A Theory of Inefficient Intrafirm Transactions

By JULIO J. ROTEMBERG*

I consider a model in which the threat of customer departures induces sellers to supply high-quality goods. Permanent attachment of buyer and seller such that transactions take place inside a firm raises the social cost of delivering high quality. Yet, such costly integration is often profitable, because prices exceed marginal cost at equilibria where market transactions provide high quality. This theory can rationalize the empirical finding that middle managers are averse to transactions between profit centers. (*JEL* 611)

Unrelated individuals work together within firms where relationships of authority mold individual behavior. At least since Ronald Coase (1937), economists have viewed the abandonment of the market implied by these relationships as a socially desirable adaptation.[1] Oliver Williamson (1985 p. 17), for instance, states: "This book advances the proposition that the economic institutions of capitalism have the main purpose of economizing on transactions costs." In this view, the routing of transactions inside firms raises profits together with social welfare.

In this paper, I follow Coase (1937) in that I assume that integration and the concomitant abandonment of market transactions occur only when this is privately profitable. However, because product market imperfections exist, their private profitability is consistent with lack of social desirability. When goods markets are distorted, firms generally earn rents. Thus, firms would willingly tolerate organizational inefficiencies if such inefficiencies help them capture these rents. Some of the economy's rents may be thus dissipated, and this may help explain why conventional measurements of profits do not uncover much monopolistic conduct.

I focus on a situation in which it is difficult to ensure that the good bought by one agent from another is of high quality. Quality is something the buyer can recognize *ex post* but not something that can be stipulated in advance when the agents enter into contracts. As shown by Benjamin Klein and Keith Leffler (1981) spot markets deal well with this problem, at least when price exceeds marginal cost. Then, sellers in spot markets try hard to provide high quality for fear that customers will change suppliers whenever they have an unsatisfactory experience. Such fear only arises if the customer's future business leads to profits and if it reasonable to suppose that the customer will indeed leave when he is dissatisfied.

Even when suppliers try hard to provide high-quality goods, the good will sometimes fall short of the customer's expectation. This is particularly true when quality refers not to a technical characteristic, but to the way the good is delivered. An example of this is the supplier's ability to alter delivery schedules and specifications to accommodate changes in the customer's needs.

As in Franklin Allen and Gerald Faulhaber (1988), quality can fall below expectation for two reasons. The first is temporary bad luck which makes it impossible for the supplier to accommodate the buyer in this particular instance. Such temporary bad luck bears no relationship to the likelihood that

*Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139. I thank three anonymous referees, Oliver Hart, Jim Poterba, Jean Tirole, Garth Saloner, and participants at seminars at Boston University, M.I.T., the University of Chicago, and Yale University for comments and the NSF and Sloan Foundations for research support.

[1] See Armen Alchian and Harold Demsetz (1972) and Sanford Grosman and Oliver Hart (1986), as well as the references in Bengt Holmstrom and Jean Tirole (1989).

the seller will offer high quality in the future. The second reason is that the supplier is sometimes genuinely incompetent. Then, he is likely to continue delivering bad quality in the future. A dissatisfied customer will typically not know whether he has faced bad luck or incompetence. Fearing the second possibility, he may well leave. As a result, the producer will indeed try to keep the customer happy.

Long-term contracts that penalize the buyer when he changes sellers reduce the incidence of these departures. This mobility is also reduced if the buyer and seller integrate and the managers of the integrated enterprise mandate internal transactions. Both of these approaches make it more difficult to ensure high quality. The seller is now less worried about losing his customer, so it is harder to induce him to improve quality.[2] In addition, both approaches make it more likely that the buyer will continue to purchase from a seller whose competence has declined. That long-term contracts and integration are at least sometimes associated with low quality should be apparent to those familiar with the food services at most American universities.

The evidence of Robert Eccles and Harrison White (1988) also suggests that low quality and internal transactions go hand in hand. They interviewed managers who buy and sell goods across profit centers in firms with multiple profit centers. The impression left by these interviews is that managers prefer external to internal transactions.[3] They would rather buy from and sell to agents outside the firm. Internal transactions occur only when they are mandated by management. Frictions are much more common in these mandated transactions than in their external counterparts. Eccles and White justly regard their findings as a challenge to Williamson's (1985) view of firms. Their evidence suggests that transactions costs actually increase when firms engage in internal transactions.

If long-term contracts and integration make quality more difficult to deliver, why do they arise? The reason is that delivery of high quality in spot markets requires that price exceed marginal cost. Reductions in mobility thus raise the combined profits of buyer and seller. When the buyer leaves, he starts paying rents to others. These rents are captured by tying the buyer to the seller.[4]

The paper proceeds as follows. Section I presents a variant of the Klein and Leffler (1981) model in which the actual level of quality is stochastic even when the supplier is trying hard to accommodate the buyer. For simplicity, I focus here only on the randomness introduced by temporary bad luck on the part of the seller. Section II presents an illustrative long-term contract. The contracts described in this section are generally efficient. They both provide a benchmark for the inefficiencies studied in later sections and illustrate the rent-capture advantages of reducing buyer mobility.

Section III focuses on inefficient outright integration of the buyer and seller. Here, I suppose that there is an alternative mechanism that can also improve quality. This more costly mechanism works even if there is no threat of customer departure. For instance, costly managerial attention might be able to ensure high quality by making seller infractions verifiable by outside agents. Alternatively, costly managers might be given a great deal of discretion which, as in Alchian and Demsetz (1972), allows them to induce the seller to perform some quality-enhancing tasks. Even this may not raise quality as much as the effort brought forth by spot markets. Because this mechanism has higher cost per unit of quality, it is socially inefficient. Nonetheless, the wedge between price and marginal cost may make integration with these features privately profitable.[5]

[2]According to Coase (1988), he discusses this loss in quality from integration in his 1934 notes for his paper "A Theory of Contract."

[3]Much as university professors prefer to eat in establishments that are not run by their employer.

[4]For other models in which socially inefficient contracts arise because they help capture rents, see Philippe Aghion and Patrick Bolton (1987), Mathias Dewatripont (1988), and Janusz Ordover et al. (1990).

[5]The wasteful monitoring by managers of their suppliers appears to be similar to the wasteful monitoring of workers in Samuel Bowles (1985). There, managers waste resources monitoring to avoid having to pay

Section IV introduces imperfect competition which raises prices. With spot markets alone, such increases in price have a socially desirable result. They make it more likely that producers will supply high quality, because they now have more to lose when the customers depart. Unfortunately, these high prices do not diminish the benefits from integration or the benefits from signing a long-term contract that eliminates all buyer departures and has low quality. As a result, increases in prices raise the likelihood that long-term contracts and integration are inefficient relative to the spot-market outcome.

Section V focuses on a different inefficiency of equilibrium long-term contracts. This inefficiency arises when the aptitude of the seller varies stochastically. Because the seller sometimes becomes inept, the immediate departures of customers when they encounter bad quality are socially beneficial. They raise the average quality that customers receive. Nonetheless, long-term contracts that reduce buyer mobility and raise the likelihood of purchases from inept suppliers can raise profits, because price exceeds marginal cost. Section VI presents my conclusions.

I. Spot Markets in the Klein-Leffler Model with Free Entry

Sellers produce a homogeneous intermediate good. The marginal cost of a bad-quality version of the good is c. After spending ε, the quality of the good sometimes improves. I assume that a particular item becomes a high-quality good with probability σ if the firm spends an additional ε on the item. With probability $1-\sigma$, the quality of the item remains poor.

There are Y customers who are, each period, willing to pay up to r for one unit of low quality. They are willing to pay an additional μ for a unit of high quality. Customers are risk-neutral: they are willing to pay $r + x\mu$ for a unit which is of high quality with probability x.

Contracts cannot specify the level of quality. This incompleteness of contracts has several possible sources. One possibility is that outsiders are unable to observe quality, so that contracts stipulating payments as a function of quality are unenforceable. Another possibility is that quality is difficult to spell out when contracts are written, although buyer and seller both recognize it when goods change hands.

Therefore, the qualities of the good I am concerned with here are not easily measurable. Rather, they are things like courteous service and prompt handling of complaints. Another important quality attribute in which these problems arise is seller flexibility in meeting changes in the buyers needs. Consider the following. The buyer and seller originally write a contract for delivery at the beginning of December. The buyer then realizes that he would much prefer delivery in the middle of November. A high-quality seller is flexible, so that he is capable of changing the delivery at a low cost. A seller of bad quality finds it very costly to change the delivery date. The courts see only the *ex post* cost of changes in delivery terms and not the effort made in being flexible.[6]

In this section, I consider spot-market transactions in which the only payments are proportional to the number of items exchanged. Each purchasing transaction has three stages. First, all upstream suppliers announce their prices. Second, buyers choose sellers and agree to receive one unit at this price. Finally, delivery occurs either

workers an efficiency wage (which represents an excess of the wage over the disutility of labor). Our papers differ not only in focus but also in structure. In Bowles's model, agent A wastefully monitors agent B to extract rents from B. Assuming that contracts can specify the level of monitoring, the contract between A and B is suboptimal for A and B taken together. Overall efficiency would be restored if B were allowed to pay A ab initio to refrain from monitoring.

By contrast, in my model, A and B benefit jointly from writing a contract that involves monitoring. They both benefit because their arrangement extracts rents from C, a third party who is not involved in the contract. This difference is partly responsible for the models' different conclusions. Whereas in my model firms grow too big, they need never grow larger than their minimum efficient scale in Bowles.

[6]This example of quality is very reminiscent of Grossman and Hart (1986).

with low or high quality. Customers know what quality they have received, though not the quality received by others.[7]

Customers' beliefs about the quality they can expect to receive depend on the price charged and on their personal experience. There are different equilibria depending on the precise form of these beliefs. The worst equilibrium for suppliers has customers believing that goods are universally of bad quality. As long as $r \geq c$, there is then an equilibrium where all active suppliers charge c and provide bad quality. For a producer to deviate by providing either good quality or lower prices represents a simple gift to customers. For him to deviate by raising his price loses him all sales.

Equilibria with high quality have different beliefs and higher prices. At these equilibria, customers believe instead that essentially all producers who charge a relatively high price P try to provide high-quality goods. In equilibrium, various firms charge P and are initially regarded as identical. Customers thus pick their initial supplier at random.

If this supplier delivers high quality and continues to charge P, they come back to this supplier in the following period. By contrast, if he delivers low quality, they choose a supplier at random from among the other producers who charge P. The assumption that customers leave suppliers from whom they have received low quality while they stay with those whose quality has been good is crucial for supporting equilibria where quality is high. As will become clear below, if customers returned no matter what quality they received, there would be no incentive to provide high quality.

The only formal justification offered in this section for the departure of unsatisfied customers is their indifference among all suppliers. In practice, it seems reasonable to suppose that customers expect suppliers who have failed to provide good quality to be worse than the average supplier. As a result, customers want to remain only with those suppliers whose quality has been high. I model this formally in Section V.

I now study the incentives to provide high quality at equilibria of this kind. In these equilibria, the value to a firm from having a single customer is V. By contrast, the value to a firm of a customer who is currently buying from another supplier is W. Therefore the value V equals

$$(1)\quad V = \max\{P - c - \varepsilon + \delta[\sigma V + (1-\sigma)W], P - c + \delta W\}$$

where δ is the discount factor. The first expression in (1) is the value when the firm tries to provide high quality, while the second expression is the value when it decides to provide low quality. If the number of firms is finite, the value to a firm of a customer who is currently buying from another firm is proportional to V. The reason is that this customer can be expected to buy from the current firm sometime in the future. Therefore, W equals ξV, where ξ is smaller than 1 since, at best, the customer comes back after one period. If customers who switch pick their firm at random, this parameter goes to 0 as the number of firms goes to infinity. Let V^h and V^l denote the values of making and failing to make the effort to improve quality. Then, (1) implies

$$V^h = \frac{P - c - \varepsilon}{1 - \delta[\sigma + \xi(1-\sigma)]}$$

$$V^l = \frac{P - c}{1 - \delta\xi}.$$

Attempting to provide high quality is worthwhile if V^h exceeds V^l or

$$(2)\qquad \varepsilon \leq \delta\sigma\frac{1-\xi}{1-\delta\xi}(P - c).$$

It is apparent from this expression that it becomes easier to sustain high-quality equi-

[7]Klein and Leffler (1981) let customers who have received bad quality communicate this to other potential buyers. These then desert the producer as well. This is less appealing if customers differ in their experiences with the producer.

libria as P becomes higher. In particular, because δ, σ, and $(1-\xi)/(1-\delta\xi)$ are each smaller than 1, high-quality goods are sold only if

$$P - c > \varepsilon \tag{3}$$

so that the firms make positive profits. Klein and Leffler (1981) and their followers have addressed what they regarded as the uncomfortable coexistence of these rents with free entry. Carl Shapiro (1983), for instance, argues that initial expenditures are necessary to establish a reputation for good quality and that the rents compensate the firm *ex post*. However, the model is equally consistent with free entry in the absence of such initial expenditures.

Free entry does prevent the equilibrium price P from exceeding c by more than $\sigma\mu$:

$$P - c \leq \sigma\mu. \tag{4}$$

Otherwise, a firm would enter, undercut the price by $\sigma\mu$, and sell low-quality goods to Y customers.

If customers expect low quality from firms that charge less than P, entrants who charge between P and $P - \sigma\mu$ get no customers, and condition (4) is the only restriction on price. Suppose instead that customers believe that any firm charging a price such that (2) holds tries to provide high quality. Such beliefs seem reasonable; firms that charge such prices will indeed try to provide high quality if customer desertions occur only in response to low quality. In the presence of such beliefs, (2) cannot hold as a strict inequality. Customers gravitate to entrants whose price makes (2) hold as an equality.

Now consider the incentives of a firm with attached buyers to charge a price different from P. If it raises its price above P, all customers will leave, and it will lose its usual profits. Moreover, its own benefit from reducing its price below P is no bigger than a new entrant's profit from undercutting P. The reason is that such cuts in P reduce revenues from the firm's attached customers. The requirement that firms with attached customers must have no incentive to deviate from the price P thus imposes no additional restrictions.

Combining (2) and (4), equilibria with high quality exist only if

$$\varepsilon \leq \delta\sigma^2\mu\frac{1-\xi}{1-\delta\xi}. \tag{5}$$

From a social point of view, the effort to provide high quality is worthwhile if and only if

$$\varepsilon < \sigma\mu. \tag{6}$$

Otherwise the benefit from the effort $\sigma\mu$ is smaller than the cost ε. Because δ, σ, and $(1-\xi)/(1-\delta\xi)$ are each smaller than 1, (5) implies (6). When it is not socially worthwhile to make the effort, it will not be provided at any equilibrium with free entry. On the other hand, (6) does not imply (5). Condition (6) holds, while (5) is violated when

$$1 < \frac{\sigma\mu}{\varepsilon} < \left(\frac{1-\delta\xi}{1-\xi}\right)\left(\frac{1}{\delta\sigma}\right). \tag{7}$$

Then, the effort to improve quality is socially worthwhile but not provided at an equilibrium with free entry. The region described by (7) disappears as δ and σ go to 1 and ξ goes to 0. Underprovision of quality (relative to the social optimum) becomes more likely as customers leave more often for accidental reasons (low σ), take longer in coming back (low δ), or come back even after having been disappointed (high ξ).

II. An Illustrative Long-Term Contract

In this section, I illustrate the advantages and disadvantages of a particular form of long-term association whose main purpose is to make separations less likely. I suppose that a single buyer–seller pair can write a long-term contract. Whatever contract this particular buyer–seller pair writes, the relationship of all other buyers with their seller is unaffected. The spot market considered in the previous section with price P continues to operate.

A general long-term contract has at least three elements. First, there is an initial transfer. Then, there is a payment $\hat{P}$ to be made from the buyer to the seller in each period that the buyer requests a unit from the seller. Finally, there is a severance payment s that the buyer makes to the seller when he changes suppliers. For simplicity, I assume that ξ is essentially zero; once this payment has been made, the buyer never returns to the seller.

The contracts in the previous section can be put in this form by letting the firm charge the same price as the others and letting the severance payment equal 0. Buyers were then indifferent between staying with their current supplier and changing suppliers. This indifference allowed me to assume that buyers always stay when they receive good quality and always depart when they receive low quality. Here, I assume instead that the contract also stipulates the probability λ that the buyer stays after he receives low quality. While there is no mechanism that enforces that buyers leave with this probability, I ensure that they are willing to do so by choosing s so that they are indifferent between staying and leaving.[8]

Suppose that the contract is devised so that the seller tries to produce high quality. In this case, all sellers produce the same quality, so the buyer must be indifferent between the payment streams associated with different sellers. In particular, the buyer involved in the long-term contract must be indifferent between severing the relationship immediately and severing the relationship only after remaining with the seller for one more period. This implies that

$$s+\frac{P}{1-\delta}=\hat{P}+\delta\left(s+\frac{P}{1-\delta}\right)$$

or

$$s=\frac{\hat{P}-P}{1-\delta} \tag{8}$$

[8]Since the aim of the long-term contract is to keep the buyer and seller together, the probability that the buyer stays after receiving good quality remains 1. This also makes it easier to induce the seller to make the effort necessary to improve quality.

so the present discounted value of payments is the same whether the buyer stays with this particular seller or pays the severance payments and then buys at the price P.

Then, the value to the seller from having a currently attached buyer is

$$V=\max\{\hat{P}-c-\varepsilon+\delta[(\sigma+(1-\sigma)\lambda)V+(1-\sigma)(1-\lambda)s],\ \hat{P}-c+\delta[\lambda V+(1-\lambda)s]\}$$

Therefore, the values of trying to provide high quality and of providing only low quality are

$$V^{h}=\frac{\hat{P}-c-\varepsilon+\delta(1-\sigma)(1-\lambda)s}{1-\delta[\sigma+(1-\sigma)\lambda]} \tag{9}$$

$$V^{l}=\frac{\hat{P}-c+\delta(1-\lambda)s}{1-\delta\lambda}. \tag{10}$$

Trying to provide high quality is profitable only if V^{h} is at least as large as V^{l} or

$$\varepsilon\le\delta\sigma\left(\frac{1-\lambda}{1-\delta\lambda}\right)\left[\hat{P}-c-(1-\delta)s\right]. \tag{11}$$

Substituting (8) in (11), the producer then actually tries to produce high quality if

$$\varepsilon\le\delta\sigma\left(\frac{1-\lambda}{1-\delta\lambda}\right)(P-c) \tag{12}$$

which is independent of s and $\hat{P}$. A one-dollar increase in s raises the present discounted value of payments if the buyer leaves by one dollar. To preserve the buyers indifference, the present discounted value of payments if the buyer stays forever must increase by one dollar as well. Similarly, seller indifference requires that a one-dollar increase in s (which reduces the payoff from trying to provide high quality) be matched by a one-dollar increase in the present discounted value of receipts if the buyer stays on. Therefore, as long as (8) is satisfied, the precise values of s and $\hat{P}$ do not affect the provision of quality.

Raising s while ensuring that $\hat{P}$ satisfies (8) has the same effect as raising the initial transfer from buyer to seller. Both make the seller better off at the expense of the buyer. For simplicity, I thus consider contracts in which s is 0. The resulting initial transfer, whose value does not affect the subsequent analysis, depends on the bargaining strength of the two agents. If the seller can make a take-it-or-leave-it offer which if rejected leads to the sequence of spot contracts, the transfer is 0. If, instead, the buyer can solicit offers from competing sellers, the transfer is negative.

A problem does arise if buyers can sign long-term contracts that give them large initial transfers not just initially but also at later points. Suppose a buyer believes that next period he can sign a long-term contract from which he receives so much initially that the total present value of his payments equals $(c+\varepsilon)/(1-\delta)$. From the buyer's perspective, this is equivalent to having access to a price of $c+\varepsilon$ from next period on; but, with P equal to $c+\varepsilon$, it is obvious from (12) that it is impossible to provide high quality in the current period. Buyer access to future contracts on good terms makes it impossible to induce the seller to provide high quality now. For equilibria with high quality to exist, buyers must believe that those who offer good terms in the future will provide only low quality.

The right-hand side of (12) falls when λ rises. Indeed, it is equal to 0 for λ equal to 1. This means that higher values of λ make it more difficult to induce the sellers to provide high quality. The highest value for λ that is consistent with attempts to provide high quality makes (12) hold with equality and equals

$$(13) \qquad \frac{\delta\sigma(P-c)-\varepsilon}{\delta\sigma(P-c)-\delta\varepsilon}.$$

A higher value for λ would require a greater incentive to provide quality in the form of a higher $\hat{P}$ or a lower s. However, either, of these changes would break the buyer's indifference and would result in immediate departure by the buyer.

There is an important case in which the maximal λ given by (13) is 0. This is the case in which customers believe that any firm whose price satisfies (2) tries to provide high quality. Then, in equilibrium, (2) holds with equality, so (13) equals 0. Greater values of λ are possible only if there is slack in condition (2). Only then are firms so eager to try to provide high quality that they continue to do so even if dissatisfied customers sometimes remain.

The next question is the value to the buyer and seller of varying λ. I compute this value assuming, as before, that the initial bargaining between buyer and seller is efficient. Therefore, only the effect of λ on the sum of buyer and seller surplus is relevant.

When the buyer purchases from the seller, the two together obtain surplus equal to $r+\sigma\mu-c-\varepsilon$. With probability $\sigma+(1-\sigma)\lambda$, the buyer returns the following period, and surplus in that period is again $r+\sigma\mu-c-\varepsilon$. With probability $(1-\sigma)\times(1-\lambda)$, the buyer leaves. He then receives a present discounted value of benefits equal to $(r+\sigma\mu-P)/(1-\delta)$ from then on. Combining these terms, the present discounted value of the benefits to buyer and seller equals

$$(14) \qquad \frac{r+\sigma\mu}{1-\delta} - \frac{c+\varepsilon+\dfrac{\delta(1-\lambda)(1-\sigma)}{1-\delta}P}{1-\delta[\sigma+(1-\sigma)\lambda]}$$

$$= \frac{r+\sigma\mu-P}{1-\delta} + \frac{P-c-\varepsilon}{1-\delta[\sigma+(1-\sigma)\lambda]}.$$

The first term in the final expression is the surplus the buyer gets when he makes all his purchases in the spot markets. Thus, the second term captures the benefits from the association. In each period that the buyer remains, these benefits equal the difference between P and the cost of production $(c+\varepsilon)$. Inequality (4) says that this difference is positive. This implies that expression (14) is increasing in λ. This is the basic problem. The buyer and seller benefit

from increases in λ because such increases capture rents from the other sellers who, collectively, get $P-c-\varepsilon$ when the buyer leaves. On the other hand, if λ rises too much, there is nothing that keeps the seller from offering bad quality.

Using the maximal value of λ in (13), this present value becomes

$$(15)\quad \frac{r+\sigma\mu-P}{1-\delta} + \frac{(P-c-\varepsilon)[\sigma(P-c)-\varepsilon]}{(1-\delta)[\sigma(P-c)-\sigma\varepsilon]} = \frac{r+\sigma\mu-c-\varepsilon/\sigma}{1-\delta}.$$

There does exist an empirically implausible method for letting λ, and thus joint profits, be bigger without sacrificing quality. This method, which I ignore here, is to have the buyer and seller write a contract with a third party. Such a third party could receive payments from the buyer (and even from the seller) when the buyer changes suppliers. Such payments tend to assure the seller of the buyer's fidelity without encouraging the seller to reduce his quality. In the current setup, reductions in quality are encouraged by the severance payments s since these go to the seller. The problem with schemes involving payments to third parties is the obvious incentive that they create for collusion between the third party and the seller.

The contracts considered so far are as efficient as the outcomes with spot markets, since they have the same quality. Within this class of contracts, inefficiency would be present only if the long-term contracts led sellers to forgo high quality. Then, joint profits when the buyer purchases from the seller equal $r-c$ per period. After the buyer leaves the seller, his net profits have a present discounted value of $(r+\sigma\mu-P)/(1-\delta)$. Since the buyer leaves with probability $1-\lambda$ in each period, the present value of joint profits equals

$$(16)\quad r-c+\delta\left[\left(\frac{\lambda}{1-\delta\lambda}\right)(r-c) + \left(\frac{1-\lambda}{1-\delta\lambda}\right)\left(\frac{r+\sigma\mu-P}{1-\delta}\right)\right] = \frac{r+\sigma\mu-P}{1-\delta} + \frac{P-c-\sigma\mu}{1-\delta\lambda}.$$

Increases in λ are worthwhile only when the price P exceeds $c+\sigma\mu$. Then, the buyer would prefer low quality at a price of c to paying P and getting high quality with probability σ. However, as I showed in Section I, free entry implies (4) and is thus inconsistent with this case. Inefficient long-term contracts in which low quality prevails cannot be optimal responses to high-quality spot-market equilibria with free entry. They can only be optimal responses if, as in Section IV, (4) is violated and the equilibrium price exceeds its maximum free-entry level.

In this section, I have illustrated the tension between rent capture and high quality. I have also shown that long-term contracts yield efficient outcomes relative to spot-market contracts when there is free entry. In the next three sections, I explore three inefficiencies relative to the spot-market outcome. The first comes from the existence of an alternative technology for providing high quality, the second comes from prices in excess of the free-entry level, and the third comes from random variations in the competence of the sellers.

III. Outright Integration

Outright integration differs from the sort of long-term contract I considered in Section II in that it changes the control rights of various agents. Integration is probably most easily thought of as involving a new agent who has (limited) authority over both the buyer and the seller. Hiring such an agent is attractive if he can improve the quality that is forthcoming when the buyer

and seller let λ equal 1. The seller and buyer would then keep all the rents this buyer would ultimately have paid to other sellers.

Suppose that the hiring of this agent at a cost of m per period ensures that the seller makes the requisite effort when he produces for a particular buyer. Assuming this payment only occurs when the customer buys from the seller, the present discounted value of the joint benefits equals

$$(17)\quad \frac{r+\sigma\mu-m-c-\varepsilon}{1-\delta\lambda} + \delta\left(\frac{1-\lambda}{1-\delta\lambda}\right)\left(\frac{r+\sigma\mu-P}{1-\delta}\right) = \frac{r+\sigma\mu-P}{1-\delta} + \frac{P-c-\varepsilon-m}{1-\delta\lambda}.$$

This expression is maximized for λ equal to one if

$$(18)\quad P \geq c+\varepsilon+m.$$

Otherwise, the optimal λ is 0, and integration is not valuable. Integration is worthwhile only if the cost of the internally produced good, $c+\varepsilon+m$, is competitive with the price of the externally supplied version.[9] When (18) holds, the maximized value of (17) is

$$(19)\quad \frac{r+\sigma\mu-c-\varepsilon-m}{1-\delta}.$$

This value exceeds (15), the value of the best long-term contract that keeps the cost of the quality effort equal to ε, if

$$(20)\quad m < \frac{1-\sigma}{\sigma}\varepsilon.$$

[9]With free entry, (4) holds, so (18) implies that $\sigma\mu \geq \varepsilon+m$. More generally, unless this condition holds, the buyer and seller together are better off forgoing quality altogether, rather than paying $\varepsilon+m$ for a benefit of $\sigma\mu$.

Condition (20) becomes easier to satisfy as σ becomes smaller. The reason is that when σ is low there are many separations, so that the rents captured by buyer and seller jointly when their contract has λ given by (13) are relatively low. Then, their collective rents can be increased substantially by letting λ equal 1 and spending m per period. By contrast, when σ is 1 (the original case in Klein and Leffler [1981]), condition (20) is violated for all strictly positive m. In this case, buyer and seller never need to separate in equilibrium, so that nothing is gained by forcing them to remain together.

The combination of (20) and (2), which is necessary for spot markets to involve high quality, implies that $\delta(P-c)$ exceeds $\varepsilon+m$. This is turn implies (17). Thus, for any σ strictly less than 1, there exist strictly positive m's such that (20) holds, and if the spot market equilibrium has high quality, integration is privately desirable and socially inefficient. If vertical integration could somehow be eliminated, quality would be the same, and society would save m.

Up to now, I have interpreted m as a managerial cost which ensures that the effort of the seller (whose cost remains ε) is actually forthcoming. I now discuss informally some other possibilities; more formal treatment is postponed for future work. One possibility is that spending m renders the seller's exertion of effort verifiable by outsiders. For instance, the manager might record the seller's actions in a way that outsiders trust. Then, a contract that stipulates that the seller be paid ε if he actually incurs the requisite effort becomes enforceable.

Along similar lines, the expenditure of m might produce verifiable evidence not on the effort expended but on the quality that the buyer actually receives. Then, the manager could write a contract with the seller in which the payments to the seller depend on the quality received by the buyer. Suppose the seller receives c when quality proves to be low and $(c+\varepsilon)/\sigma$ when it proves to be high. As long as he is risk-neutral, he is now happy to exert the necessary effort, since he on average earns $c+\varepsilon$ when he does and c when he does not.

The role of top management in the preceding two paragraphs is to spend m collecting evidence on quality (or on the effort to supply quality). Management is the monitor as in Alchian and Demsetz (1972).[10] One can also think of m rather differently. The cost m can represent a reduction in quality itself, rather than a managerial cost in obtaining the spot-markets outcome. In this interpretation, the integrated firm produces a good worth $r + \sigma\mu - m$, rather than $r + \sigma\mu$. Even more generally, m can represent any loss in utility for buyer and seller (relative to $\sigma\mu - \varepsilon$).

One particular form of this loss stressed by Eccles and White (1988) is the prevalence of conflict in integrated organizations. They suggest that conflict between buyer and seller within an organization is useful, because it helps management gather information. I now suggest a scenario in which such conflict is indeed useful. Suppose that, just before the good is produced, the buyer receives a signal. This signal tells the buyer which particular change in specification would increase buyer utility and by how much. The buyer would now like the seller to make the change. On the other hand, the seller finds any such change costly. Moreover, the seller knows something about the cost of the change that others do not know. Therefore, the seller would normally overstate the costs of the change, either to increase his compensation or to avoid the necessary effort. In the absence of any mediator between buyer and seller, the result would be costly bargaining.

The manager can now serve a useful role. He can be put in charge of deciding whether the change should be made.[11] To find out whether the change is worthwhile, he encourages buyer and seller to "fight." In this fight, both buyer and seller present evidence. They also suggest experiments for the manager to carry out. Since the manager has complete control over the tools with which buyer and seller work, he can actually carry out these experiments and learn something about the cost and value of the change. This knowledge then guides his decision. Note that the knowledge acquired by the manager is unlikely to be perfect. Therefore, it is in the seller's best interest to pretend that his costs are high to the very end.

The upshot of all this could well be a gain to the buyer g_1 which is lower than $\sigma\mu$ and a cost to the seller g_2 which exceeds ε. In this case, m is given by $\sigma\mu - \varepsilon + g_2 - g_1$. As long as (17) and (20) are met, setting λ equal to 1 and hiring the arbiter-manager is worthwhile for the buyer and seller taken together. Alternatively, an arbiter-manager will find it worthwhile to hire both a buyer and a seller. This is a more appealing description, because it emphasizes that, in the integrated organization, the buyer and the seller are employees of the arbiter-manager.

Suppose that these buying and selling employees carry out both internal and external transactions. Which will they prefer? When the buying employee carries out an external transaction, he receives $\sigma\mu$ in direct utility. This is the utility that results from courteous service and flexible specifications. Both of these facilitate the job of the buying employee. By contrast, when the buying employee carries out an internal transaction, he receives only g_1 in direct utility. Similarly, the selling employee's disutility from an internal transaction, g_2, exceeds that from an external one. It thus appears to be possible that, consistent with Eccles and White's (1988) empirical findings, both buyer and seller prefer outside transactions.

Whether buying and selling employees are averse to internal transactions obviously depends also on how compensation varies with the type of transactions that they carry out. If employees are heavily compensated every time they buy or sell internally, internal transactions will look good to them. Suppose in particular that the buying and selling employees own the integrated firm and keep its profits. Since integration raises profits, net of the utility costs of buyer and

[10] Their model assumes there is a technological reason for monitors. As a result, the monitoring that emerges is socially good. By contrast, the monitoring that takes place here could usefully be replaced by a market.

[11] This arbiter role of management is also stressed by Williamson (1975).

seller, the two would then, prefer internal transactions.

Suppose that, on the contrary, the two are only employees and the profits go to a third party. Then, at least on average, buying and selling employees must be compensated for the disutility induced by internal transactions. This implication is at least consistent with the evidence that wages rise with establishment and company size (see Charles Brown and James Medoff, 1989). In the view of this paper, employees of larger (more integrated) companies must be paid a compensating differential because they have to transact disproportionately more with other employees of the same organization.

This does not mean, however, that employees will be paid a bonus in every instance when they carry out an internal transaction. As long as agents are risk-neutral, there is little reason for compensation to vary in this manner. From an administrative point of view, such microtailoring of compensation is probably burdensome, since it requires precise definitions of the extent to which a transaction is internal. Thus, organizations may opt for flat salaries that reflect the *typical* volume of internal transactions. With this compensation structure, employees get more utility from external transactions than from internal ones. For internal transactions to take place at all, they must then be mandated by management.

One issue that arises at this point is why the cost m must be spent within an organization that also includes the buyer and the seller. Why can't these resources be spent in an arms-length transaction? The answer is that, in the situation I have in mind, arms-length transactions need very little management and supervision. The threat of dissolving the relationship is sufficient to ensure considerable effort by the seller.

Still, why can't the buyer and seller write a long-term contract that binds them to each other and then spend m on the services of an outside party? There are two reasons. The first, and principal, reason is that monitoring, policing, and verification require power. The monitor must be able to change the conditions of production to find out how the seller and buyer are actually behaving. This flexibility is essential precisely because it is difficult to predict in advance the form of the effort to provide quality. Therefore, the monitor must be in a position to dictate how the assets with which buyer and seller work are used. Such residual control over the use of assets virtually defines the role of top management. Indeed, from a legal point of view, such residual control is the purview of the owner of the asset.[12]

Second, familiarity with the buyer and seller helps in determining the cost of various quality-enhancing actions, as well as in verifying that these actions have been taken. Obtaining this familiarity takes time. There is thus a benefit to having the monitor remain in his job for a long time. By making the monitor part of the firm, his mobility costs increase, and he becomes more likely to learn the personal characteristics of buyer and seller.

The advantages of mergers that I describe are similar to those of the "incomplete contracts" theory of the firm surveyed by Holmstrom and Tirole (1987). As in that article, some individuals are put in charge, because contracts cannot specify all contingencies. These individuals get to decide how to proceed when the contract does not specify what is to follow.

Grossman and Hart (1986) also base their theory of the firm on contractual incompleteness. Yet, the implications of their model appear to be rather different from the one presented here. They focus on the contractual incompleteness that makes it difficult for individuals to garner the fruits of their human capital accumulation. In the absence of ownership rights, some of the benefits of an individual's human capital accumulation accrue to those he works with. Ownership can ameliorate this. The owner gets to decide on all aspects of the transaction not specifically contracted on in advance. As a result, he receives a disproportionate share of the payoff to both his and

[12]Grossman and Hart (1986) equate residual control over the use of assets with ownership.

his underlings human capital accumulation. Ownership of an asset by agent A alleviates the moral-hazard problem inherent in A's human capital accumulation (while worsening that of the other agents who work with this asset). They are careful to present their model as one of an owner-manager, though their theory perhaps applies also to corporate top management. In this case, one would expect top managers' compensation to be closely related to firm performance.

In my story, by contrast, moral hazard may not be an issue for top managers. Top managers are like policemen. The competence with which they dispatch these functions may be easy to determine from their reports. Top managers could still earn large sums if the skills needed to investigate whether the appropriate effort is being carried out are in short supply. In this case, one would expect larger firms, whose management problems are more complex, to hire more-competent monitors. This suggests that management compensation should increase with firm size. By contrast, their wages need not bear any particular relationship to firm performance. In practice, managerial compensation does seem more linked to firm size than to firm performance (see Michael Jensen and Kevin Murphy, 1988).

One critical question faced by all theories of the firm is the conundrum posed by Coase (1937) and Williamson (1985): what limits firm size? If top managers can intervene selectively and obtain whatever advantages there are to integration, a single firm should control the whole economy. This argument gains additional force when one recognizes that integration eases collusion in product markets.

Williamson (1985) gives a variety of reasons why managers cannot actually intervene so selectively. One reason that appears particularly germane in my context is that managers are imperfect and sometimes make mistakes. For managers to be effective at policing, they must have great authority, so these mistakes can be costly. They become more costly the larger is the manager's span of control. Giving a manager residual decision rights over the assets used by many individuals when his decisions only rarely contribute to the common good is dangerous. By contrast, giving him these rights when intervention is needed often is more attractive.

IV. Imperfect Competition

Instead of letting entry be free, I assume here that the number of sellers is fixed at N. I show that the extent to which integration is inefficient can then depend on the level of rents in the industry.

Without free entry, price can exceed $c + \sigma\mu$. It can also exceed the level that makes (2) hold with equality even when customers view all firms whose price satisfies (2) as trying to provide good quality. To sustain such high prices, firms threaten each other with punishments as in James Friedman (1971). The worse the threatened punishment, the higher the prices that can be sustained in equilibrium. For concreteness, I focus on the worst possible punishments in the style of Dilip Abreu (1986). Similar results obviously obtain for weaker punishments.

The structure of the equilibrium is the following. Each firm is expected to charge P and earn a present discounted value of $\pi = Y(P - c - \varepsilon)/N(1 - \delta)$. A firm that deviates by raising it price loses all its customers, so this deviation is unattractive.

Suppose a firm deviates by cutting its price. Then, a punishment period ensues. All other firms charge $c + \varepsilon$ and try to provide high quality in the period after the deviation. In this period, the deviating firm is expected to charge a price L so low that its losses in that period equal $\delta\pi$. If firms charge these prices, they revert to charging P in the subsequent period. If any firm charges a different price, it is itself punished in analogous fashion.

Several comments about this construction are in order. First, all firms have an incentive to provide high quality if satisfied customers return in the following period, when the price is again equal to P. Second, the firm that is being punished can expect to earn a present value of 0 whether it charges L or not. If it does charge L its present discounted value of profits is 0 by construc-

tion. If it deviates and charges a higher price, customers expect its quality to be low, so that they are willing to pay at most c for its goods. This means that the deviating firm can do no better than charge c, break even in that period, and get punished again. This again gives it 0 in present value.

By cutting its price slightly below P, a firm can, at most, capture the entire market. Then, it would earn N times its usual per-period profits in the period of the deviation. After that, it earns a present value of 0. Thus, these deviations will not take place as long as

$$N \leq \frac{\delta}{1-\delta}$$

which is independent of P and is satisfied for sufficiently high δ.

As in the case of free entry, there are both equilibria where quality is low and equilibria where quality is high. However, if (6) is met, switching from an equilibrium with low quality to one in which the firms try to supply high quality increases willingness to pay by more than it increases costs. Moreover, such a switch does not increase the equilibrium's vulnerability to deviations. Therefore, if the selling firms can pick equilibria and (6) is met, they will pick the equilibrium in which they produce high quality and charge $r + \sigma\mu$.

This ability of sellers to coordinate on equilibria with high prices is socially beneficial. It means that condition (2) becomes easier to meet, and firms find it in their interest to provide high quality. In particular, (2) can now be met even when (5) fails (so that free-entry equilibria have low quality). Moreover, in this model, there is no other distortion from high prices. With spot markets alone, their only effect on economic efficiency is to make high-quality equilibria possible in situations where, with free entry, only low-quality equilibria exist.[13]

[13]Contrast this with Michael Spence (1975), where monopoly (and thus collusion) generally distort quality away from the optimum.

I now study the effect of high prices on the attractiveness for one buyer–seller pair of both long-term contracts and outright integration. Consider first long-term contracts of the form considered in Section II. It is apparent from (14) that the derivative of joint profits with respect to λ rises with P. Higher prices make reductions in separations more attractive. As shown in (13), they also make higher values of λ feasible.

On the other hand, increases in P have a direct negative effect on the buyer. The net effect of these contradictory forces is that the joint profits at the optimal λ given by (15) are independent of P. The same is true of (16) with λ equal to 1, which gives joint profits if buyers and sellers forgo high quality altogether, and of (19), which gives joint profits under integration.

Therefore, the level of P itself has no effect on the choice of contractual form. However, it does affect the social optimality of the chosen contract. Suppose that ε is between $\sigma\mu$ and $\sigma^2\mu$ so that (5) is violated and market transactions in the presence of free entry involve low quality. If m is smaller than $\sigma\mu - \varepsilon$, buyer and seller will integrate. Relative to the spot-market outcome with free entry, this integration is efficient.

When ε is in this range, (6) is satisfied, so attempting to provide high quality is efficient. Moreover, if sellers with market power can coordinate, they choose a spot-market equilibrium with high prices in which quality is high; (16) with λ equal to 1 exceeds (15). Therefore, the contract with low quality, which avoids all separations, has higher joint profits than either spot-market transactions or any other contract of the form studied in Section II. If m is smaller than $\sigma\mu - \varepsilon$, joint profits are even higher under integration. However, both of these arrangements are socially less desirable than the spot-market equilibrium with high prices. Either they involve inefficiently low quality or inefficient integration.

The intuition for the result that the presence of market power reduces the efficiency of private contractual arrangements is the following. Market power raises prices, which makes the provision of high quality in spot markets easier. However, market power

does not reduce the private desirability of contractual forms where high quality, if it is provided at all, is costly.

V. Inefficient Separations with Long-Term Contracts

The long-term contracts of Section II feature fewer separations than the spot markets of Section I only because I have defined spot markets to have customers leave every time they encounter low quality. This definition of spot markets is, up to this point, arbitrary. If all producers are identical, customers ought to be indifferent among them. Thus, it is just as consistent with consumer maximization to assume that customers stay with a probability given by (13) when they have received low quality. This still ensures high quality and is identical to the long-term contracts of Section II.[14] Another consequence of this indifference is that, at least with free entry, long-term contracts are just as socially efficient as what I have dubbed spot-market transactions.

In this section, I provide a reason for customers to leave when they encounter low quality. The reason is that low quality can mean two things. Either the seller has been temporarily unlucky (as I have assumed until now) or the low quality is an indication of seller incompetence.[15] This possibility of seller incompetence makes buyers want to switch sellers whenever quality is low.

This possibility implies that long-term contracts that attach buyers to sellers are subject to a new form of inefficiency. This inefficiency is present even when sellers always try to provide high quality. It stems from the increased probability that buyers will purchase from incompetent sellers. In spite of this inefficiency, such attachments can be privately optimal when price is high.

I suppose that firms can be in one of two states. In state w, they are well managed while in state b they are poorly managed. Poorly managed firms are intrinsically unable to provide high quality. They spend ε, and their quality is low nonetheless.[16] The probability that a well managed firm remains well managed in the next period is $1-\phi$. The probability that a poorly managed firm becomes well managed is $1-\psi$. I assume that ψ exceeds ϕ, so that bad management has a tendency to persist. These transition probabilities imply a steady-state probability of being in state w equal to

$$\eta \equiv \frac{1-\psi}{1+\phi-\psi}. \tag{21}$$

For simplicity, I also let the initial probability of being in state w equal η. Because $\psi > \phi$, η is smaller than the probability $1-\phi$ that a previously well managed firm remains well managed.

Suppose that there is a spot-market equilibrium in which all firms charge P and in which well managed firms try to supply good quality. I first consider customer behavior, assuming such an equilibrium exists. I prove that customers come back to sellers who have given them high quality, while they desert those that have not. I then give conditions for this equilibrium to exist. These conditions ensure that well managed sellers do indeed try to provide high quality and that no seller wants to charge a different price.

Consider a customer who receives high quality in period t. He infers correctly that his current seller is well managed. His probability of obtaining high quality next period from this particular seller is thus $(1-\phi)\sigma$. By contrast, his probability of obtaining high quality from a seller chosen at random is $\eta\sigma$, which is smaller. He therefore prefers to remain with his old seller.

[14] Of course, if customers stay with higher probability, equilibrium quality must be low.

[15] The presence of two analogous possibilities also induces customer departures in the monopoly model of Allen and Faulhaber (1988). Because they focus on the case with a single supplier, the possibility of seller incompetence plays a more crucial role. Without it, customers do not react at all to low-quality realizations, so that high quality is not sustainable in pure-strategy equilibria.

[16] Letting them avoid the cost ε changes the formulas below without affecting the qualitative results.

Suppose instead that the customer receives low quality at t. The posterior probability that he assigns to having a badly managed seller depends on his prior probability. This prior probability is lowest if the customer has bought a high-quality good from this particular seller at $t-1$, so that he knows the seller was well managed at $t-1$. I consider this case first.

Let w_t represent the event that the seller is well managed at t, b_t the event that he is poorly managed at t, and L_t the event that he provides low quality at t. Then, the probability of event w_t given that the seller was previously well managed and that he provides bad quality at t is

$$(22)\quad P(w_t|L_t,w_{t-1}) = \frac{P(w_t,L_t|w_{t-1})}{P(b_t,L_t|w_{t-1})+P(w_t,L_t|w_{t-1})} = \frac{(1-\sigma)(1-\phi)}{\phi+(1-\sigma)(1-\phi)}.$$

This probability is smaller than η as long as

$$(23)\quad \sigma \geq \frac{\psi-\phi}{1-\phi}.$$

This condition is met as long as σ is sufficiently large. Then, well managed firms are sufficiently likely to deliver high quality that low quality is a signal of poor future prospects.

Now consider a buyer who has had no previous experience with this seller because he has just changed suppliers. His prior probability that the seller is well managed is then η. The posterior probability that he is well managed given that he has provided bad quality is

$$(24)\quad P(w_t|L_t) = \frac{P(w_t,L_t)}{P(b_t,L_t)+P(w_t,L_t)} = \frac{(1-\sigma)(1-\psi)}{\phi+(1-\sigma)(1-\psi)}.$$

When ψ exceeds ϕ, this probability is smaller than η, so that customers who have just arrived leave suppliers who give them bad quality. In conclusion, when ψ exceeds ϕ and (23) is met, all buyers leave firms that offer low quality.

I now turn to seller behavior. I show that, given that consumers leave those who provide bad quality, sellers find it in their best interest to supply good quality. I later turn to sellers' incentives to change prices.

For simplicity, I assume again that ξ is zero, so that customers who leave essentially never return. Then, the value to a badly managed firm of an attached customer, who is about leave, equals $P-c-\varepsilon$. Therefore, the value V to a well managed firm of having a customer is

$$V = \max\{P-c-\varepsilon+\delta[\sigma(1-\phi)V + \phi(P-c-\varepsilon)], P-c\}.$$

The values of trying to provide high quality and of providing only low quality are now

$$V^h = \frac{(1+\delta\phi)(P-c-\varepsilon)}{1-\delta\sigma(1-\phi)}$$

$$V^l = P-c$$

so that trying to provide high quality is profitable if

$$(25)\quad \varepsilon \leq \frac{\delta[\sigma+\phi(1-\sigma)]}{1+\delta\phi}(P-c).$$

As long as this modified version of (2) holds, well managed firms do not deviate by cutting quality.

For such prices to be equilibria with free entry, customers must prefer to pay P and receive high quality with probability $\eta\sigma$, rather than to pay c and receive low quality for sure. Thus,

$$(26)\quad P-c \leq \eta\sigma\mu.$$

For equilibria with high quality to exist when entry is free, prices must exist that satisfy both (25) and (26). In other words,

$\eta\sigma\mu$ must exceed

$$(1+\delta\phi)\varepsilon/\delta[\sigma+\delta(1-\sigma)].$$

If, in addition, customers believe that firms whose price satisfies (25) try to provide high quality, the unique equilibrium has (25) holding with equality.

The incentives of firms with attached customers to raise price above P are more complex in this case. The reason is that customers who have received high quality are willing to pay more to their current firm than the price charged by other firms. To simplify the exposition, I eliminate this scope. I do this by assuming that customers believe that any firm that charges a price different from P will produce low quality. Raising price then leads all customers to depart.[17]

Note that the analysis of Sections I, II, and III corresponds to the model of this section when one takes the limit as ϕ goes to zero. At this limit, the firm is almost always well managed, but if σ exceeds ψ, (23) is met, and the buyer leaves any firm whose quality proves low. Moreover, at this limit, the payoff from the long-term contracts described in Section II is well approximated by (14).

Suppose, however, that the system is not at this limit and that a certain buyer–seller pair is given the option of signing a long-term contract. Privately optimal long-term contracts are now more complex. The reason is that information about a seller's competence accumulates over time. A seller who has delivered bad quality for two periods in a row is more likely to be poorly managed than one who has delivered bad quality for only one period. These complications preclude the derivation of the privately optimal long-term contract in this paper. Rather, I focus on a very simple type of long-term contract, and I show that it will be attractive to buyer and seller taken together even though it is socially harmful.

This simple contract applies only to the first period. If the buyer receives high quality in the first period, he stays as before. If, on the other hand, he receives low quality, he now remains with probability λ. Whatever the reason for staying, his behavior in subsequent periods is the same as his behavior with spot markets; he leaves any seller whose quality is low. Thus, the contract gives the initial seller only one "second chance." If he delivers low quality after that, he loses the buyer.

To focus on the new inefficiency, I ignore contracts in which sellers cease to make the effort necessary to improve quality. In other words, I consider contracts whose λ is sufficiently small that the seller continues to try to provide high quality. By continuity, positive λ's that satisfy this condition exist when (25) and (26) are satisfied. This leaves open for further research the question of whether high quality is easier to sustain for a given λ when, as in this section, ϕ exceeds 0.

From the perspective of buyer and seller together, this contract has a cost and a benefit. The cost is that, if the seller is indeed incompetent in the first period, it is costly for the buyer to stay. The benefit is that, if the seller is only unlucky in the first period, the contract ensures the buyer–seller pair some surplus that would otherwise go to some other firm.

The contract only matters when the buyer receives low quality in the first period, as occurs with probability $1-\eta\sigma$. In the absence of any contract, the buyer would then leave and obtain good quality with probability $\eta\sigma$ from his new supplier. His total payoff after the initial period would equal

$$(27)\quad \eta\sigma(r+\mu-P+\delta G)+(1-\eta\sigma)(r-P+\delta B)$$

where G and B represent the present discounted value of the buyer's payoff after having received good and bad quality respectively.

If the buyer stays after receiving low quality in the first period, he has a probability ν

[17]This feature could be avoided by focusing on equilibria in which new customers pay lower prices (introductory offers) than do customers that remain. The basic result is not sensitive to this change in specification.

of facing a well managed seller where:

$$\nu \equiv P(w_t|L_t)(1-\phi)+[1-P(w_t|L_t)](1-\psi).$$

This expression is smaller than η, because the firm that has delivered low quality is more likely to be badly managed than a firm chosen at random. If the buyer who remains pays P, his total payoff from remaining is

$$(28)\quad \nu\sigma(r+\mu-P+\delta G) + (1-\nu\sigma)(r-P+\delta X).$$

Equation (28) is identical to (27), except that the probability of the good outcome is $\nu\sigma$ instead of $\eta\sigma$. For λ to be positive, the buyer must be indifferent between leaving and staying. Thus, if he leaves, the severance payment must equal the difference between (27) and (28).[18]

Ignoring the initial transfer, the buyer's total loss from the contract is the difference between (27) and (28) times the probability $\lambda(1-\eta\sigma)$:

$$(29)\quad \lambda(1-\eta\sigma)(\eta-\nu)\sigma[\mu+\delta(G-B)].$$

A tedious but straightforward calculation leads to a value for $G-B$:

$$G-B=\frac{\delta(1-\phi-\eta)\sigma\mu}{1-\delta\sigma(1-\phi-\eta)}.$$

Since $1-\phi-\eta$ lies between 0 and 1, this expression is positive and so is the value of (29). Note that the loss (29) is independent of P and that it vanishes as ϕ goes to 0 because, in this case, both η and ν go to 1. While I computed (29) as the loss the buyer from the lower quality he receives on average, it also equals the social loss from the contract. The reason is that the total social costs of production are the same whether the buyer stays or leaves, so the change in consumer surplus (at unchanged prices) equals the change in total surplus.

The private gain from the contract is that the price P exceeds the cost of producing these goods. After giving the buyer bad quality, there is still a positive probability ν that the seller is well managed. If the buyer stays, the buyer–seller pair therefore gains $P-c-\varepsilon+\delta V^h$ with probability ν. With 1 minus this probability, they gain only $P-c-\varepsilon$. Total gains from the excess of price over cost are thus given by

$$(30)\quad \lambda(1-\eta\sigma)(P-c-\varepsilon+\nu\delta V^h).$$

This expression is positive and increasing in P. For P sufficiently high or ϕ sufficiently low, (30) exceeds (29) for strictly positive λ. Then, it it worthwhile to make λ positive, thus incurring the social cost (29) in exchange for the rents (30). While this only demonstrates the private advantages and social costs of a very special long-term contract, similar costs and benefits attend more general long-term contracts.

VI. Conclusion

Ever since Coase (1937), economists have generally viewed the internal organizations of capitalistic firms as benign. Even Williamson (1975, 1985), who feels that vertical integration sometimes reduces competition, asserts that the main purpose of integration and long-term contracts is to reduce "transactions costs."

This paper suggests a less sanguine view. Whether the organization of firms is socially helpful is an empirical question. It is not a question economists should feel capable of answering on a priori grounds alone. Since little other information on the actual functioning of these institutions exists, economists should pay close attention to the comments of managers and to their dislike of internal transactions.

[18]This computation assumes that the price that the buyer pays if he stays is the same as the external price. Increases in the price that the buyer pays if he stays must be matched one-for-one with increases in the severance payment to maintain buyer indifference. Increases in price must also be matched one-for-one to keep constant the seller's incentive to provide high quality. Thus, once again, the breakdown between the price the buyer must pay and the severance payment does not matter.

There may, of course, exist other reasons for this aversion. A partial explanation is that there exist specialized individuals (or organizations) with great skill. These individuals are efficiently utilized only when they serve a great many different customers, so that their integration with customers is worthwhile only in the case of very large customers. A relatively small firm then has a choice of integrating with a less competent individual (to the dismay of its buyers) or having an external transaction with one of the specialized individuals. This story explains the preference of buyers for outside partners but does not explain the analogous aversion of in-house sellers.

Perhaps managers dislike intrafirm transactions because they cloud managerial compensation. Insofar as transfer prices are somewhat arbitrary, wages that depend on profits reported by profit centers become risky as intrafirm transactions become more common. Holmstrom and Tirole (1987) explain GM managers' dislike of centralized procurement along analogous lines. This story is not complete. Firms must have some reason for exposing their employees to additional risk. In the case of GM, it seems plausible that centralized procurement allows GM to use its monopsony power more effectively. Rent capture may thus lie behind GM's move.

My model has several potentially testable implications. First, in many industries, firms that provide their own inputs coexist with others that obtain theirs in the open market. These differences must involve some heterogeneity; the cost m of carrying out transactions internally must vary across firms. Insofar as these differences in m are exogenous differences in transactions costs, my theory has the same implication as Coase (1937). Suppose instead that m is the cost from a given degree of quality deterioration. Then, my theory predicts that buyers who are more concerned with quality ought to be particularly averse to integration. This suggests that, controlling for all other observable characteristics, the producers of "high-end" products within an industry ought to be less integrated than the producers of "average" products.

Second, according to my theory, downward vertical integration should be more prevalent in industries in which price exceeds marginal cost by more. Along more standard lines, the increased incentive to integrate when prices are high may be attributed to the desire to avoid the "double marginalization" problem. However, unlike my theory, this standard motivation for vertical integration loses its force if agents are able to write complex contracts such as two-part tariffs.

The theory presented here may also be helpful in understanding why the organization of transactions differs across nations. Perhaps these organizational differences are due to cultural differences. In my model, buyer mobility is less important in nations where people find it culturally less acceptable to provide low quality. Therefore, my theory predicts that long-term attachments are prevalent where these transactions costs are low.

Another possibility is that these organizational differences can be traced to other differences in the production process. Perhaps, the Japanese *kanban* system inherently leads to higher quality because of its use of small batches. If this is interpreted as a high value for σ, my model would predict relatively little centralized control of the whole production process and relatively extensive attachment via long-term contracts. While this suggests some of the features of *keiretsu*, extensive investigation of the extent to which the model can explain international differences is left for further research.

REFERENCES

Abreu, Dilip, "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, June 1986, *39*, 191–225.

Aghion, Philippe and Bolton, Patrick, "Contracts as a Barrier To Entry," *American Economic Review*, June 1987, *77*, 388–401.

Alchian, Armen and Demsetz, Harold, "Production, Information Costs and Economic Organization," *American Economic Review*, December 1972, *62* 777–95.

Allen, Franklin and Faulhaber, Gerald, "Opti-

mism Invites Deception," *Quarterly Journal of Economics*, May 1988, *103*, 397–407.

Bowles, Samuel, "The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian and Marxian Models," *American Economic Review*, March 1985, *75*, 16–36.

Brown Charles and Medoff, James L., "The Employer Size Wage Effect," *Journal of Political Economy*, October 1989, *97*, 1027–59.

Coase, Ronald H., "The Nature of the Firm," *Economica*, November 1937, *4*, 386–405. (Reprinted in G. J. Stigler and K. E. Boulding, eds., *Readings in Price Theory*, Homewood, IL: Irwin, 1952.)

_____, "The Nature of the Firm: Meaning," *Journal of Law, Economics and Organization*, Spring 1988, *4*, 19–33.

Dewatripont, Mathias, "Commitment through Renegotiation-Proof Contracts with Third Parties," *Review of Economic Studies*, July 1988, *55*, 377–91.

Eccles, Robert G. and White, Harrison C., "Price and Authority in Inter-Profit Center Transactions," *American Journal of Sociology*, 1988, *94* (suppl.), S17–S51.

Friedman, James W., "A Non-Cooperative Equilibrium for Supergames," *Review of Economic Studies*, January 1971, *38*, 1–12.

Grossman, J. Sanford and Hart, Oliver D., "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, August 1986, *94*, 691–719.

Holmstrom, Bengt R. and Tirole, Jean, "The Theory of the Firm," in R. Schmalensee and R. Willig, eds., *Handbook of Industrial Organization*, New York: North Holland, 1989, 61–133.

Jensen, Michael C. and Murphy, Kevin J., "Are Executive Compensation Contracts Structured Properly?" Harvard Business School Working Paper, February 1988.

Klein, Benjamin and Leffler, Keith B., "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, August 1981, *89*, 615–41.

Ordover, Janusz A., Saloner, Garth and Salop, Steven C., "Equilibrium Vertical Foreclosure," *American Economic Review*, March 1990, *80*, 127–42.

Shapiro, Carl, "Premiums for High Quality as Returns to Reputations," *Quarterly Journal of Economics*, November 1983, *98*, 659–80.

Spence, Michael, "Monopoly, Quality and Regulation," *Bell Journal of Economics*, Autumn 1975, *6*, 417–29.

Williamson, Oliver E., *Markets and Hierarchies*, New York: Macmillan, 1975.

_____, *The Economic Institutions of Capitalism*, New York: Macmillan, 1985.

[Footnotes]

[1] **Production, Information Costs, and Economic Organization**
Armen A. Alchian; Harold Demsetz
The American Economic Review, Vol. 62, No. 5. (Dec., 1972), pp. 777-795.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28197212%2962%3A5%3C777%3APICAEO%3E2.0.CO%3B2-Q

[1] **The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration**
Sanford J. Grossman; Oliver D. Hart
The Journal of Political Economy, Vol. 94, No. 4. (Aug., 1986), pp. 691-719.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198608%2994%3A4%3C691%3ATCABOO%3E2.0.CO%3B2-F

[4] **Contracts as a Barrier to Entry**
Philippe Aghion; Patrick Bolton
The American Economic Review, Vol. 77, No. 3. (Jun., 1987), pp. 388-401.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198706%2977%3A3%3C388%3ACAABTE%3E2.0.CO%3B2-7

[4] **Commitment Through Renegotiation-Proof Contracts with Third Parties**
Mathias Dewatripont
The Review of Economic Studies, Vol. 55, No. 3. (Jul., 1988), pp. 377-389.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28198807%2955%3A3%3C377%3ACTRCWT%3E2.0.CO%3B2-6

LINKED CITATIONS

[4] **Equilibrium Vertical Foreclosure**
Janusz A. Ordover; Garth Saloner; Steven C. Salop
The American Economic Review, Vol. 80, No. 1. (Mar., 1990), pp. 127-142.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28199003%2980%3A1%3C127%3AEVF%3E2.0.CO%3B2-K

[5] **The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models**
Samuel Bowles
The American Economic Review, Vol. 75, No. 1. (Mar., 1985), pp. 16-36.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198503%2975%3A1%3C16%3ATPPIAC%3E2.0.CO%3B2-N

[6] **The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration**
Sanford J. Grossman; Oliver D. Hart
The Journal of Political Economy, Vol. 94, No. 4. (Aug., 1986), pp. 691-719.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198608%2994%3A4%3C691%3ATCABOO%3E2.0.CO%3B2-F

[7] **The Role of Market Forces in Assuring Contractual Performance**
Benjamin Klein; Keith B. Leffler
The Journal of Political Economy, Vol. 89, No. 4. (Aug., 1981), pp. 615-641.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198108%2989%3A4%3C615%3ATROMFI%3E2.0.CO%3B2-K

[12] **The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration**
Sanford J. Grossman; Oliver D. Hart
The Journal of Political Economy, Vol. 94, No. 4. (Aug., 1986), pp. 691-719.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198608%2994%3A4%3C691%3ATCABOO%3E2.0.CO%3B2-F

[15] **Optimism Invites Deception**
Franklin Allen; Gerald R. Faulhaber
The Quarterly Journal of Economics, Vol. 103, No. 2. (May, 1988), pp. 397-407.
Stable URL:
http://links.jstor.org/sici?sici=0033-5533%28198805%29103%3A2%3C397%3AOID%3E2.0.CO%3B2-E

LINKED CITATIONS

References

Contracts as a Barrier to Entry
Philippe Aghion; Patrick Bolton
The American Economic Review, Vol. 77, No. 3. (Jun., 1987), pp. 388-401.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198706%2977%3A3%3C388%3ACAABTE%3E2.0.CO%3B2-7

Production, Information Costs, and Economic Organization
Armen A. Alchian; Harold Demsetz
The American Economic Review, Vol. 62, No. 5. (Dec., 1972), pp. 777-795.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28197212%2962%3A5%3C777%3APICAEO%3E2.0.CO%3B2-Q

Optimism Invites Deception
Franklin Allen; Gerald R. Faulhaber
The Quarterly Journal of Economics, Vol. 103, No. 2. (May, 1988), pp. 397-407.
Stable URL:
http://links.jstor.org/sici?sici=0033-5533%28198805%29103%3A2%3C397%3AOID%3E2.0.CO%3B2-E

The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models
Samuel Bowles
The American Economic Review, Vol. 75, No. 1. (Mar., 1985), pp. 16-36.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198503%2975%3A1%3C16%3ATPPIAC%3E2.0.CO%3B2-N

The Nature of the Firm
R. H. Coase
Economica, New Series, Vol. 4, No. 16. (Nov., 1937), pp. 386-405.
Stable URL:
http://links.jstor.org/sici?sici=0013-0427%28193711%292%3A4%3A16%3C386%3ATNOTF%3E2.0.CO%3B2-B

LINKED CITATIONS

Commitment Through Renegotiation-Proof Contracts with Third Parties
Mathias Dewatripont
The Review of Economic Studies, Vol. 55, No. 3. (Jul., 1988), pp. 377-389.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28198807%2955%3A3%3C377%3ACTRCWT%3E2.0.CO%3B2-6

Price and Authority in Inter-Profit Center Transactions
Robert G. Eccles; Harrison C. White
The American Journal of Sociology, Vol. 94, Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure. (1988), pp. S17-S51.
Stable URL:
http://links.jstor.org/sici?sici=0002-9602%281988%2994%3CS17%3APAAIIC%3E2.0.CO%3B2-G

A Non-cooperative Equilibrium for Supergames
James W. Friedman
The Review of Economic Studies, Vol. 38, No. 1. (Jan., 1971), pp. 1-12.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28197101%2938%3A1%3C1%3AANEFS%3E2.0.CO%3B2-2

The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration
Sanford J. Grossman; Oliver D. Hart
The Journal of Political Economy, Vol. 94, No. 4. (Aug., 1986), pp. 691-719.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198608%2994%3A4%3C691%3ATCABOO%3E2.0.CO%3B2-F

The Role of Market Forces in Assuring Contractual Performance
Benjamin Klein; Keith B. Leffler
The Journal of Political Economy, Vol. 89, No. 4. (Aug., 1981), pp. 615-641.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198108%2989%3A4%3C615%3ATROMFI%3E2.0.CO%3B2-K

Equilibrium Vertical Foreclosure
Janusz A. Ordover; Garth Saloner; Steven C. Salop
The American Economic Review, Vol. 80, No. 1. (Mar., 1990), pp. 127-142.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28199003%2980%3A1%3C127%3AEVF%3E2.0.CO%3B2-K

LINKED CITATIONS

Premiums for High Quality Products as Returns to Reputations

Carl Shapiro

The Quarterly Journal of Economics, Vol. 98, No. 4. (Nov., 1983), pp. 659-680.