The Rational Design of International Institutions

Author(s): Barbara Koremenos, Charles Lipson and Duncan Snidal

Source: *International Organization*, Autumn, 2001, Vol. 55, No. 4, The Rational Design of International Institutions (Autumn, 2001), pp. 761-799

Published by: The MIT Press

Stable URL: https://www.jstor.org/stable/3078615

The Rational Design of International Institutions

Barbara Koremenos, Charles Lipson, and Duncan Snidal

International institutions are central features of modern international relations. This is true of trade, international debt and financial restructuring, and even national security, once the exclusive realm of pure state action. It was certainly true of the two major military engagements of the 1990s, the wars in Kosovo and the Persian Gulf. As international institutions have gained prominence in the political landscape, they have increasingly become prominent topics for study. The sharpest debate among researchers has been theoretical: Do international institutions really matter? Missing from this debate is a sustained inquiry into how these institutions actually work. We shift the focus by posing researchable questions about how they operate and how they relate to the problems states face.

We begin with a simple observation: major institutions are organized in radically different ways. Some are global, essentially open to all states; others are regional, with restricted memberships. Some institutions give each state an equal vote, whereas others have weighted voting and sometimes require supermajorities. Institutions may have relatively strong central authorities and significant operating responsibilities or be little more than forums for consultation. Some arrangements—

As this project came to fruition, we received valuable input from many sources. We thank Kenneth Abbott, George Downs, James Fearon, Phillip Genschel, Charles Glaser, Lloyd Gruber, Miles Kahler, Robert Keohane, Dan Lindley, Lisa Martin, Ken Oye, Beth Yarbrough, Alexander Thompson, Mark Zacher, and especially Brian Portnoy, who participated in one or more of the conferences leading up to this volume. Jeffrey Smith, Ryan Peirce, Marc Trachtenberg, David Laitin, Joni Harlan, and Jama Adams provided other valuable comments, as did the participants at the Program on International Politics, Economics, and Security (PIPES), University of Chicago, where this project began. Students who participated in Barbara Koremenos' undergraduate seminar at UCLA, "International Cooperation," provided valuable feedback. We also thank the contributors for their efforts, not only on their individual articles but also on the design of the project as a whole. James Morrow, Ronald Mitchell, Peter Rosendorff, Robert Pahre, and especially Andrew Kydd contributed greatly to the project. We received invaluable criticism, prodding, and support from two anonymous reviewers, from the editors of *IO*, and from Lynne Bush. We thank the University of Chicago's Council on Advanced Studies on Peace and International Cooperation for funding support and the Harris Graduate School of Public Policy Studies for hosting the Rational Design conferences. Finally, we thank Loch Macdonald, Barbara Koremenos' neurosurgeon, who was there when we needed him.

International Organization 55, 4, Autumn 2001, pp. 761–799

for example, most bilateral treaties—have no formal organizational structure; these are plentiful because states have a striking tendency to codify their relationships in formal, legal arrangements.[1]

Why do these differences exist? Do they really matter, both for members and for international politics more generally? Do they affect what the institutions themselves can do? We focus on these large questions of institutional design. Our basic presumption, grounded in the broad tradition of rational-choice analysis, is that *states use international institutions to further their own goals, and they design institutions accordingly.* This might seem obvious, but it is surprisingly controversial.

One critique comes from constructivists, who argue that international institutions play a vital, independent role in spreading global norms. We agree that normative discourse is an important aspect of institutional life (though surely not the whole of it) and that norms are contested within, and sometimes propagated by, international institutions. But it is misleading to think of international institutions solely as outside forces or exogenous actors. They are the self-conscious creation of states (and, to a lesser extent, of interest groups and corporations).

The realist critique is exactly the opposite. For them, international institutions are little more than ciphers for state power. This exaggerates an important point. States rarely allow international institutions to become significant autonomous actors. Nonetheless, institutions are considerably more than empty vessels. States spend significant amounts of time and effort constructing institutions precisely because they can advance or impede state goals in the international economy, the environment, and national security. States fight over institutional design because it affects outcomes. Moreover, the institutions they create cannot be changed swiftly or easily to conform to changing configurations of international power. Japan and Germany play modest roles in the UN today because they have been unable to reverse the decision made in 1944–45 to exclude them from the Security Council. Institutions rarely adapt immediately to states' growing (or ebbing) power. For this reason, and because institutions matter, states pay careful attention to institutional design.

Our main goal is to offer a systematic account of the wide range of design features that characterize international institutions. We explore—theoretically and empirically—the *implications* of our basic presumption that states construct and shape institutions to advance their goals. The most direct implication is that design differences are not random. They are the result of rational, purposive interactions among states and other international actors to solve specific problems.

We define international institutions as *explicit arrangements, negotiated among international actors, that prescribe, proscribe, and/or authorize behavior.*[2] Explicit arrangements are public, at least among the parties themselves. According to our definition, they are also the fruits of agreement. We exclude tacit bargains and implicit guidelines, however important they are as general forms of cooperation.

1. See Abbott et al. 2000; and Koremenos 2000.
2. For related definitions of international institutions, see Keohane 1984; and Young 1994.

Institutions may require or prohibit certain behavior or simply permit it. The arrangements themselves may be entirely new, or they may build on less formal arrangements that have evolved over time and are then codified and changed by negotiation. The 1961 Vienna Law on Treaties is a good example.

Although in most arrangements negotiators are typically states, this is not part of our definition; it is an empirical observation that may vary across issues and over time. In fact, nonstate actors participate with increasing frequency in institutional design. Multinational firms, nongovernmental organizations (NGOs), and intergovernmental organizations have all shaped international institutions, solely especially those dealing with the world economy, the environment, and human rights.

Thus our definition of international institutions is relatively broad. It includes formal organizations like the World Health Organization and International Labor Organization, as well as well-defined (and explicit) arrangements like "diplomatic immunity" that have no formal bureaucracy or enforcement mechanisms but are fundamental to the conduct of international affairs.

With this definition in mind, we can begin to explore how institutions vary and, later, how that variation may be the product of rational design considerations. Our work emphasizes five key dimensions within which institutions may vary:

Membership rules (MEMBERSHIP)
Scope of issues covered (SCOPE)
Centralization of tasks (CENTRALIZATION)
Rules for controlling the institution (CONTROL)
Flexibility of arrangements (FLEXIBILITY)

These are certainly not the only significant institutional dimensions, but they have several advantages for our research. First, they are all substantively important. Negotiators typically focus on them, and so do analysts who study institutions. Second, they can be measured, allowing us to compare them within and across institutions over time. Third, they apply to the full array of international institutions, from the most formal to the least bureaucratic.

We locate our analysis in the rational regime tradition. We do not present a literature review but rather build on earlier work to develop the underlying parameters of this research project. We also do not counterpose "dueling perspectives" (realism versus institutionalism or rationalism versus constructivism, for example). Instead, we investigate the rational design approach on its own terms by developing a set of theoretically based conjectures, which are then evaluated empirically in the studies in this special issue of *International Organization.* Our view is that rational design can explain much about institutions, but not everything.[3]

3. Martin and Simmons assess past work on international institutions and propose an agenda focused on explaining causal mechanisms and institutional effects. Martin and Simmons 1998. Their framework complements ours and shows how rational choice can address other important empirical questions.

From Cooperation Theory to Rational Design

The postwar study of international institutions is coming full circle, but with a theoretical twist. The early literature focused on the operational details of international organizations. With the notable exception of neofunctionalist integration theory, it was heavily descriptive,[4] neither theorizing institutions nor clarifying their relationships to wider issues of international relations. By the 1980s the literature had turned sharply toward theory under the broad rubric of "regimes."[5] Within regime theory, one important strand built on rational, game-theoretic analysis, especially the idea that the "shadow of the future" can support "cooperation under anarchy."[6]

The study of regimes favored theoretical questions and moved the research agenda away from analyzing specific institutional arrangements.[7] Likewise, the tools of game theory were directed mainly at general theoretical questions, focusing on cooperation, not institutions, as the dependent variable. The overriding question became "How could states and other international actors produce cooperative outcomes by their own, self-interested choices?" Indirectly, however, this work laid the foundation for a renewed exploration of institutions, this time as part of a wider theory of international cooperation. In focusing on how self-interested states could cooperate, it was logical to ask what role institutions could play. Institutions could be reconceptualized and theorized as arrangements that make cooperation more feasible and durable, at least in some circumstances.

Our goal is to close the circle that began with descriptive studies by explaining major institutional features in a theoretically informed way. We first relax some key assumptions of cooperation theory and then bring in institutions directly by incorporating insights from game theory and institutional analysis. In doing so, we pay particular attention to the logic of their development.

Extending Cooperation Theory

The cooperation literature is premised on the "Folk theorem," which shows that cooperation is possible in repeated games.[8] This result has a strong theoretical foundation and can be applied empirically to a wide range of contemporary issues. The density of contemporary international interdependence creates repeated inter-

4. The early issues of *International Organization,* for example, focused on describing newly formed organizations and publicizing their rules and votes.

5. Krasner 1983.

6. See Oye 1986; and Axelrod 1984.

7. Key works are Stephen Krasner's edited volume *International Regimes* (1983) and Robert Keohane's *After Hegemony* (1984). An excellent early overview is Haggard and Simmons 1987. Several commentators have noted that the field has had less and less to say about formal international organizations. See Rochester 1986; and Abbott and Snidal 1998.

8. See Friedman 1971; and Fudenberg and Maskin 1986.

action that makes cooperation feasible.[9] In brief, the possibility of cooperation is present in most modern international issues.

If cooperation is within reach, why it is not always grasped? To answer that, we must go beyond any simple, optimistic interpretation of the Folk theorem. Although we assume that the general conditions of international interdependence are propitious, individual issues have features that make achieving and maintaining cooperation more problematic. Moreover, the standard Folk theorem conclusion needs careful refinement when applied to more realistic situations, where competing equilibria are in play, many actors are involved, and uncertainty is high.

Multiple equilibria are a major obstacle to cooperation that was downplayed by the early emphasis on 2 × 2 games. Although these simple games, especially Prisoners' Dilemma, did much to clarify our understanding of enforcement problems, their very simplicity could be misleading. In a simple 2 × 2 Prisoners' Dilemma, there is only one point of mutual cooperation, the unattainable Pareto optimum where both sides choose to cooperate rather than defect. In practice, states have a wide range of choices and many possible cooperative outcomes, often with different distributional consequences.

If actors prefer different outcomes, the range of possibilities creates bargaining problems. Which cooperative outcome should they choose? How, in other words, should they share any mutual gains from cooperation? These distributional questions do not arise in simple 2 × 2 Prisoners' Dilemma games, though they were discussed in some early work contrasting Prisoners' Dilemma and Coordination games.[10] Recent work by Stephen Krasner, James Morrow, and James Fearon goes further, showing how distributional differences can undermine cooperation in significant ways. Hence, distribution problems merit at least as much attention as enforcement problems, which we know hamper international cooperation.[11]

Large numbers also complicate cooperation. Kenneth Oye addresses the collective-action problem primarily by showing how interactions among large numbers can be decomposed into simple bilateral interactions.[12] Some issues, however, cannot be decomposed this way for technical reasons; others should not be decomposed because successful cooperation requires joint action by all (as in the provision of public goods). Large numbers raise questions about how to share both the costs and benefits of cooperation, especially when some actors are richer, bigger, or more powerful than others.

Uncertainty is a frequent obstacle to cooperation, as is "noise," the difficulty of observing others' actions clearly.[13] States are naturally reluctant to disclose vital

9. Notable exceptions are crises where immediate incentives overwhelm longer-term considerations. We set such situations aside.

10. See Snidal 1985; and Stein 1983.

11. See Krasner 1991; Morrow 1994c; and Fearon 1998.

12. See Oye 1986; and Lipson 1986 for an application.

13. This point was foreshadowed by Downs, Rocke, and Siverson in their analysis of arms races, and by Downs and Rocke in their game-theoretic analysis of the limits to cooperation. See Downs, Rocke, and Siverson 1986; and Downs and Rocke 1990.

information that could make them more vulnerable. Reducing uncertainty among participants is a major function of institutions.[14]

Taken together, these factors—distribution, enforcement, large numbers, and uncertainty—suggest that cooperation can be very brittle in the real world. As these factors vary, the prospects for cooperation can shift dramatically, making it far more difficult to manage international cooperation than earlier, simplified theories would predict.

Bringing in Institutions

In broad international relations (IR) theories institutions play only a modest role. It is, after all, cooperation *under anarchy*. The primary reason for emphasizing anarchy is to rule out centralized enforcement, but there is little consideration of the other roles institutions might play. In fact, institutions often help resolve problems of decentralized cooperation.

IR theorists have begun to address problems of cooperation in more complex and realistic settings, where there may be noise and large numbers.[15] It is generally recognized that institutions may make cooperation more likely,[16] and the compliance literature has begun to analyze empirically how regime design promotes effective cooperation.[17] So far, however, this has not developed into a more general theoretical analysis of specific institutional arrangements.

Our work departs significantly from the earlier cooperation literature. Because decentralized cooperation (supported by the Folk theorem) is difficult to achieve and often brittle, states devise institutions to promote cooperation and make it more resilient. But the form these institutions take varies widely. Often the necessary institutions are fairly minimal and simply reinforce the underlying conditions for cooperation, perhaps providing the information necessary for bilateral bargains. Other times, more complex problems may require a larger institutional role—such as when an issue involves actors with very different resources and information. Under these circumstances, institutions can play a major role in facilitating cooperation.

We argue that many institutional arrangements are best understood through "rational design" among multiple participants. This rationality is forward looking as states use diplomacy and conferences to select institutional features to further their individual and collective goals, both by creating new institutions and modifying existing ones. Even trial-and-error experiments can be rational and forward looking in this way. Although we do not argue that all institutional change is the product of conscious design, we do consider it the overriding mechanism guiding the devel-

14. See Keohane 1984; and Morrow 1994c.
15. On noise, see Downs and Rocke 1990. On large numbers, see Pahre 1994.
16. See Keohane 1984; and Axelrod and Keohane 1986.
17. See Chayes and Chayes 1995; and Mitchell 1994.

opment of international institutions.[18] Moreover, though our primary purpose is to explain institutional design, our approach also provides an appropriate foundation for prescribing policy and evaluating existing institutions.[19]

Our argument that institutional design is deliberate is reflected in the difficult process of creating an international institution. The evolution of the General Agreement on Tariffs and Trade (GATT) into the World Trade Organization (WTO) involved extensive rounds of negotiation. The Law of the Sea Treaty was the culmination of protracted debate, including the sharply contested decision not to have stronger centralized institutions. The same process is seen in the development of the UN charter, which involved extensive planning and bargaining and was designed to achieve critical goals amidst great uncertainty. Moreover, its design has been modified over the years as new members have been admitted, the Security Council has changed, and specialized agencies have been created. Continuing calls for change remind us that most institutions evolve as members learn, new problems arise, and international structures shift. But institutional evolution still involves deliberate choices made in response to changing conditions.

Institutional development frequently depends on prior outcomes ("path dependence") and evolutionary forces. As institutions evolve, rational design choices can arise in two ways. First, participants may modify institutions in stages, by making purposeful decisions as new circumstances arise, by imitating features from other institutions that work well in similar settings, or by designing explicit institutions to strengthen tacit cooperation. Second, institutions may evolve as states (and other international actors) select among them over time. States favor some institutions because they are better suited to new conditions or new problems and abandon or downplay those that are not. For example, the obvious place to handle intellectual property rights would seem to be the World Intellectual Property Organization, but the countries that generate most patents chose to move the issue to the WTO because it offered better enforcement mechanisms. Thus the institutionalization of the issue evolved significantly, not because an older institution was modified, but because another one offered a better institutional design.[20]

Even institutions that are not highly formalized and arise through informal and evolutionary processes may embody significant rational design principles. Sovereignty is clearly the result of historical and normative processes, but at important

18. Our proposed conjectures are consistent with an evolutionary perspective that treats rational designs as superior in the sense of providing greater benefits to participants, even if participants are unwitting beneficiaries. Miles Kahler provides an excellent overview and discussion of the relationship between evolutionary and rational theories of international institutions. Kahler 1999. The two approaches begin to align through such concepts as "learning" and "imitation" as key factors underlying institutional development.

19. Of course, many efforts at institutional design fail. States may misunderstand the circumstances they face or wrongly anticipate how actors will respond to institutional innovations, or simply make mistakes.

20. See Schrader 1996.

junctures (Treaty of Westphalia, Congress of Vienna, Vienna Convention) it has been the object of rational design through codification and modification.

Thus, *our basic strategy is to treat institutions as rational, negotiated responses to the problems international actors face.* We can connect our definition of institutions to the language of game theory, where institutions are aspects of equilibria, including the rules of the game and the expectations of the actors.[21] This equilibrium approach has several important implications.[22]

First, institutional rules must be "incentive compatible" so that actors create, change, and adhere to institutions because doing so is in their interests. Consider an institution that can be sustained only through sanctions and whose members must apply these sanctions themselves. This is an equilibrium institution only if the members who are supposed to apply sanctions actually have incentives to do so. Incentive compatibility does not mean that members always adhere to rules or that every state always benefits from the institutions to which it belongs. It does mean that over the long haul states gain by participating in specific institutions—or else they will abandon them.

Second, specifying independent and dependent variables requires special care. An equilibrium is a statement of consistency among its elements. Decomposing an equilibrium into causal statements connecting independent and dependent variables requires looking beyond the equilibrium itself to the sequence of, and reasons for, institutional changes.

Third, the very institutions we seek to explain as "outcomes" may also play a causal role in shaping others, now or in the future. Consider the EU. Is it a "dependent" or an "independent" variable? The answer depends on the question we ask and the time frame we use. If we want to explain why the EU was formed and the features it has, it is a dependent variable (by our own choice). If we want to explain the shape of some subsequent institution, such as the WTO or the European Monetary System, the EU plays a significant causal role as an independent variable in the institution's development. This is particularly important when we look at which actors are relevant to a particular design issue. An outcome (or dependent variable) at one stage—the membership of the EU—may become a causal factor (or independent variable) at another—the number of actors relevant in the design of the European Monetary System.

Dependent Variables

Consider an emerging international issue, such as global warming, the distribution of pirated software, or the sale of cloned human organs. If states want to promote a common interest, what kinds of institutions might they design to aid their efforts?

21. The converse is not true, and not all equilibria are institutions as we define them. In particular we exclude equilibria resulting from tacit bargains and implicit arrangements that arise without negotiation.
22. See Calvert 1995; Morrow 1994c; and Snidal 1997.

They might first ask whether they need an international institution at all. Perhaps their national capacities are more than adequate, or they are converging on tacit arrangements that require little elaboration. If they could benefit from explicit cooperation, they would ask whether current institutions could be extended to cover the issue, in whole or in part.

If the issue were novel (such as trade in cloned organs) and no existing organizations were well suited, then diplomats, executives, scientists, policy activists, and other interested parties might well consider creating a new organization. They would immediately confront several major questions. Should the new institution cover only cloned organs or should it also cover health- or trade-related issues? Should membership be limited to countries with advanced medical industries? What about other, less-developed countries? One practical reason for being inclusive is that excluded states might evade or undermine the rules. What about including scientific institutes, biotechnology companies, health advocates, medical ethicists, and other nonstate actors?

What institutional capacities are needed for success? Would a simple agreement suffice? Should the institution be centralized to collect data, monitor compliance, or even enforce some rules? Or should it be more decentralized, serving mainly as a forum for periodic bargaining? Should all actors be given equal voice and vote, or should some have only an informal, consultative role? What about the rules themselves in such a new and rapidly developing area? Should they be clear-cut and firm, or should they be more flexible, allowing easy changes by mutual agreement or opting out by dissatisfied states?

Regardless of the issue, these kinds of institutional choices zero in on our major concerns: how and why are international institutions designed as they are? To make headway on these overarching questions, we need some clear way to mark out major variations in institutional design. The simplest solution would be to use a single measure, one that describes institutions as, say, "stronger" or "weaker." Unfortunately, such measures are misleading because they collapse several important institutional features into one overly simple statement. We could measure many institutional features in great detail, yielding rich descriptions of individual institutions, but this would obscure the most important types of variation among them. We have chosen instead to focus on a few recurrent problems of institutional design, particularly those we can identify theoretically as vital aspects of cooperation and that vary in measurable ways. Our approach highlights five key dimensions: MEMBERSHIP, SCOPE, CENTRALIZATION, CONTROL, and FLEXIBILITY. These are not the only important dimensions of institutions. Others may well prove significant, theoretically and substantively. In some cases, our dimensions must be refined to clarify design issues in specific institutions. Centralization, for instance, is a broad category—perhaps too broad for some cases. Nonetheless, our first effort is to reduce the myriad elements of institutional variation to a few measurable dimensions that show up repeatedly when institutions are designed or modified. We now take a closer look at each dimension and consider how they vary in modern international institutions.

Membership

Who belongs to the institution? Is membership exclusive and restrictive, like the G-7's limitation to rich countries? Or is it inclusive by design, like the UN? Is it regional, like ASEAN, or is it universal? Is it restricted to states, or can NGOs join?

Membership has been one of the most hotly contested issues in recent years. The expansion of NATO into Eastern Europe is a key example. Expansion, for those who favor it, represents a reinvigoration of the alliance, a commitment to the joint defense of Central Europe, and a symbolic inclusion of new members in the "West." For those who oppose it, NATO's movement to the East adds nothing to the defense of Western Europe and needlessly provokes an already humiliated Russia. These issues resonate widely because NATO is such a prominent and consequential institution.

Scope

What issues are covered? In global trade institutions, for example, some of the toughest battles have been over which sectors to include in negotiations. GATT left out several key economic sectors, but the WTO has expanded to incorporate most trade issues, including agriculture and services. It may be expanded further to include cross-border investments. At the other end of the spectrum are institutions like the 1965 U.S.–Canada auto trade deal designed to cover only one or two narrowly defined issues. This agreement, too, was eventually widened when it was incorporated into NAFTA.

Sometimes two seemingly unrelated issues are linked. A trade issue, for example, may be linked to a security issue to facilitate agreement and compliance. Or a side payment may be offered, as when the Nuclear Nonproliferation Treaty offered the transfer of peaceful nuclear technology to states that agreed to forgo nuclear weapons. Such side payments are clear evidence that scope is being manipulated to facilitate cooperation.

There is a continuum of issue coverage. At one end are institutions like the Antarctic Treaty System that cover a range of scientific, economic, and political issues. At the other end are some early environmental agreements that are restricted to a few well-defined issues, such as greenhouse gas emissions.

Sometimes scope is not open to design choice because of technical considerations or shared perceptions. In the Law of the Sea negotiations, for example, jurisdiction over ocean territories could not be separated from coastal environment and fishing rights issues. Technological interactions required that these issues be dealt with together in a comprehensive settlement.[23] But other Law of the Sea issues seemed

23. A parallel and important implication within rational institutional design is that all relevant "margins" of choice must be considered. Barzel 1989. In John Richards' analysis of international airline regulation in this volume, for example, effective agreements on airline fares also require that airlines be prohibited from competing on other margins, such as food quality or seat comfort.

to have little in common. Here linkage was more cognitive—a result of how issues were framed, especially under the rubric of the "common heritage of mankind."[24]

One difficulty in analyzing scope is that the issues themselves are not clearly defined. Does trade in all commodities constitute an issue? Or should we distinguish agricultural goods from manufactures? Although there is no general answer to this difficult task of assessing issue scope, focused empirical research can reveal the extent to which actors narrow or broaden the range of matters being addressed. The problem is simplified when negotiations are expanded to cover items that could clearly be dealt with separately or were not previously linked (as occurred with the "baskets" of the Helsinki negotiations). Most important, changes in institutional issue linkage over time indicate changes in scope within an arrangement.

Centralization

Are some important institutional tasks performed by a single focal entity or not? Scholars often misleadingly equate centralization with centralized enforcement. We use the term more broadly to cover a wide range of centralized activities. In particular we focus on centralization to disseminate information, to reduce bargaining and transaction costs, and to enhance enforcement. These categories are not exhaustive, but they cover many important centralized activities found at the international level.

Centralization is controversial, politically and conceptually, because it touches so directly on national sovereignty. According to the traditional view, states reject any form of centralized international authority. International relations is seen as an immutable anarchy. This is a powerful assertion, but it is only partly right. It blends a simplifying assumption (that theory building should begin with states as independent units) with some hyperbole and errant conclusions.

States understandably guard their domestic authority and their control over foreign policy. They are suspicious of encroachments by other states and strongly resist any shift of sovereign responsibilities to superordinate bodies. But saying that states rarely devolve such authority is inaccurate, and it is a misleading basis for constructing theory. After all, European states not only signed the Treaty of Rome but also agreed to the Single European Act, which permits majority voting.[25] They went still further at Maastricht, when they abolished national controls over money.[26] The EU is uniquely powerful as an international institution, but centralized controls are important elsewhere. The dispute-resolution panels of the WTO are a particularly significant example.

The least intrusive form of centralization is information collection, and many international institutions engage in it. Members of the IMF, for instance, need not

24. Haas 1980.
25. Moravcsik 1991.
26. See Kenen 1995; and Moravcsik 1998.

gather their own data on others' balance of payments. Instead the IMF regularly collects, evaluates, and publishes itemized statistics on its members' payments.

Bargaining procedures and rule enforcement can also be more or less centralized. At the World Bank, for instance, specialists negotiate loans for economic adjustment or major infrastructure investments. These packages require collective approval from a centralized body of members. Most international organizations have relatively decentralized enforcement arrangements. They specify possible punishments for rule violations but leave it up to the members to apply them. Because these multilateral sanctions are both limited and well specified, they minimize the chances for disproportionate punishment or cycles of retaliation. Still, the members themselves must apply the decentralized punishments and bear the inevitable costs.

GATT (and now the WTO) have relied on such decentralized sanctions for decades. If a dispute panel found violations of international trade rules, it was up to the injured party to retaliate within specified limits. GATT itself had no centralized power to punish or reward, only to authorize individual members to do so. This also shows how international organizations can combine elements of centralization and decentralization. The WTO's centralized arrangements for judging trade disputes go hand-in-hand with decentralized arrangements for enforcing the judgments.

Control

How will collective decisions be made? Control is determined by a range of factors, including the rules for electing key officials and the way an institution is financed. We focus on voting arrangements as one important and observable aspect of control.

Even if membership is universal, some states may carry considerably more weight than others because of voting and decision-making rules. Two interrelated rules are especially important: whether all members have equal votes and whether a minority holds veto power. If a minority can veto, its votes inherently carry special weight. In the UN General Assembly all members have equal votes. In the Security Council they do not, since only the permanent members can veto resolutions. The IMF and World Bank have explicit weighted-voting rules; the larger economies, which provide capital to these institutions, carry disproportionate votes. Another element of control is whether a simple majority, a super-majority, or unanimity is required. If a super-majority is needed, some state (or combination of states) may be able to block new rules, members, or officers.

Finally, we distinguish control from centralization. While centralization may reduce control in some cases, the two dependent variables generally vary independently. For example, changes in the voting rules within a quasi-legislative component of an international institution represent changes in control that do not affect the level of centralization. Similarly, centralizing information collection usually has little, if any, effect on who controls an institution.

Flexibility

How will institutional rules and procedures accommodate new circumstances? Institutions may confront unanticipated circumstances or shocks, or face new demands from domestic coalitions or clusters of states wanting to change important rules or procedures. What kind of flexibility does an institution allow to meet such challenges?

It is important to distinguish between two kinds of institutional flexibility: adaptive and transformative. "Escape clauses" are a good example of adaptive flexibility. They allow members to respond to unanticipated shocks or special domestic circumstances while preserving existing institutional arrangements. The general goal is to isolate a special problem—such as a spike in steel imports from a few producing countries—and insulate the broader institution (in this case, the GATT/WTO) from its impact. This limited flexibility is designed to deal chiefly with outlying cases, to wall them off from run-of-the-mill issues.

Some institutions have built-in arrangements to transform themselves in ways that are more profound. This deeper kind of flexibility usually involves clauses that permit renegotiation or sunset provisions that require new negotiations and ratification for the institution to survive. The initial terms of commodity agreements, for example, are typically five to seven years, after which they expire and have to be renegotiated. GATT did not have such a provision, but its periodic rounds of trade negotiations facilitated planning for larger institutional changes, leading to the WTO. GATT's existing rules did nothing to block these larger changes, and its regular forums served to promote them.

Independent Variables

To explain variation in institutional design, we focus on the following independent variables: distribution problems (DISTRIBUTION); enforcement problems (ENFORCEMENT); number of actors and the asymmetries among them (NUMBER); and uncertainty about behavior, the state of the world, and others' preferences (UNCERTAINTY ABOUT BEHAVIOR, UNCERTAINTY ABOUT THE STATE OF THE WORLD, and UNCERTAINTY ABOUT PREFERENCES).

Enforcement of agreements is a cornerstone concern in international anarchy. But recent debates have increasingly stressed that to understand which, if any, international institutional bargains are struck, one must examine distributional issues. The number and relative size of key actors has been a long-standing concern in debates about international cooperation, hegemony, and, more recently, the interrelationship of regional and global politics. Finally, uncertainty is the linchpin of traditional security problems and is equally central in economic and environmental issues.

These variables also play a crucial role in game theory. Enforcement and distribution problems emerge in any strategic situation. Number is the central variable of collective-action theory, and we broaden it here to include explicitly the

asymmetries that are so important in international affairs. Finally, many important theoretical developments in game theory over the past two decades center on uncertainty.

Since we extend the existing tradition of cooperation theory, it is useful to compare our independent variables with Oye's.[27] After all, institutions to promote cooperation must be designed around the factors that affect cooperation. But we adapt the independent variables to address the particular questions raised by institutional design. Oye focuses on three independent variables. The most important is "shadow of the future." We do not focus on this as a primary source of institutional variation because the general conditions for cooperation are typically met under contemporary conditions of high interdependence.[28] Instead, we emphasize how variation in the significance of enforcement problems across different issues affects institutional design.

Oye's second independent variable is the type of 2×2 game being played, though with an emphasis on Prisoners' Dilemma. Simple games have yielded important insights and have been subjected to important criticisms.[29] The most important substantive criticism is that concentration on Prisoners' Dilemma leads to an overemphasis on enforcement and cheating and to an underemphasis on distributional conflicts.[30] This problem can be partially solved by shifting attention to another 2×2 game (Coordination, for example), but each new game misses some other salient problem (such as enforcement). We resolve this by looking at distribution problems as a second independent variable.[31]

We use a broader version of Oye's third variable, "number." Looking beyond the raw number of actors relevant to an issue, we include asymmetries that might exist among them due to different capabilities. This consideration was important in the hegemony literature and becomes even more so in understanding how different-sized actors share control in institutionalized cooperation.

Finally, and most important, driven by advances in the economics of uncertainty and game theory we add "uncertainty" as a new category of independent variable. Uncertainty can impede cooperation, but its impact can be managed through institutions. Indeed, one feature common to our independent variables is that

27. Oye 1986.

28. Alternatively, states will not waste time designing institutions that will not be enforced by their own incentives.

29. In particular, once the games are complicated even slightly, the clean distinctions among them break down. When Prisoners' Dilemma repeats through time, for example, multiple equilibria emerge, and the supergame contains distributional problems. Similarly, recurring Battle of the Sexes problems create incentives for some states to shift the prevailing equilibrium.

30. See Krasner 1991; and Grieco 1988.

31. James Fearon makes a parallel argument that, at a sufficiently general level, all problems in international relations have a common strategic structure. Fearon 1998. States must choose among the range of available cooperative arrangements and ensure that participants will adhere to the chosen arrangement. We label these the "distribution problem" and the "enforcement problem," respectively.

game-theoretic logic allows us to connect them to the dependent variables of institutional design.[32]

Distribution Problems

When more than one cooperative agreement is possible, actors may face a *distribution problem.* Its magnitude depends on how each actor compares its preferred alternative to other actors' preferred alternatives. In a pure Coordination game, where both actors prefer the same coordination point(s), there is no distribution problem. Distribution problems are greater when actors want to coordinate in a "Battle of the Sexes" game according to the intensity with which they prefer alternative coordination points. In Prisoners' Dilemma games where there are multiple efficient equilibria, the distribution problem depends on actors' differences "along the Pareto frontier."[33] Finally, the problem is most severe in a zero-sum game because a better outcome for one leaves less for the others.

Distribution problems are closely related to bargaining costs.[34] In general, where the distributional implications of a choice are small (such as when only one efficient outcome is possible or the shadow of the future is short), bargaining costs will be relatively small. In situations where the distributional implications are large (such as when there are multiple, substantially different efficient outcomes or the shadow of the future is long), bargaining costs will likely be large.

Distribution problems interact with the other independent variables, but they should be kept separate. Most important, distribution problems are not the same as uncertainty. Uncertainty arises when an actor cannot anticipate the outcome that will result from an agreement and knows only the stochastic "distribution" generating the outcome. In their collaborative venture to develop an anti-missile system, for example, Japan and the United States are uncertain whether the research will be successful even though they are sure they will both share fully in the findings. In contrast, a distribution problem refers to selecting one outcome from a range of known possible outcomes. In allocating quotas for harvesting West Coast salmon, for example, Canada and the United States know the total number of fish that will be caught; the problem is determining each country's allotment. Of course, these problems intertwine in many situations where actors choose among agreements characterized by different stochastic distributions. This is true of fishing agreements over time where both the allotments between states and the size of the fish harvest over time are at stake.

32. We asked contributors to examine these independent variables but also invited them to consider others; thus the project as a whole is open to a wider set of independent variables, albeit in a more inductive way.

33. Krasner 1991.

34. Fearon 1998.

Enforcement Problems

Enforcement problems refers to the strength of individual actors' incentives to cheat on a given agreement or set of rules. Even if an arrangement makes everyone better off, some or all actors may prefer not to adhere to it because they can do better individually by cheating—the heart of Prisoners' Dilemma and public goods problems.

The enforcement problem arises when actors find (current) unilateral noncooperation so enticing that they sacrifice long-term cooperation. It can be measured by the minimum discount factor (a state's valuation of future, as opposed to current, benefits) necessary to support cooperation. Seen this way, the necessary discount factor is a characteristic of the issue—including actors' payoffs from cooperation and defection and how frequently they interact—but not of how much actors actually value the future. Issues where actors have large incentives to break an agreement require higher discount factors to support cooperation than do issues where the immediate gains from noncooperation are smaller.

Although we focus on settings of high interdependence where cooperation is generally possible, there is significant variation across issues. At one extreme are cases with no enforcement problems, such as agreements to set technical standards where actors have no incentive to defect. Within the context of repeated Prisoners' Dilemma games, self-enforcing agreements may arise if incentives to defect are small relative to the shadow of the future. But if incentives to defect are greater, or interactions are less frequent, enforcement problems emerge.

Most situations contain both distribution and enforcement problems. In efforts to halt stratospheric ozone loss, for example, the ozone regime needed to set targets for reducing global chloro-fluorocarbon (CFC) emissions and establish rules for cutting back CFC production and use. Different rules obviously impose quite different costs on various states. Whatever rules are chosen still have to be enforced. Knowing this, states may choose particular rules partly because they are easy to monitor and enforce. In this way problems of distribution and enforcement are tightly connected.

Distribution and enforcement can be blended in differing proportions. Some problems are more squarely related to enforcement, with distributional considerations clearly secondary. If first strikes can paralyze one's opponent, enforcement of any arms control agreement overwhelms any distributional concerns about armament levels. Other issues present major distribution problems, with enforcement as a secondary issue. Macroeconomic coordination among the G-7 countries seems to have this property.[35] The same could be said of the last three GATT rounds. The critical issue was who would make what concessions, not whether the resulting agreements would be enforced.

Separating enforcement problems from distribution problems is an analytic choice, not a substantive claim. Unlike early work based on Prisoners' Dilemma or more recent work based on Coordination, it enables us to consider the more typical

35. Webb 1991.

case, where enforcement and distribution problems occur simultaneously. It does not capture more nuanced interactions between enforcement and distribution problems, but by first examining the institutional issues raised by these "main effects," we will be better situated to understand the others. Finally, it is necessary to keep enforcement problems distinct from the other independent variables. Uncertainty and large numbers usually aggravate enforcement problems, but enforcement problems can arise even in repeated-game situations with small numbers and no uncertainty.

Number of Actors

Number of actors refers to the actors that are potentially relevant to joint welfare because their actions affect others or others' actions affect them. Sulfur emissions from factories in the U.S. Midwest, for example, cause acid rain in Eastern Canada and New England, an issue involving two countries. Greenhouse gases emitted from the same factories contribute to global warming, an issue affecting more actors because of the large-scale consequences of global climate change. If firms are seen as the relevant actors, then the number of actors is significantly larger in both cases.

The number of actors involved in military issues depends on technology and on states' ability to harm or help one another militarily. Peace in the Middle East now depends on more states than it once did because technological innovations have increased the range of military aircraft and thus the number of states that can affect the military balance. Were Pakistan able to target Israel with nuclear weapons, it, too, would become a key actor.

Number does not depend solely on geographic or technological factors and is often determined by prior political and institutional arrangements. For example, a decision by the EU about monetary union is effectively a fifteen-state decision, regardless of its effects on outsiders, because EU members made a political decision to limit the number of states involved in the process, not because other states are unaffected. Similarly, when NAFTA takes up an issue, only its three members have a voice, whereas the same issue taken up within an expanded hemispheric trade arrangement would involve more states. In effect, the prior institutional membership decision has redefined the range of "potentially relevant" actors for the issue at hand.[36]

Thus it is important to distinguish between the independent variable, number, and the dependent variable, membership. Number is an exogenous feature of the issue context, including prior institutional developments, in which an institution may or may not be established. It includes the set of interested actors and their relative power in and importance to the issue. In contrast, membership is an endogenous design choice made in the course of establishing, changing, and/or operating the institution. It includes, for our purposes, the rules governing who is a member and

36. Snidal 1994.

(if relevant) different classes of membership. Over time, prior membership choices may affect number—that is, endogenous choices become exogenous constraints—because institutional settings, such as the EU or NAFTA, determine which actors will have standing in subsequent institutional negotiations.

Number also includes asymmetrical distribution of actors' capabilities. On some issues many states may be nominally involved, but only a few really drive the issue. Every state has an interest in the international economy, for example, but few have the economic power to determine its course. Similarly, many states produce some oil, copper, or bauxite, but only a few states dominate the global production of each.

The actors involved in an issue are not always the same as those who become members of the final institution. Although the entire EU membership discussed monetary union, only some met the requirements and chose to join. Similarly, while trade affects virtually all states, not all have played an active role in multilateral negotiations, and not all are members of the WTO.

Uncertainty

Uncertainty refers to the extent to which actors are not fully informed about others' behavior, the state of the world, and/or others' preferences. These distinctions correspond to three important elements of any strategic situation: choices, consequences, and preferences, respectively; and they may have different implications for institutional design. For example, uncertainty about behavior makes cooperation more difficult in many cases, but uncertainty about the state of the world may, under certain conditions, make cooperation easier. Therefore, our assertions are not about generic effects of uncertainty but about the different ways states design institutions to cope with specific types of uncertainty.

Uncertainty about behavior. States may be unsure about the actions taken by others. If states agree not to pursue technologies associated with the development of biological or chemical weapons, for example, some states may have no way of knowing whether others are abiding by the agreement. Similarly, if countries agree to restrict sulfur emissions to reduce acid rain, how can they be sure others are complying with the agreement?[37]

Uncertainty about the state of the world. Uncertainty about the state of the world refers to states' knowledge about the consequences of their own actions, the actions of other states, or the actions of international institutions. This could be scientific and technical knowledge or political and economic knowledge. Consider the dispute over the Spratly Islands, which lie off the southern coast of China and have been claimed by a number of states. Any agreement governing the dispute would have to take into account that no one knows how much oil is actually there or its future value.

37. Levy 1993.

Uncertainty about preferences. Governments are often unsure what their counterparts really want. We assume states know their own preferences, but they are often uncertain about the preferences or motivations of others. A key problem underlying arms competition is determining whether another state is simply seeking its own security or is greedy and expansive. Does India's nuclear testing reflect a desire to aggrandize itself at Pakistan's expense or to defend itself against China? Of course, a major problem in determining others' preferences is that states may have incentives to misrepresent their preferences, either verbally or through their actions.

We do not use standard game-theoretic terminology, such as *imperfect information* or *incomplete information,* because it would obscure important distinctions.[38] For example, we could capture uncertainty both about the state of the world and about preferences (or type) through games of incomplete information. But collapsing these into one category prevents us from drawing nuanced inferences about institutional design. Foreshadowing the conjectures discussed later, membership rules may mitigate uncertainty about preferences but not about the state of the world. Similarly, flexibility provisions can help states cope with uncertainty about the state of the world but have no effect on reducing uncertainty about behavior.

Although distinguishing among these kinds of uncertainty is useful conceptually, in practice they are often combined. For example, do European efforts to restrict imports of U.S. beef produced with hormone supplements reflect a concern for consumers' health or for local farmers' profits (uncertainty about others' preferences)? Scientific uncertainty (uncertainty about the state of the world) was also present initially but was resolved when a WTO-appointed panel ruled that hormones posed no health threat. An obvious solution would be to label imported beef as such and let individual Europeans make their own choices. Unfortunately, concerns about monitoring such a labeling system (uncertainty about behavior) would frustrate this solution.

Different mixes of uncertainty often characterize an issue. For example, the environmental area is plagued by enormous uncertainty (most of it scientific) about the state of the world and much less uncertainty about preferences. In contrast, there was little uncertainty about force structures during the latter years of the Cold War, but each superpower had significant uncertainty about the preferences of the other. We would expect the design of agreements in these areas to reflect their different circumstances.

Interactions Among Independent Variables

Our research design is quite simple. We have isolated a set of independent variables that we expect will determine the choice of particular institutional design features—our dependent variables. In our conjectures, we focus on "main effects"—that is, the bivariate relationships between the independent and dependent variables.

38. We do adopt standard terminology in using the term *uncertainty* instead of *risk*. See, for example, Kreps 1990; Hirshleifer and Riley 1992; and Osborne and Rubinstein 1994.

This approach has several advantages. It provides a general framework for a wide range of empirical studies and fosters comparisons across cases while allowing individual analysts to explore the implications of interactions in their particular cases. Moreover, the emphasis on bivariate relationships allows us to connect our conjectures closely to existing theoretical work—which would be possible for some but not all of the more complex interactions. Although simplicity has tremendous advantages, it ignores potential interactions among the independent variables. Enforcement problems may be combined with uncertainty about preferences or actions, as in an arms control context. Or distribution problems may be combined with large numbers, as in environmental public goods contexts. Because our independent variables may combine in many ways, we need to consider the significance of their interactions.[39] For example, when an enforcement problem occurs in a repeated Prisoners' Dilemma, cooperation is possible provided actors are sufficiently patient. But when uncertainty about actions enters the picture, the viability of cooperative strategies declines, since these strategies hinge on actors' knowledge of each other's behavior. Here the combination of two problems is substantially worse than either one alone. Similarly, uncertainty about the state of the world can interact with distributional problems, making cooperation even more challenging.[40]

The interaction of independent variables can also enhance cooperation. While both large numbers and distributional differences typically impede cooperation, sometimes large numbers mitigate distributional problems by easing relative gains concerns or by offering additional ways to balance costs and benefits across actors.

Conjectures About Rational Design

In this section we develop a series of conjectures that connect our independent and dependent variables. We call these "conjectures" to indicate that they represent generalizations based on a common rational-choice theoretical framework, although they are not formally derived here; however, in presenting the underlying logic of each conjecture we identify close variants that have been formally derived by scholars working in the rational-choice tradition. Although the conjectures follow from this general framework, individual conjectures depend on logics that may entail specific substantive assumptions. For example, public goods arguments assume that all actors share the same goals, whereas "screening" arguments suppose

39. Interaction effects may be positive, negative, or zero—that is, when two "problems" arise together in a given context, their joint effect may be less than either problem individually (a large negative effect) or more than either problem individually but less than the sum of the two (a small negative effect). Alternatively, the combined effect may equal the sum of the two individual effects (a zero interaction effect) or be greater than the sum of the individual effects (a positive interaction effect).

40. Koremenos 1999a.

that some actors do not.[41] Thus the conjectures need not be fully consistent with one another in this sense. Similarly, not all conjectures will apply to every case—something we leave to the individual case studies to determine. In the volume's conclusion we discuss the empirical and logical relationships among the conjectures. We now address four broad assumptions that underlie our conjectures.

1. *Rational design:* States and other international actors, acting for self-interested reasons, design institutions purposefully to advance their joint interests.

We thus make standard assumptions: actors have (well-behaved) preferences over various goals; and the pursuit of those goals is guided by their beliefs about each others' preferences and the relative costs and benefits of different outcomes; and actors are constrained by their capabilities.[42] Although the process of institutional design is usually contentious, we do not focus on the bargaining among the participants but on the broad characteristics of the institutional outcomes they select. These outcomes do not simply reflect the preferences of the individual actors but rather represent their joint efforts—and "compromises" among their preferences—to improve their equilibrium outcome given the strategic circumstances they face. That is to say, they concern the equilibrium outcomes that result from the strategic interaction of states, each of which has preferences. Of course, for certain sets of preferences (such as when distributional issues are absent), the strategic aspects of states' interaction are trivial, and institutional design outcomes appear to reflect only preferences.

2. *Shadow of the future:* The value of future gains is strong enough to support a cooperative arrangement.

Actors have a sufficiently high density of interaction—and a sufficiently high discount factor—that cooperation is potentially sustainable. We take a long shadow of the future to be a general condition of contemporary international interdependence, but one subject to considerable variation across issues. On some issues, actors may not interact with sufficient frequency for future incentives to be strong enough to support cooperation by themselves.[43] On other issues, such as peacekeeping, unilateral incentives to defect or distributive differences may make cooperation difficult. A variety of other circumstances—especially uncertainty and large numbers—may make cooperation not only difficult to achieve but also difficult to enforce. Therefore, general international circumstances may be propitious for cooperation, but the particular circumstances in any issue may be problematic.

41. We thank Jim Morrow for this example, which corresponds to a comparison of conjectures M1 and M2.

42. We focus on states as key actors, though most of the analysis can be generalized to nonstate actors.

43. Of course, harsher punishment strategies can be used to support greater cooperation when the shadow of the future is short; however, such strategies are subject to problems of renegotiation proofness. See Downs and Rocke 1995; and Abreu, Pearce, and Stacchetti 1986.

3. *Transaction costs:* Establishing and participating in international institutions is costly.[44]

When creating institutions, states need, for example, to acquire information about the issue, about each other, and about the likely effects of alternative institutional forms. One way they do this is through negotiations. There are other types of transaction costs as well, such as safeguards to ensure compliance and sustain cooperation.[45] As David Lake explains, these safeguards may include sanctions, hostages, and dispute-resolution arrangements.[46]

An important aspect of our independent variables is that they may raise or lower transaction costs. For example, the larger the number of actors, the slower and more cumbersome the negotiations. Likewise, greater uncertainty may make it more costly to write complete contracts to deal with every contingency. Thus, number and uncertainty operate partly through their impact on transaction costs, which is why we separate out such costs in our assumptions. We focus on these variables rather than on transaction costs directly because they are more readily observable.

4. *Risk aversion:* States are risk-averse and worry about possible adverse effects when creating or modifying international institutions.

Risk-averse actors prefer a certain outcome to a chancy one when each has the same expected value. This assumption is the bedrock of modern realism, where states' fears of destruction and keen interest in preserving their sovereignty dominate their strategic calculations. However, even realist states may trade off some sovereignty if they reap large enough gains in return.[47] Institutionalists have a broader view of what states value, but they, too, typically assume states are risk-averse.

With these four assumptions in mind, we now turn to specific conjectures about international institutional design. Because our primary purpose is to generate testable propositions that will guide the empirical analysis of international institutions, we frame the conjectures in a general way.

Each conjecture addresses the expected effect of a change in a particular independent variable, such as the level of uncertainty or the severity of the distribution problem, on one of our dependent variables. Thus our logic is that of comparative statics—that is, we ask how a (perhaps hypothetical) change in an independent variable will affect the equilibrium institutional design. For example, if uncertainty about the state of the world increases, will states design more or less

44. For a general discussion of transaction costs, see Williamson 1985. For an important application to international politics, see Lake 1996. Unlike Williamson, we do not assume that the presence of transaction costs implies bounded rationality. *Transaction costs* refers to the costs of making an agreement and operating it, not of doing what the agreement is designed to do (for example, if two states agree to jointly build a dam, the costs of negotiating and administering the agreement are transactions costs, but the costs of building the dam are not).

45. See Williamson 1985; and Yarbrough and Yarbrough 1992.

46. Lake 1996.

47. Morrow 1991.

flexibility into an international institution? In answering this question, we assume that everything else remains constant. We emphasize the "main effects" of individual independent variables rather than more complicated interactions among them. These simplifying assumptions are necessary given the level of theoretical and empirical generality to which we aspire. After presenting the conjectures we will discuss the limitations of both comparative statics and main effects approaches in terms of design interactions.

Conjectures About Membership

Membership rules determine who benefits from an institution and who pays the costs. They work in several ways beyond simply reducing or enlarging size. By setting criteria for inclusion, for example, they affect the group's homogeneity and asymmetries. Not surprisingly, such rules have important consequences for interactions.

Conjecture M1: RESTRICTIVE MEMBERSHIP INCREASES WITH THE SEVERITY OF THE ENFORCEMENT PROBLEM.

The more severe the enforcement problem, the more restricted the membership. When actors face an enforcement problem (that is, when individuals do not have an incentive to voluntarily contribute to group goals), collective action is problematic. Moreover, the severity of the enforcement problem increases with the number of actors, as Mancur Olson demonstrated.[48] For this reason, Oye argues that reducing multilateral interactions to bilateral ones will increase the incidence of cooperation.[49]

The literature on "club goods" shows that a less drastic reduction in membership may be effective in promoting cooperation among somewhat larger groups.[50] If an institutional arrangement restricts the benefits of cooperation to members, actors have an incentive to pay the price of admission to the club. One of the most important features of institutions is to define these boundaries of membership.[51] Furthermore, when uncertainty about a state's capacity to comply is at issue, inclusive membership may be suboptimal because, as George Downs and David Rocke argue, "every time the third state violates the treaty, the other two states are forced to suspend the cooperation between them to punish it."[52]

48. Olson 1965.
49. Oye 1986. Pahre points out that under strict public good conditions, such restrictions are suboptimal. Pahre 1994. He demonstrates the possibility of large-*n* multilateral cooperation under certain conditions. But unlike conjecture M1, his equilibrium is vulnerable to bad information, and it needs other institutional supports that we discuss under conjectures C1–C3.
50. Buchanan 1965.
51. Snidal 1979.
52. Downs and Rocke 1995, 126.

The effectiveness of membership restrictions depends on the specific characteristics of the issue. In issues like CFC emissions, for example, preventing free riding is virtually impossible. Alliance guarantees, however, are usually effective in restricting nonmembers from receiving security benefits. Enforcement is not always a problem, of course. Agreements on international standards are a good example. Under preference configurations like these, where everyone benefits from wider participation, free riding and enforcement are not issues, and membership tends to be inclusive.

Conjecture M2: RESTRICTIVE MEMBERSHIP INCREASES WITH UNCERTAINTY ABOUT PREFERENCES.

Membership enables states to learn about each others' preferences if the membership mechanism can distinguish cooperators from noncooperators. Ideally, a state that values the goals of an organization will want to join, whereas one that wants a free ride will find it too costly to join a regime they intend to violate. In formal terms, membership is a costly signal. Effective membership rules create a separating equilibrium where only those who share certain characteristics will bear the costs necessary to be included in an equilibrium.[53]

The WTO, for example, requires prospective members to bring key domestic economic rules in line with WTO rules—perhaps with phase-in allowances or special considerations for certain categories of states. Similarly, NATO will not accept a new member until it meets certain domestic political requirements and brings its military up to certain agreed-upon levels. By requiring concessions, these organizations ensure that prospective members are willing to bear the necessary adjustment costs and are likely to be cooperating members down the road. When the price of membership is too low, membership is not informative.

When membership rules are a significant hurdle, they say something significant about nonmembers as well. Refusal to sign the Nuclear Nonproliferation Treaty is a strong and clear signal to other states. Again, it is interesting that states unwilling to commit to this regime generally choose not to sign the treaty rather than to sign but disobey.

Conjecture M3: INCLUSIVE MEMBERSHIP INCREASES WITH THE SEVERITY OF THE DISTRIBUTION PROBLEM.

Realists argue that states care not only about their direct outcomes from cooperative interactions but also how well they fare compared with others.[54] These distributional or relative gains concerns create zero-sum considerations that seri-

53. Spence 1974 illustrates how education provides a costly signal of the quality of prospective employees to employers. Spence 1974. Fearon applies signaling models to crisis bargaining. Fearon 1994. See also Kydd 2000a,b.

54. See Waltz 1979; and Grieco 1988.

ously impair cooperation in bilateral situations. One remedy is to rearrange the terms of cooperation so that benefits are more equally balanced, but this may be difficult or costly. An alternative captured in this conjecture is to expand the number of states involved in the issue because the zero-sum properties are rapidly attenuated as membership increases.[55]

Including additional members may also mediate distributional problems by expanding the possibilities for tradeoffs among the members. Thus an agreement might give state *X* the short end of the stick compared with state *Y* but compensate state *X* with the long end of the stick compared with state *Z* and so forth. This is one advantage of multilateral trade agreements. Such possibilities often occur because new members implicitly increase the range of issues included (for example, tradable products). We deal with these considerations in the next section on issue scope.

Conjectures About Scope

International issues do not come as pre-packaged units. Instead, they are constructed and evolve in complicated ways. While the resulting issue scope partly derives from technological, cognitive-ideational, and other factors that are not analyzed here, rational institutional analysis can explain key patterns of linkage within institutions. We focus on the deliberate choices states make about which issues to include in an institutional framework. In particular, when do states bring together issues they might otherwise have dealt with separately? Our first conjecture follows from efficiency considerations:

Conjecture S1: ISSUE SCOPE INCREASES WITH GREATER HETEROGENEITY AMONG LARGER NUMBERS OF ACTORS.

When states are similarly positioned on an issue, they share common interests over a collective international policy (if any is needed), although they may well have difficulties achieving that policy. Moreover, their relative symmetry on the issue may suggest a focal resolution, especially that all adopt a similar national policy. In these cases an issue often resolves on its own.

As the number of actors increases, however, the heterogeneity within the group will typically also increase. This is especially likely in international settings where the additional actors are often qualitatively different from earlier actors (for example, less-developed countries joining a group of developed countries).[56]

55. Snidal 1991.

56. We do not claim that heterogeneity promotes cooperation; in some cases it promotes distributional differences and conflict. Our position is that linkage provides an institutional means to harness these differences in a mutually beneficial way. Also, having a larger number may promote heterogeneity in capabilities (which we do not address here). For an insightful discussion of these points that also relates heterogeneity to institutional design, see Martin 1994.

When actors have heterogeneous interests, issue linkage may generate new opportunities for resolving conflicts and reaching mutually beneficial arrangements. James K. Sebenius demonstrates how adding issues "can yield joint gains that enhance or create a zone of possible agreement."[57] The paradigmatic example is "gains from trade," both in the limited sense of exchanging commodities and in the broader sense of connecting issues. When one actor values issue *X* more than issue *Y*, and the other ranks them the opposite way, both can be made better off by exchange, that is, by agreeing to defer to each other on these issues. Environmental issues that are important to postindustrial states, for example, are often linked to issues of development and technology when less-developed states with less intrinsic interest in environmental quality are essential to the arrangement.[58]

Conjecture S2: ISSUE SCOPE INCREASES WITH THE SEVERITY OF THE DISTRIBUTION PROBLEM.

Linkage not only allows states to increase efficiency but may also allow them to overcome distributional obstacles.[59] When the benefits of an issue accrue primarily to a few, and the costs fall disproportionately on others, linkage to another issue with different distributional consequences allows cost-bearing states to be compensated by those who reap the gains.[60] When each state cares relatively more about one of two issues, linking the negotiations may be the mutually preferred option.[61] In particular, the more each state cares about "its" issue, the more essential linkage becomes in an agreement. Howard Raiffa makes an even stronger assertion, arguing that increased scope can transform a zero-sum game with no zone of agreement into a positive-sum game.[62]

Conjecture S3: ISSUE SCOPE INCREASES WITH THE SEVERITY OF THE ENFORCEMENT PROBLEM.

57. Sebenius 1983, 314.

58. In some cases, membership may act as a mediating variable through which number affects endogenous variables such as scope. Even in such cases, number may also have direct effects, perhaps due to asymmetries among the parties, for which member is not a mediating variable. This complexity is typical in a system with multiple dependent (or endogenous) and independent (or exogenous) variables. Our conjectures focus on the impact of individual independent variables' main effects and thus hold the other independent variables constant, but not the other dependent variables.

59. Tollison and Willett 1979.

60. Conjectures S1 and S2, though distinct, share a similar logic. In each case differences among the actors lead them to expand the issue set in order to find a better outcome. In this way, distributional differences (which cause conflict within issues) are the engine of efficiency gains (across issues). For an instructive analogy in the social-choice literature on logrolling, see Mueller 1989. Logrolling, however, occurs within an institutional framework and thus can lead to Pareto-inefficient moves. Riker and Brams 1973. We would not expect this in the design of new institutional arrangements.

61. Busch and Koremenos 2001a.

62. Raiffa 1982.

When the incentives on an issue are insufficient for decentralized enforcement, linkage to other issues can provide enforcement.[63] The logic here is the same as in the shadow of the future conjecture, except that this works across issues rather than over time. The United States might be unable to resist domestic pressures to impose tariffs on European wine, for example, were it not for the realization that such action would invite retaliation from the Europeans on U.S. beef. Lutz-Alexander Busch and Barbara Koremenos show formally that the higher the discount rate required to support cooperation (that is, as the enforcement problem is more severe), the greater the probability of issue linkage.[64]

Since all three conjectures point to advantages of greater scope, the question naturally arises, Why isn't everything linked to everything else? The answer is that increased scope also has costs. These include the extra bargaining costs associated with additional issues and the greater probability that some actor will "hold up" the agreement to gain additional benefits.[65] The risk of unraveling, whereby failure in one issue may lead to failure in all linked issues, is also greater. What our conjectures predict is that, all else equal, as the independent variables increase, the marginal benefits of additional scope exceed the marginal costs. This leads rational states to increase scope until the marginal cost of adding another issue roughly equals the marginal benefit.

Conjectures About Centralization

International institutions can be centralized in a variety of ways. An international agency may have centralized information-gathering capacities, for example, without having centralized adjudicative or enforcement capacities. In the conjectures that follow we emphasize general tendencies of centralization rather than specific combinations.

Conjecture C1: CENTRALIZATION INCREASES WITH UNCERTAINTY ABOUT BEHAVIOR.

The Folk theorem holds that when states interact over extended periods they can achieve cooperative outcomes on a decentralized basis through strategies of reciprocity. But when states are uncertain about others' behavior, they cannot achieve the same mutually beneficial outcomes. Greater noise lowers the joint gains they can achieve.[66] Downs and Rocke show how tacit bargaining and trigger strategies can make the best of this situation.[67] However, centralized information may offer a more

63. See Hardin 1982; McGinnis 1986; and Bernheim and Whinston 1990. A more nuanced version of this conjecture would consider the interrelationships among the issues, for example, whether they are substitutes or complements. See Spagnolo 1997.

64. Busch and Koremenos 2001a.

65. Thus our independent variables may affect the costs as well as the benefits of scope.

66. Kreps 1990.

67. Downs and Rocke 1990.

effective alternative if it can reduce uncertainty about behavior to make (otherwise) decentralized cooperation more effective.[68]

The law merchant model illustrates the value of centralization in promoting cooperation when agents are uncertain about one another's past behavior.[69] The law merchant system includes a centralized actor who serves as a repository of information about the past performance of traders. This actor makes the information available to prospective partners, thereby creating a reputational bond that facilitates current transactions. This actor plays a further centralized role in adjudicating disputes and awarding damages as warranted.

Centralized information not only lets states know how others have behaved but also can provide valuable interpretations of that behavior. States will know better whether others' noncooperation is intentional and deserves retaliation or is excusable because of extenuating circumstances. When states retaliate, their targets and third parties will better understand the action as retaliation rather than unilateral noncooperation or error. Under the WTO, for example, retaliation must be centrally authorized, making misinterpretation highly unlikely.

Conjecture C2: CENTRALIZATION INCREASES WITH UNCERTAINTY ABOUT THE STATE OF THE WORLD.

When states are uncertain about the state of the world, all may benefit from joint efforts to gather and pool information. Scientific activity in Antarctica is coordinated, and international economic organizations have substantial research capacities so that states can share the costs of collecting necessary information. In other cases states benefit from collective information sharing but have individual reasons not to share fully or honestly. James Morrow builds on the "cheap talk" literature to show how regimes can structure communication among actors to promote more efficient information sharing in such circumstances.[70]

CONJECTURE C3: CENTRALIZATION INCREASES WITH NUMBER.

As numbers increase, centralized bargaining reduces transaction costs by replacing a large number of bilateral negotiations—or even a cumbersome multilateral negotiation—with an organizational structure that reduces the costs of decision making.[71] Centralization also allows states to coordinate their operational efforts to achieve economies of scale and to ensure that they do not duplicate or work against

68. Axelrod and Keohane 1986.

69. Milgrom, North, and Weingast 1990.

70. See Morrow 1994c; and Farrell and Gibbons 1989. The parallel relationship that centralization increases to resolve uncertainty about other states' preferences or types is also likely to hold. The very willingness to allow centralized inspection by an organization like the IAEA contains useful information about a state's goals even before it generates any information about its behavior.

71. See Keohane 1984; and Martin 1992a.

each other. NATO, for example, provides these advantages through a centralized command structure that allocates tasks.[72]

Centralization of information is also increasingly valuable with larger numbers. Randall Calvert shows how with increasing group size the shadow of the future may not be sufficient to support cooperation.[73] Multilateral communication allows states to achieve decentralized cooperation through an equilibrium where noncooperation is punished by all other states, not just the one that was directly harmed. Because communication is costly, however, this can be substantially improved by a centralized arrangement where a "director" serves as an information clearinghouse. Indeed, the director can even be viewed as "a third-party enforcer . . . [who] in effect pronounces a sentence on the deviant player, a sentence that will then be carried out by rational players."[74]

The International Coffee Organization plays exactly this role in aggregating reports by importing countries on coffee shipments by exporting states.[75] Moreover, because decentralized cooperation typically entails multiple equilibria, centralization is useful in coordinating behavior on an agreeable equilibrium. An important example is standard setting, where intergovernmental organizations (such as the International Telecommunications Union) and private organizations (such as the International Accounting Standards Committee) provide valuable centralized coordination.[76]

Finally, although we are focusing on main effects, there is an interaction between independent variables that supports conjectures C1 and C3. While decentralized cooperation is theoretically possible with large numbers,[77] it becomes much more tenuous when even small levels of uncertainty are introduced. Jonathon Bendor and Dilip Mookherjee show how centralization increases cooperation under such conditions. In their model a central headquarters is effective because it monitors behavior and excludes shirkers from subsequent benefits of the institutional arrangement.[78] Such a centralized arrangement can support higher levels of cooperation than can be supported in any decentralized arrangement.

Conjecture C4: CENTRALIZATION INCREASES WITH THE SEVERITY OF THE ENFORCEMENT PROBLEM.

In the previous conjectures, centralization alleviates cooperation problems created or aggravated by uncertainty and numbers. But enforcement problems also

72. Abbott and Snidal 1998.
73. Calvert 1995.
74. Ibid., 70.
75. See Bates 1997; and Koremenos 1999a.
76. See Genschel 1997; and Abbott and Snidal 2001.
77. Fudenberg and Maskin 1986.
78. Bendor and Mookherjee 1987 and 1997. Bendor and Mookherjee offer a differentiated view of centralization and show how a combination (federalism) of centralized and decentralized arrangements is most effective for the problem they are examining. Ostrom provides evidence of how small levels of centralization can promote otherwise decentralized cooperation. Ostrom 1990.

occur with good information and small numbers. When the payoff from unilateral defection is significantly greater than from mutual cooperation, concern for the future may not guarantee reciprocity-based, self-enforcing cooperation. In such contexts states may find it optimal to delegate power to a third party to adjudicate and enforce mutually beneficial agreements.[79]

Concern for sovereignty, of course, limits the extent to which states will delegate strong coercive capacities to international organizations. But the ability of organizations like the World Bank to withhold resources gives them significant leverage over weaker states. And the informational capacities of international organizations to expose states' behavior can influence the activities of even the most powerful states by imposing international reputational costs or, sometimes, domestic audience costs. Thus states typically obey the findings of WTO dispute-settlement proceedings even though the WTO has no enforcement capacity. Such mechanisms fall far short of coercive enforcement, but they can be valuable in "topping off" the strictly decentralized incentives that support cooperation.

Expanding on Bendor and Mookherjee, Edward Schwartz and Michael Tomz show how centralized arrangements have significant advantages if the central authority has the ability to expel shirkers from the group. High levels of monitoring will encourage contributions from all actors because shirkers are too likely to be detected and expelled and the value of remaining in the group will increase.[80]

Even centralized institutions that have no enforcement or even adjudicative capacities may be effective in resolving enforcement problems. Eric Posner shows that even if courts are "radically incompetent" in determining fault—that is, they can determine only whether a legal agreement existed but cannot verify whether actors obeyed it—formalized agreements can create reputational incentives that enable parties to solve commitment problems.[81] The reason is that the incentive for each party to cheat is reduced by the increased reputational costs of the breakdown of the agreement regardless of who is at fault. In a similar vein Lisa Martin shows that international organizations are instrumental in maintaining support for sanctions partly because states do not want to undermine the other benefits provided through these organizations.[82]

Finally, modest international centralization is sometimes effective because it harnesses domestic enforcement capacities. The 1998 OECD Anti-Bribery Convention relies on domestic legislation for implementation and on domestic court systems for enforcement, but a centralized inspection system ensures that states

79. Using similar logic, Lake argues that "the probability that the partner will engage in opportunistic behavior decreases with relational hierarchy." Lake 1996, 14. In other words, as the expected costs of opportunism increase, hierarchy will be the preferred governance structure.

80. Schwartz and Tomz show that the value of centralization does not always increase monotonically with the *capacity* of the central agent. Schwartz and Tomz 1997. In their model, an intermediate level of monitoring means that some shirking will occur so that less talented actors are detected and excluded from the group.

81. Posner 1999.

82. Martin 1992b.

police their own firms. This reinforces the point that centralization does not require international agents to have an independent coercive capacity to effectively promote cooperation.

Despite the advantages of centralization captured in the conjectures, states retain deep-seated concerns, intensified by their risk aversion, about how international institutions might behave. Will resources be squandered in bureaucratic excess? Even more important, will international agencies expand their authority over time? Consequently, states view centralization warily, and its overall baseline level may remain quite low. Our conjectures only express conditions under which states will increase (or decrease) centralization in response to their environment. For the same reasons, states also are concerned about maintaining tight control over the institutional arrangements, as indicated in the next set of conjectures.

Conjectures About Control

Two conjectures are relevant to the rules chosen to govern institutions:

Conjecture V1: INDIVIDUAL CONTROL DECREASES AS NUMBER INCREASES.

Conjecture V2: ASYMMETRY OF CONTROL INCREASES WITH ASYMMETRY AMONG CONTRIBUTORS (NUMBER).

The first conjecture seems obvious: as the number of actors increases, the control of any one actor or subgroup of actors decreases.[83] For example, as the EU has expanded, the leverage of individual members has steadily decreased.[84] This is because when the number of actors is large, states must sacrifice individual control to achieve collective benefits. Each state may be adversely affected on occasion, and without the veto a state has no unilateral protection—although its ability to withdraw from the institution ultimately limits its vulnerability. States agree to such a scheme because they benefit from others' inability to veto and strategically block group decisions. An important example is the EU's move toward "qualified majority" voting as membership has expanded.[85]

This conjecture follows directly from the social choice literature on voting rules. Brian Barry, for example, shows that for issues that are recurrent and symmetric in

83. *Number* here refers to members of the institution who are eligible to have a say in its operations. This is a good example of our earlier observation that a prior institutional decision may be treated as exogenous in considering the adoption of other rules. Alternatively, membership and control rules may be determined together such that, for example, a decision to have a large membership is compatible with one set of control rules, and a decision to have a small membership is compatible with another set of control rules.

84. Hosli 1993.

85. A more sophisticated analysis would also consider the policy preferences of governments. Garrett and Tsebelis show how this leads to a consideration of a broader set of control institutions (for example, the Commission and the Council of Ministers) and to rules regarding other forms of control, such as agenda setting. Garrett and Tsebelis 1996.

several senses, majority voting maximizes expected utility.[86] Similarly, the conjecture is supported by analogy to the theory of the core and noncooperative solution concepts, where increased power to subgroups (such as through vetoes) leads to paralysis by eliminating mutually agreeable outcomes.[87]

The second conjecture follows from an intuition that an actor's control over an institution relates to the actor's importance to the institution. This corresponds to cooperative game-theoretic solution concepts such as the Shapley value, which relates what an actor (potentially) brings to different coalitions to the pay-off the actor receives. When some states contribute more to an institution than others—perhaps because they pay more dues or their behavior is vital to the institution's success—they will demand more sway over the institution. Other states will grant this control to ensure their participation—as the UN did to the permanent members of the Security Council, whose military and financial support was considered essential to the enforcement of resolutions.[88] Membership and voting rules typically formalize this control in some way, as is the case in the UN Security Council and in the weighted voting in the IMF.

Conjecture V3: INDIVIDUAL CONTROL (TO BLOCK UNDESIRABLE OUTCOMES) INCREASES WITH UNCERTAINTY ABOUT THE STATE OF THE WORLD.

Because states are risk-averse, they design institutions that protect them from unforeseen circumstances. Veto power is a standard design feature that provides such protection, either to individual states or, in the case of super-majority requirements, to groups of states. A parallel in U.S. politics is the institutional norm of universalism, where legislators place a project in every member's district rather than risk being excluded from a (minimum winning) majority program.[89] The "theoretical engine" behind the universalistic result is uncertainty and legislators' risk aversion.[90]

Conjectures C2 and V3 illustrate quite different institutional responses to the problem of uncertainty. For example, centralization of information can be increased to remedy uncertainty about the state of the world, with the level of control unaffected. Or super-majority voting may mitigate uncertainty about the state of the world without changing the level of centralization. In short, control and centralization can be varied independently or together to deal with uncertainty.

86. Barry 1979. See also the Rae-Taylor theorem in Rae 1969; and Taylor 1969. Mueller provides an excellent overview of the issues and a comparison of majority/unanimity rules. Mueller 1989. Buchanan and Tullock argue for the virtues of unanimity in promoting efficient outcomes when there are no transaction costs. Buchanan and Tullock 1962. As decision-making costs increase—including the costs of preference revelation (which corresponds to uncertainty about preferences)—the case for smaller majorities grows.

87. Shubik 1982.

88. Winter 1996.

89. Weingast 1979.

90. Collie 1988.

Other institutional arrangements provide different forms of protection against uncertainty. Escape clauses in effect allow a state to "veto" some institutional dictates only for themselves. Withdrawal clauses allow the more dramatic step of leaving an institution entirely to avoid undesired outcomes. Such control features blend into what we call flexibility.[91]

Conjectures About Flexibility

Uncertainty about the current or future state of the world presents states with a dilemma. Becoming locked into an institution may lead to unanticipated costs or adverse distributional consequences. But by not making a bargain, states might pass up significant benefits from cooperation.

If uncertainty is high and anticipated benefits are low, risk-averse states will avoid committing themselves to rigid institutions. But what if the uncertainty is lower and the potential benefits are higher? Under these more benign conditions, institutional flexibility becomes important. The possibility of adjusting the agreement when adverse shocks occur allows states to gain from cooperation without tying themselves to an arrangement that may become undesirable as conditions change.[92]

Conjecture F1: FLEXIBILITY INCREASES WITH UNCERTAINTY ABOUT THE STATE OF THE WORLD.

Similarly, states may be uncertain about the distributional implications of particular aspects of an agreement. Koremenos develops a model where states plan to renegotiate all or part of an agreement once they have learned from experience which states benefit the most.[93] The desirability of renegotiation (versus a single, longer agreement) increases with uncertainty about the distribution of gains and decreases with the degree of "noise" in the environment from which the effects of the agreement must be distinguished. An example is the Antarctic Treaty. Although it has no expiration date, the treaty was designed to allow states to learn from their experience and modify the agreement over time. One procedure for modification operated during the first thirty years, another during the subsequent period. In the first learning phase, the parties met biannually for consultations, and the agreement could be changed only by unanimous consent. Some changes and extensions were made, such as the follow-on arrangement to ban resource extraction. Now that the initial period has ended, individual states can press for renegotiation, this time under

91. We proposed but later dropped the related conjecture that "individual control (to block undesirable outcomes) increases with the severity of the distributional problem" because it was logically equivalent to conjecture V3. The impact of distribution flowed fundamentally from *uncertainty* about the distribution rather than from *known* distributional consequences, which could be dealt with in other institutional ways. The deleted conjecture was strongly supported in the empirical studies, so dropping it does not bias the results in our favor.

92. Downs and Rocke 1995.

93. Koremenos 2001.

majority rule. They do so with more certainty about how the agreement operates and a better understanding of its costs and benefits.[94]

Flexibility need not be so formalized. For example, "soft" international law allows states to respond to uncertainty by designing arrangements that are less formalized than full legalization. Although often seen as a "failure" of international law, soft law may represent a superior institutional adaptation because of its flexibility.[95]

Even when states face no uncertainty about proposed agreements, flexibility may resolve distributional problems:

Conjecture F2: FLEXIBILITY INCREASES WITH THE SEVERITY OF THE DISTRIBUTION PROBLEM.

Fearon argues that when states lengthen the shadow of the future to solve enforcement problems, distributional concerns become increasingly severe. States bargain harder because the results will affect them for a longer period.[96] Koremenos suggests that in this case states may reduce distributional problems, and bargaining costs, by adopting a more flexible agreement structure.[97] Busch and Koremenos show that under certain conditions, a series of shorter agreements still embodies the shadow of the future required for enforcement while avoiding the bargaining costs associated with a single, long agreement in Fearon's model.[98]

Flexibility has a downside. Renegotiation of treaty terms, as well as dealing with unilateral invocations of flexibility such as escape clauses, is costly. Moreover, individual states have incentives to free ride on an agreement by developing self-serving interpretations of escape clauses that are broader than intended. And renegotiation provides an opportunity for states to "hold up" the cooperative bargain in an effort to increase their own share. Such incentives become greater as more states are party to an agreement—for the familiar reasons associated with collective action.[99] Even without these strategic considerations, as more states become involved, modification becomes more difficult and time consuming. This reasoning leads to our final conjecture.

Conjecture F3: FLEXIBILITY DECREASES WITH NUMBER.

All else equal, states will introduce less flexibility into institutions with larger numbers because larger numbers increase the costs associated with flexibility more than they increase its benefits. For example, where flexibility takes the form of

94. This kind of flexibility also solved important distributional issues, the subject of conjecture F2.
95. Abbott and Snidal 2000.
96. Fearon 1998.
97. Koremenos 2001.
98. Busch and Koremenos 2001b.
99. Hardin 1992.

periodic renegotiation of the agreement, larger numbers will increase the associated bargaining costs. Koremenos shows formally that as renegotiation costs increase, rational parties to an agreement will renegotiate less often or not at all.[100] Thus commodity agreements involving forty or so countries are renegotiated significantly less often than are monetary agreements involving the G-7. As renegotiation costs rise, other forms of flexibility become *relatively* less expensive. For example, states may switch to more centralized forms of flexibility, such as escape clauses combined with a centralized monitoring institution to keep the moral hazard problem in check or the creation of a quasi-legislative institution empowered to adjust the terms of an agreement.[101] Such changes are consistent with conjecture C3, that centralization increases with number, which brings up the question of design interactions. Finally, note that for some types of flexibility, such as withdrawal clauses, the effects of number on the form or incidence of the provisions may be minimal.

Design Interactions

Our simple research design has considerable advantages, but it also has limitations. Because our definitions are broad, they encompass significant institutional variation. The best example is centralization, which includes everything from rudimentary forums for bargaining, through information and monitoring functions, to centralized adjudication and enforcement. Such general conceptions are essential for assessing similarities across cases, but finer conceptual distinctions are needed to understand the more detailed workings and differences among institutions. The volume's contributors begin to do precisely that in the empirical studies that follow.

Our bivariate relationships cannot capture more complex interactions among the variables. For example, while both large numbers and increased uncertainty promote centralization, the interaction of their effects may be most significant of all. The most interesting complexities are those that (may) arise because the dependent variables interact among themselves—as "substitutes," "complements," or "conflicts." Institutional features may substitute for one another by offering alternative ways to solve a particular problem. Escape clauses, for example, introduce flexibility to allow hard-pressed states to avoid the full burden of their treaty obligations on a decentralized basis. An alternative arrangement would be to require states facing special difficulties to seek relief from a centralized institution that can decide how rules apply to new situations. Thus institutional design can enable choice among different means toward the same ends—that is, a choice among multiple institutional equilibria.

Design features may also complement one another. Membership rules, for example, provide one means to deal with enforcement problems (conjecture M1),

100. Koremenos 1999a.
101. For a theoretical analysis with corresponding empirical support, see Koremenos 2000.

but these can be enhanced by centralization when incentives to defect are especially large. Centralization may work either directly as a separate source of enforcement capacity (conjecture C4) or interactively in making the membership mechanism more effective by providing information on members' performance.[102]

Design principles may conflict with one another. Consider an issue with both distribution and enforcement problems. When enforcement is problematic, membership needs to be restricted (conjecture M1), but when there are distributional problems, it needs to be more inclusive (conjecture M3). Obviously, membership rules cannot remedy both problems simultaneously. The only way to circumvent this conflict is to move to a more complex design (such as addressing the enforcement problem with membership rules and the distribution problem by increasing scope).[103] Our bivariate analysis cannot fully capture such complex interactions.[104]

Finally, our analysis looks at individual institutional arrangements in isolation. Substitutabilities, complementarities, and conflicts arise not only in the design of individual institutions but also in *relationships among them.* Just as individual features of institutions can complement each other, so too can different institutions. One way is by vertical nesting, where institutions that deal with one issue or region are situated within a larger global institution. Vinod Aggarwal has analyzed exactly this kind of relationship between GATT and various textile arrangements.[105] Likewise, the policymakers who planned NAFTA made sure it conformed to GATT trading rules, an issue that will remain important as both NAFTA and the WTO evolve.

We have embraced these challenges by asking the authors of the empirical studies to *begin* from our concepts and conjectures. We also asked them to be critical of the concepts and on the lookout for ways to refine and improve the conjectures. The ultimate value of our conjectures lies less with their individual veracity than with whether they spur our collective effort to systematize and refine our knowledge of institutional design.

Roadmap to the Rational Design Project

The wide range of conjectures (summarized in Table 1) represents our effort to understand the design of international institutions from a rationalist perspective. The ultimate value of our framework depends on its ability to explain phenomena across a range of substantive issues. The articles that follow take up this challenge by

102. The choice among alternatives may also depend on interactions with other independent variables. Thus, the WTO's move toward more centralized dispute resolution was related to the large number of states involved.

103. This problem has been central to the analysis of macroeconomic policy in open economies, especially the relationship between the number of policy goals and the number of policy instruments. Mundell 1962.

104. This problem would bias the empirical results *against* our bivariate conjectures.

105. Aggarwal 1985.

TABLE 1. *Summary of Rational Design conjectures*

M1:	Restrictive MEMBERSHIP increases with the severity of the ENFORCEMENT problem
M2:	Restrictive MEMBERSHIP increases with UNCERTAINTY ABOUT PREFERENCES
M3:	MEMBERSHIP increases with the severity of the DISTRIBUTION problem
S1:	SCOPE increases with NUMBER
S2:	SCOPE increases with the severity of the DISTRIBUTION problem
S3:	SCOPE increases with the severity of the ENFORCEMENT problem
C1:	CENTRALIZATION increases with UNCERTAINTY ABOUT BEHAVIOR
C2:	CENTRALIZATION increases with UNCERTAINTY ABOUT THE STATE OF THE WORLD
C3:	CENTRALIZATION increases with NUMBER
C4:	CENTRALIZATION increases with the severity of the ENFORCEMENT problem
V1:	CONTROL decreases with NUMBER
V2:	Asymmetry of CONTROL increases with asymmetry of contributors (NUMBER)
V3:	CONTROL increases with UNCERTAINTY ABOUT THE STATE OF THE WORLD
F1:	FLEXIBILITY increases with UNCERTAINTY ABOUT THE STATE OF THE WORLD
F2:	FLEXIBILITY increases with the severity of the DISTRIBUTION problem
F3:	FLEXIBILITY decreases with NUMBER

evaluating our conjectures in the context of many different areas of international politics.

The empirical articles all share our rationalist approach, taken broadly, but they vary widely in other respects. The institutions examined cover the full spectrum of international politics, from environmental protection to national security. Some institutions are highly articulated organizations; others are much more informal arrangements. The cases exhibit considerable variation in key institutional dimensions, such as centralization of information or breadth of membership.

We have deliberately included methodological diversity. Case studies and quantitative approaches are represented. Some analysts develop our conjectures further by using a formal deductive approach to explain the design of institutions that affect specific issues; others use a more inductive and empirical approach to evaluate and extend the theoretical framework. While most of the studies treat states or international organizations as their central actors, others focus on private international actors, such as firms and private courts, or relax the unitary actor assumption to incorporate key domestic political factors. Most of the studies treat institutional design as a deliberate rational choice; one, however, focuses on "indirect" rational design driven by actors' selection among available institutional alternatives. The first three articles develop the theory in specific contexts and enrich it by connecting it to specific empirical cases. The next five articles use the theory as the basis for intensive empirical analysis of a specific issue-area.

Andrew Kydd looks at NATO enlargement and investigates the causes and consequences of NATO's membership criteria. NATO enlargement has built trust among the potential entrants but weakened it between NATO and Russia. The membership criteria are fairly restrictive: new members must have firmly entrenched democracies, civilian control of the military, and no ethnic or border

disputes with their neighbors. These restrictive criteria build trust among new members by diminishing uncertainty about their preferences; they also mitigate the distrust generated in Russia, by showing that NATO is not just expanding willy-nilly to include any state that wants to join.

Peter Rosendorff and Helen Milner look at one of the most common and controversial features of trade agreements: escape clauses. This design feature allows states to enter into agreements they might not otherwise accept because of unforeseeable contingencies. But escape clauses must be costly, or else countries might use them cynically to abandon agreements that are merely inconvenient. Rosendorff and Milner develop a formal model that shows how states design escape clauses to balance these considerations and facilitate agreement.

Robert Pahre asks why states often "cluster" negotiations with multiple states at the same time. He develops a model of clustering, which he tests on nineteenth-century trade relations. But his analysis is equally insightful for understanding the use of negotiating rounds in the postwar GATT/WTO. Clustering occurs in other issue areas as well. It is especially important when states are committed to most-favored-nation policies because these exacerbate distributional problems by linking every bilateral trade negotiation to every other negotiation. Clustering is important because it helps states resolve these distributional problems.

Ronald Mitchell and Patricia Keilbach use their study of environmental issues to investigate institutional design when asymmetric relationships exist among actors. Sometimes "upstream" states create pollution, and "downstream" states are its victims. Polluters have no incentive to join an institution to reduce pollutants unless the institution's scope includes issues they might benefit from. Asymmetry occurs in another way as well. Polluting states can be stronger or weaker than the victims. Mitchell and Keilbach show that weak victims seek institutional designs with positive linkages or rewards, whereas strong victims prefer negative linkage or sanctions.

Walter Mattli highlights the growth of private institutions to arbitrate international business disputes. Private tribunals are often faster, more discreet, and less expensive than public courts. They can be designed to focus closely on specific commercial practices within an industry, a kind of expertise courts rarely possess. The demand for arbitration has been so strong that business groups have produced a multitude of arbitration tribunals. The strengths and weaknesses of different designs lead business partners to select a tribunal to handle disputes as part of commercial contracts. Their choice, Mattli argues, depends on the number of parties involved and their uncertainty about the future state of the world and each other's behavior.

Thomas Oatley deals with a very public institution, the system of multilateral trade and payments for Europe's postwar reconstruction. Two major design problems faced Europeans. One was distributional: who would bear the costs of adjustment to trade imbalances? The second was hard-currency reserves. The United States was willing to provide dollars through the Marshall Plan but feared it might lead to bloated debts rather than disciplined development. Oatley shows how the

payments union begun in 1950 resolved these issues with a series of interrelated design features: centralized trade and credit balances, flexible administration, and relatively weak enforcement.

When fighting breaks out, enemy soldiers are frequently seized as prisoners of war. States have joint treaties to ensure that prisoners are treated humanely and modify them to cope with new types of war and imprisonment. James Morrow notes that a workable treaty design must affect the behavior of front-line troops who actually capture prisoners; twentieth-century treaties are designed with that in mind. Moreover, because these treaties entail some costs, ratifying them sends signals about national intentions. Standards for treatment are generally straightforward, partly to make them easily understood by soldiers, partly to resolve any wrangling over the distribution of burdens.

John Richards deals with the institutional design of the global aviation regime. States had to decide whether markets or regulation would govern air routes and fares. Their choice of regulation was prompted by national security concerns, which were closely tied to aeronautics and to states' desire to promote high-technology industries at home. Once on the regulatory path, states faced the complicated task of building effective international institutions. Richards shows how the regulatory institutions that emerged were profoundly shaped by the particular features of the industry, including the large number of states involved and their uncertainty about one another's behavior and future conditions.

The volume concludes with two articles. We invited Alexander Wendt to comment on the project from an "external" perspective. Wendt is both sympathetic to our enterprise and skeptical of it. He questions our decision to focus on rational choice explanations without directly engaging either competing approaches or what he believes are complementary but "deeper" explanations. Wendt further argues that our analysis is insufficiently "forward looking" to address important normative concerns. While we do not fully agree with Wendt's critique, his article provides insight for both insiders and outsiders about the limitations of our approach.

In the final article we summarize the findings. We also combine internal and external critiques of what the volume has accomplished and consider how our rationalist approach can be improved by addressing questions raised by alternative perspectives.