INFORMATION, INCENTIVES, AND ORGANIZATIONAL MODE*

MICHAEL H. RIORDAN AND DAVID E. M. SAPPINGTON

We examine the choice of organizational mode for a two-stage production process wherein cost realizations at each stage are observed only by the producing party. When these costs are positively correlated, the principal prefers to undertake second-stage production herself. When the correlation is negative and sufficiently small, she will prefer that the agent who performs the first stage also perform the second. For large negative correlation, either mode might be preferred. When costs are uncorrelated, the principal is indifferent between modes.

I. INTRODUCTION

An important concern in organizations is the assignment of tasks to individuals. In the absence of private information, the obvious assignment criterion is productive efficiency. But when individual decision-makers are privately informed, it also matters how the organization of tasks affects incentives. Moreover, the structure of private information itself may depend on the chosen mode of organization, especially when production and information-monitoring are complementary activities. For example, a decision-maker undertaking a particular productive task often is able to discern the difficulty of that task better than any other member of the organization. Consequently, the problem of creating proper incentives differs fundamentally under different organizational modes.

We investigate the relationship between private information and organizational design by considering a production process with two stages: for example, development and production. The Stackelberg leader in our model, the principal, must decide whether to perform the second stage herself, or purchase the second-stage output from the same agent who performs the first stage. The party that carries out production at any stage privately observes the realization of a random cost parameter. Cost realizations at the two

*We are grateful to the National Science Foundation for financial support (Grant Nos. IST83-17249 and IST83-15690); and to Chester Spatt; participants in seminars at Birkbeck College, Stanford University, and the Universities of Pennsylvania and Southern California; anonymous referees; and the Editorial Board for helpful comments. The views of this paper are not necessarily those of Bell Communications Research.

The Quarterly Journal of Economics, May 1987

stages may be correlated. The principal and agent are equally adept at performing the second stage of production, and both are risk neutral. Thus, the choice of organizational mode depends solely on considerations of private information, and not on technological advantage or differing attitudes toward risk.[1]

To illustrate, suppose that the principal is an entrepreneur contemplating production of a new product. Production requires a component input, which we call a "chipp." The performance of the final product, and hence its value to consumers, depends upon the purity of the chipp. Suppose further that the agent is the sole upstream manufacturer with the specialized expertise required to design a machine that produces chipps. The principal must decide whether to purchase this machine from the agent and produce chipps herself, or whether to subcontract with the agent for the production of chipps also.

We derive three main conclusions. First, the principal is indifferent between the two modes of organization if cost realizations at the two stages are independent. Second, the principal prefers to perform the second stage herself ("partial delegation") when the cost realizations are positively correlated. Third, the principal prefers to delegate second-stage production to the agent ("complete delegation") when the degree of negative correlation between the costs is sufficiently small; she might prefer either partial or complete delegation when the degree of negative correlation is large.

These conclusions stem from the interplay of two effects: the "linkage effect" and the "monitoring effect." The linkage effect arises because control of the second stage compounds (mitigates) the agent's incentive to overstate privately known first-stage costs when cost realizations at the two stages are positively (negatively) correlated. The monitoring effect arises because the output at the second stage can serve as a signal about the agent's private information. The precise nature of these effects is detailed in Section IV.

First, though, we describe the formal model in Section II, and state our results in Section III. Our main conclusions concern the choice of organizational mode. We also characterize optimal payment schedules under each mode. While our analysis of complete delegation resembles Baron and Besanko's [1984b] analysis of multiperiod regulation, our analysis of partial delegation is, to our

1. Following Coase [1937], we consider only the marginal transaction, i.e., second-stage production. We take as given that the agent performs the first stage and that the principal designs the contract.

knowledge, novel. In Section V we draw conclusions and discuss some extentions.

II. Description of the Model

Consider a profit-maximizing entrepreneur, the principal, who markets a single product. Revenues (r) depend upon the quantity (X) sold and the quality (Q) of the product according to the relationship $r = R(X,Q)$, satisfying the following assumption. (A subscript i hereafter denotes a partial derivative with respect to the ith argument.)

Assumption 1.

(a) $R(X,Q)$ is a strictly concave function of X and Q;

(b) $R_1(0,Q) = \infty$ and $R_1(\infty,Q) = 0, \quad \forall\, Q \geqq 0$;

(c) $R_2(X,0) = \infty$ and $R_2(X,\infty) = 0, \quad \forall\, X > 0$.

The revenue function is known by the principal and may or may not be known by the agent.

The product in question is manufactured in two stages, corresponding to quality and quantity choices. By assumption, the first-stage quality choice is delegated to the agent. The agent initially possesses perfect, private knowledge of the opportunity cost (q) of supplying each unit of quality. The principal's initial beliefs about q are described by a smooth distribution function $G(q)$. The corresponding density function $G'(q)$ has strictly positive support on the interval $[\underline{q},\overline{q}]$.

The level of quality supplied by the agent is observed costlessly by all parties, as is the ultimate quantity of output. The unit cost (c) of output becomes known only to the party who ultimately makes the output choice, and only after the quality is chosen. Prior to the agent's quality choice, it is common knowledge that c might take on any value in the interval $[\underline{c},\overline{c}]$, and that the cumulative distribution of c given q is $F(c|q;\alpha)$. $F(\cdot)$ is smoothly differentiable, and α parameterizes the statistical dependence of q and c. When $\alpha = 0$, q and c are independent. When $\alpha > 0$, smaller realizations of q imply that smaller realizations of c are systematically more likely in the sense of stochastic dominance ("positive correlation"). When $\alpha < 0$, smaller realizations of q are similarly associated with larger realizations of c ("negative correlation").[2]

2. We ignore the special case of perfect correlation, considered in Riordan and Sappington [1986].

ASSUMPTION 2. For all $(c,q) \in \text{int}\,[\underline{c},\bar{c}] \times [\underline{q},\bar{q}]$:

$$\text{(a)} \quad F_2(c|q;0) = 0;$$
$$\text{(b)} \quad F_{23}(c|q;\alpha) < 0.$$

(Hereafter we suppress α in our notation whenever possible.)

Either positive or negative correlation can arise under plausible scenarios. To illustrate, consider the chipp example again. Quality indexes the purity of the chipps produced. The amount of design effort required per unit of chipp quality is measured by q, and is privately known by the agent. The unit production cost for chipps, c, depends on the speed with which the machine that produces chipps can be operated. (This speed is not immediately apparent because it takes some time to learn how to operate the machine most efficiently.) There are three reasonable scenarios: speed is unrelated, positively related, or negatively related to the design effort required per unit of chipp quality. These possibilities correspond to $\alpha = 0$, $\alpha < 0$, and $\alpha > 0$, respectively.

The principal, acting as the Stackelberg leader, designs a contract that the agent can either accept or refuse. The contract states which party will make the quantity decision, and also specifies the magnitude of a monetary transfer, M, from principal to agent for any level of quality and output. Thus, under each mode, the contracting problem is essentially a "delegation problem" [Holmstrom, 1984].[3]

We now formalize the principal's problem in two different scenarios: (1) when the principal makes the quantity choice (partial delegation), and (2) when the agent makes the quantity choice (complete delegation). Then we compare the principal's expected profits under the two optimal contracts and characterize the principal's preferred mode of organization.

The principal's problem under partial delegation (PD) is given by

$$(PD) \quad \underset{M,Q,X}{\text{maximize}} \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - aX(a,b) - M(a,b)\} F_1(a|b) G'(b)\, da\, db$$

subject to, $\forall\ c,\hat{c} \in [\underline{c},\bar{c}]$, and $q,\hat{q} \in [\underline{q},\bar{q}]$:

$$\text{(i)} \quad \int_{\underline{c}}^{\bar{c}} M(a,q) F_1(a|q)\, da - qQ(q) \geqq 0;$$

3. Formalizing the problem as a delegation problem simplifies our analysis, but is restrictive in that it rules out explicit communication between the principal and the agent which may be valuable. See Section V.

(ii) $\int_{\underline{c}}^{\bar{c}} M(a,q)F_1(a|q)\,da - qQ(q)$

$$\geqq \int_{\underline{c}}^{\bar{c}} M(a,\hat{q})F_1(a|q)\,da - qQ(\hat{q});$$

(iii) $R(X(c,q), Q(q)) - cX(c,q) - M(c,q)$

$$\geqq R(X(\hat{c},q), Q(q)) - cX(\hat{c},q) - M(\hat{c},q).$$

This formalization of the principal's problem identifies the agent's choice of quality with a report $\hat{q} \in [\underline{q},\bar{q}]$, and the principal's subsequent choice of quantity with a report $\hat{c} \in [\underline{c},\bar{c}]$. Invoking the revelation principle [Myerson, 1979], we focus on contracts that induce truthful reports.

The individual rationality constraints (i) under partial delegation guarantee a nonnegative level of expected profits for all feasible first-stage costs. The self-selection constraints (ii) on the agent identify $Q(q)$ as the chosen quality level when first-stage unit costs are q. The self-selection constraints (iii) on the principal identify $X(c,q)$ as the chosen quantity level for costs q and c. The program also determines equilibrium payments $M(c,q)$, from the principal to the agent, as a function of costs.

The principal's problem under complete delegation (CD) is the following:

$$(CD) \quad \underset{M,Q,X}{\text{maximize}} \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - M(a,b)\}F_1(a|b)G'(b)\,da\,db$$

subject to $\forall\ c,\hat{c} \in [\underline{c},\bar{c}]$, and $q,\hat{q} \in [\underline{q},\bar{q}]$:

(i) $\int_{\underline{c}}^{\bar{c}} (M(a,q) - aX(a,q))F_1(a|q)\,da - qQ(q) \geqq 0;$

(ii) $\int_{\underline{c}}^{\bar{c}} (M(a,q) - aX(a,q))F_1(a|q)\,da - qQ(q)$

$$\geqq \int_{\underline{c}}^{\bar{c}} (M(a,\hat{q}) - aX(a,\hat{q}))F_1(a|q)\,da - qQ(\hat{q});$$

(iii) $M(c,q) - cX(c,q) - qQ(q) \geqq M(\hat{c},q) - cX(\hat{c},q) - qQ(q).$

The individual rationality (i), the self-selection at stage 1 (ii), and the self-selection at stage 2 (iii) constraints under complete delegation are analogous to those under partial delegation. The main difference is that, under complete delegation, the agent privately observes both cost realizations; consequently, all three constraints reflect the agent's incentives.[4]

If all information were public, there would be no incentive problem, as the principal could ensure a *first-best outcome* in which the agent's expected profits are zero and the following two conditions are satisfied $\forall\ q \in [\underline{q},\overline{q}]$ and $c \in [\underline{c},\overline{c}]$:

(i) $$\int_{\underline{c}}^{\overline{c}} R_2(X(a,q),\, Q(q))F_1(a\,|\,q)\; da = q;$$

(ii) $$R_1(X(c,q),\, Q(q)) = c.$$

The first-best outcome specifies a quality level that equates the marginal expected benefit with the marginal cost of quality (condition (i)), and a quantity level at which marginal benefits and costs of output are equal (condition (ii)). The principal is indifferent between the modes of organization here since she and the agent are equally adept at the second stage.[5]

Before proceeding to our findings in Section III, we record a standard regularity assumption on the distribution of q. (For a discussion see Baron and Besanko [1984a].)

ASSUMPTION 3. $q + G(q)/G'(q)$ is strictly in q.

III. STATEMENT OF FINDINGS

We now state our major conclusions regarding the choice of organizational mode in Propositions 1, 2, and 3. In Propositions 4 and 5 we characterize optimal contracts under partial and complete delegation, respectively, for a particular class of environments. An explanation of these results follows in Section IV. All proofs are in the Appendix.

The following definitions are helpful: $\Pi^{PD}(\alpha^0)$ (respectively, $\Pi^{CD}(\alpha^0)$) are the expected profits of the principal in the solution to (PD) (respectively, (CD)) when $\alpha = \alpha^0$.

PROPOSITION 1. $\Pi^{PD}(0) = \Pi^{CD}(0)$. That is, the principal is indifferent between organizational modes when the cost realizations at the two stages of production are statistically independent.

4. If the support of c is degenerate, i.e., $\overline{c} = \underline{c}$, so that there is no private information at the second stage, then problems (PD) and (CD) are equivalent, reducing to the one-stage regulation problem studied by Baron and Myerson [1982].

5. The first-best outcome is also feasible if only q is public or if the agent observes q only after the contract is designed [Riordan, 1984], and may also be feasible if c is publicly observed and $\alpha \neq 0$ [Baron and Besanko, 1984a; Riordan and Sappington, 1986].

Proposition 1 reports that, when the agent's private observation of q is uninformative about c, the principal's maximal expected profit is not affected by the choice of organizational mode. Proposition 2 reports that this conclusion does not extend to the case where the cost parameters are positively correlated.

PROPOSITION 2. $\Pi^{PD}(\alpha) > \Pi^{CD}(\alpha)$ for $\alpha > 0$. That is, when the costs parameters at the two stages are positively correlated, the principal prefers partial delegation to complete delegation.

Moreover, the proof of Proposition 2 (in the Appendix) shows that the principal prefers partial to complete delegation even if she cannot condition the agent's payment on the quantity produced under partial delegation. Thus, the principal will not allow the agent to control both quantity and quality when costs are positively correlated across stages. With negative correlation, though, this conclusion may be reversed, as reported next.

PROPOSITION 3. $\Pi^{CD}(\alpha) > \Pi^{PD}(\alpha)$ for $\alpha < 0$ and $|\alpha|$ sufficiently small. That is, when the cost parameters at the two stages are negatively correlated but the degree of correlation is small, the principal prefers complete delegation to partial delegation.

When α is negative but not in the neighborhood of zero, the principal's preferred mode of organization is difficult to specify in general. It is possible, though, to construct examples where, for sufficiently high degrees of negative correlation, the principal prefers partial delegation.[6]

The explanation for all of these findings follows from an examination of the solution to the principal's problem under partial and complete delegation. The propositions below characterize optimal contracts under each mode for sufficiently small degrees of correlation. The proofs in the Appendix make clear that the characterizations may also apply to cases of large correlation.

6. Suppose that q and c are binary random variables with $q_1 < q_2$ and $c_1 < c_2$. Let ϕ_{ij} represent the joint probability that $q = q_i$ and $c = c_j$; Q_i^* the first-best quality when $q = q_i$; X_{ij}^* the first-best quantity when $(q,c) = (q_i,c_j)$; $\alpha \equiv \phi_{11}\phi_{22} - \phi_{12}\phi_{21}$; and $Z \equiv [R(X_{21}^*,Q_2^*) - c_1 X_{21}^*] - [R(X_{22}^*,Q_2^*) - c_1 X_{22}^*]$. The following results are straightforward to verify: (1) the first-best outcome is feasible under complete delegation with $\alpha < 0$ if and only if

$$\frac{q_2 - q_1}{X_{21}^*[c_2 - c_1]} \leqq |\alpha| \leqq \frac{Q_1^*[q_2 - q_1]}{X_{12}^*[c_2 - c_1]};$$

(2) the first-best outcome is feasible under partial delegation with $\alpha < 0$ if and only if $[Q_1^*(q_2 - q_1)]/Z \leqq |\alpha|$. Consequently, for sufficiently high negative correlation, $[q_2 - q_1]$ small, and $[c_2 - c_1]$ large, the principal prefers partial to complete delegation.

PROPOSITION 4. If $|\alpha|$ is sufficiently small, the necessary and sufficient conditions for a solution to (PD) are, $\forall\ (c,q) \in [\underline{c},\bar{c}] \times [\underline{q},\bar{q}]$:

$$(4.1)\quad \int_{\underline{c}}^{\bar{c}} R_2(X(a,q),Q(q))F_1(a|q)\,da = q + \frac{G(q)}{G'(q)} - \int_{\underline{c}}^{\bar{c}} R_2(X(a,q),Q(q))F_{12}(a|q)\frac{G(q)}{G'(q)}\,da;$$

$$(4.2)\quad R_1(X(c,q),Q(q)) = c + \frac{F_2(c|q)G(q)}{F_1(c|q)G'(q) + F_{12}(c|q)G(q)};$$

and

$$(4.3)\quad M(c,q) = R(X(c,q),Q(q)) - cX(c,q) - \int_c^{\bar{c}} X(a,q)\,da - \int_{\underline{c}}^{\bar{c}} \{R(X(a,q),Q(q)) - aX(a,q)\}F_1(a|q)\,da + \int_{\underline{c}}^{\bar{c}} X(a,q)F(a|q)\,da - \int_q^{\bar{q}}\int_{\underline{c}}^{\bar{c}} \{R(X(a,b),Q(b)) - aX(a,b)\}F_{12}(a|b)\,da\,db + \int_q^{\bar{q}}\int_{\underline{c}}^{\bar{c}} X(a,b)F_2(a|b)\,da\,db + \int_q^{\bar{q}} Q(b)\,db + qQ(q).$$

PROPOSITION 5. If $|\alpha|$ is sufficiently small, then necessary and sufficient conditions for a solution to (CD) are, $\forall\ (c,q) \in [\underline{c},\bar{c}] \times [\underline{q},\bar{q}]$:

$$(5.1)\quad \int_{\underline{c}}^{\bar{c}} R_2(X(a,q),Q(q))F_1(a|q)\,da = q + \frac{G(q)}{G'(q)};$$

$$(5.2)\quad R_1(X(c,q),Q(q)) = c - \frac{F_2(c|q)G(q)}{F_1(c|q)G'(q)};$$

and

$$(5.3)\quad M(c,q) = cX(c,q) + \int_c^{\bar{c}} X(a,q)\,da - \int_{\underline{c}}^{\bar{c}} X(a,q)F(a|q)\,da - \int_q^{\bar{q}}\int_{\underline{c}}^{\bar{c}} X(a,b)F_2(a|b)\,da\,db + qQ(q) + \int_q^{\bar{q}} Q(b)\,db.$$

Propositions 4 and 5 enable a direct comparison with the first-best outcome. Under both modes, with $|\alpha|$ small so that $F_{12}(\cdot)$ can be ignored, it is apparent from equations (4.1) and (5.1) that too little quality is chosen relative to the first-best. There are also distortions in the quantity decision (see equations (4.2) and (5.2)), except in the special case of zero correlation. With positive correlation, too much quantity is chosen under partial delegation, while too little is chosen under complete delegation. With negative correlation, the converse is true.

The optimal contract under complete delegation with positive correlation is qualitatively similar to that characterized by Baron and Besanko [1984b] for a two-period regulation problem.[7] As they discuss, the first-stage distortion is essentially the same as that characterized by Baron and Myerson [1982] for a one-period regulation problem, while the second-stage distortion depends on the "informativeness" of the agent's privately known first-stage cost (q) about second-stage cost (c). The distortions in the first- and second-stage decisions are introduced to limit the rents the agent commands from his private information.

No changes in the first-stage distortions are introduced under complete delegation with negative correlation, but the second-stage distortions now involve too much rather than too little quantity. This qualitative difference arises from the "monitoring effect," discussed in the next section.

Different distortions at both stages arise under partial delegation. The quality distortion is particularly interesting, involving an additional term,

$$-\int_{\underline{c}}^{\bar{c}} R_2(X(a,q),\, Q(q))F_{12}(a|q)\; da\, \frac{G(q)}{G'(q)}.$$

Note that if $R_{12}(\cdot) = 0$ (so that, for example, quality does not affect marginal revenue), the entire term vanishes, and the quality distortions are the same as under complete delegation. If, on the other hand, $R_{12}(\cdot) > 0$ (so that higher quality levels increase marginal revenue), the optimal quality distortion is relatively greater with positive correlation and smaller with negative correlation. The converse is true if $R_{12}(\cdot) < 0$. In either case, if the degree of correlation is small, then $F_{12}(\cdot)$ is small, and so the quality distortion is "close" to the distortion under complete delegation.

7. Baron and Besanko impose additional individual rationality constraints at the second stage, assume that the principal's benefit function is additively separable ($R_{12}(\cdot) = 0$), and consider only positive correlation.

The second-stage quantity distortions also differ under partial delegation. In particular, the distortions are always in opposite directions under partial and complete delegation when the degree of correlation is small. This too follows from the monitoring effect discussed below. As under complete delegation, the distortions are introduced to limit the agent's information rents, and payments are structured to induce self-selection while satisfying individual rationality for the agent. The payment schedule under partial delegation (equation (4.3)) differs from that under complete delegation (equation 5.3)) because it is the principal's incentives that are being controlled at the second stage, rather than the agent's.

IV. Explanation of the Findings

The key to our results is that the agent's private information may be valuable in two respects. An observation of q reveals the unit cost of quality. It is also a signal about the realization of c, and as such provides information about the ultimate value of revenues.

Consider first the case in which c and q are independent. Since q provides no information about the distribution of c, the principal and agent agree on the expected value of a title to all revenues conditional upon any observed level of quality. Consequently, any incentive problem that arises after c is observed can be avoided under complete delegation by selling the title to all revenues to the agent for a prearranged fee which varies only with the chosen quality. Such an arrangement induces an efficient quantity decision.[8] Furthermore, because the fixed fee is set to ensure the agent zero expected rents from the title he receives, the agent has no incentive to misrepresent q other than what would prevail if the first stage of production were the only one.[9] Finally, the principal could provide exactly the same incentives if she were to retain the title to revenues and produce the output herself, basing the agent's compensation solely on the selected level of quality. Therefore, the principal is indifferent between modes.

The analysis is more complicated when the realizations of c and q are correlated. In this case, the agent's private observation of q endows him with private information about the value of a title to

8. This observation follows directly from Harris and Raviv's [1979] analysis of "franchise" contracts, and Theorem 2 in Baron and Besanko [1984b]. It is also evident from conditions (4.2) and (5.2) in Propositions 4 and 5; $F_2(c|q;0) = 0$, so marginal revenue and marginal cost are equated.

9. If there were only a single stage of production, the agent would be induced to produce too little quality relative to the first-best level for all realizations of $q > \underline{q}$ [Baron and Myerson, 1982].

revenues. Hence, the principal can no longer design a fee for the title that is sure to appropriate all of the expected rents from second-stage production. While the agent retains an incentive to overstate q in order to justify greater reimbursement for the first stage, he now has an additional incentive to misrepresent q in order to justify a lower fixed fee for the title to revenues at the second stage. When $\alpha > 0$, for example, this "linkage effect" under complete delegation provides the agent with further incentive to exaggerate q, because a large q indicates that c is also likely to be large, implying that the value of a title to revenues is relatively small. Hence, if cost parameters are positively correlated, the principal's task of controlling the agent's adverse incentives becomes more onerous when two stages of production are linked under complete delegation than when there is only a single stage.

On the other hand, if c and q are negatively correlated (i.e., if $\alpha < 0$), the linkage effect of complete delegation mitigates the agent's incentives to exaggerate q. Under negative correlation, the realization of c is more likely to be small if q is relatively large. Thus, when the agent exaggerates the value of q in order to increase his compensation for producing quality, he implicitly understates the likely realization of c. A small realization of c implies that the agent's compensation for performing the second stage of production need not be great.

On the basis of the linkage effect alone, the principal prefers partial delegation when the cost parameters are positively correlated, and complete delegation when the cost parameters are negatively correlated.[10] However, there is also a "monitoring effect."

Under partial delegation, the principal observes the realization of c directly, while under complete delegation, she learns of c only through the agent's report. Under both modes of organization, the optimal strategy for the principal is to treat implicitly the agent's report of q as a forecast about the subsequent report of c, and then to reward or penalize the agent based on the accuracy of that forecast. For example, the agent is penalized if his report of q indicates that c is likely to be high, and the report of c turns out to be low. By using the report of c as a signal about the actual value of q, the principal is able to lower the cost of inducing a truthful report of q under both modes.

There are moral hazard concerns under each mode of organization when the report on c is used as a monitor, since it is impossible

10. This fact is apparent from the proof of Proposition 2.

for both parties to verify directly whether the report is truthful. Under complete delegation, the agent might misrepresent c in order to conceal a previous false claim about q. Similarly, under partial delegation, the principal might misrepresent c in order to justify a forecast penalty being paid by the agent. In both cases, these adverse incentives create distortions from the first-best quantity. For example, under complete delegation with $\alpha > 0$, it is apparent from condition (5.2) that the agent has an incentive to exaggerate c, producing too little quantity relative to the first-best, in order to conceal a previous exaggeration of q.[11]

In general, the net benefits of employing the reported value of c as a monitor can be greater under either complete or partial delegation, depending on the relative extents of the second-stage moral hazard problem. However, when q and c are positively correlated, the debilitating linkage effect dominates any advantage of the monitoring effect under complete delegation, so the principal always prefers partial delegation (Proposition 2). The monitoring effect can at best only mitigate the linkage effect.

With negative correlation, even though the linkage effect under complete delegation is favorable to the principal, effective use of the monitor may be less costly under partial delegation (see note 6). However, if $|\alpha|$ is small, then the realization of c is not a very reliable monitor of whether the agent has truthfully reported q. Consequently, very large penalties for "incorrect" forecasts and large bonuses for "correct" ones will be necessary to eliminate the agent's rents while inducing truthful reporting of q. But when payments to the agent vary dramatically with the report on c, the second-stage moral hazard problem becomes particularly severe (under both modes). Consequently, when $|\alpha|$ is small, the costs of employing the monitor (i.e., second-stage distortions) are large relative to potential benefits, and the linkage effect dominates the monitoring effect.[12]

11. With $\alpha < 0$ under complete delegation, quantity is below the first-best level for analogous reasons: the agent understates c in order to conceal an earlier exaggeration of q. Similarly, with $\alpha > 0$, the quantity level will generally exceed the first-best level under partial delegation (though strong monitoring effects ($|F_{12}(\cdot)|$ large) might alter this conclusion). Since the agent's incentives are to overstate q, the principal's incentives will generally be to contradict the agent's report and understate c. Analogous arguments explain the optimal "underproduction" of quantity with $\alpha < 0$ under partial delegation.

12. Technically, the monitoring effect (i.e., the value of information) is of second-order in the neighborhood of $\alpha = 0$. See Singh [1982] and Radner and Stiglitz [1984].

V. CONCLUSIONS

We have analyzed how the choice of organizational mode is affected by the presence of private information. Our essential argument is that a "linkage effect" determines the principal's preferred mode when privately observed costs at vertically related production stages are either positively correlated or weakly negatively correlated. The linkage effect encourages the principal to integrate forward under positive correlation, and not to do so under negative correlation.

We have maintained the assumption that either the principal or the agent must undertake second-stage production, i.e., the principal cannot subcontract with a third party. Implicitly, we are assuming that any third party is at a substantial technological disadvantage. If this were not the case, the principal could achieve the first-best outcome by selling the title to revenues to the third party, and using the third party's report on c as an independent monitor of the agent's report on q. The agent would be assessed forecast penalties for reports of q that were inconsistent with the third party's subsequent report of c. But these penalties would be paid to the principal, not to the third party. Consequently, the third party would have no incentive to misrepresent c, and the second-stage moral hazard problem would be finessed entirely. Inasmuch, the third party's report of c is equivalent to a public monitor of the agent's private information q; thus, the principal can achieve the first-best outcome by appropriately designed forecast penalties paid by the agent [Baron and Besanko, 1984b; Riordan and Sappington, 1986].[13]

We have also maintained the assumption that the agent's payment function must depend implicitly only on the observed quantity and quality. In particular, we have ruled out additional (direct or indirect) communication between the principal and the agent. Conceivably, some form of "mediated communication" could be valuable under partial delegation (as in Myerson [1986] and Demski and Sappington [1987]). The principal might employ a

13. The first-best outcome is feasible because the principal breaks the budget-balance constraint between the agent and the third party [Holmstrom, 1982]. A related possibility, suggested by a referee, is that the first-best might be achieved by selling the entire enterprise to a third party, who then subcontracts with the agent and principal, respectively, for first- and second-stage production. This argument abstracts from the possibilities that the principal has private information about the revenue function or that the principal and agent have common private information about the probability distribution over costs—possibilities that our model allows.

third party to receive the agent's first-stage cost report privately and to "recommend" a quality choice according to a prespecified rule. If the recommended quality choice were the same for different reports of q (i.e., pooling were induced), then the principal would not be perfectly informed about q when she made her report on c. Consequently, the principal's incentives to truthfully report c might be improved by reducing her net payoff (possibly through output distortions) for reports of c that are inconsistent with the third party's subsequent (honest) revelation of the agent's original report. And with the second-stage moral hazard problem relieved to some extent, the principal's preferences might be tipped more heavily in favor of partial delegation.

We do not think that mediated communication by itself would alter our main conclusions. Clearly, mediated communication is not valuable if costs are independent at each stage. For positive correlation, partial delegation would still be preferred to complete delegation, as mediated communication at best strengthens preferences for partial delegation. And for small degrees of negative correlation, we conjecture that any advantages of mediated communication under partial delegation would be of second-order importance, so long as the mediator is not also empowered to make and receive transfer payments. Thus, the linkage effect would still dominate, and complete delegation would remain the preferred mode.[14]

In closing, we note briefly a few other elements that may be relevant in the choice of organizational mode. Aside from the roles of third parties and mediated communication, one might explore the possibility that the principal can choose the organizational mode after she observes the quality decision. Also, the complications introduced by risk aversion, bankruptcy, or renegotiation might be considered.[15]

14. If all payments by the principal must flow to the agent, then any attempt to resolve the second-stage moral hazard problem by imposing monetary penalties on the principal for forecast errors would compound the agent's adverse incentives to misrepresent q. And if quantity distortions were employed to control the principal, these would have to be large to be effective when the correlation between q and c is small; thus, the second-order importance of the gains from mediated communication. However, if the third party can also receive and make payments (thereby relieving the balanced-budget constraint), mediated communication might enable an outcome arbitrarily close to the first-best outcome under partial delegation. As a referee has suggested to us, the principal might be induced to truthfully reveal c by pooling reports of q over very small intervals while imposing large penalties on the principal (which are paid to the third party) when her report is inconsistent with the agent's report. With correlated costs, first-stage incentives might be resolved in similar fashion.

15. Williamson [1975] emphasizes that the optimal choice of organizational mode hinges on a broad comparative assessment of transaction costs, of which considerations arising from private information are a critical component.

APPENDIX

We prove Propositions 1 through 4. (The proof of Proposition 5 is analogous to that of Proposition 4.) The proofs are facilitated by modifying (*PD*) and (*CD*) by replacing the individual rationality and self-selection constraints with a single constraint on the payment function $M(\cdot)$.

The principal's modified problem under partial delegation (PD′) is

$$(PD') \quad \underset{X,Q,M}{\text{maximize}} \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - aX(a,b) - M(a,b)\} F_1(a \mid b) G'(b)\, da db$$

subject to (4.3).

The principal's modified problem under complete delegation (*CD′*) is

$$(CD') \quad \underset{X,Q,M}{\text{maximize}} \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - M(a,b)\} F_1(a \mid b) G'(b)\, da db$$

subject to (5.3).

The arguments underlying the proof of Proposition 1 are presented as a series of lemmas, and draw closely on those developed by Baron and Myerson [1982] for a single-stage contracting problem. The five lemmas jointly imply that the solution to (*PD′*) will be identical to the solution to (*PD*) if $|\alpha|$ is sufficiently small.

LEMMA A1. Constraint (iii) of (*PD*) is satisfied if and only if the following two conditions hold $\forall\, c,\hat{c} \in [\underline{c},\bar{c}]$, $\forall\, q \in [\underline{q},\bar{q}]$, and for some function $\phi(q)$:

$$\text{(A1.1)} \quad M(c,q) = R(X(c,q), Q(q)) - cX(c,q) + \int_{\underline{c}}^{c} X(a,q)\, da - \phi(q);$$

$$\text{(A1.2)} \quad \text{If } \hat{c} \geqq c, \text{ then } X(c,q) \geqq X(\hat{c},q).$$

The proof of Lemma A1 is analogous to the proof of Lemma 1 in Baron and Myerson [1982] and is omitted.

LEMMA A2. If constraint (ii) of (*PD*) is satisfied, then

$$\text{(A2.1)} \quad \int_{\underline{c}}^{\bar{c}} M(a,q) F_1(a \mid q)\, da = \int_{\underline{q}}^{q} \int_{\underline{c}}^{\bar{c}} M(a,b) F_{12}(a \mid b)\, da db - \int_{\underline{q}}^{q} Q(b)\, db + qQ(q) + M_0$$

for some constant, M_0, and $\forall\, q \in [\underline{q},\bar{q}]$.

Proof. Define $m(\hat{q}|q) \equiv \int_{\underline{c}}^{\bar{c}} M(a,\hat{q})F_1(a|q)\,da$, $\tilde{m}(q) \equiv m(q|q)$, $n(\hat{q}|q) \equiv m(\hat{q}|q) - qQ(\hat{q})$, and $\tilde{n}(q) \equiv n(q|q)$. Constraint (ii) of (PD) implies that $\tilde{n}(q) \geqq n(\hat{q}|q)\ \forall\ q,\hat{q} \in [\underline{q},\bar{q}]$. It is straightforward that

$$\tilde{n}(q) = \int_{\underline{q}}^{q} [m_2(b|b) - Q(b)]\,db + M_0 \tag{A2.2}$$

for some constant M_0. The lemma follows immediately from (A2.2) and the definitions of $m_2(b|b)$ and $\tilde{n}(q)$.

LEMMA A3. If constraints (ii) and (iii) of (PD) both hold, then

$$\begin{aligned} M(c,q) = {} & R(X(c,q), Q(q)) - cX(c,q) + qQ(q) \\ & - \int_{\underline{q}}^{q} Q(b)\,db + M_0 - \int_{\underline{c}}^{\bar{c}} \{R(X(a,q), Q(q)) \\ & - aX(a,q)\}F_1(a|q)\,da + \int_{\underline{q}}^{q}\int_{\underline{c}}^{\bar{c}}\{R(X(a,b), Q(b)) \\ & - aX(a,b)\}F_{12}(a|b)\,da\,db - \int_{c}^{\bar{c}} X(a,q)\,da \\ & + \int_{\underline{q}}^{q}\int_{\underline{c}}^{\bar{c}} [X(a,q) - X(a,b)]F_2(a|b)\,da\,db. \end{aligned} \tag{A3.1}$$

Proof. Substituting (A1.1) into (A2.1) yields

$$\begin{aligned} \tilde{m}(q) = {} & \int_{\underline{q}}^{q}\int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) \\ & - aX(a,b)\}F_{12}(a|b)\,da\,db - \int_{\underline{q}}^{q}\int_{\underline{c}}^{\bar{c}} X(a,b)F_2(a|b)\,da\,db \\ & - \int_{\underline{q}}^{q} Q(b)\,db + qQ(q) + M_0. \end{aligned} \tag{A3.2}$$

The derivation of (A3.2) entails reversing the order of integration and noting that $\int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \phi(q)F_{12}(a|b)\,da\,db = 0$. Also, using integration by parts, we have from (A1.1),

$$\begin{aligned} \tilde{m}(q) = {} & \int_{\underline{c}}^{\bar{c}} \{R(X(a,q), Q(q)) - aX(a,q)\}F_1(a|q)\,da \\ & + \int_{\underline{c}}^{\bar{c}} X(a,q)(1 - F(a|q))\,da - \phi(q). \end{aligned} \tag{A3.3}$$

The lemma then follows from (A1.1), after equating the right-hand sides of (A3.2) and (A3.3).

Lemma A4 refers to the following condition: $\forall\ q,\hat{q} \in [\underline{q},\bar{q}]$,

$$\text{(A4.0)}\quad 0 \geq \int_{q}^{\hat{q}} \int_{\underline{c}}^{\bar{c}} [\{R(X(a,b), Q(b)) - aX(a,b)\} - \{R(X(a,\hat{q}), Q(\hat{q})) - aX(a,\hat{q})\}]F_{12}(a|b)\, dadb + \int_{q}^{\hat{q}} \int_{\underline{c}}^{\bar{c}} [X(a,\hat{q}) - X(a,b)]F_2(a|b)\, dadb + \int_{q}^{\hat{q}} (Q(\hat{q}) - Q(b))\, db.$$

LEMMA A4. Suppose that (A1.2) and (A4.0) hold. If (A3.1) is satisfied, then constraints (ii) and (iii) of (PD) are also satisfied.

Proof. The proof follows from manipulating the definitions in the proof of Lemma A2, and from (A3.1), (A4.0), and Lemma A1.

Note that if $|\alpha|$ is sufficiently small, then (A4.0) is equivalent to

$$\text{(A4.1)}\qquad \text{If } \hat{q} \geq q, \qquad \text{then } Q(q) \geq Q(\hat{q}).$$

Therefore, Lemma A4 implies that, for $|\alpha|$ sufficiently small, (A3.1) is sufficient for constraints (ii) and (iii) of (PD) to hold if monotonicity conditions (A1.2) and (A4.1) also hold.

Lemma A5 refers to the following two conditions:

$$\text{(A5.0)}\quad M_0 \geq \int_{\underline{q}}^{\bar{q}} Q(b)\, db + \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} X(a,b)F_2(a|b)\, dadb - \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - aX(a,b)\}F_{12}(a|b)\, dadb;$$

$$\text{(A5.1)}\qquad Q(q) > 0 \qquad \forall\ q \in [q,\bar{q}].$$

LEMMA A5. If conditions (A5.0) and (A5.1) are satisfied, then constraint (i) of (PD) is satisfied for $|\alpha|$ sufficiently small.

Proof. If (A5.0) holds as an equality, then $\tilde{n}(\bar{q}) = 0$. It remains to show that $\tilde{n}'(\bar{q})$ is negative. From (A2.2) and (A5.1), $\tilde{n}'(q) = m_2(q|q) - Q(q) < 0$, since $m_2(q|q)$ goes to zero for $|\alpha|$ sufficiently small.

In summary, Lemmas A1–A5 imply that if $|\alpha|$ is sufficiently small and if conditions (A1.2), (A4.1), and (A5.1) hold, then the principal's problem under partial delegation (PD) is equivalent to choosing $X(\cdot)$, $Q(\cdot)$, and $M(\cdot)$ functions to maximize her expected profits subject to $M(\cdot)$ being determined by (A3.1) and (A5.0).

(A5.0) will hold as an equality at an optimum. Substituting this expression into (A3.1) yields the constraint on $M(\cdot)$ in the formulation of (PD'). Hence, if (A1.2), (A4.1), and (A5.1) hold, (PD) is equivalent to (PD') for $|\alpha|$ sufficiently small.

Proof of Proposition 1. Suppose that $\alpha = 0$. Then at the solution to (PD'), we have

$$\text{(A6.1)} \quad \int_{\underline{c}}^{\bar{c}} R_2(X(a,q), Q(q))F_1(a|q)\, da = q + \frac{G(q)}{G'(q)},$$

and

$$\text{(A6.2)} \quad R_1(X(c,q), Q(q)) = c.$$

Differentiating (A6.1) and (A6.2) with respect to q and combining the two results yields

$$\text{(A6.3)} \quad \frac{dQ}{dq} = R_{11}(\cdot)\frac{d}{dq}\left\{q + \frac{G(q)}{G'(q)}\right\} \Bigg/ \int_{\underline{c}}^{\bar{c}} \{R_{22}(\cdot)R_{11}(\cdot) - (R_{12}(\cdot))^2\}F_1(a|q)\, dc.$$

Differentiating (A6.2) with respect to c yields

$$\text{(A6.4)} \quad X_1(c,q) = (R_{11}(\cdot))^{-1}.$$

Hence, from (A6.3) and Assumption 1(a), (A4.1) holds strictly at $\alpha = 0$. And, from (A6.4) and Assumption 1(a), (A1.2) also holds at $\alpha = 0$. Therefore, by the smooth differentiability of $F(\cdot)$, and since $Q(\bar{q}) \geqq 0$, conditions (A1.2), (A4.1), and (A5.1) all hold at the solution to (PD') for $|\alpha|$ sufficiently small, and (PD) is equivalent to (PD'). A similar analysis establishes that (CD) is equivalent to (CD') for $|\alpha|$ sufficiently small.

Substituting the expression for $M(\cdot)$ into the objective function in (PD'), the principal's expected profits are equal to

$$\text{(A7.1)} \quad \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - aX(a,b)\}F_1(a|b)G'(b)\, dadb - \int_{\underline{q}}^{\bar{q}} Q(b)(bG'(b) + G(b))\, db - \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} X(a,b) \times F_2(a|b)G(b)\, dadb + \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - aX(a,b)\}F_{12}(a|b)G(b)\, dadb.$$

Similarly, substituting into the objective function in (CD'), the principal's expected profits are equal to

$$\text{(A7.2)} \quad \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - aX(a,b)\} F_1(a \mid b) G'(b)\, da db - \int_{\underline{q}}^{\bar{q}} Q(b)(bG'(b) + G(b))\, db + \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} X(a,b) \times F_2(a \mid b) G(b)\, da db.$$

We have already established that at $\alpha = 0$, the solutions to (PD) and (PD') are identical, as are the solutions to (CD) and (CD'). Also, from Assumption 2, $F_2(c \mid q;0) = F_{12}(c \mid q;0) = 0$. Hence, at $\alpha = 0$, (A7.1) and (A7.2) are identical, so $\Pi^{PD}(0) = \Pi^{CD}(0)$.

Q.E.D.

REMARK 1. The proof of Proposition 1 establishes that the constraint on (PD') is a necessary condition for constraints (i)–(iii) of (PD) to hold for all α. Therefore, the maximized value of (PD') places an upper bound on the maximized value of (PD) even for values of α that are not in the neighborhood of zero. For analogous reasons, the solution to (CD') places an upper bound on the solution of (CD). This fact plays a role in the proof of Proposition 2 that follows.

REMARK 2. The proof of Proposition 1 actually establishes that (PD') is equivalent to (PD) if (A1.2), (A4.0), and $\tilde{n}'(q) \leq 0$ all hold at the solution to (PD'). While this is certainly satisfied for α sufficiently close to zero, it may also be satisfied elsewhere.

Proof of Proposition 2. Let Π^{RPD} represent the expected profits of the principal in the solution of (PD') when the payment to the agent is restricted to be of the form $M(q,c) = \tilde{M}(q)$. (Note that Π^{RPD} is not a function of α.) We shall argue that $\Pi^{RPD} > \Pi^{CD}(\alpha)$, $\forall\ \alpha > 0$, and that the solution to our restricted partial delegation problem is a feasible solution to (PD).

The expression for Π^{RPD} can be shown to be given by the maximized value of the first three terms in (A7.1). The profits to the principal under (CD') are equal to the maximized value of (A7.2). Since the term $\int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} X(a,b)F_2(a \mid b)G(b)\, dadb$ is negative when $\alpha > 0$ by Assumption 2, it follows that the expected profits under (CD') must be less than Π^{RPD}.

As noted in Remark 1 above, the solution to (CD') places an upper bound on the principal's profits under (CD). Consequently

$\Pi^{RPD} > \Pi^{CD}(\alpha)$. Finally, it is straightforward to verify that the solution to the restricted partial delegation problem satisfies the monotonicity conditions (described in Remark 2) that must be satisfied in order to be feasible for (PD). Therefore, $\Pi^{PD}(\alpha) \geq \Pi^{RPD} > \Pi^{CD}(\alpha)$ for $\alpha > 0$.

Q.E.D.

Proof of Proposition 3. From equations (A7.1) and (A7.2), we have for α in the neighborhood of zero

$$\begin{aligned}\Delta(\alpha) &\equiv \Pi^{PD}(\alpha) - \Pi^{CD}(\alpha)\\ &= \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \{R(X(a,b), Q(b)) - aX(a,b)\} F_{12}(a|b;\alpha) G(b)\, da db\\ &\quad - 2\int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} X(a,b) F_2(a|b;\alpha)\, G(b)\, da db > 0.\end{aligned}$$

At $\alpha = 0$, $R_1(X(c,q), Q(q)) = c$, as shown in the proof of Proposition 1. Hence,

$$\begin{aligned}\Delta'(0) &= \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} \left[- \int_{\underline{c}}^{a} X(s,b)\, ds \right] F_{123}(a|b;0) G(b)\, da db\\ &\quad - 2\int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} X(a,b) F_{23}(a|b;0) G(b)\, da db\\ &= - \int_{\underline{q}}^{\bar{q}} \int_{\underline{c}}^{\bar{c}} X(a,b) F_{23}(a|b;0) G(b)\, da db > 0.\end{aligned}$$

The second equality is derived from reversing the order of integration. The last inequality follows from Assumption 2.

Q.E.D.

Proof of Proposition 4. The proof of Proposition 1 established that (PD) is equivalent to (PD') for $|\alpha|$ sufficiently small. Hence $M(\cdot)$ must satisfy (4.3). Conditions (4.1) and (4.2) are obtained by differentiating (A7.1).

Q.E.D.

STANFORD UNIVERSITY
BELL COMMUNICATIONS RESEARCH

REFERENCES

Baron, D., and D. Besanko, "Regulation, Asymmetric Information, and Auditing," *Rand Journal of Economics*, XV (1984a), 447–70.

——, and ——, "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, I (1984b), 267–302.

Baron, D., and R. Myerson, "Regulating a Monopolist with Unknown Costs," *Econometrica*, L (1982), 911–30.

Coase, R., "The Nature of the Firm," *Economica,* IV (1937), 386–405.
Demski, J., and D. Sappington, "Delegated Expertise," *Journal of Accounting Research,* (1987), forthcoming.
Harris, M., and A. Raviv, "Optimal Incentive Contracts with Imperfect Information," *Journal of Economic Theory,* XX (1979), 231–59.
Holmstrom, B., "Moral Hazard in Teams," *Bell Journal of Economics,* XIII (1982), 324–40.
——, "On the Theory of Delegation," in M. Boyer and R. H. Kihlstrom, eds., *Bayesian Models in Economic Theory* (Amsterdam: North-Holland Publishing Company, 1984).
R. Myerson, "Incentive Compatibility and the Bargaining Problem," *Econometrica,* LXVII (1979), 61–74.
——, "Multistage Games with Communication," *Econometrica,* LIV (1986), 323–58.
Radner, R., and J. Stiglitz, "Nonconcavity in the Value of Information," In M. Boyer and R. E. Kihlstrom, eds., *Bayesian Models in Economic Theory* (Amsterdam: North-Holland Publishing Company, 1984).
Riordan, M., "Uncertainty, Asymmetric Information and Bilateral Contracts," *Review of Economic Studies,* LI (1984), 83–93.
——, and D. Sappington, "Optimal Contracts with Public and Private Ex Post Information," Working Paper at Bell Communications Research and Stanford University, revised January 1986.
Singh, N., "The Effects of Asymmetric Information on Resource Allocation," Ph.D. thesis, University of California, Berkeley, 1982.
Williamson, O., *Markets and Hierarchies: Analysis and Antitrust Implications* (New York: The Free Press, 1975).