

ROBERT O. KEOHANE

***From After Hegemony: Cooperation and Discord  
in the World Political Economy*****Realism, Institutionalism,  
and Cooperation**

Impressed with the difficulties of cooperation, observers have often compared world politics to a "state of war." In this conception, international politics is "a competition of units in the kind of state of nature that knows no restraints other than those which the changing necessities of the game and the shallow conveniences of the players impose" (Hoffmann, 1965, p. vii). It is anarchic in the sense that it lacks an authoritative government that can enact and enforce rules of behavior. States must rely on "the means they can generate and the arrangements they can make for themselves" (Waltz, 1979, p. 111). Conflict and war result, since each state is judge in its own cause and can use force to carry out its judgments (Waltz, 1959, p. 159). The discord that prevails is accounted for by fundamental conflicts of interest (Waltz, 1959; Tucker, 1977).

Were this portrayal of world politics correct, any cooperation that occurs would be derivative from overall patterns of conflict. Alliance cooperation would be easy to explain as a result of the operation of a balance of power, but system-wide patterns of cooperation that benefit many countries without being tied to an

alliance system directed against an adversary would not. If international politics were a state of war, institutionalized patterns of cooperation on the basis of shared purposes should not exist except as part of a larger struggle for power. The extensive patterns of international agreement that we observe on issues as diverse as trade, financial relations, health, telecommunications, and environmental protection would be absent.

At the other extreme from these "Realists" are writers who see cooperation as essential in a world of economic interdependence, and who argue that shared economic interests create a demand for international institutions and rules (Mitrany, 1975). Such an approach, which I refer to as "Institutionalist" because of its adherents' emphasis on the functions performed by international institutions, runs the risk of being naive about power and conflict. Too often its proponents incorporate in their theories excessively optimistic assumptions about the role of ideals in world politics, or about the ability of statesmen to learn what the theorist considers the "right lessons." But sophisticated students of institutions and rules have a good deal to teach us. They view institutions not simply as formal organizations with headquarters buildings and specialized staffs, but more broadly as "recognized patterns of practice around which expectations converge" (Young, 1980, p. 337). They regard these patterns of practice as significant because they affect state behavior. Sophisticated institutionalists do not expect cooperation always to prevail, but they are aware of the mal-

From Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, N.J.: Princeton University Press, 1984), Chaps. 1, 6, 7. Some of the author's notes have been omitted.

leability of interests and they argue that interdependence creates interests in cooperation.

During the first twenty years or so after World War II, these views, though very different in their intellectual origins and their broader implications about human society, made similar predictions about the world political economy, and particularly about the subject of this [discussion], the political economy of the advanced market-economy countries. Institutionalists expected successful cooperation in one field to "spill over" into others (Haas, 1958). Realists anticipated a relatively stable international economic order as a result of the dominance of the United States. Neither set of observers was surprised by what happened, although they interpreted events differently.

Institutionalists could interpret the liberal international arrangements for trade and international finance as responses to the need for policy coordination created by the fact of interdependence. These arrangements, which we will call "international regimes," contained rules, norms, principles, and decisionmaking procedures. Realists could reply that these regimes were constructed on the basis of principles espoused by the United States, and that American power was essential for their construction and maintenance. For Realists, in other words, the early postwar regimes rested on the *political hegemony* of the United States. Thus Realists and Institutionalists could both regard early postwar developments as supporting their theories.

After the mid-1960s, however, U.S. dominance in the world political economy was challenged by the economic recovery and increasing unity of Europe and by the rapid economic growth of Japan. Yet economic interdependence continued to grow, and the pace of increased U.S. involvement in the world economy even accelerated after 1970. At this point, therefore, the Institutional and Realist predictions began to diverge. From a strict Institutional standpoint, the increasing need for coordination of policy, created by interdependence, should have led to more cooperation. From a Realist perspec-

tive, by contrast, the diffusion of power should have undermined the ability of anyone to create order.

On the surface, the Realists would seem to have made the better forecast. Since the late 1960s there have been signs of decline in the extent and efficacy of efforts to cooperate in the world political economy. As American power eroded, so did international regimes. The erosion of these regimes after World War II certainly refutes a naive version of the Institutional faith in interdependence as a solvent of conflict and a creator of cooperation. But it does not prove that only the Realist emphasis on power as a creator of order is valid. It might be possible, after the decline of hegemonic regimes, for more symmetrical patterns of cooperation to evolve after a transitional period of discord. Indeed, the persistence of attempts at cooperation during the 1970s suggests that the decline of hegemony does not necessarily sound cooperation's death knell.

International cooperation and discord thus remain puzzling. Under what conditions can independent countries cooperate in the world political economy? In particular, can cooperation take place without hegemony and, if so, how? This [project] is designed to help us find answers to these questions. I begin with Realist insights about the role of power and the effects of hegemony. But my central arguments draw more on the Institutional tradition, arguing that cooperation can under some conditions develop on the basis of complementary interests, and that institutions, broadly defined, affect the patterns of cooperation that emerge.

Hegemonic leadership is unlikely to be revived in this century for the United States or any other country. Hegemonic powers have historically only emerged after world wars; during peacetime, weaker countries have tended to gain on the hegemon rather than vice versa (Gilpin, 1981). It is difficult to believe that world civilization, much less a complex international economy, would survive such a war in the nuclear age. Certainly no prosperous hegemonic power is likely to emerge from such a cataclysm. As long as a

world political economy persists, therefore, its central political dilemma will be how to organize cooperation without hegemony.

\* \* \*

### A Functional Theory of International Regimes

\*\*\* [I]nternational regimes could be created and emphasized their value for overcoming what could be called "political market failure." \* \* \* [Following is a] detailed examination of this argument by exploring why political market failure occurs and how international regimes can help to overcome it. This investigation will help us understand both why states often comply with regime rules and why international regimes can be maintained even after the conditions that facilitated their creation have disappeared. The functional theory developed in this chapter will therefore suggest some reasons to believe that even if U.S. hegemonic leadership may have been a crucial factor in the creation of some contemporary international economic regimes, the continuation of hegemony is not necessarily essential for their continued viability.

#### Political Market Failure and the Coase Theorem

Like imperfect markets, world politics is characterized by institutional deficiencies that inhibit mutually advantageous cooperation. \* \* \* [I]n this self-help system, [there are] conflicts of interest between actors. In economic terms, these conflicts can be regarded as arising in part from the existence of externalities: actors do not bear the full costs, or receive the full benefits, of their own actions.<sup>1</sup> Yet in a famous article Ronald Coase (1960) argued that the presence of externalities alone does not necessarily prevent effective coordination among independent actors.

Under certain conditions, declared Coase, bargaining among these actors could lead to solutions that are Pareto-optimal regardless of the rules of legal liability.

To illustrate the Coase theorem and its counter-intuitive result, suppose that soot emitted by a paint factory is deposited by the wind onto clothing hanging outdoors in the yard of an old-fashioned laundry. Assume that the damage to the laundry is greater than the \$20,000 it would cost the laundry to enclose its yard and install indoor drying equipment; so if no other alternative were available, it would be worthwhile for the laundry to take these actions. Assume also, however, that it would cost the paint factory only \$10,000 to eliminate its emissions of air pollutants. Social welfare would clearly be enhanced by eliminating the pollution rather than by installing indoor drying equipment, but in the absence of either governmental enforcement or bargaining, the egoistic owner of the paint factory would have no incentive to spend anything to achieve this result.

It has frequently been argued that this sort of situation requires centralized governmental authority to provide the public good of clean air. Thus if the laundry had an enforceable legal right to demand compensation, the factory owner would have an incentive to invest \$10,000 in pollution control devices to avoid a \$20,000 court judgment. Coase argued, however, that the pollution would be cleaned up equally efficiently even if the laundry had no such recourse. If the law, or the existence of a decentralized self-help system, gave the factory a right to pollute, the laundry owner could simply pay the factory owner a sum greater than \$10,000, but less than \$20,000, to install anti-soot equipment. Both parties would agree to some such bargain, since both would benefit.

In either case, the externality of pollution would be eliminated. The key difference would not be one of economic efficiency, but of distribution of benefits between the factory and the laundry. In a self-help system, the laundry would have to pay between \$10,000 and \$20,000 and the fac-

factory would reap a profit from its capacity to pollute. But if legal liability rules were based on "the polluter pays principle," the laundry would pay nothing and the factory would have to invest \$10,000 without reaping a financial return. Coase did not dispute that rules of liability could be evaluated on grounds of fairness, but insisted that, given his assumptions, efficient arrangements could be consummated even where the rules of liability favored producers of externalities rather than their victims.

The Coase theorem has frequently been used to show the efficacy of bargaining without central authority, and it has occasionally been applied specifically to international relations (Conybeare, 1980). The principle of sovereignty in effect establishes rules of liability that put the burden of externalities on those who suffer from them. The Coase theorem could be interpreted, therefore, as predicting that problems of collective action could easily be overcome in international politics through bargaining and mutual adjustment—that is, through cooperation \* \* \* The further inference could be drawn that the discord observed must be the result of fundamental conflicts of interest rather than problems of coordination. The Coase theorem, in other words, could be taken as minimizing the importance of [Mancur] Olson's [1965] perverse logic of collective action or of the problems of coordination emphasized by game theory. However, such a conclusion would be incorrect for two compelling sets of reasons.

In the first place, Coase specified three crucial conditions for his conclusion to hold. These were: a legal framework establishing liability for actions, presumably supported by governmental authority; perfect information; and zero transaction costs (including organization costs and the costs of making side-payments). It is absolutely clear that none of these conditions is met in world politics. World government does not exist, making property rights and rules of legal liability fragile; information is extremely costly and often held unequally by different actors; transaction costs, including costs of organization and side-payments, are often very high. Thus an *inversion*

of the Coase theorem would seem more appropriate to our subject. In the absence of the conditions that Coase specified, coordination will often be thwarted by dilemmas of collective action.

Second, recent critiques of Coase's argument reinforce the conclusion that it cannot simply be applied to world politics, and suggest further interesting implications about the functions of international regimes. It has been shown on the basis of game theory that, with more than two participants, the Coase theorem cannot necessarily be demonstrated. Under certain conditions, there will be no stable solution: any coalition that forms will be inferior, for at least one of its members, to another possible coalition. The result is an infinite regress. In game-theoretic terminology, the "core" of the game is empty. When the core is empty, the assumption of zero transaction costs means that agreement is hindered rather than facilitated: "in a world of zero transaction costs, the inherent instability of all coalitions could result in endless recontracting among the firms" (Aivazian and Callen, 1981, p. 179; Veljanovski, 1982).

What do Coase and his critics together suggest about the conditions for international cooperation through bargaining? First, it appears that approximating Coase's first two conditions—that is, having a clear legal framework establishing property rights and low-cost information available in a roughly equal way to all parties—will tend to facilitate cooperative solutions. But the implications of reducing transaction costs are more complex. If transaction costs are too high, no bargains will take place; but if they are too low, under certain conditions an infinite series of unstable coalitions may form.

Inverting the Coase theorem allows us to analyze international institutions largely as responses to problems of property rights, uncertainty, and transaction costs. Without consciously designed institutions, these problems will thwart attempts to cooperate in world politics even when actors' interests are complementary. From the deficiency of the "self-help system" (even from the perspective of purely self-interested national actors) we

derive a need for international regimes. Insofar as they fill this need, international regimes perform the functions of establishing patterns of legal liability, providing relatively symmetrical information, and arranging the costs of bargaining so that specific agreements can more easily be made. Regimes are developed in part because actors in world politics believe that with such arrangements they will be able to make mutually beneficial agreements that would otherwise be difficult or impossible to attain.

This is to say that the architects of regimes anticipate that the regimes will facilitate cooperation. Within the functional argument being constructed here, these expectations explain the formation of the regimes: the *anticipated effects* of the regimes account for the actions of governments that establish them. Governments believe that *ad hoc* attempts to construct particular agreements, without a regime framework, will yield inferior results compared to negotiations within the framework of regimes. Following our inversion of the Coase theorem, we can classify the reasons for this belief under the categories of legal liability (property rights), transaction costs, and problems of uncertainty. We will consider these issues in turn.

**LEGAL LIABILITY** Since governments put a high value on the maintenance of their own autonomy, it is usually impossible to establish international institutions that exercise authority over states. This fact is widely recognized by officials of international organizations and their advocates in national governments as well as by scholars. It would therefore be mistaken to regard international regimes, or the organizations that constitute elements of them, as characteristically unsuccessful attempts to institutionalize centralized authority in world politics. They cannot establish patterns of legal liability that are as solid as those developed within well-ordered societies, and their architects are well aware of this limitation.

Of course, the lack of a hierarchical structure of world politics does not prevent regimes from developing bits and pieces of law (Henkin, 1979, pp. 13–22). But the principal significance of international regimes does not lie in their formal legal status, since any patterns of legal liability and property rights established in world politics are subject to being overturned by the actions of sovereign states. International regimes are more like the “quasi-agreements” that William Fellner (1949) discusses when analyzing the behavior of oligopolistic firms than they are like governments. These quasi-agreements are legally unenforceable but, like contracts, help to organize relationships in mutually beneficial ways (Lowry, 1979, p. 276). Regimes also resemble conventions: practices, regarded as common knowledge in a community, that actors conform to not because they are uniquely best, but because others conform to them as well (Hardin, 1982; Lewis, 1969; Young, 1983). What these arrangements have in common is that they are designed not to implement centralized enforcement of agreements, but rather to establish stable mutual expectations about others’ patterns of behavior and to develop working relationships that will allow the parties to adapt their practices to new situations. Contracts, conventions, and quasi-agreements provide information and generate patterns of transaction costs: costs of renegeing on commitments are increased, and the costs of operating within these frameworks are reduced.

Both these arrangements and international regimes are often weak and fragile. Like contracts and quasi-agreements, international regimes are frequently altered: their rules are changed, bent, or broken to meet the exigencies of the moment. They are rarely enforced automatically, and they are not self-executing. Indeed, they are often matters for negotiation and renegotiation. As [Donald] Puchala has argued, “attempts to enforce EEC regulations open political cleavages up and down the supranational-to-local continuum and spark intense politicking along the cleavage lines” (1975, p. 509).

**TRANSACTION COSTS** Like oligopolistic quasi-agreements, international regimes alter the relative costs of transactions. Certain agreements are forbidden. Under the provisions of the General Agreement on Tariffs and Trade (GATT), for instance, it is not permitted to make discriminatory trade arrangements except under specific conditions. Since there is no centralized government, states can nevertheless implement such actions, but their lack of legitimacy means that such measures are likely to be costly. Under GATT rules, for instance, retaliation against such behavior is justified. By elevating injunctions to the level of principles and rules, furthermore, regimes construct linkages between issues. No longer does a specific discriminatory agreement constitute merely a particular act without general significance; on the contrary, it becomes a "violation of GATT" with serious implications for a large number of other issues. In the terms of Prisoners' Dilemma, the situation has been transformed from a single-play to an iterated game. In market-failure terms, the transaction costs of certain possible bargains have been increased, while the costs of others have been reduced. In either case, the result is the same: incentives to violate regime principles are reduced. International regimes reduce transaction costs of legitimate bargains and increase them for illegitimate ones.

International regimes also affect transaction costs in the more mundane sense of making it cheaper for governments to get together to negotiate agreements. It is more convenient to make agreements within a regime than outside of one. International economic regimes usually incorporate international organizations that provide forums for meetings and secretariats that can act as catalysts for agreement. Insofar as their principles and rules can be applied to a wide variety of particular issues, they are efficient: establishing the rules and principles at the outset makes it unnecessary to renegotiate them each time a specific question arises.

International regimes thus allow governments to take advantage of potential economies of scale. Once a regime has been established, the marginal cost of dealing with each additional issue will be lower than it would be without a regime. \* \* \* [I]f a policy area is sufficiently dense, establishing a regime will be worthwhile. Up to a point there may even be what economists call "increasing returns to scale." In such a situation, each additional issue could be included under the regime at lower cost than the previous one. As [Paul] Samuelson notes, in modern economies, "increasing returns is the prime case of deviations from perfect competition" (1967, p. 117). In world politics, we should expect increasing returns to scale to lead to more extensive international regimes.

In view of the benefits of economies of scale, it is not surprising that specific agreements tend to be "nested" within regimes. For instance, an agreement by the United States, Japan, and the European Community in the Multilateral Trade Negotiations to reduce a particular tariff will be affected by the rules and principles of GATT—that is, by the trade regime. The trade regime, in turn, is nested within a set of other arrangements, including those for monetary relations, energy, foreign investment, aid to developing countries, and other issues, which together constitute a complex and interlinked pattern of relations among the advanced market-economy countries. These, in turn, are related to military-security relations among the major states.<sup>2</sup>

The nesting patterns of international regimes affect transaction costs by making it easier or more difficult to link particular issues and to arrange side-payments, giving someone something on one issue in return for her help on another.<sup>3</sup> Clustering of issues under a regime facilitates side-payments among these issues: more potential *quids* are available for the *quo*. Without international regimes linking clusters of issues to one another, side-payments and linkages would be difficult to arrange in world politics; in the absence of a price system for the exchange of

favors, institutional barriers would hinder the construction of mutually beneficial bargains.

Suppose, for instance, that each issue were handled separately from all others, by a different governmental bureau in each country. Since a side-payment or linkage always means that a government must give up something on one dimension to get something on another, there would always be a bureaucratic loser within each government. Bureaus that would lose from proposed side-payments, on issues that matter to them, would be unlikely to bear the costs of these linkages willingly on the basis of other agencies' claims that the national interest required it.

Of course, each issue is not considered separately by a different governmental department or bureau. On the contrary, issues are grouped together, in functionally organized departments such as Treasury, Commerce, and Energy (in the United States). Furthermore, how governments organize themselves to deal with foreign policy is affected by how issues are organized internationally; issues considered by different regimes are often dealt with by different bureaucracies at home. Linkages and side-payments among issues grouped in the same regime thus become easier, since the necessary internal tradeoffs will tend to take place within rather than across bureaus; but linkages among issues falling into different regimes will remain difficult, or even become more so (since the natural linkages on those issues will be with issues within the same regime).

Insofar as issues are dealt with separately from one another on the international level, it is often hard, in simply bureaucratic terms, to arrange for them to be considered together. There are bound to be difficulties in coordinating policies of different international organizations—GATT, the IMF [International Monetary Fund], and the IEA [International Energy Agency] all have different memberships and different operating styles—in addition to the resistance that will appear to such a move within member governments. Within regimes, by contrast, side-payments are facilitated by the fact that regimes bring together negotiators to consider sets of issues that

may well lie within the negotiators' bureaucratic bailiwicks at home. GATT negotiations, as well as deliberations on the international monetary system, have been characterized by extensive bargaining over side-payments and the politics of issue-linkage (Hutton, 1975). The well-known literature on "spillover" in bargaining, relating to the European Community and other integration schemes, can also be interpreted as concerned with side-payments. According to these writings, expectations that an integration arrangement can be expanded to new issue-areas permit the broadening of potential side-payments, thus facilitating agreement (Haas, 1958).

We conclude that international regimes affect the costs of transactions. The value of a potential agreement to its prospective participants will depend, in part, on how consistent it is with principles of legitimacy embodied in international regimes. Transactions that violate these principles will be costly. Regimes also affect bureaucratic costs of transactions: successful regimes organize issue-areas so that productive linkages (those that facilitate agreements consistent with the principles of the regime) are facilitated, while destructive linkages and bargains that are inconsistent with regime principles are discouraged.

**UNCERTAINTY AND INFORMATION** From the perspective of market-failure theories, the informational functions of regimes are the most important of all. \*\*\* [W]hat Akerlof [1970] called "quality uncertainty" was the crucial problem in [a] "market for lemons" example. Even in games of pure coordination with stable equilibria, this may be a problem. Conventions—commuters meeting under the clock at Grand Central Station, suburban families on a shopping trip "meeting at the car"—become important. But in simple games of coordination, severe information problems are not embedded in the structure of relationships, since actors have incentives to reveal information and their own preferences fully to one another. In these games the problem is to reach some point of agreement; but it may not

matter much which of several possible points is chosen (Schelling, 1960/1978). Conventions are important and ingenuity may be required, but serious systemic impediments to the acquisition and exchange of information are lacking (Lewis, 1969; Young, 1983).

Yet as we have seen in \* \* \* discussions of collective action and Prisoners' Dilemma, many situations—both in game theory and in world politics—are characterized by conflicts of interest as well as common interests. In such situations, actors have to worry about being deceived and double-crossed, just as the buyer of a used car has to guard against purchasing a "lemon." The literature on market failure elaborates on its most fundamental contention—that, in the absence of appropriate institutions, some mutually advantageous bargains will not be made because of uncertainty—by pointing to three particularly important sources of difficulty: *asymmetrical information*; *moral hazard*; and *irresponsibility*.

**ASYMMETRICAL INFORMATION** Some actors may know more about a situation than others. Expecting that the resulting bargains would be unfair, "outsiders" will be reluctant to make agreements with "insiders" (Williamson, 1975, pp. 31-33). This is essentially the problem of "quality uncertainty" as discussed by Akerlof. Recall that this is a problem not merely of insufficient information, but rather of *systematically biased* patterns of information, which are recognized in advance of any agreement both by the holder of more information (the seller of the used car) and by its less well-informed prospective partner (the potential buyer of the "lemon" or "creampuff," as the case may be). Awareness that others have greater knowledge than oneself, and are therefore capable of manipulating a relationship or even engaging successful deception and double-cross, is a barrier to making agreements. When this suspicion is unfounded—that is, the agreement would be mutually beneficial—it is an obstacle to improving welfare through cooperation.

This problem of asymmetrical information only appears when dishonest behavior is possi-

ble. In a society of saints, communication would be open and no one would take advantage of superior information. In our imperfect world, however, asymmetries of information are not rectified simply by communication. Not all communication reduces uncertainty, since communication may lead to asymmetrical or unfair bargaining outcomes as a result of deception. Effective communication is not measured well by the amount of talking that used-car salespersons do to customers or that governmental officials do to one another in negotiating international regimes! The information that is required in entering into an international regime is not merely information about other governments' resources and formal negotiating positions, but also accurate knowledge of their future positions. In part, this is a matter of estimating whether they will keep their commitments. As the "market for lemons" example suggests, and as we will see in more detail below, a government's reputation therefore becomes an important asset in persuading others to enter into agreements with it. International regimes help governments to assess others' reputations by providing standards of behavior against which performance can be measured, by linking these standards to specific issues, and by providing forums, often through international organizations, in which these evaluations can be made.<sup>4</sup> Regimes may also include international organizations whose secretariats act not only as mediators but as providers of unbiased information that is made available, more or less equally to all members. By reducing asymmetries of information through a process of upgrading the general level of available information, international regimes reduce uncertainty. Agreements based on misapprehension and deception may be avoided; mutually beneficial agreements are more likely to be made.

Regimes provide information to members, thereby reducing risks of making agreements. But the information provided by a regime may be insufficiently detailed. A government may require precise information about its prospective partners' internal evaluations of a particular situation,

their intentions, the intensity of their preferences, and their willingness to adhere to an agreement even in adverse future circumstances. Governments also need to know whether other participants will follow the spirit as well as the letter of agreements, whether they will share the burden of adjustment to unexpected adverse change, and whether they are likely to seek to strengthen the regime in the future.

The significance of asymmetrical information and quality uncertainty in theories of market failure therefore calls attention to the importance not only of international regimes but also of variations in the degree of closure of different states' decisionmaking processes. Some governments maintain secrecy much more zealously than others. American officials, for example, often lament that the U.S. government leaks information "like a sieve" and claim that this openness puts the United States at a disadvantage vis-à-vis its rivals.

Surely there are disadvantages in openness. The real or apparent incoherence in policy that often accompanies it may lead the open government's partners to view it as unreliable because its top leaders, whatever their intentions, are incapable of carrying out their agreements. A cacophony of messages may render all of them uninterpretable. But some reflection on the problem of making agreements in world politics suggests that there are advantages for the open government that cannot be duplicated by countries with more tightly closed bureaucracies. Governments that cannot provide detailed and reliable information about their intentions—for instance, because their decisionmaking processes are closed to the outside world and their officials are prevented from developing frank informal relationships with their foreign counterparts—may be unable convincingly to persuade their potential partners of their commitment to the contemplated arrangements. Observers from other countries will be uncertain about the genuineness of officials' enthusiasm or the depth of their support for the cooperative scheme under consideration. These potential partners will there-

fore insist on discounting the value of prospective agreements to take account of their uncertainty. As in the "market for lemons," some potential agreements, which would be beneficial to all parties, will not be made because of "quality uncertainty"—about the quality of the closed government's commitment to the accord.<sup>5</sup>

**MORAL HAZARD** Agreements may alter incentives in such a way as to encourage less cooperative behavior. Insurance companies face this problem of "moral hazard." Property insurance, for instance, may make people less careful with their property and therefore increase the risk of loss (Arrow, 1974). The problem of moral hazard arises quite sharply in international banking. The solvency of a major country's largest banks may be essential to its financial system, or even to the stability of the entire international banking network. As a result, the country's central bank may have to intervene if one of these banks is threatened. The U.S. Federal Reserve, for instance, could hardly stand idly by while the Bank of America or Citibank became unable to meet its liabilities. Yet this responsibility creates a problem of moral hazard, since the largest banks, in effect, have automatic insurance against disastrous consequences of risky but (in the short-run at least) profitable loans. They have incentives to follow risk-seeking rather than risk-averse behavior at the expense of the central bank (Hirsch, 1977).

**IRRESPONSIBILITY** Some actors may be irresponsible, making commitments that they may not be able to carry out. Governments or firms may enter into agreements that they intend to keep, assuming that the environment will continue to be benign; if adversity sets in, they may be unable to keep their commitments. Banks regularly face this problem, leading them to devise standards of creditworthiness. Large governments trying to gain adherents to international agreements may face similar difficulties: countries that are enthusiastic about cooperation are

likely to be those that expect to gain more, proportionately, than they contribute. This is a problem of self-selection, as discussed in the market-failure literature. For instance, if rates are not properly adjusted, people with high risks of heart attack will seek life insurance more avidly than those with longer life expectancies; people who purchased "lemons" will tend to sell them earlier on the used-car market than people with "cream-puffs" (Akerlof, 1970; Arrow, 1974). In international politics, self-selection means that for certain types of activities—such as sharing research and development information—weak states (with much to gain but little to give) may have more incentive to participate than strong ones, but less incentive actually to spend funds on research and developments.<sup>6</sup> Without the strong states, the enterprise as a whole will fail.

From the perspective of the outside observer, irresponsibility is an aspect of the problem of public goods and free-riding; but from the standpoint of the actor trying to determine whether to rely on a potentially irresponsible partner, it is a problem of uncertainty. Either way, informational costs and asymmetries may prevent mutually beneficial agreement.

**REGIMES AND MARKET FAILURE** International regimes help states to deal with all of these problems. As the principles and rules of a regime reduce the range of expected behavior, uncertainty declines, and as information becomes more widely available, the asymmetry of its distribution is likely to lessen. Arrangements within regimes to monitor actors' behavior \* \* \* mitigate problems of moral hazard. Linkages among particular issues within the context of regimes raise the costs of deception and irresponsibility, since the consequences of such behavior are likely to extend beyond the issue on which they are manifested. Close ties among officials involved in managing international regimes increase the ability of governments to make mutually beneficial agreements, because intergovernmental relationships characterized by ongoing com-

munication among working-level officials, informal as well as formal, are inherently more conducive to exchange of information than are traditional relationships between closed bureaucracies. In general, regimes make it more sensible to cooperate by lowering the likelihood of being double-crossed. Whether we view this problem through the lens of game theory or that of market failure, the central conclusion is the same: international regimes can facilitate cooperation by reducing uncertainty. Like international law, broadly defined, their function is "to make human actions conform to predictable patterns so that contemplated actions can go forward with some hope of achieving a rational relationship between means and ends" (Barkun, 1968, p. 154).

Thus international regimes are useful to governments. Far from being threats to governments (in which case it would be hard to understand why they exist at all), they permit governments to attain objectives that would otherwise be unattainable. They do so in part by facilitating intergovernmental agreements. Regimes facilitate agreements by raising the anticipated costs of violating others' property rights, by altering transaction costs through the clustering of issues, and by providing reliable information to members. Regimes are relatively efficient institutions, compared with the alternative of having a myriad of unrelated agreements, since their principles, rules, and institutions create linkages among issues that give actors incentives to reach mutually beneficial agreements. They thrive in situations where states have common as well as conflicting interests on multiple, overlapping issues and where externalities are difficult but not impossible to deal with through bargaining. Where these conditions exist, international regimes can be of value to states.

We have seen that it does not follow from this argument that regimes necessarily increase global welfare. They can be used to pursue particularistic and parochial interests as well as more widely shared objectives. Nor should we conclude that all potentially valuable regimes

will necessarily be instituted. \* \* \* [E]ven regimes that promise substantial overall benefits may be difficult to invent.

\* \* \*

### Bounded Rationality and Redefinitions of Self-Interest

The perfectly rational decisionmaker \* \* \* may face uncertainty as a result of the behavior of others, or the forces of nature, but she is assumed to make her own calculations costlessly. Yet this individual, familiar in textbooks, is not made of human flesh and blood. Even the shrewdest speculator or the most brilliant scientist faces limitations on her capacity for calculation. To imagine that all available information will be used by a decisionmaker is to exaggerate the intelligence of the human species.

Decisionmakers are in practice subject to limitations on their own cognitive abilities, quite apart from the uncertainties inherent in their environments. Herbert Simon has made this point with his usual lucidity (1982, p. 162):

Particularly important is the distinction between those theories that locate all the conditions and constraints in the environment, outside the skin of the rational actor, and those theories that postulate important constraints arising from the limitations of the actor himself as an information processor. Theories that incorporate constraints on the information-processing capacities of the actor may be called *theories of bounded rationality*.

Actors subject to bounded rationality cannot maximize in the classical sense, because they are not capable of using all the information that is potentially available. They cannot compile exhaustive lists of alternative courses of action, ascertaining the value of each alternative and accurately judging the probability of

each possible outcome (Simon, 1955/1979a, p. 10). It is crucial to emphasize that the source of their difficulties in calculation lies not merely in the complexity of the external world, but in their own cognitive limitations. In this respect, behavioral theories of bounded rationality are quite different from recent neoclassical theories, such as the theories of market failure \* \* \*, which retain the assumption of perfect maximization:

[In new neoclassical theories] limits and costs of information are introduced, not as psychological characteristics of the decision maker, but as part of his technological environment. Hence, the new theories do nothing to alleviate the computational complexities facing the decision maker—do not see him coping with them by heroic approximation, simplifying and satisficing, but simply magnify and multiply them. Now he needs to compute not merely the shapes of his supply and demand curves, but in addition, the costs and benefits of computing those shapes to greater accuracy as well. Hence, to some extent, the impression that these new theories deal with the hitherto ignored phenomena of uncertainty and information transmission is illusory. (Simon, 1979b, p. 504)

In Simon's own theory, people "satisfice" rather than maximize. That is, they economize on information by searching only until they find a course of action that falls above a satisfactory level—their "aspiration level." Aspiration levels are adjusted from time to time in response to new information about the environment (Simon, 1972, p. 168). In view of people's knowledge of their own cognitive limitations, this is often a sensible strategy; it is by no means irrational and may well be the best way to make most decisions.

In ordinary life, we satisfice all the time. We economize on information by developing habits, by devising operating rules to simplify calculation in situations that repeat themselves, and by adopting general principles that we expect, in the long run, to yield satisfactory results. I do not normally calculate whether to brush my teeth in the morning, whether to hit a tennis ball

directed at me with my backhand or my forehand; or whether to tell the truth when asked on the telephone whether Robert Keohane is home. On the contrary, even apart from any moral scruples I might have (for instance, about lying), I assume that my interests will be furthered better by habitually brushing my teeth, applying the rule "when in doubt, hit it with your forehand because you have a lousy backhand," and adopting the general principle of telling the truth than by calculating the costs and benefits of every alternative in each case. I do not mean to deny that I might occasionally be advantaged by pursuing a new idea at my desk rather than brushing my teeth, hitting a particular shot with my backhand, or lying to an obnoxious salesman on the telephone. If I could costlessly compute the value of each alternative, it might indeed be preferable to make the necessary calculations each time I faced a choice. But since this is not feasible, given the costs of processing information, it is in my long-run interest to eschew calculation in these situations.

Simon's analysis of bounded rationality bears some resemblance to the argument made for rule-utilitarianism in philosophy, which emphasizes the value of rules in contributing to the general happiness.<sup>7</sup> Rule-utilitarianism was defined by John Austin in a dictum: "Our rules would be fashioned on utility; our conduct, on our rules" (Mackie, 1977, p. 136). The rule-utilitarian adopts these rules, or "secondary principles," in John Stuart Mill's terms, in the belief that they will lead, in general, to better results than a series of *ad hoc* decisions based each time on first principles.<sup>8</sup> A major reason for formulating and following such rules is the limited calculating ability of human beings. In explicating his doctrine of utilitarianism, Mill therefore anticipated much of Simon's argument about bounded rationality (1861/1951, p. 30):

Nobody argues that the art of navigation is not founded on astronomy, because sailors cannot wait to calculate the Nautical Almanack. Being rational creatures, they go to sea with it ready calculated;

and all rational creatures go out upon the sea of life with their minds made up on the common questions of right and wrong, as well as on many of the far more difficult questions of wise and foolish. And this, as long as foresight is a human quality, it is to be presumed they will continue to do.

If individuals typically satisfice rather than maximize, all the more so do governments and other large organizations (Allison, 1971; Steinbruner, 1974; Snyder and Diesing, 1977). Organizational decision-making processes hardly meet the requirements of classical rationality. Organizations have multiple goals, defined in terms of aspiration levels; they search until satisfactory courses of action are found; they resort to feedback rather than systematically forecasting future conditions; and they use "standard operating procedures and rules of thumb" to make and implement decisions (Cyert and March, 1963, p. 113; March and Simon, 1958).

The behavioral theory of the firm has made it clear that satisficing does not constitute aberrant behavior that should be rectified where possible; on the contrary, it is intelligent. The leader of a large organization who demanded that the organization meet the criteria of classical rationality would herself be foolish, perhaps irrationally so. An organization whose leaders behaved in this way would become paralyzed unless their subordinates found ways to fool them into believing that impossible standards were being met. This assertion holds even more for governments than for business firms, since governments' constituencies are more varied, their goals more diverse (and frequently contradictory), and success or failure more difficult to measure. Assumptions of unbounded rationality, however dear they may be to the hearts of classical Realist theorists (Morgenthau, 1948/1966) and writers on foreign policy, are idealizations. A large, complex government would tie itself in knots by "keeping its options open," since middle-level bureaucrats would not know how to behave and the top policymakers would be overwhelmed by minor problems. The search for complete flexibility is as

quixotic as looking for the Holy Grail or the fountain of youth.

If governments are viewed as constrained by bounded rationality, what are the implications for the functional argument \*\*\* about the value of international regimes? \*\*\* [U]nder rational-choice assumptions, international regimes are valuable to governments because they reduce transaction costs and particularly because they reduce uncertainty in the external environment. Each government is better able, with regimes in place, to predict that its counterparts will follow predictably cooperative policies. According to this theory, governments sacrifice the ability to maximize their myopic self-interest by making calculations on each issue as it arises, in return for acquiring greater certainty about others' behavior.

Under bounded rationality, the inclination of governments to join or support international regimes will be reinforced by the fact that the alternatives to regimes are less attractive than they would be if the assumptions of classical rationality were valid. Actors laboring under bounded rationality cannot calculate the costs and benefits of each alternative course of action on each issue. On the contrary, they need to simplify their own decisionmaking processes in order to function effectively at all. The rules of thumb they devise will not yield better, and will generally yield worse, results (apart from decisionmaking costs) than classically rational action—whether these rules of thumb are adopted unilaterally or as part of an international regime. Thus a comparison between the value of a unilateral rule of thumb and that of a regime rule will normally be more favorable to the regime rule than a comparison between the value of costless, perfectly rational calculation and the regime rule.

When we abandon the assumption of classical rationality, we see that it is not international regimes that deny governments the ability to make classically rational calculations. The obstacle is rather the nature of governments as large, complex organizations composed of human beings with limited problem-solving capabilities. The

choice that governments actually face with respect to international regimes is not whether to adhere to regimes at the expense of maximizing utility through continuous calculation, but rather on what rules of thumb to rely. Normally, unilateral rules will fit the individual country's situation better than rules devised multilaterally. Regime rules, however, have the advantage of constraining the actions of others. The question is whether the value of the constraints imposed on others justifies the costs of accepting regime rules in place of the rules of thumb that the country would have adopted on its own.

Thus if we accept that governments must adopt rules of thumb, the costs of adhering to international regimes appear less severe than they would be if classical rationality were a realistic possibility. Regimes merely substitute multilateral rules (presumably somewhat less congenial *per se*) for unilateral ones, with the advantage that other actors' behavior thereby becomes more predictably cooperative. International regimes neither enforce hierarchical rules on governments nor substitute their own rules for autonomous calculation; instead, they provide rules of thumb in place of those that governments would otherwise adopt.

\*\*\* [W]e can see how different our conception of international regimes is from the self-help system that is often taken as revealing the essence of international politics. In a pure self-help system, each actor calculates its interests on each particular issue, preserving its options until that decision has been made. The rational response to another actor's distress in such a system is to take advantage of it by driving a hard bargain, demanding as much as "the traffic will bear" in return for one's money, one's oil, or one's military support. Many such bargains are in fact struck in world politics, especially among adversaries; but one of the key features of international regimes is that they limit the ability of countries in a particularly strong bargaining position (however transitory) to take advantage of that situation. This limitation, as we have stressed, is not the result of altruism but of the fact that

joining a regime changes calculations of long-run self-interest. To a government that values its ability to make future agreements, reputation is a crucial resource; and the most important aspect of an actor's reputation in world politics is the belief of others that it will keep its future commitments even when a particular situation, myopically viewed, makes it appear disadvantageous to do so. Thus even classically rational governments will sometimes join regimes and comply with their rules. To a government seeking to economize on decisionmaking costs, the regime is also valuable for providing rules of thumb; discarding it would require establishing a new set of rules to guide one's bureaucracy. The convenience of rules of thumb combines with the superiority of long-run calculations of self-interest over myopic ones to reinforce adherence to rules by egoistic governments, particularly when they labor under the constraints of bounded rationality.

\* \* \*

#### NOTES

1. For an elaborated version of this definition, see Davis and North (1971, p. 16).
2. For the idea of "nesting," I am indebted to Aggarwal (1981). Snidal (1981) also relies on this concept, which was used in a similar context some years ago by Barkun (1968, p. 17).
3. On linkage, see especially the work of Kenneth A. Oye (1979, 1983). See also Stein (1980) and Tollison and Willett (1979).
4. This point was suggested to me by reading Elizabeth Colson's account of how stateless societies reach consensus on the character of individuals: through discussions and gossip that allow people to "apply the standards of performance in particular roles in making an overall judgement about the total person; this in turn allows them to predict future behavior" (1974, p. 53).
5. In 1960 Thomas Schelling made a similar argument about the problem of surprise attack. Asking how we would prove that we were not planning a surprise attack if the Russians suspected we were, he observed that "evidently it is not going to be enough just to tell the truth. . . . There has to be some way of authenticating certain facts, the facts presumably involving the disposition of forces" (p. 247). To authenticate facts requires becoming more open to external monitoring as a way of alleviating what Akerlof [1970] later called "quality uncertainty."
6. Bobrow and Kudrle found evidence of severe problems of collective goods in the IEA's energy research and development program, suggesting that "commercial interests and other national rivalries appear to have blocked extensive international cooperation" (1979, p. 170).
7. In philosophy, utilitarianism refers to an ethical theory that purports to provide generalizable principles for moral human action. Since my argument here is a positive one, seeking to explain the behavior of egoistic actors rather than to develop or criticize an ethical theory, its relationship to rule-utilitarianism in philosophy, as my colleague Susan Okin has pointed out to me, is only tangential.
8. John Mackie argues that even act-utilitarians "regularly admit the use of rules of thumb," and that whether one follows rules therefore does not distinguish act- from rule-utilitarianism (1977, p. 137). Conversely, Joseph Nye has pointed out to me that even rule-utilitarians must depart at some point from their rules for consequentialist reasons. The point here is not to draw a hard-and-fast dichotomy between the two forms of utilitarianism, but rather to point out the similarities between Mill's notion of relying on rules and Simon's conception of bounded rationality. If all utilitarians have to resort to rules of thumb to some extent, this only strengthens the point I am making about the

importance of rules in affecting, but not determining, the behavior of governments. For a succinct discussion of utilitarianism in philosophy, see Urmson (1968).

#### BIBLIOGRAPHY

\* \* \* Where a date is given for an original as well as a later edition, the latter was used; page references in the text refer to it.

- Aggarwal, Vinod, 1981. *Hanging by a Thread: International Regime Change in the Textile/Apparel System, 1950-1979* (Ph.D. dissertation. Stanford University).
- Aivazian, Varouj A., and Jeffrey L. Callen, 1981. The Coase theorem and the empty core. *Journal of Law and Economics*, vol. 24, no. 1 (April), pp. 175-81.
- Akerlof, George A., 1970. The market for "lemons." *Quarterly Journal of Economics*, vol. 84, no. 3 (August), pp. 488-500.
- Allison, Graham, 1971. *Essence of Decision: Explaining the Cuban Missile Crisis* (Boston: Little, Brown).
- Arrow, Kenneth J., 1974. *Essays in the Theory of Risk-Bearing* (New York: North-Holland/American Elsevier).
- Barkun, Michael, 1968. *Law without Sanctions: Order in Primitive Societies and the World Community* (New Haven: Yale University Press).
- Bobrow, Davis W., and Robert Kudrle, 1979. Energy R & D: in tepid pursuit of collective goods. *International Organization*, vol. 33, no. 2 (Spring), pp. 149-76.
- Coase, Ronald, 1960. The problem of social cost. *Journal of Law and Economics*, vol. 3, pp. 1-44.
- Colson, Elizabeth, 1974. *Tradition and Contract: The Problem of Order* (Chicago: Aldine Publishing Company).
- Conybeare, John A.C., 1980. International organization and the theory of property rights. *International Organization*, vol. 34, no. 3 (Summer), pp. 307-34.
- Cyert, Richard, and James G. March, 1963. *The Behavioral Theory of the Firm* (Englewood Cliffs, N.J.: Prentice-Hall).
- Davis, Lance, and Douglass C. North, 1971. *Institutional Change and American Economic Growth* (Cambridge: Cambridge University Press).
- Fellner, William, 1949. *Competition among the Few* (New York: Knopf).
- Gilpin, Robert, 1981. *War and Change in World Politics* (Cambridge: Cambridge University Press).
- Haas, Ernst B., 1958. *The Uniting of Europe* (Stanford: Stanford University Press).
- Hardin, Russell, 1982. *Collective Action* (Baltimore: The Johns Hopkins University Press for Resources for the Future).
- Henkin, Louis, 1979. *How Nations Behave: Law and Foreign Policy*, 2nd edition (New York: Columbia University Press for the Council on Foreign Relations).
- Hirsch, Fred, 1977. The Bagehot problem. *The Manchester School*, vol. 45, no. 3 (September), pp. 241-57.
- Hoffmann, Stanley, 1965. *The State of War: Essays on the Theory and Practice of International Politics* (New York: Praeger).
- Hutton, Nicholas, 1975. The salience of linkage in international economic negotiations. *Journal of Common Market Studies*, vol. 13, nos. 1-2, pp. 136-60.
- Krasner, Stephen D., ed., 1983. *International Regimes* (Ithaca: Cornell University Press).
- Lewis, David K., 1969. *Convention: A Philosophical Study* (Cambridge: Harvard University Press).
- Lowry, S. Todd. 1979. Bargain and contract theory in law and economics. In Samuels, 1979, pp. 261-82.
- Mackie, J. L., 1977. *Ethics: Inventing Right and Wrong* (Harmondsworth, England: Penguin Books).
- March, James G., and Herbert Simon, 1958. *Organizations* (New York: John Wiley & Sons).
- Mill, John Stuart, 1861/1951. *Utilitarianism* (New York: E. P. Dutton).

Mitrany,  
Politics  
London  
Science  
Morgenth  
Nation  
Olson, M  
Action  
Oye, Ker  
Oye et  
Oye, Ker  
ing an  
Econo  
vard U  
Oye, Ker  
J. Liek  
eign I  
Longr  
Puchala  
region  
muni  
pp. 49  
Samuel  
tem o  
tion E  
Samuel  
comp  
Mon  
John  
Schelli  
Confli  
Schelli  
Mach  
Simon  
ratio  
ics,  
Repr  
Simon  
ratio  
161-  
Simon  
(New  
Simon  
mak  
Eco

- Mitrany, David, 1975. *The Functional Theory of Politics* (London: St. Martin's Press for the London School of Economics and Political Science).
- Morgenthau, Hans J., 1948/1966. *Politics among Nations*, 4th edition (New York: Knopf).
- Olson, Mancur, 1965. *The Logic of Collective Action* (Cambridge: Harvard University Press).
- Oye, Kenneth A., 1979. The domain of choice. In Oye et al., 1979, pp. 3-33.
- Oye, Kenneth A., 1983. *Belief Systems, Bargaining and Breakdown: International Political Economy 1929-1934* (Ph.D. dissertation, Harvard University).
- Oye, Kenneth A., Donald Rothchild, and Robert J. Lieber, eds., 1979. *Eagle Entangled: U.S. Foreign Policy in a Complex World* (New York: Longman).
- Puchala, Donald J., 1975. Domestic politics and regional harmonization in the European Communities. *World Politics*, vol. 27, no. 4 (July), pp. 496-520.
- Samuels, Warren J., 1979. *The Economy as a System of Power* (New Brunswick, N.J.: Transaction Books).
- Samuelson, Paul A., 1967. The monopolistic competition revolution. In R. E. Kuenne, ed., *Monopolistic Competition Theory* (New York: John Wiley & Sons).
- Schelling, Thomas C., 1960/1980. *The Strategy of Conflict* (Cambridge: Harvard University Press).
- Schelling, Thomas C., 1978. *Micromotives and Macrobehavior* (New York: W. W. Norton).
- Simon, Herbert A., 1955. A behavioral model of rational choice. *Quarterly Journal of Economics*, vol. 69, no. 1 (February), pp. 99-118. Reprinted in Simon, 1979a, pp. 7-19.
- Simon, Herbert A., 1972. Theories of bounded rationality. In Radner and Radner, 1972, pp. 161-76. Reprinted in Simon, 1982, pp. 408-23.
- Simon, Herbert A., 1979a. *Models of Thought* (New Haven: Yale University Press).
- Simon, Herbert A., 1979b. Rational decision making in business organizations. *American Economic Review*, vol. 69, no. 4 (September), pp. 493-513. Reprinted in Simon, 1982, pp. 474-94.
- Simon, Herbert A., 1982. *Models of Bounded Rationality*, 2 vols. (Cambridge: MIT Press).
- Snidal, Duncan, 1981. *Interdependence, Regimes and International Cooperation* (unpublished manuscript).
- Snyder, Glenn H., and Paul Diesing, 1977. *Conflict among Nations: Bargaining, Decision making, and System Structure in International Crises* (Princeton: Princeton University Press).
- Stein, Arthur A., 1980. The politics of linkage. *World Politics*, vol. 33, no. 1 (October), pp. 62-81.
- Steinbruner, John D., 1974. *The Cybernetic Theory of Decision: New Dimensions of Political Analysis* (Princeton: Princeton University Press).
- Tollison, Robert D., and Thomas D. Willett, 1979. An economic theory of mutually advantageous issue linkages in international negotiations. *International Organization*, vol. 33, no. 4 (Autumn), pp. 425-49.
- Tucker, Robert W., 1977. *The Inequality of Nations* (New York: Basic Books).
- Urmson, J. O., 1968. Utilitarianism. *International Encyclopedia of the Social Sciences* (New York: Macmillan), pp. 224-29.
- Veljanovski, Cento G., 1982. The Coase theorems and the economic theory of markets and law. *Kyklos*, vol. 35, fasc. 1, pp. 53-74.
- Waltz, Kenneth, 1959. *Man, the State and War* (New York: Columbia University Press).
- Waltz, Kenneth, 1979. *Theory of World Politics* (Reading, Mass.: Addison-Wesley).
- Williamson, Oliver, 1975. *Markets and Hierarchies: Analysis and Anti-Trust Implications* (New York: The Free Press).
- Young, Oran R., 1980. International regimes: problems of concept formation. *World Politics*, vol. 32, no. 3 (April), pp. 331-56.
- Young, Oran R., 1983. Regime dynamics: the rise and fall of international regimes. In Krasner, 1983, pp. 93-114.