Credible Commitments: Using Hostages to Support Exchange

By Oliver E. Williamson*

Credible commitments and credible threats share the following common attribute: both appear mainly in conjunction with irreversible, specialized investments. But whereas credible commitments are undertaken in support of alliances and to promote exchange, credible threats appear in the context of conflict and rivalry. The former involve reciprocal acts designed to safeguard a relationship, while the latter are unilateral efforts to preempt an advantage. Efforts to support exchange generally operate in the service of efficiency; preemptive investments, by contrast, are commonly antisocial. Both are plainly important to politics and economics, but the study of credible commitments is arguably the more fundamental of the two.

Interest in credible threats is much more widespread and the credible threat literature is more fully developed, however, than is the interest and economic literature dealing with credible commitments. This disparity is consistent with the treatment accorded to each in Thomas Schelling’s classic essay (1956) on bargaining, where the main emphasis is placed on tactics by which one party can realize an advantage in relation to a rival by credibly “tieg ones hands.” But Schelling also, albeit briefly, addresses the matter of promise. He observes in this connection that “Bargaining may have to concern itself with an ‘incentive’ system as well as the division of gains” (p. 300) and adds in a footnote that the exchange of hostages served incentive purposes in an earlier age (p. 300, fn. 17).

That the study of credible commitments has been relatively neglected is explained by the assumption, common to both law and economics, that the legal system enforces promises in a knowledgeable, sophisticated, and low-cost way. Albeit instructive, this convenient assumption is commonly contradicted by the facts—on which account additional or alternative modes of governance have arisen. Bilateral efforts to create and offer hostages are an interesting and, as it turns out, economically important illustration. Absent a recognition and appreciation for the merits of “private ordering,” the suggestion that hostages are used to support contemporary exchange is apt to be dismissed as fanciful. I submit, however, that not only are the economic equivalents of hostages widely used to effect credible commitments, but failure to recognize the economic pur-

*Department of Economics, Yale University, New Haven, CT 06520. This paper has benefited from the comments of participants in the Sloan Foundation Workshops in Transaction Cost Economics at the University of Pennsylvania and the University of Minnesota, and from seminar presentations at Oxford University, Williams College, and Yale University. Comments by Dennis Carlton, Frank Easterbrook, Benjamin Klein, Scott Masten, Michael Riordan, David Sappington, Richard Stewart, and Lester Telser, and from the referees of this journal are gratefully acknowledged. Research support from the National Science Foundation is gratefully acknowledged.

1It should be noted that I use the terms threat and commitment differently than do Curtis Eaton and Richard Lipsey (1981). They distinguish between empty and credible threats and use the term commitment to refer to the latter. I submit that the language of rivalry is well serviced by reference to threats; and I suggest that the term commitment be reserved to describe exchange. Thus both credible and noncredible threats would be distinguished in assessing rivalry. Similarly, credible and noncredible commitments are distinguished in evaluating exchange. Alliances complicate matters in that these are organized in relation to another party. This could be wholly beneficial, but it need not be. Thus suppliers could form an alliance in relation to buyers, with possible antisocial results. Credible commitments which simultaneously support exchange and promote alliances thus sometimes pose tradeoffs.

poses served by hostages has been responsible for repeated policy error.

The private ordering approach to the study of contract is sketched in Section I. A simple model for assessing the efficiency ramifications of alternative contracts, one of which involves hostages, is developed in Section II. The model demonstrates that the investments made by suppliers are influenced by the incentives experienced by buyers. Incentive complications that reach beyond the model are discussed in Section III. Applications of the argument to unilateral and to bilateral exchange are set out in Sections IV and V. Some evidence bearing on petroleum exchanges and public policy attitudes that relate thereto are examined in Section VI.

I. Private Ordering

A. Contracting Traditions

Most studies of exchange assume that efficacious rules of law regarding contract disputes are in place and that these are applied by the courts in an informed, sophisticated, and low-cost way. These assumptions are convenient, in that lawyers and economists are relieved of the need to examine the variety of ways by which individual parties to exchange “contract out of or away from” the governance structures of the state by devising private orderings. A division of effort thus arises whereby economists are preoccupied with the economic benefits that accrue to specialization and exchange, while legal specialists focus on the technicalities of contract law.³

The “legal centralism” tradition reflects this orientation. It maintains that “disputes require ‘access’ to a forum external to the original social setting of the dispute...[and that] remedies will be provided as prescribed in some body of authoritative learning and dispensed by experts who operate under the auspices of the state” (Marc Galanter, 1981, p. 1). The facts, however, disclose otherwise. Most disputes, including many of those which, under current rules, could be brought to a court, are resolved by avoidance, self-help, and the like (Galanter, p. 2).

The unreality of the assumptions of legal centralism can be defended by reference to the fruitfulness of the pure exchange model. This is not disputed here. My concern with this tradition is that the law and economics of private ordering have been pushed into the background as a consequence. This is unfortunate, since “In many instances the participants can devise more satisfactory solutions to their disputes than can professionals constrained to apply general rules on the basis of limited knowledge of the dispute” (Galanter, p. 4).⁴

Four distinct, albeit related, literatures within economics⁵ have developed over the past decade in which private ordering is expressly or implicitly featured: the incentive compatibility literature (Leonid Hurwicz, 1972); the literature on the economics of internal organization (Ronald Coase, 1937; Kenneth Arrow, 1963, 1974; myself, 1971, 1975, 1979; Benjamin Klein et al., 1978; David Teece, 1982; Eugene Fama and Michael Jensen, 1983); the financial economics literature dealing with bonding (Joseph Stiglitz, 1974; Jensen and William Meckling, 1976; Sanford Grossman and Oliver Hart, 1982); and the study of self-enforcing agreements (Lester Telser, 1981; Klein and Keith Leffler, 1981). The first two of these have been reviewed elsewhere (Hurwicz, 1973; myself, 1981, 1982b). The third takes issue

³Lawyers do not have a monopoly on refining contractual rules. For recent contributions by economists, see Peter Diamond and Eric Maskin (1979) and Steven Shavell (1980). Such an economic approach to contract focuses on the technicalities of legal rules. The approach taken here holds that even refined rules of law are costly to implement, whence private ordering is widely employed.

⁴Galanter elaborates as follows: “The variability of preferences and of situations, compared to the small number of things that can be taken into account by formal rules..., and the loss of meaning in transforming the dispute into professional categories suggest limits on the desirability of conforming outcomes to authoritative rules” (1981, p. 4).

⁵There is also a long legal tradition in which contract as legal rule is disputed. Karl Llewellyn’s views regarding “contract as framework” (1931) are especially important. Recent significant contributions include Stewart Macaulay (1963) and Ian Macneil (1974). For a discussion, see my 1979 article.
with the Modigliani-Miller theorem that the firm's production plan is independent of its financial structure. The fourth deals with contracting in intermediate product markets and expressly relies on private ordering.

Telser characterizes a self-enforcing agreement as one which, if "one party violates the terms the only recourse of the other is to terminate the agreement" (p. 27). Contrary to legal centralism, the courts and other third parties are assumed away. Klein and Leffler are explicit on this: "we assume throughout... that contracts are not enforceable by the government or any third party" (p. 616). Commercial contract law in late nineteenth-century Taiwan evidently approximated this condition (Rosser Brockman, 1980). Stewart Macaulay's remarks about the informality of contract in business are likewise in this spirit: "Often businessmen do not feel they have 'a contract'—rather they have 'an order.' They speak of 'cancelling the order' rather than 'breaching our contract'" (1963, p. 61).

To be sure, pure private ordering is extreme. As Robert Mnookin and Lewis Kornhauser state, private ordering invariably operates in "the shadow of the law" (1979). It suffices for my purposes to argue that the incentives of private parties to devise bilateral contractual safeguards is a function of the efficacy of court adjudication, and that this varies with the attributes of transactions. Specifically the courts experience serious disabilities with respect to the transactions of the kinds herein described.

B. Some Attributes of this Article

This article examines self-enforcing agreements in transaction cost terms. Contracting agents are thus assumed to be subject to bounded rationality and, where circumstances permit, are given to opportunism. Although hostages can have both ex ante (screening) and ex post (bonding) effects, the ex post contract execution consequences are of principal interest here.9 This is also the focus of the self-enforcing agreement literature. Additionally, like both Telser and Klein-Leffler, the intertemporal contracts of concern here feature both uncertainty and transaction-specific capital. But in other respects there are important differences.

Thus, whereas Telser deals with "a sequence of transactions over time such that the ending date is unknown and uncertain" (p. 30), because any finite sequence of transactions using his model will unravel (p. 29), the transactions that I consider can be (indeed, normally are) finite. Furthermore, the role of transaction specific capital is more explicit and fully developed in this article than in Telser's.

The self-enforcing contracts studied by Klein and Leffler are likewise of indefinite rather than finite duration. The hostage model is further distinguishable from Klein-Leffler in that 1) they deal with quality un-
certainty in final goods markets, whereas I assume that quality is known and focus on stochastic demand in intermediate product markets; their "fundamental theoretical result" involves the assurance of quality through the sacrifice of "minimum-cost production techniques" (pp. 618, 628–29), while the hostage model involves no such sacrifice (indeed, the use of hostages to support exchange encourages investment in specific asset technologies which have lower expected costs);1) suppliers in their model are confronted neither with expropriation nor adaptation hazards, while these are both of concern to me. More generally, the hostage model and imagery have application to a quite different set of contractual circumstances than the analysis of quality assurance contemplates.

II. The Hostage Model

The simple hostage model serves to illuminate both unilateral and bilateral exchange, permits the concept of specific capital to be extended beyond earlier uses, and clarifies how costs should be described in assessing exchange. While it is primitive and suggestive, rather than refined and definitive, it serves as a paradigmatic wedge by which the importance of private ordering is exposed and is easily made the vehicle for further analysis.

A. Technologies and Costs

The assessment of alternative contracts will be facilitated by assuming that the product in question can be produced by either of two technologies. One is a general purpose technology; the second is a special purpose technology. The special purpose technology requires greater investment in transaction specific durable assets and, as described below, is more efficient for servicing steady-state demands.

Costs that are highly specific to a transaction have two attributes: they are incurred in advance of the contemplated exchange; and their value in alternative uses, or by alternative users, is greatly reduced.12 As Klein and Leffler put it, the irreversible, nonsalvageable part of an advance commitment is sunk (p. 619). It is common to think of this as applying to physical plant or accounting costs that are reported as fixed, but this is not the critical distinction. Thus investments in labor (transaction specific human capital) can be highly specific. And many costs that for accounting purposes are reported to be fixed are in fact nonspecific, hence can be recovered (salvaged) by redeployment. Durable but mobile assets such as general purpose trucks or airplanes are illustrations.

The two technologies in question will thus be described in value realization terms. The value that can be realized by redeploying variable and fixed costs will be given by \( v \). The nonsalvageable value of advance commitments will be denoted by \( k \). The two technologies can thus be described as \( T_1 \); the general purpose technology, all advance commitments of which are salvageable, the redeployable unit operating costs of which are \( v \); and \( T_2 \); the special purpose technol-

10My discussion of franchising in Section IV.C parallels Klein-Leffler and assumes that quality uncertainty is responsible for a demand externality. The hostage model developed in Section II does not apply directly to this case, but the spirit carries over in three respects: 1) franchisees, like buyers, are given a choice among alternative contracts; 2) the decision to expose specific assets is deliberately taken because this has superior incentive effects; and 3) the sunk cost technology is more efficient, which vitiates the inefficiency tradeoff that is central to the Klein-Leffler paper.

12Contrary to the argument advanced by Klein and Leffler, total costs need not increase in a quality assurance model because investments are shifted from a reversible (fixed cost) to an irreversible (sunk cost) technology. Thus instead of a general purpose building of nondescript design, the producer could construct a building with a distinctive "signature." The durable investments could be the same, but the alternative value that can be realized from the second building might be much lower. The long-term commitments that are signaled by this second design relieve customers of quality shading hazards, which is the central issue with which Klein and Leffler are concerned.

Klein et al. use the term "appropriable quasi rent" to describe this condition. Use vs. user distinctions are relevant in this connection: "The quasi-rent value of the asset is the excess of its value over its salvage value, that is, its value in its next best use to another renter. The potentially appropriable specialized portion of the quasi-rent is the portion, if any, in excess of its value to the second highest-valuing user" (p. 298).
ogy, the nonsalvageable value of advance commitments of which are \( k \) and the redeployable unit operating costs of which are \( v_2 \).

### B. Contracting

There are two periods. Orders are placed in the first, and production if any occurs in the second. Buyers can either take delivery or refuse it. Demand is stochastic. The gross value to buyers is assumed to be uniformly distributed over the interval \([0, 1]\), and the quantity demanded at every price will be assumed to be constant, which it will be convenient to set equal to unity. Sunk costs, if any, are incurred in the first period. Inasmuch as sunk costs are incurred for certain while the decision to incur redeployable costs is contingent on the buyer's decision to confirm or cancel an order, a choice between technologies is interesting only if \( k + v_2 < v_1 \).

The demand and cost relations are set out in Figure 1.

#### 1. Net Benefits

The criterion by which decisions to take or refuse delivery will be evaluated is that of joint profit maximization. Feasibility and/or bureaucratic disabilities aside, vertical integration assuredly accomplishes the joint profit-maximization result. Thus the reference condition for evaluating contracts will be an integrated firm with two divisions, a producing division and a marketing division. The producing division has access to the same two technologies described above, one of which involves specific assets, the other of which does not. Whichever technology is employed, product is transferred between divisions at marginal cost.

That \( k + v_2 < v_1 \) does not establish that the special purpose technology \( (T_2) \) is the more efficient. Whether it is or not depends on a net benefit calculation. The expected net benefits of using the general purpose technology \( (T_1) \) are given by the product of the probability that the integrated firm will decide to produce and the average net benefits that are realized when product is supplied. The integrated firm will decide to produce only if the realized demand price exceeds marginal costs, whence the probability of production under \( T_1 \) is \( 1 - v_1 \). The mean net benefits during production periods are \( (1 - v_1)/2 \), whence the expected net benefits for technology \( T_1 \) are

\[
(1) \quad b_1 = (1 - v_1)(1 - v_1)/2 = (1 - v_1)^2/2.
\]

The expected net benefits for the specific asset technology \( (T_2) \) are found similarly.

Expected net receipts, however, must be reduced by the amount of the earlier investment in specific assets, \( k \), in computing expected net benefits. Thus we have

\[
(2) \quad b_2 = (1 - v_2)(1 - v_2)/2 - k = (1 - v_2)^2/2 - k,
\]

where the first term is the expected excess of revenue over out-of-pocket costs.

The specific asset technology will be selected only if \( b_2 > b_1 \), which requires that

\[
(3) \quad k < (1 - v_2)^2/2 - (1 - v_1)^2/2.
\]

#### 2. Autonomous Contracting

Assume that the inequality in (3) holds and consider the case of autonomous contracting between a buyer, who services final demand, and a producer, who manufactures the product. Assume that demand and production technologies are as described above. Efficient contracting relations are those that replicate the vertical integration result,\(^{13}\)

\(^{13}\)Here and throughout the remainder of this article I assume that exchange is governed by contract rather than by vertical integration. Vertical integration is thus used merely as a reference condition. That manufacturers do not integrate forward or distributors integrate backward can be explained on a number of
namely, 1) select the specific asset technology, and 2) produce and sell product whenever realized demand price exceeds \( v_2 \). Assume that both parties are risk neutral and that the production side of the industry is competitively organized. Whatever contracting relation is described, producers will be willing to supply if a break-even condition (expressed in expected value terms) can be projected.\(^{14}\)

Recall that orders are placed in the first period. Specific assets, if any, are committed in the first period in anticipation of second-period supply. Whether second-period production actually occurs, however, is contingent on demand realizations. Buyers have the option of confirming or cancelling orders in the second period. Consider three contracting alternatives:

I. The buyer purchases specific assets and assigns them to whichever seller submits the lowest bid, \( p \);

II. The producer makes the specific asset investment himself and receives a payment of \( \bar{p} \) in the second period if the buyer confirms the order but nothing otherwise; and

III. The producer makes the specific asset investment himself and receives \( \alpha h \), \( 0 < \alpha < 1 \), if the order is cancelled while the buyer pays \( \bar{p} \) upon taking delivery and experiences a reduction in wealth of \( h \) if second-period delivery is cancelled.

The third scenario can be thought of as one where the buyer posts a hostage that he values in amount \( h \), which hostage is delivered to the producer, who values it in amount \( \alpha h \) if the order is cancelled.

The producer will break even under contracting relation I if he is compensated in amount \( v_2 \), which is his out of pocket cost, for each unit demanded. The low bidder will thus offer to supply product for \( p = v_2 \). Since the buyer's net benefits are maximized if he invests in the specific assets, and since product is transferred on marginal cost terms, this contract replicates the vertical integration relation. Contracts of type I are feasible, however, only if the specialized assets are mobile and the specificity is attributable to physical features (for example, specialized dies). Market procurement can then service the needs of the parties without posing hold-up problems by concentrating the ownership of the specific assets on the buyer (who then assigns them to the low bidder). Inasmuch as the buyer can reclaim the dies and, without cost, solicit new bids should contractual difficulties develop, type I contracts yield an efficient result.\(^{15}\)

Attention hereafter will be focused on contracts II and III, the assumption being that asset specificity is of the human or dedicated asset kinds (see Part C, below). The autonomous buyer will confirm an order under contract II whenever realized demand price exceeds \( \bar{p} \) but not otherwise. The producer will thus break even if

\[
(1 - \bar{p})\bar{p} - [(1 - \bar{p})v_2 + k] = 0,
\]

whence

\[
\bar{p} = v_2 + k / (1 - \bar{p}).
\]

Product will thus be exchanged at a price that exceeds marginal cost under this contracting scenario.\(^{16}\) Plainly if \( \bar{p} \geq v_1 \), the

\(^{14}\)There is no problem in principle in allowing suppliers to extract positive profits as a condition of supply. The salient features of the hostage model are all preserved if, instead of an expected break-even condition, the supplier was assumed to realize expected profits of \( \bar{\pi} > 0 \) on each contract. Although final demands will be choked off as a consequence, the main features of the contractual argument survive.

\(^{15}\)This ignores the possibility that suppliers will abuse the dies if ownership resides with the buyer.

\(^{16}\)Conceivably \( \bar{\pi} \) will exceed \( v_1 \), in which event the buyer who is contemplating contract II will prefer instead to purchase from sellers who use the general purpose technology. The comparison in the text implicitly assumes that \( \bar{\pi} < v_1 \). Also note that a standby technology that can be costlessly switched into and out of the product in question could effectively truncate demand at \( v_1 \). This would be true if potential middlemen could place orders to take product at \( v_1 \) from general purpose manufacturers, which orders could be costlessly cancelled (and general purpose assets redeployed) if demands fell below this value. I will arbitrarily assume that this is not feasible. The problem could, however, be reformulated by describing demand as uniformly distributed over the interval \( 0 \) to \( v_1 \), with \( v_1 \) having measure \( 1 - \bar{\pi} \).
buyer is better off to scuttle contract II and purchase instead from producers who utilize the (inferior) variable cost technology $T_1$ (and will break even by supplying product on demand for a price of $v_1$).

The buyer will confirm an order under contract III whenever the realized demand price exceeds $\beta - h$. Let $\beta - h$ be denoted by $m$. The seller will then break even when $(1 - m)\beta + mah - [(1 - m)v_2 + k] = 0$, whence

$$\beta = v_2 + (k - mah)/(1 - m).$$

The case where $h = k$ and $\alpha = 1$ is one where the buyer gives up wealth in amount of the investment in specific assets in cancellation states and this is delivered to the producer who values it in amount $k$. Under these circumstances, (5) becomes

$$\beta = v_2 + k.$$

Since the buyer places an order whenever demand exceeds $m = \beta - h$, this yields the result that $m = v_2$, whence orders will be placed whenever demand exceeds $v_2$, which is the efficient (marginal cost) supply criterion.

The buyer's net benefits under contracting scheme III are

$$b_3 = (1 - m)\left[\left(m + \frac{1 - m}{2}\right) - \beta\right] - mh,$$

where $(1 - m)$ is the probability of placing an order, $m + (1 - m)/2$ is the expected demand price for all orders that are placed, $\beta$ is the payment in demand confirmation states to the producer, and $h$ is the wealth sacrifice in cancellation states (which occur with probability $m$). Under the assumptions that $h = k$ and $\alpha = 1$, this reduces to

$$b_3 = (1 - v_2)^2/2 - k,$$

which is identical to the net benefit calculation for technology $T_2$ under the vertical integration reference condition (see equation (2)).

Accordingly, contracting scheme III accompanied by the stipulations that $h = k$ and $\alpha = 1$ replicates the efficient investment and supply conditions of vertical integration. Problems arise, however, if $h < k$ or $\alpha < 1$. The disadvantage, moreover, accrues entirely to the buyer—since the seller, by assumption, breaks even whatever contracting relation obtains. Thus although after the contract has been made, the buyer would prefer to offer a lesser-valued hostage and cares not whether the hostage is valued by the producer, at the time of the contract he will wish to assure the producer that a hostage of $k$ for which the producer realizes full value ($\alpha = 1$) will be transferred in nonexchange states. Failure to make this commitment will result in an increase in the contract price. Thus, whereas producers who are concerned only with ex ante screening can tolerate values of $\alpha$ less than one—see the discussion of ugly princesses in Section III.A, below—this is not the case at all when ex post opportunism is the concern. If the producer is not indifferent, as between two princesses, each of whom is valued identically by the buyer, the producer's preferences now need to be taken into account.\textsuperscript{17}

To summarize, therefore, it can be observed that contract I mimics vertical integration, but only under special asset specificity conditions; contract II is inferior; and contract III yields the vertical integration result if $h = k$ and $\alpha = 1$. Furthermore, note that an important feature of contract III is that the buyer takes delivery in all demand states for which realized demand exceeds $m = \beta - h$. Since the supplier is always paid $\beta$ upon execution, the buyer sometimes takes delivery when his realized receipts (upon resale of the product) are less than $\beta$. This does not, however, signal inefficiency, since orders are never confirmed when realized demand price falls below marginal cost ($v_2$). Indeed, it is precisely because of the hostage feature that efficiency is realized and contract III is superior to contract II.

\textsuperscript{17}Placing an upper bound of unity on $\alpha$ precludes the possibility that the supplier values the hostages more than does the buyer. Potential gains from trade would exist for all hostages for which $\alpha$ exceeds unity. A case for negatively valued hostages could be made in the context of ugly princesses (see Section III, Part A).
C. Dedicated Assets

Three types of asset specificity have been recognized in earlier discussions: site specificity—as where successive stations are located in a cheek-by-jowl relation to each other so as to economize on inventory and transportation expenses; physical asset specificity—such as specialized dies that are required to produce a component; and human asset specificity that arises in a learning-by-doing fashion. A fourth type, previously unremarked but nevertheless important for some transactions, needs to be added: dedicated assets. Such assets represent a discrete investment in plant. Although these assets add to the firm’s generalized (as contrasted with special purpose) production capability, the investment would not be undertaken but for the prospect of selling a significant amount of product to a specific customer. As with other types of asset specificity, dedicated assets lose value if employed in alternative uses (or by or to service alternative users). Dedicated assets thus are those that are put in place contingent upon particular supply agreements and, should such contracts be prematurely terminated, would result in significant excess capacity.

III. Engaging the Supplier

Suppliers are passive instruments in this model. They are indifferent among contracts, since their expected profits are the same (zero) whichever choice the buyer makes. What drives the argument is that buyers can secure better terms only by relieving producers of demand cancellation penalties. Buyers cannot have their cake (product supplied by the efficient technology at a price of \( p \)) and eat it too (cancel without cost). Inasmuch as optimality is realized if \( h = k \) and \( a = 1 \), the ideal hostage would appear to be an offer of generalized purchasing power: money. A security bond in amount \( h = k \) would serve this purpose. That the argument does not terminate here is because such an arrangement does not assuredly engage the interests and cooperation of the supplier. Three reasons can be adduced for this condition: contrived cancellation, uncertain valuation, and incomplete contracting. All are a consequence of joining bounded rationality with opportunism.

A. Supplier Opportunism

1. Contrived Cancellation

The issue of contrived cancellation has been addressed by Kenneth Clarkson, Roger Miller, and Timothy Muris in their discussion of refusal of the courts to enforce stipulated damage clauses where breach has been deliberately induced (1978, pp. 366–72). Induced breach could arise where a party intentionally withholds relevant information, yet complies with the letter of the contract. Or it might involve perfunctory fulfillment of obligations where more resourceful cooperation is needed (Clarkson et al., pp. 371–72). In either case, induced breach is costly to detect and/or prove (Clarkson et al., p. 371).

This explanation for selective enforcement of liquidated damage clauses has troubled other legal scholars (Richard Posner, 1979, p. 290), but a more satisfactory explanation has yet to be advanced. At the very least, the Clarkson et al. treatment reflects a sensitivity to the subtleties of opportunism—on which account private ordering is more complicated than the bare bones hostage model would suggest. Among other things, the expropriation hazard to which they refer may explain the use of ugly princesses.

Thus suppose that demand uncertainties are negligible, whence order cancellation hazards can be disregarded. Suppose further, however, that buyers differ in credit risk respects, and that producers would, if they could, refuse sales to poor risks. Assuming that the difference between good and poor risks is sufficiently great that a separating equilibrium is feasible, producers could demand hostages (or, put differently, good risks could offer hostages) as a way by which to

18Common ownership is the preponderant response to site specificity. Close proximity is desired because of inventory or related processing cost (for example, thermal economy) savings. Once sited, the assets in question are highly immobile—which is to say that the setup and/or relocation costs are great.

19See fn. 9.
screen. Given, moreover, that the only use to which hostages are put is as a screen, a value of \( \alpha = 0 \) would accomplish this purpose without exposing the buyer to an expropriation hazard (based, say, on a legal technicality). Specifically, a king who is known to cherish two daughters equally and is asked, for screening purposes, to post a hostage is better advised to offer the ugly one.

2. Uncertain Valuation

The model assumes that the value of the specific investment \((k)\) is well specified. This need not be the case. Indeed, it may be difficult for buyers to ascertain whether the investments made in response to first-period orders are of the amount or of the kind that producers claim. This is not a serious problem if the production side of the market is competitively organized and fly-by-night concerns can be disregarded. Where, however, this cannot be presumed, the possibility that buyers will be expropriated arises. Producers may feign delivery competence (claim to have invested in specific assets in amount \(k\) but only committed \(k' < k\)) and expropriate bonds for which \(h = k\) by contriving breach or invoking a technicality.

This hazard is especially great if the producer, who retains possession of the assets for which specificity is claimed, can preserve asset values by integrating forward into the buyer's market upon taking possession of the hostage. Even though the producer is poorly suited to perform successor stage functions, the possession of specialized stage I assets effectively reduces the costs that would otherwise attend de novo stage II entry.

To be sure, the buyer who offers a hostage and recognizes a risk of contrived expropriation will adjust the original terms to reflect this. Specifically, contracts supported by hostages for which expropriation risks are believed to be great will command less than those where these same hazards are believed to be lower. But this is to concede that, absent additional safeguards, neither the transfer of product on marginal cost terms nor the efficient level and kind of investment will assuredly attend contracts of type III. Deeper governance issues than those contemplated by the simple model are evidently posed.

3. Incomplete Contracts/Haggling

For the reasons and in the ways described elsewhere (see my 1975 study, pp. 20–36; 91–94), complex contracts are invariably incomplete and many are maladaptive. The reasons are two: many contingencies are unforeseen (and even unforeseeable); and the adaptations to those contingencies that have been recognized and for which adjustments have been agreed to are often mistaken—possibly because the parties acquire deeper knowledge of production and demand during contract execution than they possessed at the outset (Richard Nelson and Sidney Winter, 1982, pp. 96–136). Instrumental gap filling, thus, is an important part of contract execution. Whether this is done easily and effectively, or if instead reaching successive agreements on adaptations and their implementation is costly, makes a huge difference in evaluating the efficacy of contracts.

Thus even if contrived breach hazards could be disregarded, producers who are entirely open and candid about contract execution may nevertheless be in a position to haggle—thereby to expropriate sellers—because contracts are incomplete or maladaptive. Specialized governance structures that have the purpose and effect of promoting harmonious adaptations and preserving the continuity of exchange relations arise in response to this condition. Knowledgeable third parties and reciprocal exposure of specialized assets are two possibilities.

B. Protective Governance Structures

1. Arbitration

Institutions that have the capacity to evaluate disputes in a more knowledgeable way than the courts may arise in this way. The parties, for example, may agree to submit disputes over contract execution to arbitrators who have specialized knowledge of the industry. Lon Fuller's remarks concerning procedural differences between arbitration and litigation are instructive:

...there are open to the arbitrator... quick methods of education not open to the courts. An arbitrator will frequently interrupt the examination of witnesses with a request that the parties
educate him to the point where he can understand the testimony being received. The education can proceed informally, with frequent interruptions by the arbitrator, and by informed persons on either side, when a point needs clarification. Sometimes there will be arguments across the table, occasionally even within each of the separate camps. The end result will usually be a clarification that will enable everyone to proceed more intelligently with the case. [1963, pp. 11–12]

Many agreements which, were it not for arbitration, would be regarded as excessively hazardous can, in this way, be reached and implemented.  

2. Reciprocal Exposure

An alternative way by which to protect contracts against expropriation is to expand the contractual relation. One way of accomplishing this is for buyer and seller to devise a mutual reliance relation. Thus suppose that the buyer does not post a hostage as such, but himself invests in specific capital that has value only in conjunction with servicing final demands for the product in question. Assume that these are valued in amount $k''$. The buyer then has the incentive to take delivery as long as realized demand exceeds $\hat{p} - k''$. If $k'' = k$, this yields the marginal cost supply result, and the producer will be satisfied with the buyer's incentives. Or suppose that producer and buyer engage in reciprocal trade. Specifically, suppose that the producer contracts to procure product from the buyer, the supply of which requires the buyer to invest in specific assets in amount $k''$. Each party to this reciprocal trade will experience appropriate incentives if 1) $k'' = k$, 2) demand variation in the two markets is perfectly correlated, and 3) each party has the option to cancel an order if a cancellation notice is received from his opposite. As discussed in Section V, bilateral trades (reciprocity; swaps) can sometimes be made to approximate these conditions.

IV. Unilateral Trading

The argument that buyers can affect the terms and manner of supply by offering (or refusing to offer) hostages has ramifications for Robinson-Patman price discrimination and to an understanding of franchising and two-part pricing.

A. Robinson-Patman

The Robinson-Patman Act has been interpreted as an effort “to deprive a large buyer of [discounts] except to the extent that a lower price could be justified by reason of a seller’s diminished costs due to quantity manufacture, delivery, or sale, or by reason of the seller’s good faith effort to meet a competitor’s equally low price.” Plainly, that $\hat{p}$ is less than $p_{r}$ in the hostage model has neither quantity nor meeting competition origins. Neither is it contrary to the public interest. Indeed, it would be inefficient and unwarranted for a producer to charge the same price to two customers who order an identical amount of product, but only one of which offers a hostage, if 1) investments in specialized assets are required to support the transactions in question, or 2) if, because of a refusal to make a credible commitment, transactions of the second kind are produced

20 Labor unions can help to assure integrity in contractual relations where workers are asked to accept assignments that involve considerable investments in human capital. Not only can the union intercede on behalf of the worker(s) where an expropriation effort is suspected, but it provides an institutional memory whereby reputation effects can be communicated to successor generations of workers. For both of these reasons, the firm is deterred from attempting expropriation. Setting aside the possibility that unions will attempt to negotiate monopoly wages, perceptive firms will prefer and actively assist in the creation of unions if these serve to attenuate expropriation risks—since otherwise workers may refuse to make (or will need to be bribed to make) mutually beneficial investments in human capital.

21 The buyer must, of course recover his full costs if he is to place specialized marketing assets at hazard. This will obtain if final demand is uniformly distributed over the interval $k''$ to $1 + k''$ and derived demand is as described earlier.

22 This last condition protects each against a prisoner’s dilemma result.

23 FTC v. Morton Salt Co., 334 U.S. 37 (1948); emphasis added.
with a general purpose (but high cost) technology.

The missing ingredients, plainly, are the differential commitment to buy (as reflected by the willingness to offer hostages) and the differential incentives to breach once hostages have been posted. The confusion is explained by the propensity to employ conventional (steady state) microtheory to the neglect of transaction cost aspects. Rectifying this involves examination of the microanalytics of transactions, with special reference to asset specificity and the hazards thereby posed, and evaluating alternative contracts with respect to a common reference condition—prospective break even being a useful standard. Once this is done, a different understanding of many nonstandard or unfamiliar contracting practices, many of which are held to be presumptively unlawful, frequently emerges.

B. Franchising

Klein and Leffler argue that franchisees may be required to make investments in transaction specific capital as a way by which to safeguard the franchise system against quality shading. As Klein puts it, franchisers can better

...assure quality by requiring franchisee investments in specific...assets that upon termination imply a capital loss penalty larger than can be obtained by the franchisee if he cheats. For example, the franchiser may require franchisees to rent from them short term (rather than own) the land upon which their outlet is located. This lease arrangement creates a situation where termination can require the franchisee to move and thereby impose a capital loss on him up to the amount of his initial non-salvageable investment. Hence a form of collateral to deter franchisee cheating is created.

[1980, p. 359]

The arrangement is tantamount to the creation of hostages to restore integrity to an exchange.

This logic notwithstanding, the use of hostages to deter franchisees from exploiting demand externalities is often regarded as an imposed (top down) solution. Franchisees are "powerless"; they accept hostage terms because no others are available. Such power arguments are often based on ex post reasoning. That the use of hostages to support exchange can be and often is an efficient systems solution, hence is independent of who originates the proposal, can be seen from the following revised sequence.

Suppose that an entrepreneur develops a distinctive, patentable idea that he sells outright to a variety of independent suppliers, each of which is geographically dispersed and is assigned an exclusive territory. Each supplier expects to sell only to the population located within its territory, but all find to their surprise (and initially to their delight) that sales are also made to a mobile population. Purchases by the mobile population are based not on the reputation of individual franchisees but on customers' perceptions of the reputation of the system. A demand externality arises in this way.

Thus, if sales were made only to the local population, each supplier would fully appropriate the benefits of its promotional and quality enhancement efforts. Population mobility upsets this; since the costs savings that result from local quality debasement accrue to the local operator while the adverse demand effects are diffused throughout the system, suppliers now have an incentive to free ride off of the reputation of the system. Having sold the exclusive territory rights outright, the entrepreneur who originated the program is indifferent to these unanticipated demand developments. It thus remains for the collection of independent franchisees to devise a correction themselves—lest the value of the system deteriorate to their individual and collective disadvantage.

24Note that the argument applies only to \( \hat{p} \) vs. \( \hat{P} \) comparisons in trades where specific assets are involved. The efficiency properties of customer price differentials that do have these origins are not reached by the argument in this paper.

25That this is a useful way to pose the franchise issue evolved out of discussions that I had with Jeffrey Goldberg. For a more complete development, see Goldberg's dissertation, 1982.
The franchisees, under this revised scenario, thus create an agent to police quality or otherwise devise penalties that deter quality deterioration. One possibility is to return to the entrepreneur and hire him to provide these services. Serving now as the agent of the franchisees, the entrepreneur may undertake a program of quality checks (certain purchasing restraints are introduced, whereby franchisees are required to buy only from qualified suppliers; periodic inspections are performed). The incentive to exploit demand externalities may further be discouraged by requiring each franchisee to post a hostage and by making franchises terminable.26

This indirect scenario serves to demonstrate that it is the system that benefits from the control of externalities. But this merely confirms that the normal scenario in which the franchiser controls the contractual terms is not an arbitrary exercise of power. Indeed, if franchisees recognize that the demand externality exists from the outset, if the franchiser refuses to make provision for the externality in the original contract, and if it is very costly to reform the franchise system once initial contracts are set, franchisees will bid less for the right to a territory than they otherwise would. It should not therefore be concluded that perceptive franchisers, who recognize the demand externality in advance and make provision for it, are imposing objectionable ex ante terms on unwilling franchisees. They are merely taking steps to realize the full value of the franchise. Here as elsewhere, contracts need to be examined in their entirety.

C. Two-Part Pricing

Victor Goldberg and John Erickson (1982) describe an interesting two-part pricing scheme that they observed in the sale of coke. The producer both sold coke to the calciner, and owned and leased the land upon which the plant of the calciner was built. Inasmuch as the coke was sold for “about one-quarter the current market price of equivalent quality coke” (p. 25), Goldberg and Erickson conjecture that “the rental rate was above the fair market rate and that the contract was designed to ensure that [the calciner] would continue to perform” (p. 25). Assuming that marginal costs are much less than average, such an arrangement can be interpreted as one by which the parties are attempting to strike efficient pricing terms that approximate those of the hostage model.

The pricing of utility services, whereby ex ante installation fees are paid by subscribers, also have interesting two-part pricing attributes.27 The risk that sellers will expropriate buyers upon receipt of advance payment can be mitigated by creating a specialized third party, which for convenience may be referred to as a regulatory commission (Goldberg, 1976). Utilization of utility services can then be priced so as to more nearly approximate marginal cost.

More generally, Goldberg and Erickson conjecture that nonlinear pricing schemes are much more widespread than is commonly believed. They further point out that such arrangements are often very subtle and will require detailed knowledge of contracts to investigate (pp. 56–57).

V. Bilateral Applications

As indicated, the offer of hostages poses a hazard of expropriation. One way to deter this is to expand the contracting relationship from one of unilateral to bilateral exchange. Credible commitments are signaled without exposing assets to expropriation hazards. Reciprocal trades, especially those that involve product exchanges (swaps), sometimes come about in this way.

A. Reciprocity, General

Reciprocity transforms a unilateral supply relation—whereby A sells X to B—into a bilateral one, whereby A agrees to buy Y

26Termination is a credible threat only if the franchisee who cheats on the system bears a capital loss. This is the basic Klein and Leffler message. It would not do, therefore, if the terminated franchisee were permitted to sell the franchise to a highest bidder unless the investment in specific capital took the form of the franchisee's specialized knowledge of the system, and the terminated franchisee were thereafter prohibited from participating in owner, adviser, or employee status.

27This possibility was suggested to me by Alvin Klevorick.
from B as a condition for making the sale of X and both parties understand that the transaction will be continued only if reciprocity is observed. Although reciprocal selling is widely held to be anticompetitive (George Stocking and Willard Mueller, 1957; Harlan Blake, 1973), others regard it more favorably. George Stigler offers the following affirmative rationale for reciprocity:

The case for reciprocity arises when prices cannot be freely varied to meet supply and demand conditions. Suppose that a firm is dealing with a colluding industry which is fixing prices. A firm in this collusive industry would be willing to sell at less than the cartel price if it can escape detection. Its price can be reduced in effect by buying from the customer-seller at an inflated price. Here reciprocity restores flexibility of prices. [1969, p. 39]

Inasmuch as, however, many industries do not satisfy the prerequisites for oligopolistic price collusion (Posner; myself, 1975, ch. 12) and as reciprocity is sometimes observed among these, reciprocity presumably has other origins as well. Tie breaking is one of these. A second is that reciprocity can have advantageous governance structure benefits. These two can be distinguished by the type of product being sold.

The tie-breaker explanation applies where firm B, which is buying specialized product from A, asks that A buy standardized product from B on the condition that B meets market terms. Other things being equal, procurement agents at A are apt to accede. F. M. Scherer notes that “Most of the 163 corporation executives responding to a 1963 survey stated that their firms’ purchases were awarded on the basis of reciprocity only when the price, quality, and delivery conditions were equal” (1980, p. 344).

The more interesting case is where reciprocity involves the sale of specialized product to B conditioned on the procurement of specialized product from B. The argument here is that reciprocity can serve to equalize the exposure of the parties, thereby reducing the incentive of the buyer to defect from the exchange—leaving the supplier to redeploy specialized assets at greatly reduced alternative value. Absent a hostage (or other assurance that the buyer will not defect), the sale by A of specialized product to B may never materialize. The buyer’s commitment to the exchange is more assuredly signaled by his willingness to accept reciprocal exposure of specialized assets. Defection hazards are thereby mitigated.

Lest the argument be uncritically considered to be a defense for reciprocal trading quite generally, note that it applies only where specialized assets are placed at hazard by both parties. Where only one or neither invests in specialized assets, the practice of reciprocity plainly has other origins.28

B. Exchanges

Although reciprocal trading among non-rivals may occasionally be justified, the exchange of product among nominal rivals is surely more puzzling and troublesome. Firms that are presumed to be in head-to-head competition ought to be selling product against one another rather than to one another. What explains the reverse?

Several distinctions are useful in considering exchanges. First, trade among rivals—short term or long term, unilateral or bilateral—is feasible only if product is fungible. This is not true for many differentiated goods and services, whence the issue of trade among rivals never arises for these. Second, short-term supply agreements are usefully distinguished from long term. The former may be explained as an “occasional exception,” whereby one rival will sell product to another on a short-term, gap-filling basis so as to provide temporary relief against unanticipated product shortfalls (occasioned by either demand or supply changes). Recognizing that the shoe may be on the other foot next time, otherwise rivalrous firms may assist one another for stop-gap purposes. Public policy can presumably recognize merit in such trades and, so long as they lack a pat-

28 Possible trading objections are discussed by Scherer (pp. 344–45). Another objection is that reciprocity becomes a bureaucratic habit that salesmen and purchasing agents find convenient and that outsiders are thereby disadvantaged in attempting to secure sales. See my 1975 study, pp. 163–64.
tern, hence do not give rise to a "web of interdependence," will regard these as unobjectionable. Long-term trading among rivals is, however, much less consistent with the notion of effective head-to-head rivalry. At the very least, such arrangements warrant scrutiny.

Whether there are efficiency incentives for rivals to supply product to one another on a long-term basis turns initially on prospective realization of production cost savings. The realization of production cost savings through long-term trade between rivals requires that economies of scale be large in relation to the size of geographic markets and, if they are, that firm-specific reputation effects extend across geographic market boundaries. The former is obvious since, absent economies of scale, every firm would presumably supply everywhere to its own long-term needs. Where, however, scale economies are significant, each market will support only a limited number of plants of minimum efficient size. But fungibility and scale economies do not establish that gains from trade will be realized from such sales. This will obtain only if the value of (identical) product sold by rivals exceeds that sold by the local supplier. The issue here is whether valued reputation effects will go unrealized if rivals are unable to secure local product on favorable terms. Firms that possess valued reputations that extend beyond their local market to include distant markets are thus the ones for which long-term supply by rivals will be attractive.\(^29\)

Even supposing that fungibility, scale economy, and reputation effect conditions are satisfied, this merely establishes that unilateral long-term trade among rivals can yield economies. A justification for bilateral (exchange) agreements is not reached by these arguments. Indeed, the usual defense for exchanges—that inefficient cross-hauling will occur if every firm is required to supply everywhere to its own needs—conveniently suppresses the obvious alternative, which is not zero trade, but rather unilateral long-term trade. Failure to address these matters directly and demonstrate wherein exchanges enjoy comparative institutional advantages over more standard and familiar forms of unilateral trade presumably explains the suspect or hostile attitude with which exchanges are typically regarded. The argument that emerges from this paper is that bilateral exchanges offer prospective advantages over unilateral trade if the resulting exposure of transaction specific assets effects a credible commitment without simultaneously posing expropriation hazards.

The type of specific asset that is placed at hazard by unilateral long-term trade, but which a reciprocal long-term exchange agreement serves to protect, is that of a dedicated asset. Recall that dedicated assets were described as discrete additions to generalized capacity that would not be put in place but for the prospect of selling a large amount of product to a particular customer. Premature termination of the contract by the buyer would leave the supplier with a large excess of capacity that could be disposed of only at distress prices. Requiring buyers to post a bond would mitigate this hazard, but only by posing another: the supplier may contrive to expropriate the bond. More generally, the interests of the supplier in adapting efficiently to new circumstances are not fully engaged. Reciprocal trading supported by separate but concurrent investments in specific assets provides a mutual safeguard against this second class of hazards. The hostages that are thereby created have the interesting property, moreover, that they are never exchanged. Instead, each party retains possession of its dedicated assets should the contract be prematurely terminated.

The usual argument that exchanges are justified because they avoid costly cross-hauling does not get to these issues and, by itself, is not an adequate justification for widespread use of exchanges. Were it only that transportation cost savings were realized, unilateral trading would suffice. Indeed, petroleum firms should be expected to create a central exchange in which supplies and demands were brought into correspondence by an auctioneer. Firms would end up selling to each other only by accident in

\(^29\)Reputation effect valuations may be illusory or real. Those that are real take the form of customer convenience (billing, contracting) or assured knowledge of product characteristics.
these circumstances. Where dedicated assets are exposed, however, the identity of the parties clearly matters. Trades of this kind will not go through an auction market but will be carefully negotiated between the parties. Reciprocity in these circumstances is thus a device by which the continuity of a specific trading relation is promoted with risk attenuation effects.

VI. Petroleum Exchanges

"The task of linking concepts with observations demands a great deal of detailed knowledge of the realities of economic life" (Tjalling Koopmans, 1957, p. 145). The phenomenon of petroleum exchanges has puzzled economists for a long time. It routinely comes up in antitrust cases and investigations. The 1973 case brought by the United States Federal Trade Commission against the major petroleum firms maintained the view that exchanges were instrumental in maintaining a web of interdependencies among these firms, thereby helping to effect an oligopolistic outcome in an industry that was relatively unconcentrated on normal market structure criteria. The more recent study on The State of Competition in the Canadian Petroleum Industry likewise regards exchanges as objectionable. The Canadian Study, moreover, produces documents—contracts, internal company memoranda, letters, and the like—as well as deposition testimony to support its views that exchanges are devices for extending and perfecting monopoly among the leading petroleum firms. Such evidence on the details and purposes of contracting is usually confidential and hence unavailable. But detailed knowledge is clearly germane—and sometimes essential—to a microanalytic assessment of the transaction cost features of contract.

A. The Evidence from the Canadian Study

Volume V of the Canadian Study deals with the refining sector. Arguments are advanced and supporting evidence is developed that interfirm supply arrangements permit the major refiners to perfect oligopolistic restrictions in the following four respects: 1) valuable knowledge about investment and marketing plans of rivals are disclosed by such agreements (p. 56); 2) leading firms are able to control lesser firms by exercising discretionary power through the terms of exchange (pp. 49–50); 3) competition is impaired by conditioning supply on the payment of an "entry fee" (pp. 53–54); and 4) exchange agreements impose limits on growth and supplementary supply (pp. 51–52).

The first two of these fail to pass scrutiny of the most rudimentary comparative institutional kind. Thus assuming that trade between rivals is efficient and that unilateral supply agreements (if not exchange) will be permitted, the objectionable information disclosures attributed to exchanges would presumably continue—since investment and marketing plans will be unavoidably disclosed in the process. Accordingly, evaluated in comparative institutional terms, the information disclosure objection is properly regarded as an objection to long term trade of any kind. Exchanges are not uniquely culpable.

The suggestion that exchanges are anti-competitive because they permit firms to realize bargaining advantages is similarly misplaced. The correct view is that firms should always be expected to realize such bargaining advantages as their positions lawfully permit. Absent a showing that exchanges are different from unilateral trades in bargaining respects, this objection is properly disregarded also.

31 Robert J. Bertrand, Q. C., Director of Investigation and Research, Combines Investigation Act, coordinated the eight-volume study, The State of Competition in the Canadian Petroleum Industry (Quebec, 1981). All references in this paper are to Vol. V, The Refining Sector. This study will hereinafter be referred to as the Canadian Study.

32 The Canadian Study contends that "a close examination of the interest of the [major refiners] and their actions shows that refining arrangements were meant to restrict competition. The collection of information, the intent to control lesser firms, the imposition of an 'entry fee,' the use of restrictions on downstream growth are not characteristics that would be expected normally from a competitive market" (Vol. V, p. 76).
The entry fee and marketing restraint objections are more substantial, however, and warrant elaboration.

1. Entry Fees

The entry fee objection to exchanges is that this has foreclosure consequences. That such fees are required as a precondition for trade, or at least the sale of product at favorable prices, is set out in the Canadian Study as follows:

Evidence of an understanding that a fee relating to investment was required for acceptance into the industry can be found in the following quotation from Gulf: “We do believe that the oil industry generally, although grudgingly, will allow a participant who has paid his ante, to play the game; the ante in this game being the capital for refining, distributing and selling products” (Document #71248, undated, Gulf). The significance of the quotation lies equally in the notion that an “entry fee” was required and in the notion that the industry set the rules of the “game.” The meaning of the “entry fee” as well as the rules of the “game” as understood by the industry can be found in the actual dealings between companies where the explicit mention of an “entry fee” arises. These cases demonstrate the rules that were being applied—the rules to which Gulf was referring. Companies which had not paid an “entry fee,” that is, companies which had not made a sufficient investment in refining capacity or in marketing distribution facilities would either not be supplied or would be penalized in the terms of the supply agreement.

[pp. 53–54, emphasis added]

2. Marketing Restraints

The Canadian Study notes that exchanges were made conditional on growth and territorial restraints and regards both as objectionable. The Imperial-Shell exchange agreement, under which Imperial supplied product to Shell in the Maritimes and received product in Montreal, is cited in both connections.

The agreement between Imperial and Shell, originally signed in 1963, was renegotiated in 1967. In July 1972, Imperial did this because Shell had been growing too rapidly in the Maritimes. In 1971/72, Imperial had expressed its dissatisfaction with the agreement because of Shell’s marketing policies. Shell noted:

“There [sic] [Imperial’s] present attitude is that we have built a market with their facilities, we are aggressive and threatening them all the time, and they are not going to help and in fact get as tough as possible with us” (Document #23633, updated, Shell). [Vol. 5, p. 51]

Imperial renewed the agreement with Shell only after imposing a price penalty if expansion were to exceed “normal growth rates” and furthermore stipulated that “Shell would not generally be allowed to obtain product from third party sources” to service the Maritimes (p. 52).

Gulf Oil likewise took the position that rivals receiving product under exchange agreements should be restrained to normal growth: “Processing agreements (and exchange agreements) should be entered into only after considering the overall economics of the Corporation and should be geared to providing competitors with volumes required for the normal growth only.” It furthermore sought and secured assurances that product supplied by Gulf would be used only by the recipient and would not be diverted to other regions or made available to other parties (p. 59).

B. Interpretations

These practices are subject to several interpretations. One is that the entry fees and marketing restraints are both anticompetitive. A second is that efficiency purposes are arguably served, especially by the former. A third is that there are mixed effects.

33The Canadian Study (p. 59) identifies the source as Document #73814, January 1972, Gulf.
1. The Inhospitality Tradition

The two polar contracting traditions for evaluating nonstandard or unfamiliar contracting practices are the common law tradition, and the antitrust or inhospitality tradition. Donald Turner makes reference to both of them in expressing his views about vertical market restrictions: “I approach territorial and customer restrictions not hospitably in the common law tradition, but inhospitably in the tradition of antitrust law.”

Thus whereas contractual irregularities are presumed to serve affirmative economic purposes under the common law tradition, a deep suspicion of anticompetitive purposes is maintained by the antitrust (or inhospitality) tradition.

The inhospitality tradition is supported by the widespread view that economic organization is technologically determined. Economies of scale and technological nonseparabilities explain the organization of economic activity within firms. All other activity is appropriately organized by market exchanges. Legitimate market transactions will be mediated entirely by price; restrictive contractual relations signal anticompetitive intent.

The authors of the Canadian Study are evidently persuaded of the merits of this tradition. Long-term trade among rivals of any kind is suspect. And exchanges, which represent an irregular if not unnatural contracting form, are especially objectionable. Not only do exchanges facilitate information disclosure and permit bargaining strength, but they are used punitively against nonintegrated independents who, because they have not paid an entry fee, are denied product on parity terms. Furthermore, the marketing restraints that are associated with exchanges are patently offensive.

2. An Efficiency Assessment

Unlike the inhospitality tradition, the transaction cost approach is in the common law tradition. A comparative institutional orientation (Coase, 1964) is maintained. “Defects” are thus objectionable only where superior feasible alternatives can be described. Inasmuch as the information disclosure and bargaining concerns raised by the authors of the Canadian Study continue under unilateral trading, these are set aside and attention is focused on other matters.

(a) Entry Fees. The entry fee issue is a matter of special interest to this paper. Long-term exchange agreements permit firms to secure product in geographic markets where own-production is not feasible because economies of scale are large in relation to their own needs. The amount of product in question may nevertheless be substantial. Firms with whom exchange agreements are reached will thus construct and maintain larger plants than they otherwise would. Specific investments in dedicated assets are made as a consequence of such agreements.

Were it that supply agreements were of a unilateral kind and the buyer was unable or unwilling to offer a hostage, contracts of type II would presumably be negotiated—whence the trading price would be \( \bar{p} = v_2 + \frac{k}{(1 - \bar{p})} \). If, instead, the contract is extended to include bilateral rather than unilateral trade, the contract is converted to one of type III. Although exchange agreements stipulate the physical flows of product, the effective price is \( p = v_2 + k \), which is less than \( \bar{p} \). Moreover, the parties have the incentive to exchange product so long as realized demand price in both regions exceeds \( v_2 \), which is the marginal cost supply criterion. Assuming that demands in the two regions are highly correlated, the parties will normally reach common decisions on the desirability of trade.

34 The quotation is attributed to Turner by Stanley Robinson, 1968, N.Y. State Bar Association, Antitrust Symposium, p. 29.

35 This assumes common costs, which condition will normally be approximated in exchanges of product between firms within a single country where factor prices are very similar.

36 The possibility that the contract will drift out of alignment nevertheless needs to be recognized. Should one of the firms in an exchange agreement operate much closer to its capacity limits than the other, the latter party would incur much higher costs of termination than would the former. Recognition of this may explain
(b) Marketing Restraints. The supply and growth restraints discussed by the Canadian Study can be looked at in three ways. First, these can be viewed as a means by which to protect the exchange agreement against unilateral defection. Second, such restraints may serve strategic market division purposes. Third, restraints may serve to regularize markets. These are not mutually exclusive.

Only the first of these purposes is consonant with an efficiency interpretation. The argument here is that marketing restraints help to preserve symmetrical incentives. Such symmetry could be upset if one of the firms were to receive product in its deficit region from third parties. Such a firm might then be in a position to play one supplier off against the other. Or symmetry could be placed under strain if one party were to receive product from the other such that it began to grow "in excess of normal"—in which event it might be prepared to construct its own plant and scuttle the exchange agreement. Marketing restraints which help to forestall such outcomes encourage parties to participate in exchanges that might otherwise be unacceptable.

3. A Mixed View

Monopoly explanations are commonly advanced when economists, lawyers, or other interested observers come across contractual practices that they do not understand. Inasmuch as "we are very ignorant [in this field], the number of ununderstandable prac-37

tices tends to be very large, and the reliance on a monopoly explanation frequent" (Coase, 1972, p. 67). A rebuttable presumption that nonstandard contracting practices are serving affirmative economic purposes, rather than monopoly purpose, would arguably serve antitrust law and economics better than the inhospitality presumption which, until recently, has prevailed.

The presumption that exchanges have efficiency purposes could be challenged on any or all of three grounds. First, it might be argued that exchanges are merely a clever device by which to deny product to non-integrated rivals. Refusals to sell to non-integrated firms on F terms would support this contention. (It is plainly unrealistic, however, for buyers that have not made credible commitments to expect to receive product at $p$.) Second, the market in question could be shown to have troublesome structural properties. The issue here is whether the requisite preconditions for market power—mainly high concentration coupled with high barriers to entry—are satisfied. A third would be that the preconditions for efficiency are not satisfied. Factors favorable to the efficiency interpretation are the following: the exchange should be of a long-term kind; the amount of product exchanged should represent a significant fraction of plant capacity; and economies of plant scale should be large in relation to the amount of product traded. Exchanges for a small quantity of product where economies of scale are insubstantial are much more problematic.

To be sure, exchanges might simultaneously serve efficiency and anticompeti-

37To be sure, this is an oversimplification. Antitrust, has been loath to declare contractual constraints to be per se illegal. It came perilously close to taking this step in U.S. v. Arnold Schwinn & Co., 388 U.S. 365 (1967), however. The prevailing enforcement view toward contractual restraints in the 1960's is accurately characterized as inhospitable.

38There is growing agreement that the structural preconditions that must be satisfied before claims of strategic anticompetitive behavior are seriously entertained are very high concentration coupled with barriers to entry (my 1977 article, pp. 292–93; P. L. Joskow and A. K. Klevorick, 1979, pp. 225–31; Janusz Ordover and Robert Willig, 1981, pp. 307–08).
tive purposes. Here as elsewhere, where tradeoffs are posed, they need to be evaluated.

VII. Concluding Remarks

The study of contract in both law and economics has mainly emphasized legal rules and technicalities. Such an orientation is supported by the implicit assumption that the courts "work well." Whether they work well or poorly, however, requires a comparative institutional assessment. There is growing awareness that the (comparative) limitations of the courts are more severe than the legal centralism tradition admits.

The severity of these limitations is not uniform but varies with the circumstances. A discriminating approach to the study of contract will necessarily make provision for this. Specifically, if different transactions have different governance needs, these will be expressly recognized. Accordingly, the study of contract is appropriately extended from legal rules to include an assessment of alternative governance structures, of which the courts are only one. Of special interest in this connection is the use of bilateral governance structures (private ordering) to implement nonstandard contracts where the adaptation and continuity needs of the parties are especially great.

This paper is an effort to deepen the understanding of private ordering. The central points are these:

1) Hostages: Contrary to the prevailing view that hostages are a quaint concept with little or no practical importance to contemporary contracting, the use of hostages to support exchange is widespread and economically important. But hostage creation is only part of the story. Expropriation hazards and prospective maladaptation conditions also need to be considered. Complex governance structures, of which reciprocal trading is one, arise in response to such conditions.

2) Asset Specificity: The organization of economic activity is massively influenced by the degree to which the transactions under examination are supported by assets that are specific to the parties. This paper 1) reaffirms the basic proposition that governance structures need to be matched to the underlying attributes of transactions in a discriminating way if the efficiency purposes of economic organization are to be realized, 2) extends the scope of asset specificity to include dedicated assets, and 3) establishes that, as between two buyers, one of whom posts a hostage in support of specific asset investments by suppliers while the other does not, suppliers will offer better terms to the former, ceteris paribus.

3) Microanalytics: The relevant unit of analysis for studying exchange relations of the kinds discussed in this article is the transaction. Assessing transactions and assigning them to governance structures in a discriminating (mainly transaction cost economizing) way requires much more microanalytic knowledge of economic activity and organization than is customary within economics. Empirical work will necessarily reflect this. Price and quantity of course remain relevant, but the contractual devices by which prices are made to track costs, the manner in which adaptations are effected, and the safeguards that are provided are not only germane but are sometimes decisive.

4) Contracting in its Entirety: Not every transaction poses defection hazards, and it may not be possible to safeguard all that do. Where the potential hazards that beset contracts are evident to the parties from the outset, however, studies of contracts and of contracting institutions arguably start "at the beginning." This has ramifications for assessing the importance of the prisoners' dilemma and for understanding the administration of justice.

(a) Prisoners' Dilemma: The benefits of cooperation notwithstanding, the achievement of cooperation is widely thought to be frustrated by the relentless logic of the prisoners' dilemma. To be sure, it has always been evident that defection can be deterred if payoffs are appropriately altered. But this stratagem is held to be infeasible or is other-

39Examples of microanalytic studies of contract include myself (1976), Thomas Palay (1981), Goldberg and Erickson (1982), and Scott Masten (1982).
wise dismissed—on which account the dilemma persists or appeal is made to “exogenous norms of cooperative behavior [that are] adhered to by the actors” (Albert Hirschman, 1982, p. 1470). I submit that the feasibility of crafting superior \textit{ex ante} incentive structures warrants more attention. A leading reason for its neglect is because the study of the institutions of contract has occupied such a low place on the research agenda. Subtle incentive features that are incorporated in nonstandard contracting practices have gone undetected as a consequence of this nonchalance—on which account the practical significance of the prisoners’ dilemma to the study of exchange has been vastly exaggerated.

(b) Justice: The notion that hostages are demanded as a condition for supplying product on favorable terms has the appearance of an arbitrary exercise of power: the stronger party “demands” a hostage from the weaker, who accedes it because it has no other choice. In fact, a comparative institutional assessment of contractual alternatives discloses that efficiency purposes are often served by hostages and that it is in the mutual interest of the parties to achieve this result. Not only can producers be induced to invest in the most efficient technology, but buyers can be induced to take delivery whenever demand realizations exceed marginal cost. More generally, contracts need to be examined in their entirety, with special attention to their governance features. Principles of justice or competition that look at the relation between the parties at the execution stage without examining the \textit{ex ante} bargaining relation are at best incomplete and are frequently mistaken.\footnote{Robert Nozick’s views on justice are apposite: “whether a distribution is just depends upon how it came about. In contrast, current-time-slice principles of justice hold that the justice of a distribution is determined by how things are distributed (who has what)” (1975, p. 153). What he refers to as the current-time-slice approach to justice neglects \textit{ex ante} bargaining and evaluates justice in terms of outcomes alone. Upon realization that justice is administered in this way, initial bargains will be struck on different terms than they would if the parties were given assurance that the complete contract would be subject to review in evaluating the merits of a contracting relation. Two difficult issues nevertheless remain if the comprehensive bargain orientation to justice is adopted: the initial distribution of resources; and the competence of the parties to evaluate complex contracts. The relative importance of these varies with the circumstances.} Parties to a contract should not expect to have their cake (low price) and eat it too (no hostage).

REFERENCES


Stigler, George J., in President’s Task Force on Productivity and Competition, reprinted in CCH Trade Regulation Reporter, No. 419, June 24, 1969.


