

Understanding Institutional Designs Within Marketing Value Systems

The authors show how institutional arrangements, which consist of contracting, ownership, and social elements, tie together the joint profits, or efficiency, of the marketing system. They use a criterion of remediable efficiency to develop a design framework for institutional arrangements and illustrate the process by designing a recycling system for newspapers. The authors close with an empirical research agenda and extract five institutional design principles for managers.

...without the appropriate institutions, no market economy of any significance is possible.

—Ronald H. Coase, Alfred Nobel Memorial Prize
Lecture in Economic Science, 1991

Coase's 1991 Nobel lecture portrays institutions as vital elements of a modern economy. Early research in marketing focused on institutions as central to the marketing process (e.g., Alderson 1965). Briefly, institutions are the "rules of the game" by which players, both individuals and organizations,¹ interact in exchange ties, be they social or economic. The largely exogenous institutional environment (IE) surrounding an exchange is distinguished from the largely endogenous institutional arrangement (IA) that the parties deploy to support an exchange (Davis and North 1971). The IE consists of the formal and informal rules emanating from the macrolevel aspects of a society, including the polity, the judicial system, cultural norms, and kinship patterns. It is slow to change and defines the world in which firms and people interact. The IA consists of the formal and informal microlevel rules of exchange devised by specific parties to a specific exchange. It represents the arrangements firms and people make to facilitate specific exchanges.

An IA can consist of various formal and informal components and will possess contractual, ownership, and social characteristics. The commission component of a sales force incentive compensation plan would represent a contractual element to the IA between a firm and its sales employees. Property rights and decision-making authority claims, such as those possessed by a partner in a law firm,

would represent the ownership character of IAs. Finally, the most intangible and complex part of the IA is its social components, which are made up of relational and reputational elements. The private and personal networks in Asian business society, known as *guanxi*, are a classic example of a relational IA, whereas Microsoft's less than gentle handling of software firms not fully endorsing the Windows standard is an example of the reputational side of an IA.

As we noted, marketing scholars long have studied the impact of these rules on the exchange process. However, this early institutional school was largely descriptive and, in common with more recent work on institutions, did not lead cumulatively to a productive stream of scholarly work. Instead, during the past two decades, economics, strategy, and marketing scholars have developed productive research programs centered primarily around the firm's resource allocation decisions (the decision-theoretic view) and its interactions with other firms (the game-theoretic view). Recently, at least three significant changes have refocused attention on the role of institutions: the limitations of existing theoretical approaches, the globalization of markets, and the blurring of boundaries between public and private organizations.

There is a growing realization that both the decision- and game-theoretic views provide a limited picture of contemporary management practice in which an emphasis on changing the rules of the game has become more prevalent (D'Aveni 1994). A simple illustration is the shift in thinking about inventory policy in supply chains. Much of the post-war work focused on developing increasingly more sophisticated inventory models in a decision-theoretic tradition. More recent approaches, by emphasizing designing supply chains with (virtually) no inventory through rapid replenishment practices such as "efficient consumer response" and "quick response," turn this idea around completely. These practices do not rest on more advanced inventory modeling but instead involve changing the rules of the game about such things as who controls inventory ordering and releases. Bergen and Iyer (1997) show that such changes do not occur automatically or transparently. Rather, cooperation between self-interested parties must be achieved by redesigning the IA. Institutions are particularly germane to such

¹Institutions are distinct from organizations. Although organizations are one form of institution, more often organizations are players (such as individuals) who play the game according to the rules.

Stephen Carson is a doctoral student, and George John is Research Professor, Marketing Department, Carlson School of Management, University of Minnesota. Timothy Devinney is a professor, and Grahame Dowling is Professor of Marketing, Australian Graduate School of Management, University of New South Wales, Sydney.

modern management practices because their central role is to induce cooperation (North 1990). However, there is little systematic knowledge about how to choose the right IA. Most critically, firms lack a criterion by which designers of institutional structures can compare alternative IAs in terms of firm profitability and marketing system value.

Returning to the preceding list, as businesses continue to generate larger portions of their profits from global and international markets, the range and diversity of the IEs in which they operate continue to increase. The transition experiences of the Eastern European economies and the current economic difficulties of the East Asian "tigers" are forceful reminders of the impact of IEs on management practices. Both the collapse of the socialist system and the difficulties of economic and social transition, as well as the problems with "crony capitalism," are attributed largely to the poor state of IEs (Devinney 1998). Although marketing scholars have argued in general terms about the importance of macrolevel institutions in achieving economic progress (e.g., Hunt 1997), little headway has been made into incorporating the IE into the choice of specific IAs, beyond exhortations that context and cultures matter. Most critically, firms lack analytic typologies that disclose the IE's constraints on IA choices.

Finally, public-private partnerships are blurring the distinctions between for-profit and nonprofit marketing and management. Privatization and market-driven initiatives are being introduced into these highly politicized settings, yet little work has focused on the interactions between the IE and the IA in these settings. Firms find themselves grappling with an expanding group of stakeholders who emphasize socially responsible issues such as human rights, environmentalism, child labor restrictions, and workplace rights. Within this environment, IAs that are economically efficient nevertheless may fail to account for the plethora of claimants to the social rents of the firm. Most critically, firms lack an understanding of institutional efficiency as it pertains to settings in which profits are less important.

Goals of the Article

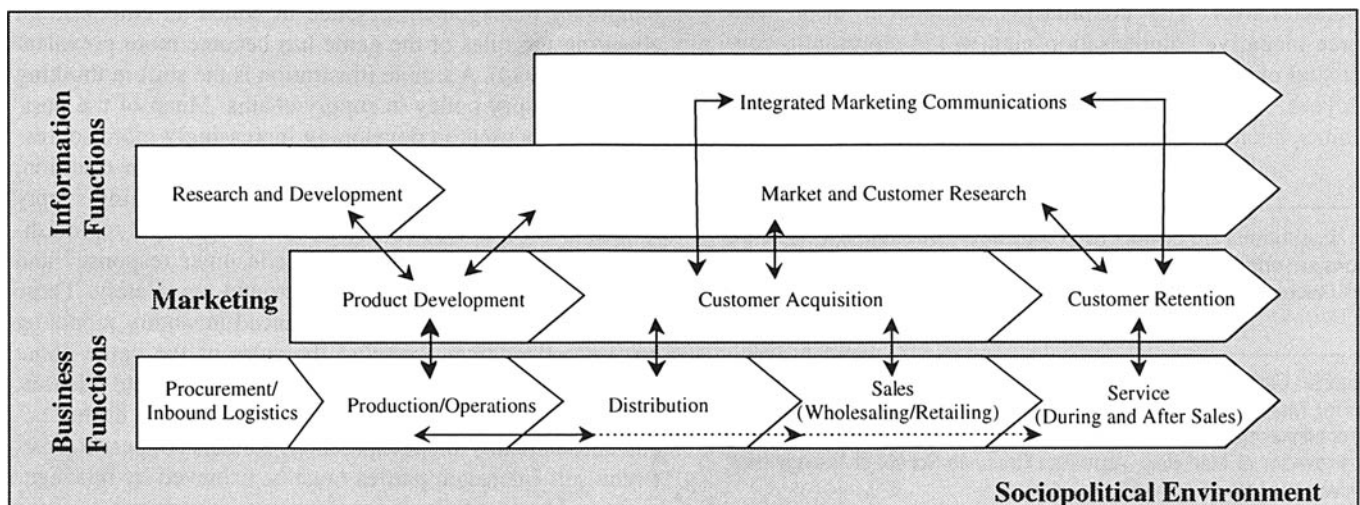
We integrate and synthesize recent developments from organizational analysis, economics, and political science to create a framework to guide institutional design, particularly as it applies to the marketing activities necessary to deliver end-customer value. To accomplish this, we place institutional design within the context of a marketing value system (MVS). In Figure 1, we illustrate one such system that consists of all the activities (and firms) that create and deliver value to the end customer. The bottom panel of the figure (labeled "business functions") is a traditional Porter value system that is linked to the three major dynamic activities of marketing—product development, customer acquisition, and customer retention—which are represented by the middle panel. The top panel describes information flows required in the fulfillment of marketing and business tasks—research and development (R&D), market and customer research, and integrated marketing communications.

The key observation that accompanies the MVS is that certain activities are internal to the "focal" firm, whereas its partners undertake others. Therefore, it is the system that delivers value. However, the cost of each activity typically is borne by a firm individually. Likewise, though value is derived from the total margin produced by the system, the rules of the game in the system determine how this total margin is split into margins for each firm. As we demonstrate subsequently, this translation of total system margin and costs into own-firm profit is crucial to designing institutions that result in the most efficient activities being implemented independently by each firm in the system.

Also apparent in Figure 1 is the large number of interfaces critical to marketing (indicated by arrows²), which has

²There could be more linkages. Figure 1 is meant to be illustrative rather than exhaustive as to the nature of the marketing and information interfaces. The dashed arrows signify that there are vertical interfaces between the business functions that are critical and must be recognized; however, because our focus is on marketing institutions, they are secondary to our discussion.

FIGURE 1
Exchange Interactions in the Marketing Value System



always been considered a boundary-spanning function. Each solid arrow represents an exchange of product, resources, information, and effort that is governed by an IA. An IE that surrounds the entire MVS also affects each interface. The design of these IAs, in relevant IEs, is the heart of our article.

Organization of the Article

The subsequent sections pose and answer a series of questions. In the next section, we develop the conceptual background and explore the role of institutional arrangements in creating efficient MVSs that simultaneously maximize the joint profit created in the system and the own-firm profits of each firm. We observe that our institutional orientation is more comprehensive than either a decision- or a game-theoretic orientation. Next, we suggest a remediable efficiency criterion for comparing rival institutional arrangements on the basis of the activity sets that they allow to be implemented.

We then assemble these conceptual building blocks into an operational framework for institutional design. Our approach directs managers to start simply and increase the complexity of the institutions they devise as the complexity of the exchange and deficiencies in the institutional environment mount. We close with some thoughts about empirical research agendas and guidelines for managers.

Conceptual Background

Efficiency, Joint Profits, and Own-Firm Profits

Fundamentally, the strategic marketing process consists of selecting the most efficient set of activities to perform, given the needs of targeted customers. Efficiency often is misunderstood to refer solely to minimization of out-of-pocket costs, but it defines the level of joint profits in the MVS relative to an upper bound or “fully efficient” state. In this construction, efficiency is enhanced by a reduction in out-of-pocket costs, as well as by an increase in revenues (i.e., effectiveness), which may be viewed as reducing the opportunity costs of forgone trade. Our usage of the term is common in economics, but differs from the “efficiency versus effectiveness” language found in managerial literature.

What complicates the implementation of efficient activities in the strategic marketing process is that what is “best” usually is judged by each firm according to its own profit. Because the activity set chosen by each firm affects the profits of all the firms in the system, each firm has an incentive to improve its own profit by offering to reallocate profits to other firms to compensate or motivate them to adopt activity sets in its interest. This is trivial if the new activity set in question improves all firms’ profits without reallocation. Such Pareto-improving choices will be adopted without exception but for issues of differential perceptions and interpretations of the same data by different firms.

But consider the scenario in which a focal firm wishes to get the other firms to adopt activity sets that increase its own profit but that decrease other firms’ profits under the existing rules of the game. For example, consider a move to a quick response supply system. Here, manufacturers are

asked to respond faster, thus enabling retailers to order closer in time to customer demand. This lowers retailers’ costs by reducing their markdowns and unsold goods, but more responsive, frequent deliveries increase manufacturers’ distribution and billing costs. Should such activity sets be implemented? According to the traditional economic view, the answer is yes as long as the joint profits in the new activity set are greater than in the extant activity set. Economists assume that side payments or other monetary reallocation schemes will emerge automatically so that each and every firm benefits. Thus, Coase (1960) sets forth the straw man (later dubbed the “Coase Theorem”) that the incumbent positions of the parties are immaterial to the final choices, provided that reallocation can be conducted without cost.

But will these activity sets be implemented when reallocation is not automatic and must be managed (as noted by Coase)? The core argument of our thesis is that such efficient activity sets will be implemented only when the IE and IA support the reallocation of profits such that joint profit maximization is aligned with own-firm profit maximization for each firm. To illustrate this in the context of quick response supply systems, Bergen and Iyer (1997) show that if the retailer would agree to a new contractual element in the IA, for example, to commit to order volumes under a take-or-pay contract, both parties can be made better off in a quick response system.

Notice the key elements that emerge as relevant. First, the new set of activities to be implemented by each firm must be identified and described (in this case, the manufacturer must reengineer to deliver faster). Second, the new set of activities must increase joint profits (in this case, this arises from being able to match supply and demand better). Third, each firm actually must be compensated in some manner for implementing these activities (in this case, this is accomplished by the switch to contractually committed purchase volumes by the retailer).

The crucial role of institutions is at the third point, the sharing of profits. Thus, the relevant question in this example is: Are the prospective new contractual elements feasible? That depends on the IE, in this case, the availability of the Uniform Commercial Code (UCC) and an independent, impartial judiciary to give the parties confidence that they can enforce the contractual element of the new IA, that is, the committed purchase volumes. However, in fragile IEs, such as those in the developing countries of Asia, Africa, and Eastern Europe, the ability to implement an efficient contract is seriously compromised by the failure of the IE to create an environment in which the parties are confident about their claims to the rents from the arrangement. Notice that both the IA and IE are responsible for making reallocations feasible.

The Role of Power

In seeming contrast to the reallocation/efficiency view, sociologists and political scientists aver strongly that powerful incumbents will act to protect their positions regardless of the consequences for efficiency. (For example, see McGuire, Granovetter, and Schwartz’s 1993 analysis of the origins of the U.S. electric power industry.) How do we accommodate the power of incumbents into the efficiency view?

The distribution of power in the incumbent IA influences the division of the gains from proposed new activity sets. Powerful partners will be in a position to just compensate other firms for implementing desired activities, whereas systems composed of equally powerful partners will share gains more evenly. For example, a powerful retailer such as Wal-Mart may just compensate a small supplier for adopting efficient inventory control systems, but it surely will retain most of the gains beyond this minimal compensation.

However, such an unequal distribution of gains does not break the tie between efficiency and own-firm profits as long as each firm is not compelled forcibly to implement activities. In other words, each firm's participation constraint must be satisfied, or equivalently, the firm must be no worse off in the long run for choosing to engage in a particular set of activities. This is true of voluntary business ties, even when the power of the parties is highly unbalanced. In the Wal-Mart example, the small supplier would not cooperate if it did not expect to benefit at least marginally in the long run. Thus, joint profit-increasing activities will not reduce *ex ante* expected own-firm profits if implemented in a voluntary exchange regime. That said, there is a major exception to this main case argument.

Powerful incumbents in fragile IEs often can manipulate or change IAs so easily that virtually any promised reallocations can be undone or blocked *ex post*. Two contemporary examples illustrate this issue. With crony capitalism, it is not unusual to find strong firms being required to bail out weaker but politically well-connected firms, as happened recently in Malaysia. In the United States, Brickley, Dark, and Weisbach (1991) show that state-level legislation reduced the market value of franchisers in their states by an average of 6.4% by diminishing their ability to enforce preexisting contractual elements of their IAs that dealt with franchisee termination. With perfect foresight, firms would factor these possibilities into their calculus, but this is true only in the limit.

Efficiency and Sustainable Competitive Advantage

Managing IAs properly leads to greater system efficiency and own-firm profits—essentially a momentary advantage. What about the sustainability of these gains? Sustainable competitive advantage is believed to arise from valuable positional or resource advantages that are difficult to replicate (*i.e.*, imitate). Nonreplicability has been associated with complex configurations of resources, as well as with the ability to dynamically adapt activities and resources more rapidly than rivals (for example, see Teece, Pisano, and Shuen 1997). We consider these in turn.

Replication of advantages emanating from institutional design must involve either (1) replicating the entire efficient MVS (*i.e.*, entering with a rival system) or (2) supplanting the focal firm in the MVS. For the former, developing a rival system with efficient IAs in place is clearly more complex than simply replicating the activities of a single firm, because it involves assembling an entire rival configuration of resources (*i.e.*, partners and IAs). Such a configuration is difficult to pull together, especially when partners are scarce

or efficient relationships are the result of complex IAs developed over time (such as social norms). The replication of such institutions is a complex, time-consuming, and in many ways uncertain task. Thus, efficient systems both offer customers greater value (limiting their desire to do business with firms from a rival system) and erect a barrier due to the difficulty of forming a rival configuration of resources, which must include efficient institutions.

Turning to the replacement of a given firm in an established system, we find that it is also more difficult to supplant a focal firm in its system when complex IAs have been developed to govern exchange between firms. As previously, these IAs must be formed anew on replacement. Furthermore, regardless of the complexity of IAs, it becomes more and more difficult to offer the other firms in a system an incentive to replace a focal firm as the system efficiency increases. To demonstrate this, note that, should a firm not cooperate in creating an efficient marketing system, its partners will have strong incentives to actively seek a replacement firm. However, as the system's efficiency approaches its upper bound (given the IE and characteristics of the transactions), there is less and less to gain by replacing a partner, especially because of the costs and risks when complex institutions are involved. Thus, competition between firms essentially is shifted to competition between systems, because individual firms are difficult to supplant.

With regard to dynamic adaptation, we observe that developing IAs, which allow efficient exchanges between firms in a MVS, also tends to facilitate efficient adaptation. Essentially, IAs permit reallocations that improve efficiency in a given state and enhance the feasibility of reallocations associated with changing activity sets over time.

The Remediable Efficiency Criterion

We formalize the preceding ideas with the criterion of remediable efficiency for institutional design. To understand this concept, consider a firm in an MVS with a given IE and some incumbent IA in place to support an existing activity set. The firm is contemplating a new IA that supports a new activity set that promises to deliver greater customer value (and/or reduced costs). How should it evaluate this transition between IAs and corresponding activity sets? We offer remediable efficiency as the appropriate criterion. Drawing on Wernerfelt's (1994) efficiency criterion, we also accommodate concerns about incumbency by incorporating Williamson's (1996) notion of remediability.

Definition: An IA (and the activity set that it allows) is remediable efficiency if it maximizes the joint profit created in an MVS subject to the IA's feasibility given (1) the IE and characteristics of the proposed activity set and (2) switchover costs associated with transitioning into and out of the IA.

As an operational matter, this amounts to asking the following questions about a new IA and the activity set it allows relative to an incumbent IA and activity set:

1. The joint profit requirement: Does the proposed activity set allowed by the IA increase joint profits (with reallocation, if needed)?

2. The reallocation feasibility requirement: Is the new IA required to support reallocation feasible given (a) the IE and (b) the characteristics of the proposed activity set?
3. The switchover feasibility requirement: Does the new IA remain feasible after accounting for switchover (setup and takedown) costs?

Although stated with reference to a redesign case, these tests also apply to a clean-slate design if we substitute null values for the incumbent activity set and IA. We trace the ramifications of each of the three questions next.

The Joint Profit Requirement

The first requirement is the minimum joint profit standard. Unless joint profits increase, there is no possible reallocation such that all firms in the MVS would approve the change. Furthermore, if the proposed activity set is a Pareto-improving action, there is still no need to consider a new IA. Only when joint profit-enhancing actions require reallocation to align efficiency with own-firm profits does a new IA come into consideration. Assuming that a joint profit-increasing IA has been identified for consideration, we now turn to its feasibility.

The Reallocation Feasibility Requirement

IE constraints. What types of constraints do IEs place on IA design? Recent advances in political science on the role of polities in facilitating efficiency are used to develop Table 1, which depicts the judicial, political, and social elements of IEs. Each affects IA feasibility.

The judicial elements of IEs vary along a de facto versus de jure dimension. De facto judiciaries are subordinated to extant political power, whereas de jure judiciaries possess greater independence. Democracies usually possess more de jure judiciaries, whereas authoritarian regimes tend to feature more de facto judiciaries.

Compared with de facto judiciaries, de jure judiciaries support a greater range of contractual elements in an IA. Devices such as the UCC and an independent judiciary strengthen the predictability of the law and make contract enforcement easier. In turn, this enlarges the scope of contracting possibilities, particularly those complex and fragile types of contracts such as franchising and intellectual property contracts. That said, the links from the IE's judicial elements to the IA are not always so clear-cut. For example, de jure IEs nevertheless may yield poor intellectual property rights because of a political consensus within a nation that, as a net user of intellectual property for example, favors weak intellectual property protection, which benefits domestic parties (as has been true historically of India in the case of pharmaceuticals).

The political elements of IEs vary along a unitary versus federal dimension. Political power in unitary polities is organized hierarchically with a strong central government, whereas federal IEs feature multipolar loci of power that constrain central authority. Parliamentary democracies such as the United Kingdom and communist countries such as the former Soviet Union are examples of unitary polities, whereas the United States and China represent federal polities. The Chinese system deserves comment. Although not widely understood, the Chinese central government, its regional governments, and its military all represent centers of political power with some ability to check one another.

When the polity itself is an exchange partner in an MVS (as often occurs in emerging markets), private parties' weak economic property rights are subject to expropriation by the sovereign. Federal polities are more protective against expropriation in such exchanges. This expands the feasible set of IAs, especially fragile ones. China's economic success relative to Russia is attributed to its multipolar polity. In China, the central and regional governments and the military provide de facto checks and balances on one another to protect private investors in joint ventures, despite the lack of an independent judiciary to enforce contractual elements of IAs (Montignola, Qian, and Weingast 1995). In contrast, the unitary polity in Russia makes foreign joint venture investors more vulnerable to unilateral changes by the polity.

The third dimension in Table 1 places social norms within the IE along a continuum of trust. Trust is defined in various, contentious ways, but the core idea is that greater trust lowers transaction costs because of greater concerns for the trading partner (Hunt 1997) and more predictability in behavior (Bhattacharya, Devinney, and Pillutla 1998; Williamson 1996). Although it is recognized that trust, trustworthiness, and cooperation vary within and across societies, there is considerable debate about the meaning of these concepts (e.g., Fukuyama 1995). However, that said, the clearer the norms of cooperative behavior within an IE, the greater is the scope for beneficial exchanges to go through. From our perspective, greater trust in the IE supports a larger range and diversity of relational and reputational elements of IAs. We discuss a typology of trust-based social rules in the subsequent framework.

Exchange constraints. The characteristics of the specific exchange activities on IA feasibility have been studied extensively in the transaction cost literature (Williamson 1996). The principal insight from this work is that specific investment by the contracting parties and uncertainty about allocations and outcomes make contracting increasingly infeasible as an IA. However, there appear to be several issues pertaining to the feasibility of IAs that are not explored fully in this literature. Consider two such issues: seemingly inefficient IAs and interactions among different elements of an IA.

We often observe IAs that are seemingly inefficient in one or another respect. In one case that has been studied, Heide and John (1988) describe manufacturers' agents, with heavy specific investments in market development activities, engaging in additional expenses to strengthen customer bonds beyond those required to develop the market for the product. The agents do this to deter their principals from ex-

TABLE 1
Dimensions of the Institutional Environment

Dimension	Range of Dimension
Judiciary	De jure → De facto
Polity	Unitary → Federal
Social norms	Opportunistic → Trusting

propriating their market development investments. The primary threat is unilateral termination of representatives in favor of direct sales.

The agents' activities seem quite inefficient on the surface. Sales managers often complain that their salespeople spend time bonding with clients rather than engaging in "productive" activities such as developing new business and communicating customer needs back to their principals. A seemingly more efficient alternative might envision the firm posting an interest-bearing bond that is forfeited if the salesperson is dismissed inappropriately. However, because of the imprecise nature of marketing activity sets, enforcement of such a bond would be problematic. In other words, the seemingly inefficient IA is nevertheless remedially efficient when the characteristics of the exchange are considered. It is the "best we can do."

Turning to interactions among the contractual, ownership, and social elements of IAs, we find that these are often pervasive sources of infeasibility. For example, desired contractual elements often clash with desired ownership elements. To illustrate, ostensibly similar contracts often have disparate effects because of interactions with the ownership elements. Contrast an independent insurance agent with a direct-writing agent. Both typically are paid on an identical, commission-only schedule. However, in ownership respects, the independent agent owns (*de jure*) the client list, whereas the insurance company owns the client list in the direct-writing IA. A direct-writing firm that owns this list cannot commit credibly to the agent that it will not make deleterious changes to the territory. So, for example, if a direct-writing agent grows too large, the company can split the territory. Thus, the "identical" pay plan will have a smaller incentive effect in the direct-writing IA compared with the independent agent IA.

The Switchover Feasibility Requirement

A key aspect of the remediable approach is the contention that transitions between IAs and the characteristics of the incumbent IA matter. Thus, both the setup costs of switching to the new IA and the anticipated future costs of taking down that IA must be included in the remediable efficiency calculus. Feasibility of IAs is not just a matter of external factors, such as the IE and exchange characteristics, but also of extant IAs.

Switchover problems are pervasive in marketing and can be traced to two principal drivers: path dependence and switching-out costs. The simplest impact of path dependence is exhibited when the incumbent IA affects the feasibility of a proposed new IA. For example, a new IA involving common ownership may be feasible or not depending on extant cultural aspects of the incumbent IA.

Broader manifestations of path dependence occur when incumbent IAs result from a series of past choices. Choices in the initial stages of the development of IAs between parties often are constrained, if only by limited information and a failure to consider alternatives. As time passes, specific investments undertaken to increase the efficiency of interaction within a given IA create a lock-in condition, and future IAs are constrained. For example, with the small volume of export sales, firms often cannot contemplate using any IA

other than distributors or export agents in the early stages of international market development. Increasing returns to scale in technology-intensive markets also create such lock-ins. Firms may jump on a product-standard bandwagon that turns out to be technologically inferior *ex post* (e.g., the VHS tape format). As markets mature, other IAs may be more profitable if the firm did not need to take down the incumbent IA.

Thus, North (1990) has argued that (path-dependent) lock-in effects preclude us from assuming that institutions that persist must be efficient. However, we cannot conclude from this extant inefficiency that new IAs (and activity sets) are necessarily better on remediable efficiency grounds. On the contrary, when switchover costs are factored in, the net gains may be negative. Liebowitz and Margolis (1995) show that switching tape formats today is not warranted given existing VHS-formatted tape libraries. In our language, the locked-in incumbent IA is remedially efficient (i.e., the best we can do). High-tech markets are replete with such examples.

Lock-in sometimes is designed purposely into income redistribution programs and nonprofit social service agencies that serve nonpaying clients. The IAs in these situations often are designed in convoluted and seemingly inefficient ways. Straightforward cash payouts are often orders of magnitude cheaper than noncash setups. The classic example is the often criticized U.S. sugar program. Stigler (1992) estimates that it costs \$3 billion annually to raise the intended beneficiaries' incomes by a mere quarter of this amount. These additional costs arise from a convoluted program with multiple agencies and oversight bodies and an emphasis on indirect transfers. Would the often-touted alternative of a direct cash transfer payment to sugar farmers be remedially efficient?

We contend that these IAs are designed purposely to protect weak political rights by creating costs of switching out of it. Although redistribution programs obviously are favored by its beneficiaries, they represent fragile political rights subject to revision in the next election. In this regard, it is far easier to undo a direct cash payment than it is to undo a Byzantine program with multiple constituencies and indirect noncash subsidies. Thus, a cash transfer sugar program is not remedially efficient compared with the incumbent design because of the switching-out costs designed into the current design. Put differently, the proposed alternative does not pass the switchover feasibility test. The seemingly inefficient IA design protects U.S. sugar farmers' weak political rights (to the redistributed money), just as multipolar IEs protect weak economic rights of joint venture investors in China.

Switchover costs also can arise from the power of incumbents. Recall the general argument that incumbency power shapes IAs in profound ways (McGuire, Granovetter, and Schwartz 1993). In line with this expectation, Weiss and Anderson (1992) show that firms with incumbent sales forces that they consider less efficient than a proposed company sales force nevertheless often do not convert to the latter IA. Why not? A principal reason uncovered in their work is the firm's perception of the large switching costs associated with conversion. Among these costs is retaliation from

the terminated distributors, either individually or collectively. Switching sometimes requires buying out entrenched incumbents at a price that renders the new IA unattractive. In our terminology, when these additional costs are factored in, the proposed superior alternative is not remediablely efficient. When evaluating a proposed IA, managers often are able to anticipate the lock-in effects that follow. If these are not desirable (as opposed to the preceding examples), the anticipated costs of switch-out and loss of flexibility lessen the remediable efficiency of the IA.

In summary, switchover costs and incumbent IA characteristics loom large in the remediable efficiency calculus. We demonstrate subsequently that much of the prospective gain from designing IAs carefully and with foresight is the avoidance of lock-in traps, as well as making (or attempting in vain) unwarranted switches out of best-we-can-do, locked-in positions.

A Framework for Designing Institutions

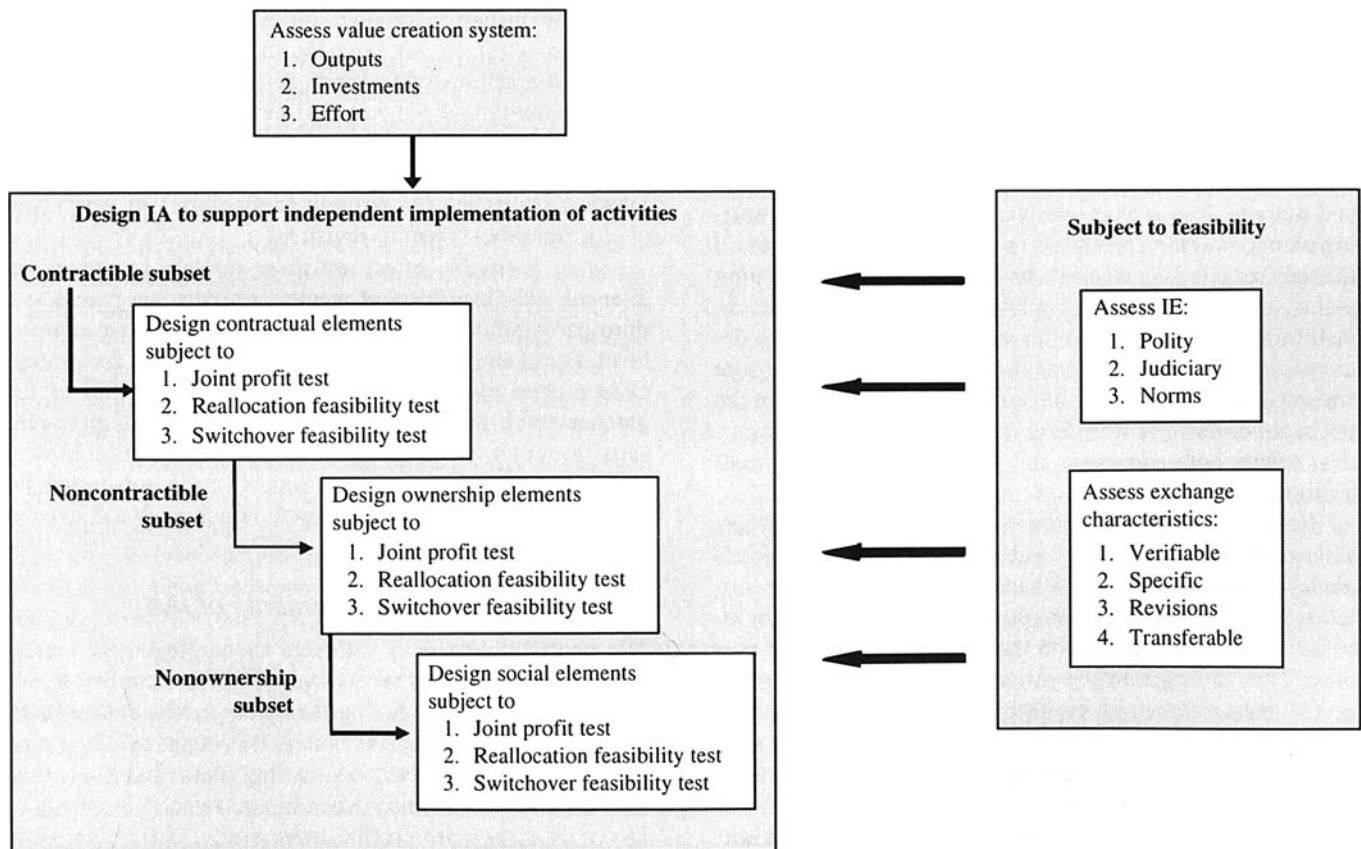
In this section, we address the important practical problem of designing appropriate institutional structures by proposing a framework for managers. The framework begins with a consideration of the desired outputs (i.e., customer benefits) and the activity sets required to bring about these outcomes. Then, we design contractual, ownership, and social

elements of IAs that support these joint profit-maximizing (JPM) activity sets according to our remediable efficiency tests.

We proceed in a staged manner, moving from contractual to ownership to social (relational and reputational) elements of the IA. This sequencing is deliberate. We hold to the idea that institutional design proceeds from the simple to the complex, that complexity should be just enough and no more. When feasible, contracting poses the least complex IA design problem because it uses fine-grained support from the IE judiciary and polity to bind parties to JPM activity sets. Ownership is more complex than contracting because it only relies on coarse support from the IE judiciary and polity and requires IA-specific realignment of ownership with investment to support JPM activity sets. Social elements of IAs are even more complex to develop because they depend minimally on the IE polity and judiciary and almost entirely on norms in the IE and IA to support JPM activity sets. In keeping with our overall approach, simpler social rules are considered before more complex ones.

As a running example in this section, we illustrate the design of a recycling system for newspapers. Such an endeavor involves a mix of profit, nonprofit, and governmental actors, so it affords an opportunity to illustrate the process across nontraditional contexts. In Figure 2, we provide an overview of the formulation.

FIGURE 2
A Design Framework



The Value Creation System

Outputs. The MVS in Figure 1 delivers a set of desired outputs that can be represented by a vector of end-user valued attributes. This is standard practice for marketers, and they possess powerful choice modeling tools to accomplish this by segment or even individual customers. In the extant example of the recycling system, the ease of use, the reliability of pickup, the psychological comfort felt by the "green" customer, and so forth are the vector of valued outputs for the individual household customer. To the city government, relevant outputs of a vendor might include the ability to accommodate growth, adherence to sound environmental practices, and so forth.

The key for institutional design is to identify the degree to which each output is (1) difficult to measure and verify (ex ante or ex post) and (2) foreseeably subject to revision. Both economic and political perspectives matter in this assessment. From an economic perspective, attributes such as pickup times are relatively easy to measure and monitor, but adherence to environmentally friendly methods of disposal and handling of the trash mixed in with the recyclables is harder to ascertain. Similarly, desired capacity is an attribute that is likely to require revision into the future.

From a political perspective, an attribute is foreseeably subject to revision if political consensus is weak regarding rights to that output. Thus, though the beneficiaries of a cash grant may support the transfer payment strongly, other people may not do so and rescind this grant in the future. In our recycling system example, an output such as adherence to sound environmental practices is a weak political right subject to revision because consensus often is lacking on this issue. Controversies over preferences for activities such as burning versus landfill disposal of unsold newsprint highlight weak political rights.

Investments. Investments are expenditures in people, information, equipment, and procedures that are incurred ahead of the actual operational activities that generate valued outputs. In our example, garbage-hauling trucks, sorting plants, training specialists in auction markets (to get rid of the recycled newsprint and such), garbage-burning plants, and landfills are all relevant investments. From an institutional design standpoint, the key is to identify the degree to which each investment is (1) difficult to measure and verify, (2) specific to the intended task or parties to the transaction, and (3) transferable in ownership across parties. Again, both economic and political perspectives matter in this assessment.

Readily verifiable investments include garbage burners and trucks, whereas training and quality processes and R&D are less verifiable investments in our example. More specific assets are defined as those possessing lower value in alternative uses compared with their originally intended purpose. They are open to the sort of opportunism highlighted in the transaction cost economics literature (Williamson 1996). In the current example, garbage trucks are fairly non-specific assets, but a garbage-burning plant located next to a recycling facility is much more specific.

Finally, more readily transferable (or alienable) assets are those for which ownership (de jure or de facto) can be

transferred from one party to another without much difficulty or loss in value, given the IE. For example, the garbage trucks can be transferred readily from one party (e.g., a fleet operator) to the recycling operator. In contrast, tacit assets, such as a process for separating glossy coated paper from newsprint, that were acquired through years of experimentation and learning by doing are much more difficult to transfer at low cost from one party (e.g., the inventor) to the recycling plant operator. Likewise, a plant operator's expertise is an asset whose ownership cannot be transferred readily. We caution that these are not fixed exogenous characteristics. Expert systems and process documentation are means by which such knowledge-based assets (or know-how) are made more alienable (Brynjolfsson 1994).

Political aspects of investments are also important in this assessment. Investments made by the polity magnify specificity and alienability problems for private investments. Normally, state investments are made when either private parties do not have sufficient funds (e.g., welfare clients) or market prices do not suffice to cover costs (e.g., newsprint recycling). The problem is that when the state is a party to the exchange, its sovereign power makes it more difficult to commit credibly not to expropriate its exchange partners' investments by changing the terms of trade unilaterally. Studies of public utilities in the United States (Troesken 1997) and overseas (Levy and Spiller 1994) show compelling evidence of this problem. In our example, a privately financed garbage burner whose client is the municipal government faces ex post hazards of downward revision in the prices paid for its services.

Effort. Effort represents costly activities undertaken by the relevant actors in connection with the investments previously identified to produce the attributes desired by the user segment(s). In our example, the diligence of the truck drivers, the actions of the garbage-plant operators, and the selling activities of the recycled newsprint salespeople are all relevant effortful activities. From an institutional design standpoint, the key is to identify the degree to which these relevant efforts are (1) difficult to measure and verify and (2) foreseeably subject to revision.

More verifiable effort are those matters for which an after-the-fact inspection of results or audits can reveal to a third-party referee that an action was undertaken at some level. For example, the time taken to complete a fixed recycling pickup route is a good measure of the effort of the garbage-truck driver. In contrast, the separation of glossy inserts from newsprint by employees at the collection facility is less verifiable because the precollection separation by households has a large impact on the final result achieved at the recycling station.

Designing Contracting Elements of IAs

The contractible subset refers to those attributes, investments, and efforts that are not subject to verifiability, specificity, ownership transfer, and revision problems (see shaded areas of Table 2). As such, they can be supported through suitably devised contracting elements. Such contractual elements trump ownership and social elements of IAs, in that they are erected more easily, if feasible. Both standard-form contracts and complex, contingent contracts

TABLE 2
Assessing the Value Creation System

Outputs	Investments	Effort
Measurement/verification	Measurement/verification	Measurement/verification
Subject of revision		Subject to revision
	Specificity	
	Transferability	

can be written to support this subset. The primary design challenge is to ensure that the contracts provide the parties with the proper incentives to implement desired activities independently, as discussed. There is a wealth of research on mechanism design to inform us about the design of such contractual elements.

Potential contractual elements are assessed with our three-part test for remediable efficiency. Recall that the first test is a joint profit-increasing requirement for the proposed activity set. This is fairly straightforward, with one twist. In a nonprofit setting such as our recycling example, the joint profit test becomes a joint cost test because the absence of well-functioning markets for recycled newsprint makes it inevitable that the program will lose money (absent subsidies). Assuming that net gains are available, the reallocation and switchover feasibility tests are applied. We cannot discuss the impact of all aspects of IEs and exchanges on IA feasibility exhaustively. Instead, we offer some illustrations of each.

The first part of our reallocation feasibility test assesses IE support for required reallocations. One implication of this test is as follows: Given IE support for property rights of nominal owners, arm's-length bidding is the preferred contractual form to support required investments. This is true even if increasing economies of scale result in a natural monopoly. For example, an investor contemplating buying a fleet of garbage trucks to bid for the business of picking up recyclables would be confident of reselling his or her equipment in the resale market at fair value should his or her bid fail or contract later be terminated. Efficient bidding formats can be quite complex. For example, we know that "second-price" (Vickery³) auctions are more efficient than lowest-price auctions and are coming into use in sophisticated firms and government agencies.

The support from the IE is particularly fragile when the polity itself is a direct participant in the MVS. In these cases, a world of pseudo contracts may arise. In our example, a contract may allow a municipality to reprice its payments to a recycling operation on the basis of the threat of external bidding, even though such bidding may never have taken place. Although such a contract may appear flawed, this pseudo contractual IA still may be remedially efficient given the reasons for the direct involvement of the polity (such as an inadequate market price for recycled newsprint to cover the costs of the operation). The IA allows a recycling op-

eration to be supported yet benchmarks payments to the operator on the basis of a marketlike mechanism.

The second part of the second test assesses feasibility given the exchange characteristics themselves. We consider outputs, efforts, and investments. Verifiable output, such as the time to complete a pickup route, offer obvious opportunities to devise contracting elements. More important, advances in contract design can be used to devise complex, contingent contracts for unverifiable effort, provided the output is verifiable. The core idea is to devise pay-for-performance elements in the IA.

However, when unverifiable effort is paired with unverifiable output, this forces transactions outside the realm of pay-for-performance contracts. For example, the observable output of operators who separate glossy paper from newsprint is not decomposable from the efforts of more diligent presorting by households. Paying plant operators on a performance basis is infeasible.

Investment transferability is another exchange-level constraint on contractual feasibility. For example, the operator of the recycling facility would be nervous about investing time and effort in R&D to improve the sorting process. The ownership of this tacit know-how is relatively difficult to transfer or sell separately from the plant itself, should such a sale become necessary (perhaps due to the contract with the municipality coming up for rebid). However, if the process was separable from the specific site and could be licensed to other operators, the operator would be more responsive to contracts rewarding such investment (e.g., pay-for-performance).

Our third test assesses switchover costs. Proposed contractual elements may not be feasible because entrenched incumbents may capture most, if not all, of the gains. For example, a firm might purchase a service on a bid basis and require multiple bids on each repurchase occasion. Yet, nonincumbent bidders might not bid aggressively because they fear the buyer will factor in switching costs that favor the incumbent vendor. If there is any uncertainty attached to the gains, the net expected gains after overcoming this inertia might not be sufficiently attractive to the nonincumbent. Industry observers suggest that many outsourced information services purchases fit this scenario. In remediable efficiency language, the original bidding system design (an IA) did not take into account the takedown costs on future occasions. When this is factored in, bidding at repurchase no longer improves the remediable efficiency. If adopted in the first place, the bidding system should be replaced with a more efficient institution to protect the buyer's interests.

³The winner in a second-price Vickery auction is the highest bidder, but he or she only pays the amount of the highest losing bid.

Improving measurement. Effective measurement of value received and cost incurred is a sine qua non of contractual exchange, and efforts to improve measurement can expand the contractible subset. Historically, contractual exchange has been fostered by the development of specialized monitors in the form of private and public/legal rules. Commonly agreed on weights and measures and repositories of reputation such as the "Law Merchants" of medieval Europe are examples of institutions designed to improve measurement. Contemporary institutions devoted to measurement and information brokering include the various standards organizations that are so vital to the electronics industry. Although the significance of standards is well recognized, it is less clear how to order the relative capabilities of alternative types of standards bodies. For example, mandated versus voluntary standards organizations often coexist but rarely are studied.

The feasibility of contracting is linked directly to the parties' inability to measure costs and values. Therefore, the development of more complex, contingent contracting elements can be traced to parties' abilities to resolve the measurement problem. Thus, it is likely not a coincidence that the relatively recent movement toward external contracting instead of internal bureaucracy has gone hand-in-hand with the development of information technology, communication systems, and flexible manufacturing technologies. Such process innovations allow for both separation and focus and guarantees and control.

For example, the measurement of advertising agency performance (a neglected topic) is a classic example in which IAs had been deficient before improvements in measurement technologies led to a separation of validation from execution. Historically, advertisers relied on the agency to both develop the advertising and place it in the media. However, the development of sophisticated advertising monitoring systems, the advent of real-time customer tracking programs, and the development of independent media purchasing agencies have remedied the inability to measure agency performance, at least to some extent.

How do we judge measurement improvement? Although more precise measurement appears to be an unqualified good (how could better data hurt?), our remediable efficiency criterion evokes a more stringent test for moving to an IA with more precise measurements. The principal roadblock in incorporating a more precise measurement technology is the switchover feasibility test applied to the new IA.

For example, the widespread use of point-of-sale scanners in grocery stores makes it technically feasible to control inventories and stock-outs better. This is the premise behind efficient consumer response (ECR) in the food industry. However, to realize these gains, it is necessary to shift the responsibility for store-level inventories to the vendor. Some retailers and wholesalers view vendor-managed inventory systems with deep suspicion, and overcoming this problem has proven to be the key hurdle in implementing ECR. From our viewpoint, unless the prospective joint gains can be reallocated through a suitable IA, a move to ECR is not remedially efficient.

Another arena where better measurement does not always represent a remedially efficient improvement is the nonprofit sector. There is a growing tendency to introduce pay-for-performance contracting elements into nonprofit IAs with the aim of making them more efficient, on the grounds that the nonprofit motivations of these agencies are insufficient to motivate cost-efficient behavior. Actually, the choice of a nonprofit IA is indicative of strong noncontractibility problems in the first place, so we might expect tensions between the pay-for-measured-performance aspect of the revised IA and the special goals of nonprofits. One study (Cragg 1997) finds that introducing pay-for-performance for employment training providers has the unintended effect of shifting the providers' focus to easier-to-place clients. Put differently, pay-for-performance is not a remedially efficient remedy for the apparent inefficiency of the provider agency (i.e., because the performance itself is noncontractible). Treating nonprofit agencies as if the noncontractual elements of the IA did not matter is wrong. Improvements rest on identifying remediable efficiency improvements.

In summary, exchanges from the contractible subset can and should be organized using contractual elements. Mechanism designs (of which agency theory is one variant) offer useful, though complex, solutions to this design challenge. Improvements in measurement increase the contractible subset and should be sought, provided they are tested against the feasibility standards for remediable efficiency, not just against the joint profit standard.

Designing Ownership Elements of IAs

The next level of complexity in IA design is used to cope with those outputs, investments, and efforts in the noncontractible subset for which feasible contractual elements cannot be specified satisfactorily. In this section, we study the possibility of (re)organizing ownership as a means to redistribute profits in support of activities that generate greater joint profits.

Ownership reorganization. The incomplete contracting work of Grossman and Hart (1986) defines IAs solely on the basis of the ownership pattern and traces all the remediable efficiency differences between IAs to their ability to offer investment incentives through ownership of the relevant investment or asset. The general argument is that an IA that features an ownership pattern that confers stronger property rights on a particular investing party will evoke higher levels of investment from that party, especially for nonredeployable investments. Although this may appear an innocuous rule, it is more complex when set in our framework, which considers also the influence of the IE (as emphasized by North 1990) and alternative (contractual and social) IAs (as articulated by Williamson 1996).

To illustrate the interaction of these elements, we note that many branded product companies engage in taper integration (partial vertical ownership) in less-developed countries (e.g., China, India), with the express purpose of controlling and avoiding expropriation by self-interest-seeking distribution companies. This is done solely because of the inability of the owner of the brand to enforce distribution

contracts in these IEs. There is little doubt that these companies view ownership as a second-best response to poor contract enforcement in the IE (for example, see Vanhonacker 1997).

Ownership rights, as we consider them, consist of far more than a simple claim to residual income (as is traditionally the case in the agency theory and contracting literature). Ownership of an asset is the (de facto) ability to redirect its use and can be separated from claims to residual income. Residual income (or at least some fraction of it) can be contracted to another person (as would be the case with mortgage-backed securities, for example). To illustrate, we reconsider the independent insurance agent versus the direct-writing insurance representative. Although both are paid in a nominally identical fashion with a share of income generated (a pure commission plan), the former owns the client list. The direct-writing representative does not own the list and thus cannot redirect its use. Instead, this discretion belongs to the firm.

Giving the ownership of the list to the agent secures the agent's investments in developing and servicing his or her client base but evokes less investment by the firm in complementary assets (e.g., building brand image). In the noncontractible subset, the value created by investments and efforts of the parties in the MVS is divided through bargaining. If the firm were to own the list, the representative would be in a decidedly disadvantageous bargaining position. Thus, absent effective ownership of the list (or asset, more generally), the representative would reduce its investment preemptively; the same is true of the firm. In reaching this conclusion, Grossman and Hart (1986) argue that individuals are equally self-interested in alternative institutions, and no superior auditing and/or measurement properties attach to different institutional forms. Thus, firms are no more capable of supervising their employees than their independent representatives. Therefore, it is the ability to direct the use of the asset *ex post* that matters.

Historically, as alluded to previously, contract theorists put all their emphasis on the granting of claims on residual income (a contracting IA) as a solution to the investment problem, rather than on the reallocation of ownership. It should be clear that this is insufficient in the noncontractible subset. For example, a U.S. firm involved in a Chinese joint venture found that though it had residual claims to more than 50% of the operation's profits, it did not have sufficient board control to exercise effective ownership. Its attempt to expand the joint venture's scope by linking with a Shanghai firm was blocked consistently by its North Chinese partner (in this case, the effective owner).

Remediable efficiency and ownership reorganization. We assess proposed ownership patterns with our three-part test for remediable efficiency. The first test is a joint profit-increasing requirement for the proposed activity set. Designers contemplating the joint profit consequences of ownership reorganization can apply the following heuristic:

Joint profits increase when asset ownership accrues to the party whose investments are relatively more important to generating value. Specifically, the value loss were the actual owner

not to own the asset would be greater than the actual loss from the other firm not owning the asset.

Notice that some loss occurs in all cases. Ownership is a blunt instrument, and it dulls the nonowner's incentives (unlike contracting in the contractible subset). Consider applying this rule to ownership of advertisements.

Devinney and Dowling (in press) show that, in the case of the agency-advertiser relationship, both the agency and client would be better off if the agency could claim greater ownership in the relationship—in other words, if the agency bought out the client (and its brands). However, this is infeasible because of the switchover costs. However, transferring copyrights to the intellectual property (i.e., advertisement) from the creator to the client is quite feasible, as witnessed in book publishing.

Although ownership favoring the more "important" investor improves value creation and passes our first test for remediable efficiency improvement (the joint profit criterion), such rearrangements might not always pass the second and third tests (the reallocation and switchover feasibility tests). Regarding reallocation feasibility, it may be impossible to transfer ownership credibly given the nature of the exchange and/or the IE. Regarding switchover feasibility, the path to the new IA may be too costly to one of the partners.

Ownership reallocation feasibility is especially an issue when knowledge or information assets are the subject matter of the exchange. These assets are difficult to transfer even when efficiency considerations dictate it. At one level, transferring ownership to the "correct" entity is difficult because of the "public good" nature of some know-how. For example, Michelin has refused to engage in anything but direct overseas investment and maintains an almost fanatical level of secrecy. Its claim is that production technologies that took decades to develop can be stolen with little effort. In contrast, corporate knowledge can be so complex and embedded that causal ambiguity makes transference impossible without time and effort. For example, Chaparral Steel found its expansion options limited because it could find no way, other than through personal experience and training, to create a new facility (Leonard-Barton 1996).

These are not immutable, exogenous characteristics. First, the nature of the asset may be changed to make it more transportable. Documentation and high bandwidth communications are one example. Another response is a private ordering solution in which the formal IE support from the UCC and such is supplemented by judicial support or forbearance toward nonstandard and seemingly one-sided agreements. The success of business format franchising as an institution demonstrates that a fragile, intangible asset (a brand name) can be effectively "rented out" in return for local managerial effort, despite all the problems inherent in such an exchange. The hallmark of franchise systems is the seemingly one-sided and "unfair" control of franchisees' operations and fragile ownership rights of franchisees. These features are key to making franchising a remediable efficiency-improving alternative to corporate ownership. Because such fragile ownership patterns are unlikely to be supported in poorly endowed IEs, franchising companies expanding into such nations will find more robust owner-

ship patterns to be the IAs of choice under the remediable efficiency criterion.

Ownership reallocation feasibility is especially complex when one party is the state or polity. Consider the alternative ownership arrangements for the noncontractible investments of our newsprint recycling example, that is, investor-owned, regulated, and publicly owned facilities. With investor-owned facilities, the polity asks for bids from private investors to serve the market as an exclusive territory franchise. In other words, the market is owned by the winning bidder for a fixed term.

Such a transfer is not feasible in this case. It is especially difficult for a local government to claim credibly not to expropriate the franchise asset's value, for example, by refusing to allow price increases as energy costs rise (e.g., Troesken 1997). It is more credible for a multipolar polity, such as the state and municipal governments, to make such claims jointly. The Enron utility plant situation in India illustrates this point. There, electric power is a purely central government matter, and the original agreements were voided by a newly elected central government when a more potent provincial government might have prevented it (Rangan 1998).

Second, another ownership alternative relies on the creation of a quasi-judicial regulatory board with some independence from the polity. Whether this is feasible depends on the IE. Federal polities and *de jure* judiciaries are necessary to create credible regulatory regimes, according to the evidence across nations (Levy and Spiller 1994).⁴

Third, public ownership of the immobile noncontractible resources coupled with bids from private operators to operate the facilities on a performance-contracting basis could occur. However, as we noted previously, different institutional forms possessing similar incentive structures will lead to similar outputs. Thus, public or private ownership in this case does little to solve the incentive problem. Matching public ownership of the fixed assets with private ownership of their operation resolves the monopoly-pricing incentive created by the single operator but does little to alleviate the investment incentive problem. This is identical to the problem of the direct-writing insurance firm that wished to use commission-only pay as an incentive. Ownership of the relevant assets blunts the incentive effect of the pay contract.

In summary, ownership favoring the more important investor typically passes the first remediable efficiency test, but the feasibility and switchover tests are more stringent. Applying the feasibility test to our recycling case, the regulated utility is the least complex but still feasible alternative.

Designing the Social Elements of IAs

We have shown that, when contracting becomes infeasible, ownership structure adjustment is the least complex alternative IA to support more efficient proposed activity sets.

⁴We offer the following (bold?) conjecture regarding the current wave of electric utility reorganization away from regulated setups toward more franchise-type bids: The change will be remediable efficiency-improving only in larger jurisdictions with multiple polities (e.g., counties rather than cities, states rather than counties).

However, the desired ownership pattern also may be infeasible in the nonownership subset, as was previously shown. We turn now to the consideration of social elements in these circumstances.

There are bewildering arrays of social rules that can be found in different parts of the world. We focus on efficient social norms that support JPM activity sets. Efficiency of social rules is explained by invoking the threat of being shunned by the network. This provides a powerful incentive for self-policing because the actions of one transgressor is a "public bad" borne by all with whom the transgressor is associated. The desire of the participants to continue to expand trading leads to an incentive to protect reputation.

However, entry into the network will be difficult and can lead to an inefficient overreliance on less qualified insiders. For example, the famed Benetton system of subcontracting is highly flexible within the European and North Italian context from which it sprung but has proven to be a serious constraint on the company's ability to expand effectively outside Europe.

Prominent scholars (e.g., North 1990) argue that social rules are more likely than not to be inefficient because the path dependence of social rules tends to perpetuate incumbency. North points to long-lived social rules of the Suq markets of North Africa to illustrate this. In the contemporary marketing context, we find that advertising self-regulatory bodies serve less to control quality and more to protect agencies from clients and prying outsiders. Thus, social rules are complex to set up as well as take down.

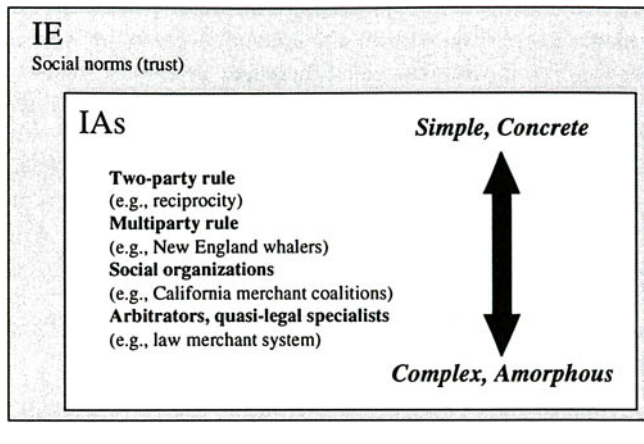
To isolate plausibly efficient social rules, we focus on trust-based norms that can bind parties to JPM activity sets. Trust can operate at the level of a specific exchange (relational) or derive from a broader (reputational) level. Thus, we consider network arrangements (e.g., the theory of clubs; Ellickson 1991), reputation bonding (Devinney and Dowling, in press), personal ties (e.g., the Chinese tradition of *guanxi*), and other trust-based social rules as informal next-best alternatives to contractual and ownership-based economic guarantees.⁵ To incorporate these trust-based social elements into our design framework, we devise the following typology of trust-based social rules (see Figure 3).

A typology of trust-based social rules. Trust-based social rules run the gamut from simple and concrete rules to complex and amorphous norms. At one end, two entities involved in repeated transactions can readily create reciprocity rules to favor jointly valued outcomes, even if such actions are not legally binding. Thus, prisoners' dilemmas are (almost) never dilemmas in repeat interactions. There is considerable controversy about interpreting these behaviors as farsighted self-interest (calculative trust, in Williamson's [1993] language) versus altruism (noncalculative trust), but either interpretation serves our purpose.

If we scale up from dyadic ties, multiple persons in a tight network sometimes devise complex but efficient social rules, for example, those developed by the New England whaling community in the nineteenth century for determin-

⁵We note that some marketing scholars place primacy on these social arrangements over economic elements (e.g., Hunt 1997).

FIGURE 3
A Typology of Trust-Based Social Rules



ing disputed ownership of whales caught on the high seas. Quaint expressions such as the “fast fish” and “iron holds the whale” rules have been analyzed by Ellickson (1991) to show how they supported efficient whale hunting.

The efficient social rules developed by more amorphous communities are even more complex. Clay (1997) describes the merchant coalitions among U.S. traders in Mexican California in the 1830s, after its independence from Spain. The core feature of the coalition is reciprocal dealing among loosely defined members (i.e., taking turns as agents and principals). Enforcement for rule breaking was completely informal and confined to a partial loss of future business.

This example is particularly revealing about the power of efficient social rules because the preexisting Mexican legal system did not allow for enforcement of civil contracts. Thus, only the coalition’s informal rules governed the trade between resident merchants who bought hides (and other products) from ranchers and manufactured goods from ships. Ships’ captains sold on credit, as did ranchers. Due to fluctuating prices and the lack of a legal system, the adaptation needs were severe. Clay (1997) concludes that these rules were efficient by observing the huge increases in trade concomitant with the growth of the coalition. Her analysis also highlights the ambiguous antecedents of the developments of these efficient rules. Path dependence could have developed them along inefficient lines (such as the Mexican system that preceded it).

Even more complex than the reputation-based rules are the efficient quasi-legal social rules devised to support exchange involving multiple parties separated by place, time, and circumstances in communities that are even looser than the geographically constrained California merchant coalitions. The merchant law system of private courts and judges organized by European merchants in the Middle Ages and later is an example of such a form. Unlike the merchant coalition, this system used specialists who served as arbitrators, adjudicators, and enforcers of the rules (North 1990). Present-day arbitrators whose decisions are binding on the parties represent contemporary examples of such private law. They apply principles uniquely crafted to the specific

industry or context and are thus capable of supporting more activity sets.

Remediable efficiency and social elements. To begin, we reiterate that formal contracting rules and/or ownership setups tend to be more powerful institutional forms compared with these informal social forms (on the limits of social rules, see North 1990). The available evidence shows that informal social rules typically are trumped by formal legal systems. For example, the United States’s annexation of California in 1846, which brought an enforceable contract law with it, led to the demise of the merchant coalition previously described. Likewise, the European merchant law systems ultimately evolved into codified law. We are considerably more cautious about the roles and influence of social institutions than recent articles in marketing on trust (e.g., Morgan and Hunt 1994), especially when norms supporting social exchange must be developed purposefully. Our view softens somewhat when the IE is endowed with preexisting norms. This alleviates complexities associated with actually developing norms to support a given exchange, leaving only complexities involved in social enforcement. Note that though the impact of norms and trust have been studied considerably, there has been far less attention given to the important issue of how these are to be developed intentionally by managers.

The general principle about informal social rules under the remediable efficiency criterion is that a more complex rule is warranted only when a simpler form is demonstrably insufficient and the more complex form is implementable. Proposed social rules can be assessed with our familiar three-part test for remediable efficiency. The first test is the joint profit-increasing requirement for the proposed activity set. This test was quite innocuous for contractual and ownership elements, and this is true here as well.

The other test has two parts, feasibility given support from the IE and feasibility given characteristics of the exchange. We consider them in turn. The support from the IE for using an informal social rule rests entirely on the preexisting norms in the IE. Thus, we cannot simply bring the merchant coalition into existence. Instead, we must start with the simplest social rules in our typology (dyadic reciprocity) and scale up as time goes by. Thus, relational dyads eventually coalesce into a network. When formed, however, preexisting networks always can be used to good advantage, as we alluded to previously.

The other part of the feasibility test deals with the characteristics of the exchange. Are there particular characteristics that make proposed social elements ill-suited to support the associated activity sets? Outputs, investments, and effort that are foreseeably subject to substantial revision do not sit well with the highly path-dependent nature of informal social rules. Rapid obsolescence requires taking down existing activity sets, which is not easy to do with social rules. For example, in our recycling case, the preferred technology for environmentally sound disposal is subject to frequent change. Therefore, investments in a garbage burner are not well supported by informal social rules. Research and development is the generic category most affected by this pressure of obsolescence.

Switchover problems are likely to be the principal roadblock with the proposed social elements of IAs. Moving away from one social rule to another may evoke anger and some loss of solidarity. This is compounded by the path dependence of social rules, whose legitimacy is wholly dependent on precedence. An added difficulty is that we know little about the process of designing and implementing a wholesale social institution; thus, incremental changes tend to prevail over wholesale changes in social rules.

Even though social institutions offer greater flexibility than contracting and ownership in many circumstances, observed social rules are not always, or even mostly, efficient. Although trust generally improves support for JPM activity sets, the remediable efficiency tests require us to emphasize social rules as a last resort to more simple, manageable IAs.

Conclusions

The Value of Institutional Thinking for Researchers

Fruitful speculation on the nature of the evolution of marketing institutions in the next century will alter the way marketing academics approach the study of marketing. We focus on motivating empirical work, which is conspicuously rare in this area. Absent empirical follow-up, our framework will be no more successful than previous efforts to take institutions seriously. One advantage is that we have been able to capitalize on several strands of work that have emerged recently.

The remediable efficiency criterion at the heart of our approach emphasizes the need to attend to several issues in designing institutional arrangements. They must be efficient in a joint profit sense, but they also must distribute gains so that each party has incentives to implement its assigned activities. They also must be feasible given the constraints of the broader IE and the characteristics of the exchange. Finally, they must be achievable, in that firms can "get there" from incumbent arrangements. These are detailed requirements that diverge significantly from prior approaches, including those based on efficiency. We offer a few avenues for research that follow from our discussion.

Comparing multiple institutional forms. There has been extensive research in the marketing literature on transaction cost analysis into when firms (ownership) replace contractual arrangements. We have expanded the field by comparing these IAs with social rules. The most important question emanating from our framework is whether contracts trump social rules generally, or are there circumstances in which we might observe social institutions predominating in practice, even though effective contracts could be written? Researchers must be especially vigilant in answering such questions; for example, in comparing contracts and social rules, we are interested only in whether contracts trump social institutions within the contractible subset in which both are feasible. Current research on social institutions tends not to focus on their efficiency relative to other forms.

Switching between IAs. Transitions between IAs are not automatic, and the costs of switching are a function of both the new and incumbent institutions. What are the switching

costs associated with specific pairwise transitions between IAs? What incumbent institutions make switches to contracts, ownership, or social institutions more or less costly? Also interesting are the interactions between institutional elements during transitions. For example, are there certain macrolevel institutions that exacerbate or elevate costs of switching between more microlevel institutions? To illustrate, can common ownership serve as a transition institution to make a common culture easier to establish, rather than be an institution selected simply for its direct role on coordinating exchange? Or, viewed in reverse, do firms with strong common cultures require common ownership, or should they revert back to arm's-length transactions strengthened by both market incentives and their cultural bonds when established?

Interactions between IAs in place. We have provided several examples in which IAs in place interact in unintended ways. Our insurance sales and our pay-for-performance in nonprofit organizations examples both illustrate that IAs can interact in desirable or undesirable ways. Little research has been devoted to such interactions, despite their importance.

Managing social institutions. An important point we make is the difficulty of managing and purposefully developing social ties when these are not endowed on the relationship or IE. Relatively little is known about how complex social institutions are developed and managed. Are strong norms and culture galvanized by unique events in the history of the marketing system or organization, by forceful leadership, or by other institutions (e.g., gainsharing or employee ownership as a form of compensation)? Or, are any of these necessary precursors?

The effects of IAs on market responsiveness. What are the comparative effects of social and economic institutions on adaptation? Do social institutions possess superior capabilities to adapt to small but not large environmental changes? Are firms with strong cultures or that use social coordination at a disadvantage in perceiving and adapting to differences across markets and IEs (as in the Benetton example)? More research is needed.

Changing the value creation system. One important insight from our discussion is that the characteristics of the production system may be partially endogenous (e.g., information assets may be transformed to more alienable forms). To what extent is this a viable option, and when is it easier to modify characteristics of the production system rather than the IA?

Shifting competition to the system level. We have discussed how a network of firms linked by a set of IAs can enjoy certain sustainable advantages due to the difficulty faced by entrants in developing similar IA-linked networks and the flexibility garnered from efficient institutional ties. An interesting question is the degree to which particular IAs are more or less imitable than others, and whether managers strategically can select institutions that, though more difficult to establish, are also more difficult to imitate as a means of establishing a positional advantage in the realm of institutions.

Institutional arrangements with powerful partners. Our efficiency focus emphasizes the need for proposed changes to increase the profits of each affected firm in the MVS. However, power plays an important role in terms of both how the gains in excess of this participation constraint are divided and the credibility of certain IAs when powerful partners can change the rules to their advantage. For example, we illustrate the hazards of contracting with the state, especially under unitary polities. This is especially interesting because the state typically is viewed as having weak self-interests in profits. We also hint that similar issues might arise in any social relationship. Interesting research questions include the degree to which social institutions are more or less susceptible to power influences than contractual and ownership forms are.

Overcoming deficient IEs. To what extent do the IEs of international markets, or different IEs in general, require different IAs and strategies, as we have argued? Can strong culture within the marketing system overcome deficiencies in the IE? Put differently, are firms that struggle with deficient IEs at fault for not adapting to the IE or attempting to establish surrogate social institutions (and other institutions) that are absent in the IE?

What matters in the IE and why. Finally, we believe there is considerable investigation needed into which exact characteristics of the IE matter for institutional design. Issues such as unitary versus federal polity do not appear (to our knowledge) in texts on international marketing but are important nonetheless. Many other features of IEs probably are being overlooked as well.

The Value of Institutional Thinking for Managers

Practicing marketing managers have had a far greater appreciation for the value of institutional structures and their influence than have marketing theorists and academics. This may be attributed partially to the willingness of practitioners to accept ambiguity they cannot explain and the tendency of academics to avoid messy details that fall outside the power of elegant theory. However, the understanding that managers have about institutional structures is arbitrary and superstitious in the sense that it is adaptive and contextual. For example, most managers historically have overemphasized the importance of joint venture and alliance structures for global market expansion and failed to understand the role that more sophisticated ownership and contractual arrangements can play in mitigating risk and ensuring long-run success (e.g., Vanhonacker 1997).

Perhaps the most critical recommendation we can make is to emphasize the importance that managers must place on total system maximization. It is fundamentally important to select strategies with an eye toward maximizing joint value across the MVS. All too often, managers are concerned sole-

ly with the implications of a set of activities on their bailiwick. This is perhaps best illustrated by the “more control or ownership is good” mentality. At one level, this is completely rational. However, at the strategic level, top management must separate, to the greatest extent possible, the issue of system maximization—what we call the JPM—from the distribution of the net rents from the activity set. Placing themselves in the shoes of the firm without control or ownership is a simple means for managers to understand why that firm too may need to share in control or ownership to maximize the size of the pie created in the system.

A pervasive tendency in managerial decision making is to reject options as infeasible too early in the decision-making process. The application of remedial efficiency separates institutional thinking into three separate and ordered component parts: the system’s profitability, the allocation of that profitability, and transition from the old to the new institutional configuration. The framework in Figure 2 expands these basic points into a process model for designing institutions. We end our discussion with some illustrative design principles for managers.

First, contracting trumps ownership, which trumps informal social rules. There is an ordered hierarchy of the internalization of institutional structures that should be considered in sequence. Social rules are less complex, however, when they preexist between two firms or in the IE.

Second, devise pay-for-performance structures for contractible subsets but not for other sets. Incentive contracts fail to the degree that the measurement is weak, and to the extent that measurement is weak, ownership and social rules dominate contracting. Therefore, firms should avoid using monetary incentives to motivate culture and ownership.

Third, ownership dulls incentives, so it should go to more important investors. Ownership presents its own moral hazard problem in situations in which there is more than one owner. Therefore, managerial control should reside with the set of most important investors.

Fourth, devise designed-in inefficiencies to protect weak economic and political property rights. In situations in which weak IEs exist, institutional inefficiency can serve as a protective mechanism. Combined with the preceding point, it can be demonstrated how ownership control can be ceded to a dominant group or owners while other stakeholders rely on alternative IA mechanisms to make their influence felt.

Fifth, use simple informal social rules first and scale up over time. Social rules are the most adaptive of institutional arrangements, in the sense that they fill in cracks left by more formal contractual and ownership structures, but their path-dependent and causally ambiguous natures make them the most difficult to change. Managers must appreciate the full effects of social institutions on efficiency and flexibility.

REFERENCES

- Alderson, Wroe (1965), *Dynamic Marketing Behavior*. Homewood, IL: Richard D. Irwin.
- Bergen, Mark and Ananth V. Iyer (1997), “Quick Response in Manufacturer-Retailer Channels,” *Management Science*, 43 (April), 559–70.
- Bhattacharya, Rajeev, Timothy Devinney, and Madan Pillutla (1998), “A Formal Model of Trust Based on Outcomes,” *Academy of Management Review*, 23 (July), 459–72.
- Brickley, James A., Frederick H. Dark, and Michael S. Weisbach (1991), “The Economic Effects of Franchise Termination

- Laws," *Journal of Law and Economics*, 34, 101–32.
- Brynjolfsson, Erik (1994), "Information Assets, Technology and Organization," *Management Science*, 40 (December), 1645–62.
- Clay, Karen (1997), "Trade Without Law: Private-Order Institutions in Mexican California," *Journal of Law, Economics and Organization*, 13 (April), 201–31.
- Coase, Ronald H. (1960), "The Problem of Social Costs," *Journal of Law and Economics*, 3 (October), 1–44.
- Cragg, Michael (1997), "Performance Incentives in the Public Sector: Evidence from the Job Training Partnership Act," *Journal of Law, Economics and Organization*, 12 (April), 147–68.
- D'Aveni, Richard A. (1994), *Hypercompetition: Managing the Dynamics of Strategic Maneuvering*. New York: The Free Press.
- Davis, Lance and Douglass C. North (1971), *Institutional Change and American Economic Growth*. Cambridge: Cambridge University Press.
- Devinney, Timothy M. (1998), "The Asia Crisis: *Causa Sine Qua Non*," *Policy*, 14 (Summer) 35–37.
- and Grahame Dowling (in press), "Paying the Piper an Incentive to Play a Better Tune: Understanding and Resolving Advertiser–Agency Conflicts," *Journal of Business-to-Business Marketing*, (forthcoming).
- Ellickson, Robert C. (1991), *Order Without Law: How Neighbors Settle Disputes*. Cambridge, MA: Harvard University Press.
- Fukuyama, F. (1995), *Trust: The Social Virtues and the Creation of Prosperity*. New York: The Free Press.
- Grossman, Sanford and Oliver Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94 (August), 691–719.
- Heide, Jan B. and George John (1988), "The Role of Dependence Balancing in Safeguarding Transaction-Specific Assets in Conventional Channels," *Journal of Marketing*, 52 (January), 20–35.
- Hunt, Shelby D. (1997), "Resource-Advantage Theory and the Wealth of Nations: Developing the Socio-Economic Research Tradition," *Journal of Socio-Economics*, 26 (4), 335–57.
- Leonard-Barton, Dorothy (1996), *Well-Springs of Knowledge: Building and Sustaining the Sources of Innovation*. Cambridge, MA: Harvard Business School Press.
- Levy, Brian and Pablo Spiller (1994), "The Institutional Foundations of Regulatory Commitment," *Journal of Law, Economics and Organization*, 9 (Fall), 201–46.
- Liebowitz, Stephen J. and Stephen E. Margolis (1995), "Path Dependence, Lock-In and History," *Journal of Law, Economics and Organization*, 11 (April), 205–26.
- McGuire, Patrick, Mark Granovetter, and Michael Schwartz (1993), "Thomas Edison and the Social Construction of the Early Electricity Industry in America," in *Explorations in Economic Sociology*, Richard Swedberg, ed. New York: Russell Sage Foundation.
- Montignola, Gabriella, Yingyi Qian, and Barry Weingast (1995), "Federalism, Chinese Style: The Political Basis for Economic Success in China," *World Politics*, 48 (October), 50–81.
- Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38.
- North, Douglass C. (1990), *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Rangan, Kasturi (1998), personal communication, (July).
- Stigler, George J. (1992), "Law or Economics?" *Journal of Law and Economics*, 35 (October), 455–68.
- Teece, David J., Gary Pisano, and Amy Shuen (1997), "Dynamic Capabilities and Strategic Management," *Strategic Management Journal*, 18 (7), 509–33.
- Troesken, Werner (1997), "The Sources of Public Ownership: Historical Evidence from the Gas Industry," *Journal of Law, Economics and Organization*, 13 (April), 1–38.
- Vanhonacker, Wilfried (1997), "Entering China: An Unconventional Approach," *Harvard Business Review*, 75 (March/April), 130–40.
- Weiss, Allen M. and Erin Anderson (1992), "Converting from Independent to Employee Sales Forces: The Role of Perceived Switching Costs," *Journal of Marketing Research*, 29 (February), 101–15.
- Wernerfelt, Birger (1994), "An Efficiency Criterion for Marketing Design," *Journal of Marketing Research*, 31 (November), 462–70.
- Williamson, Oliver E. (1993), "Calculativeness, Trust and Economic Organization," *Journal of Law and Economics*, 36 (April), 453–86.
- (1996), *The Mechanisms of Governance*. New York: Oxford University Press.