$ increases the rents earned by all types higher than z_i by an amount equal to $-C_z(\mu, z_i)$. We can now interpret the expression $A(\mu, z_i)g(z_i)$. Consider the definition of the A function, equation (5). The first two terms, multiplied by $g(z_i)$, represent the direct increase in net revenue of increasing $P_i(z_i)$ while holding all of the other $P(z)$'s constant. The remaining term in (5), multiplied by $g(z_i)$, is $(1 - G(z))C_z(\mu, z)$. By the envelope-theorem result just stated, this term subtracts the increase in rents for all types higher than z_i, weighted by the frequency of such types. The sum of the direct effect and the incentive effect forms the coefficient for $P_i(z_i)$. Since this incentive term represents the rents the principal must pay the selected agent because of the potential agents' private information, we shall refer to it as an informational rent.

Given the interpretation that A represents the expected net return to the principal, one might predict that the principal's optimal choice of contract selects the best agent whose type exceeds $\underline{z}$ and induces expected output μ^*. Theorem 1 shows that this is indeed the case.[8]

Theorem 1. If first, $dA(\mu^*(z), z)/dz \geq 0$; second, the agent with the highest type is selected if his type exceeds $\underline{z}$, and otherwise no agent is selected; and third, the selected agent, with type z, is induced to choose $\mu = \mu^*(z)$, then the contract is optimal for the principal.

Note several implications of Theorem 1. First, the efficient level of effort is defined by $\mu_e(z)$ satisfying

$$1 - C_\mu(\mu_e(z), z) = 0. \tag{10}$$

Thus, since

$$1 - C_\mu(\mu^*(z), z) = -\frac{1 - G(z)}{g(z)}C_{\mu z}(\mu^*(z), z) > 0, \tag{11}$$

we have less effort than the full-information optimum for all but the highest type z_1.[9] The reason for this inefficient level of effort is that, as noted, the principal's information cost—the cost of the incentive-compatibility constraint—is $-(1 - G(z))C_z(\mu, z)/g(z)$, which is an increasing function of effort, μ. Thus, by lowering the induced effort level below the first-best, the principal captures some of the informational rents.

A second source of inefficiency arises because the principal will not contract with any agent whose type is less than $\underline{z}$. In a full-information world, in contrast, the principal would contract with any agent satisfying $\mu_e(z) - C(\mu_e(z), z) \geq 0$; that is, where there are positive

---

[8] The proofs of all of the following results are in the Appendix.

[9] Holmström (1983) and Sappington (1983, 1984) showed that the combination of adverse selection and moral hazard can lead to an inefficient outcome despite the risk neutrality of the agent.

gains from trade. The cut-off value $\underline{z}$ is excessive for two reasons. The principal wishes to reduce the informational rent that he gives the selected agent; in this respect $\underline{z}$ works exactly as a reserve price in an auction (Harris and Raviv, 1981; Laffont and Maskin, 1980; McAfee and McMillan, 1987; Milgrom and Weber, 1982; Myerson, 1981; Riley and Samuelson, 1981). In addition, there is an effect not present in the auction model, namely the agent's choice of effort. An insufficient amount of effort is evoked from a $\underline{z}$ agent and, as a result, $\underline{z}$ is larger than it would be if the principal were constrained to offer an efficient contract to any agent selected.

Third, the desired effort $\mu^*(z)$ and the cut-off $\underline{z}$ are invariant to the number of bidders n. Thus, the only effect of increasing n is to alter the distribution of selected types (being the largest order statistic from a sample of n); increasing n does not affect the extent of inefficiency, given the best agent.

Fourth, by equations (5) and (6),

$$0 = 1 - C_\mu(\mu^*(z), z) + \frac{1 - G(z)}{g(z)} C_{\mu z}(\mu^*(z), z). \tag{12}$$

Thus,

$$\frac{d\mu^*(z)}{dz} = \frac{C_{\mu z}\left(1 - \frac{\partial}{\partial z} \frac{1 - G(z)}{g(z)}\right) - \frac{1 - G(z)}{g(z)} C_{\mu z z}(\mu^*(z), z)}{-C_{\mu\mu}(\mu^*(z), z) + \frac{1 - G(z)}{g(z)} C_{\mu\mu z}(\mu^*(z), z)}. \tag{13}$$

Hence, a sufficient condition for expected output to increase with the selected agent's type $(d\mu^*(z)/dz \geq 0)$ is

$$\frac{\partial}{\partial z}\left(z - \frac{1 - G(z)}{g(z)}\right) \geq 0, \tag{14}$$

and

$$C_{\mu z z}(\mu, z) \geq 0. \tag{15}$$

4. Linear contracts

■ From Theorem 1 the principal maximizes profit if he finds any contract inducing honest reports, inducing a choice of effort $\mu = \mu^*(z)$, and selecting agents only of type $z \geq \underline{z}$. The main result of this article is that in a wide variety of circumstances this can be accomplished with a contract that is linear in output. We assume throughout this section that $A(\mu^*(z), z)$ is increasing in z (compare with Theorem 1).

Theorem 2. A contract of the form $S(\hat{z}, x) = \alpha(\hat{z})x + \beta(\hat{z})$ implements the optimum for all $n \geq 1$ if and only if

$$\frac{d}{dz} C_\mu(\mu^*(z), z) \geq 0. \tag{16}$$

If so,

$$S(\hat{z}, x) = C_\mu(\mu^*(\hat{z}), \hat{z})(x - \mu^*(\hat{z})) + C(\mu^*(\hat{z}), \hat{z}) - G(\hat{z})^{-(n-1)} \int_{\underline{z}}^{\hat{z}} G(s)^{n-1} C_z(\mu^*(s), s)ds. \tag{17}$$

The proof (in the Appendix) first shows that, if a contract is linear, it must have the form given in the theorem. Then it shows that the agent will select an action $\hat{\mu}$ that is a function of both his reported type $\hat{z}$ and his true type z. Finally, it shows that, given the condition (16), the contract (17) induces any potential agent to report his type z honestly.

The sharing term α equals C_μ. This has the commonsense interpretation that the agent equates his marginal cost of effort, C_μ, to the expected marginal benefit that accrues to him (since the agent keeps a fraction α of the expected output μ).

An unsurprising corollary of the theorem is that, since the contract allows a more able agent to keep a larger share of his marginal output, a more able agent will work harder.

Corollary 1. If $d\alpha/dz \geq 0$, then $d\mu^*(z)/dz \geq 0$.

The converse does not hold: only if $d\mu^*/dz$ is sufficiently positive does it follow that $d\alpha/dz \geq 0$ and that therefore a linear contract works.

Why is the contract linear? As shown by Theorem 1, the principal's problem reduces to finding an incentive-compatible contract that evokes the desired effort $\mu^*(z)$. In general, many different contracts will achieve this. The point of Theorem 2 is that, in the stated circumstances, among those many contracts is the linear contract.

Thus, for a contract linear in output to be optimal for the principal, the sharing term α must increase with type. In particular, a more able agent faces a larger α and thus bears a greater responsibility for his own actions. Incentive compatibility is the reason α must increase with ability. If α increases with ability, then it does not pay a potential agent with low ability to claim that he is more able than he really is, because he is then penalized by having a contract that makes his payment sensitive to his output; since he actually has low ability, his output, and therefore his payment, will be low. The same argument in reverse shows why an able potential agent would not claim to have low ability when α increases with ability.[10]

An agent who reports his ability level to be $\hat{z}$ and is selected is paid an amount $\beta(\hat{z}) + \alpha(\hat{z})x$ in exchange for his output x. The lump-sum term $\beta(\hat{z})$ can be interpreted as a salary or fixed payment; the sharing term $\alpha(\hat{z})$ can be interpreted as a royalty or commission rate or an incentive scheme. Although $\alpha(\hat{z})$ and $\beta(\hat{z})$ depend upon the agent's reported ability, they are fixed for the selected agent. The sharing term $\alpha(\hat{z})$ satisfies $0 < \alpha(\hat{z}) \leq 1$. It increases with the agent's ability. Thus, the more able is an agent, the larger the fraction of his marginal output he is allowed to keep. As a result, a more able agent is induced to work harder.[11] If the agent has the highest possible ability, then the principal sets his sharing term equal to unity. (From (5), (6), and (17), $\alpha(z_1) = 1$.) Thus, the most able possible agent keeps all of his marginal output and works at his first-best rate.

As the market for agents' services becomes perfectly competitive ($n \to \infty$), the selected agent's expected profit goes to zero. With an infinity of potential agents, the selected agent has the highest possible ability z_1 with probability one. In this case the agent keeps all of his marginal output. Then $\beta = C(\mu^*(z_1), z_1) - \mu^*(z_1)$. Thus, β is in this case negative: the agent pays for the right to do the work. The size of this payment is such that in expectation the agent produces just enough to cover his initial payment to the principal of β plus his cost.

Some further insight into when linear contracts work is given by the following result.

Corollary 2. The linear contract (17) is optimal if and only if

$$\frac{d}{dz}\left(\frac{1 - G(z)}{g(z)}\right)C_{\mu z}(\mu^*(z), z) \geq 0. \tag{18}$$

To understand this condition, recall that the principal's information cost is $-[1 - G(z)]C_z(\mu, z)/g(z)$. Thus, (18) says that the marginal information cost declines when,

---

[10] That the sharing provisions in contracts can serve as a screening device has been discussed before in the context of labor contracts by Pencavel (1977) and in the context of sharecropping contracts by Hallagan (1978).

[11] The size of the combined self-selection and effort effects can be estimated by using data from labor markets in which incentive schemes ($\alpha > 0$) and fixed salaries ($\alpha = 0$) coexist. Pencavel (1977), using data on punch-press operators, found that piece-rate workers earned on average 7% more than time-rate workers; Seiler (1984), using data from the clothing and shoe industries, estimated the earnings difference to be 14%. Marriott (1961, Ch. 6) presents a variety of statistical, case-study, and experimental evidence on the extent to which incentive payments increase output. He reports output gains of 50% or more.

adjusting effort optimally, ability is increased. Information costs become relatively less important as the ability of the selected agent increases.

Condition (18) involves third derivatives such as $C_{\mu zz}$ and $C_{\mu\mu z}$, so that little more can be said about it in general. For a simple special case, suppose that $C(\mu, z) = (1 - z)\mu - \mu^2$. Then (18) reduces to the relatively mild requirement that the inverse hazard rate $[1 - G(z)]/g(z)$ be decreasing in z. For this example, one can show that the difference between the equilibrium output and the first-best output is equal to the inverse hazard rate; and the deviation of equilibrium effort from the first-best declines as ability increases.[12]

5. Nonlinear contracts

■ We digress now to examine the nature of the contract when the principal is forced to use a nonlinear contract because (16) is not satisfied. The advantage of a nonlinear contract is that it allows the principal to use more of his knowledge about the distribution of x in inducing the agent to choose the desired action $\mu^*(z)$.

For the purposes of illustration, we consider in this section only the extreme case in which the outcome x is nonstochastic: thus $x = \mu$. In this case the principal can condition payment on the choice of effort: thus, he pays the agent only when $\mu = \mu^*(\hat{z})$, given the agent's report $\hat{z}$. This forces an agent of type z who reported his type to be $\hat{z}$ to choose an effort $\hat{\mu} = \mu^*(\hat{z})$.

Theorem 3. Incentive compatibility is satisfied regardless of the number of potential agents n if and only if

$$\mu^{*\prime}(z) \geq 0. \tag{19}$$

Thus (by comparison with Theorem 2), when

$$0 \leq \mu^{*\prime}(z) < -\frac{C_{\mu z}(\mu^*(z), z)}{C_{\mu\mu}(\mu^*(z), z)}, \tag{20}$$

there may be nonlinear contracts that implement μ^*, but no linear contracts. Even when x is stochastic, contracts such as forcing contracts (Holmström, 1982) may implement μ^* when linear contracts are not feasible.

The uncertainty in x strengthens the case for using linear contracts, since the informational requirements (the principal need only know the mean of x, and not the entire distribution) are reduced relative to other contracts.

6. Conclusion

■ This model predicts that a range of different contracts, optimally linear, will coexist in the market for agents' services. The sharing provision in a contract both screens the potential agents and evokes effort from the selected agent. There is, however, a tradeoff: the greater

---

[12] Laffont and Tirole (1986) analyzed the case of a single agent with adverse selection and moral hazard and showed that the contract is linear in the agent's type. Laffont and Tirole (1985) analyzed competition for contracts in a model similar to the present model, except that the payment to the winning bidder depends not only on the winning bidder's reported ability, but also on the second-highest reported ability: this scheme has the advantage that, like the Vickrey auction, the bidding game has a dominant equilibrium. In the model of Riordan and Sappington (1987), the potential agents bid for the contract after having observed imperfect signals of their types; only after the contract is awarded and fixed costs are incurred does the winning agent learn his type. Results similar to those of Laffont and Tirole and the present article are obtained, except that the two-stage structure of the adverse selection and the absence of moral hazard result in the optimal contract being nonlinear. Rogerson (1987) gave an alternative derivation of the linear-contracts result. Melumad and Reichelstein (1986) exhibited circumstances under which having the agent report his type does or does not improve the efficiency properties of the contract. Baron and Besanko (1987) examined the effects of the agent's risk aversion in a contracting model with both adverse selection and moral hazard.

the diversity in the potential agents' abilities, the less successful is the contract in inducing effort. In an incentive-compatible contract, the more able the agent, the larger the fraction of his marginal output he keeps, and thus the harder he works. Thus, both earnings and output increase with the sharing term.[13]

The striking feature of the contract is that in a broad range of circumstances there is an optimal bidding-contracting arrangement in which the contractual payments are linear functions of the output. Linear contracts have obvious practical advantages in ease of use. For this reason, and because of the analytical simplicity of linear contracts, contracting models sometimes impose linearity as an *ad hoc* assumption (McAfee and McMillan, 1986; Stiglitz, 1974; Weitzman, 1980). This article shows that there are cases in which the optimal contract is linear in the output (although it is not linear in the agent's ability). Payment schemes used in practice that involve royalties or commissions are usually linear in the outcome.

Appendix

■ Proofs of Theorems 1–3 and Corollaries 1 and 2 follow.

Proof of Theorem 1. From the envelope theorem and (2)

$$\frac{d\pi}{dz} = \frac{\partial \pi}{\partial z}\bigg|_{\hat{z}=z} = -C_z(\mu, z)P_i(z). \tag{A1}$$

Thus, the principal's expected profit is

$$v = \sum_{i=1}^{n} \int_{z_0}^{z_1} [\mu(z) - ES(z, x)]P_i(z)g(z)dz$$

$$= \sum_{i=1}^{n} \int_{z_0}^{z_1} [\mu(z) - C(\mu(z), z)]P_i(z)g(z)dz - \sum_{i=1}^{n} \int_{z_0}^{z_1} \pi_i(z)g(z)dz$$

$$= \sum_{i=1}^{n} \int_{z_0}^{z_1} [\mu(z) - C(\mu(z), z)]P_i(z)g(z)dz + \sum_{i=1}^{n} \left[\pi_i(z)(1 - G(z))\bigg|_{z_0}^{z_1} - \int_{z_0}^{z_1} (1 - G(z))\frac{d\pi_i}{dz}dz \right]$$

$$= \sum_{i=1}^{n} \int_{z_0}^{z_1} \left[\mu(z) - C(\mu(z), z) + \frac{1 - G(z)}{g(z)} C_z(\mu(z), z) \right] P_i(z)g(z)dz - \sum_{i=1}^{n} \pi_i(z_0)$$

$$= \sum_{i=1}^{n} \int_{z_0}^{z_1} A(\mu(z_i), z_i)P_i(z_i)g(z_i)dz_i - \sum_{i=1}^{n} \pi_i(z_0)$$

$$= \int_{z_0}^{z_1} \cdots \int_{z_0}^{z_1} [\sum_{i=1}^{n} A(\mu(z_i), z_i)p_i(z_1, \ldots, z_n)] \prod_{j=1}^{n} g(z_i)dz_1 \ldots dz_n - \sum_{i=1}^{n} \pi_i(\mu(z_0), z_0). \tag{A2}$$

From simple pointwise optimization

$$\pi_i(\mu(z_0), z_0) = 0,$$

$$p_i(z_1, \ldots, z_n) = \begin{cases} 1 & \text{if} \quad (\forall_j)z_i \geq z_j \quad \text{and} \quad A(\mu(z_i), z_i) \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

and $\mu^*(z) = \arg\max_{\mu} A(\mu, z)$.   Q.E.D.

---

[13] A system used by IBM for remunerating salesmen reported by Gonik (1978) resembles this menu of linear contracts. Different sales territories have different potentials (z in our notation); also, actual sales vary with the salesman's effort. Each salesman is asked to forecast sales in his territory. The commission rate and the fixed payment vary discontinuously with the forecast. The commission rate is set at one of two levels, with the higher rate applying when the forecast is relatively high; with the higher commission rate, the fixed payment is relatively small.

Proof of Theorem 2. First, consider the implications of linearity. The agent who reports $\hat{z}$, has type z, and chooses effort μ obtains rents

$$\pi = [ES(\hat{z}, x) - C(\mu, z)]G(\hat{z})^{n-1}$$

$$= [\alpha(\hat{z})\mu + \beta(\hat{z}) - C(\mu, z)]G(\hat{z})^{n-1}. \tag{A3}$$

Thus, necessarily, if the agent is honest and the contract induces effort $\mu = \mu^*(z)$,

$$\alpha(z) - C_\mu(\mu^*(z), z) = 0. \tag{A4}$$

By the envelope theorem,

$$\frac{d\pi}{dz} = -C_z(\mu^*(z), z)G(z)^{n-1}. \tag{A5}$$

Integration of (A5) with $\pi|_{z} = 0$ and (A3) and (A4) yield (17). Thus, we may restrict attention to contracts of the form (17). We first show that (16) is sufficient, and then necessary.

Sufficiency. Assume that $d(C_\mu(\mu^*(\hat{z}), \hat{z}))/d\hat{z} \geq 0$. If the contract (19) is offered, the agent's rents are

$$\pi = [C_\mu(\mu^*(\hat{z}), \hat{z})(\mu - \mu^*(\hat{z})) + C(\mu^*(\hat{z}), \hat{z}) - C(\mu, z)]G(\hat{z})^{n-1} - \int_{\underline{z}}^{\hat{z}} G(s)^{n-1}C_z(\mu^*(s), s)ds \tag{A6}$$

and

$$\frac{\partial\pi}{\partial\mu} = [C_\mu(\mu^*(\hat{z}), \hat{z}) - C_\mu(\mu, z)]G(\hat{z})^{n-1}$$

$$\frac{\partial^2\pi}{\partial\mu^2} = -C_{\mu\mu}(\mu, z)G(\hat{z})^{n-1} < 0. \tag{A7}$$

Thus, the agent chooses $\hat{\mu}$ satisfying

$$C_\mu(\mu^*(\hat{z}), \hat{z}) = C_\mu(\hat{\mu}, z), \tag{A8}$$

$$\frac{d\hat{\mu}}{d\hat{z}} = \frac{\frac{d}{d\hat{z}}C_\mu(\mu^*(\hat{z}), \hat{z})}{C_{\mu\mu}(\hat{\mu}, z)} \geq 0, \tag{A9}$$

$$\frac{d\hat{\mu}}{dz} = -\frac{C_{\mu z}(\hat{\mu}, z)}{C_{\mu\mu}(\hat{\mu}, z)} \geq 0, \tag{A10}$$

$$\frac{\partial\pi}{\partial\hat{z}}\bigg|_{\mu=\hat{\mu}} = G(\hat{z})^{n-1}\left[\left(\frac{\partial}{\partial\hat{z}}C_\mu(\mu^*(\hat{z}), \hat{z})\right)(\hat{\mu} - \mu^*(\hat{z})) + C_\mu(\mu^*(\hat{z}), \hat{z})\left(\frac{d\hat{\mu}}{d\hat{z}} - \mu^{*\prime}(\hat{z})\right) + \frac{\partial}{\partial\hat{z}}C(\mu^*(\hat{z}), \hat{z}) - C_\mu(\hat{\mu}, z)\frac{d\hat{\mu}}{d\hat{z}}\right]$$

$$+ (n-1)g(\hat{z})G(\hat{z})^{n-2}[C_\mu(\mu^*(\hat{z}), \hat{z})(\hat{\mu} - \mu^*(\hat{z})) + C(\mu^*(\hat{z}), \hat{z}) - C(\hat{\mu}, z)] - G(\hat{z})^{n-1}C_z(\mu^*(\hat{z}), \hat{z})$$

$$= G(\hat{z})^{n-1}\left[\left[\frac{\partial}{\partial\hat{z}}C_\mu(\mu^*(\hat{z}), \hat{z})\right](\hat{\mu} - \mu^*(\hat{z}))\right] + (n-1)g(\hat{z})G(\hat{z})^{n-2}[C_\mu(\mu^*(\hat{z}), \hat{z})(\hat{\mu} - \mu^*(\hat{z})) + C(\mu^*(\hat{z}), \hat{z}) - C(\hat{\mu}, z)]. \tag{A11}$$

Clearly,

$$\frac{\partial\pi}{\partial\hat{z}}\bigg|_{\hat{z}=z} = 0, \tag{A12}$$

$$\frac{\partial^2\pi}{\partial\hat{z}\partial z} = G(\hat{z})^{n-1}\left[\frac{\partial}{\partial\hat{z}}C_\mu(\mu^*(\hat{z}), \hat{z})\frac{\partial\hat{\mu}}{\partial z}\right] + (n-1)g(\hat{z})G(\hat{z})^{n-2}\left[[C_\mu(\mu^*(\hat{z}), \hat{z}) - C_\mu(\hat{\mu}, z)]\frac{d\hat{\mu}}{dz} - C_z(\hat{\mu}, z)\right] \geq 0. \tag{A13}$$

Thus,

$$\frac{\partial\pi}{\partial\hat{z}} \gtreqless 0 \quad \text{as} \quad \hat{z} \lesseqgtr z,$$

and this forces $\hat{z} = z$ to be optimal. Given $\mu = \mu^*(z)$ and $\hat{z} = z$, we then have

$$\pi = -\int_{\underline{z}}^{z} G(s)^{n-1}C_z(\mu^*(s), s)ds \gtreqless 0 \quad \text{as} \quad z \gtreqless \underline{z},$$

so that the correct agents are selected.

Necessity. It is clearly necessary that

$$\frac{\partial^2 \pi}{\partial \hat{z} \partial z}\bigg|_{\hat{z}=z} \geq 0$$

and that this hold for $n = 1$, since we required it to hold for all n. By (A13) this entails

$$\frac{\partial}{\partial \hat{z}} C_\mu(\mu^*(\hat{z}), \hat{z}) \frac{\partial \hat{\mu}}{\partial z}\bigg|_{\hat{z}=z} \geq 0$$

and by (A9) and (A10), we have

$$\frac{\partial}{\partial \hat{z}} C_\mu(\mu^*(\hat{z}), \hat{z}) \geq 0$$

is necessary.   *Q.E.D.*

Proof of Corollary 1.

$$0 \leq \frac{d}{dz} C_\mu(\mu^*(z), z) = C_{\mu\mu}\mu^{*\prime}(z) + C_{\mu z}.$$

Since $C_{\mu z} \leq 0$, $C_{\mu\mu} \geq 0$, $\mu^{*\prime} \geq 0$.   *Q.E.D.*

Proof of Corollary 2. From (6)

$$1 - C_\mu(\mu^*(z), z) + \frac{1 - G(z)}{g(z)} C_{\mu z}(\mu^*(z), z) = 0$$

so that

$$\alpha(z) = C_\mu(\mu^*(z), z) = 1 + \frac{1 - G(z)}{g(z)} C_{\mu z}(\mu^*(z), z)$$

is increasing if and only if

$$\frac{d}{dz} \frac{1 - G(z)}{g(z)} C_{\mu z}(\mu^*(z), z)) \geq 0.   \text{Q.E.D.}$$

Proof of Theorem 3. An agent who reports $\hat{z}$ obtains profits of

$$\pi(\hat{z}, z) = G(\hat{z})^{n-1}[C(\mu^*(\hat{z}), \hat{z}) - C(\mu^*(\hat{z}), z)] - \int_{\underline{z}}^{\hat{z}} G(s)^{n-1} C_z(\mu^*(s), s) ds$$

by (4) and integration. We have

$$\frac{\partial \pi}{\partial \hat{z}} = G(\hat{z})^{n-1}[[(C_\mu(\mu^*(\hat{z}), \hat{z}) - C_\mu(\mu^*(\hat{z}), z)]\mu^{*\prime}(\hat{z}) + C_z(\mu^*(\hat{z}), \hat{z})] - G(\hat{z})^{n-1} C_z(\mu^*(\hat{z}), \hat{z})$$

$$+ (n-1)G(\hat{z})^{n-2} g(\hat{z})[C(\mu^*(\hat{z}), \hat{z}) - C(\mu^*(\hat{z}), z)].$$

Clearly, $\partial \pi / \partial \hat{z}|_{\hat{z}=z} = 0$. Moreover,

$$\frac{\partial^2 \pi}{\partial z \partial \hat{z}} = -G(\hat{z})^{n-1}[C_{\mu z}(\mu^*(\hat{z}), z)\mu^{*\prime}(\hat{z})] + (n-1)G(\hat{z})^{n-2} g(\hat{z}) C_z(\mu^*(\hat{z}), z).$$

To obtain necessity, note that $\partial^2 \pi / \partial z \partial \hat{z}|_{\hat{z}=z} \geq 0$, for $n = 1$ is necessary, so that $\mu^{*\prime}(z) \geq 0$.

For sufficiency, note that if $\mu^{*\prime}(\hat{z}) \geq 0$, then $\partial^2 \pi / \partial z \partial \hat{z} \geq 0$ for all $(\hat{z}, z)$, so that $\partial \pi / \partial \hat{z} \gtrless 0$ as $\hat{z} \lessgtr z$, which forces $\hat{z} = z$.   *Q.E.D.*

References

ARROW, K.J. "The Economics of Agency" in J.W. Pratt and R.J. Zeckhauser, eds., *Principals and Agents,* Cambridge: Harvard Business School Press, 1985.

BARON, D.P. AND BESANKO, D. "Monitoring, Moral Hazard, Asymmetric Information, and Risk Sharing in Procurement Contracting." Mimeo, Stanford University, 1987.

ELLIOT, J.H. *Imperial Spain 1469–1716.* London: Penguin, 1970.

GONIK, J. "Tie Salesmen's Bonuses to Their Forecasts." *Harvard Business Review,* Vol. 56 (1978), pp. 16–23.

HALLAGAN, W. "Self-Selection by Contractual Choice and the Theory of Sharecropping." *Bell Journal of Economics,* Vol. 9 (1978), pp. 344–354.

HARRIS, M. AND RAVIV, A. "Optimal Incentive Contracts with Imperfect Information." *Journal of Economic Theory,* Vol. 20 (1979), pp. 231–259.

——— AND ———. "Allocation Mechanisms and the Design of Auctions." *Econometrica,* Vol. 49 (1981), pp. 1477–1500.

HOLMSTRÖM, B. "Moral Hazard and Observability." *Bell Journal of Economics,* Vol. 10 (1979), pp. 74–91.

———. "Moral Hazard in Teams." *Bell Journal of Economics,* Vol. 13 (1982), pp. 324–340.

———. "Managerial Incentive Problems: A Dynamic Perspective" in *Vetenskap och Foretagsledning,* Stockholm: Swedish School of Economics and Business Administration, 1983.

——— AND MILGROM, P. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica,* Vol. 55 (1987), pp. 303–328.

LAFFONT, J.-J. AND MASKIN, E. "Optimal Reservation Price in the Vickrey Auction." *Economics Letters,* Vol. 6 (1980), pp. 309–313.

——— AND TIROLE, J. "Auctioning Incentive Contracts." Mimeo, Massachusetts Institute of Technology, October 1985.

——— AND ———. "Using Cost Observation to Regulate Firms." *Journal of Political Economy,* Vol. 94 (1986), pp. 614–641.

MCAFEE, R.P. AND MCMILLAN, J. "Bidding for Contracts: A Principal-Agent Analysis." *RAND Journal of Economics,* Vol. 17 (1986), pp. 326–338.

——— AND ———. "Auctions and Bidding." *Journal of Economic Literature,* Vol. 25 (1987), pp. 699–738.

MACDONALD, G.M. "New Directions in the Economic Theory of Agency." *Canadian Journal of Economics,* Vol. 17 (1984), pp. 415–440.

MARRIOTT, R. *Incentive Payment Systems,* 2nd ed. London: Staples Press, 1961.

MELUMAD, N. AND REICHELSTEIN, S. "Value of Communication in Agencies." Research Paper 895, Graduate School of Business, Stanford University, August 1986.

MILGROM, P. AND WEBER, R.J. "A Theory of Auctions and Competitive Bidding." *Econometrica,* Vol. 50 (1982), pp. 1089–1122.

MYERSON, R.B. "Optimal Auction Design." *Mathematics of Operations Research,* Vol. 6 (1981), pp. 58–73.

———. "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems." *Journal of Mathematical Economics,* Vol. 10 (1982), pp. 67–81.

PENCAVEL, J. "Work Effort, On-the-Job Screening, and Alternative Methods of Remuneration" in R.G. Ehrenberg, ed., *Research in Labor Economics,* Vol. 1, Greenwich, Conn.: JAI Press, 1977, pp. 225–258.

RILEY, J.G. AND SAMUELSON, W. "Optimal Auctions." *American Economic Review,* Vol. 71 (1981), pp. 381–392.

RIORDAN, M.H. AND SAPPINGTON, D. "Awarding Monopoly Franchises." *American Economic Review,* Vol. 77 (1987).

ROGERSON, W.P. "On the Optimality of Menus of Linear Contracts." Mimeo, Northwestern University, January 1987.

SAPPINGTON, D. "Limited Liability Contracts between Principal and Agent." *Journal of Economic Theory,* Vol. 29 (1983), pp. 1–21.

———. "Incentive Contracting with Asymmetric and Imperfect Precontractual Knowledge." *Journal of Economic Theory,* Vol. 34 (1984), pp. 52–70.

SEILER, E. "Piece Rate vs. Time Rate: The Effect of Incentives on Earnings." *Review of Economics and Statistics,* Vol. 66 (1984), pp. 363–376.

SHAVELL, S. "Risk Sharing and Incentives in the Principal and Agent Relationship." *Bell Journal of Economics,* Vol. 10 (1979), pp. 55–73.

STIGLITZ, J.E. "Incentives and Risk Sharing in Sharecropping." *Review of Economic Studies,* Vol. 41 (1974), pp. 219–255.

WEITZMAN, M.L. "Efficient Incentive Contracts." *Quarterly Journal of Economics,* Vol. 44 (1980), pp. 719–730.