

---

# Situation Structure and Institutional Design: Reciprocity, Coercion, and Exchange

Ronald B. Mitchell and Patricia M. Keilbach

---

States create international institutions in attempts to resolve problems they cannot solve alone. Yet states vary in their desire to form and join such institutions and in their incentives to defect from those they do join. These obstacles to cooperation have produced considerable variation in the mechanisms institutions use to deter defection without deterring participation. Some rely on narrow issue-specific reciprocity, whereas others rely on broader linkages involving coercive sanctions or positive rewards. This diversity in institutional scope is neither meaningless variation nor simple experimentation. Instead, states tend to base institutions on issue-specific reciprocity when possible but incorporate positive or negative linkage to other issue-areas when the distribution and enforcement problems within an issue-area appear more severe.

In an interdependent world one state's behavior often imposes unintended costs on other states. Yet, though all such negative externalities create demands for their resolution, all externalities are not alike.<sup>1</sup> Some are symmetric, with all states being simultaneously victims and perpetrators. Others are asymmetric, with "downstream" states being victims of, or dissatisfied with, the externality and "upstream" states being perpetrators of it.<sup>2</sup> Dissatisfied states may accept both types of externalities,

We thank Frank Alcock, Jeffrey Berejikian, Thomas Bernauer, Liliana Botcheva-Andonova, Robert Darst, Walter P. Falcon, James Fearon, Robert O. Keohane, Barbara Koremenos, Charles Lipson, Lisa Martin, James Morrow, Rosamond Naylor, Thomas Oatley, Kenneth Oye, John Richards, Duncan Snidal, Alexander Thompson, Michael Zürn, and the other contributors to this volume for helpful comments on earlier drafts of this article. Insightful suggestions by David Lake, Peter Gourevitch, and two anonymous reviewers proved particularly helpful in refining the argument. Valuable research assistance was provided by Hannah Fairbank, Elyce Hues, Aaron Knott, and Jonah Spiegelman. Ronald Mitchell thanks the University of Oregon's Department of Political Science and Stanford University's Center for Environmental Science and Policy for research support that contributed to completion of this article. Patricia Keilbach thanks the University of Oregon's Department of Political Science for research support that contributed to completion of this article.

1. Milner 1997, 44.

2. See the discussion of the number of actors variable, NUMBER, in Koremenos, Lipson, and Snidal, this volume, 777–78.

or they may try to resolve them by force. But they often create international institutions to resolve them.<sup>3</sup>

In symmetric externalities the fact that all states prefer mutual cooperation to the status quo predisposes states toward narrow institutions that rely on issue-specific reciprocity. Although coercion or side payments could also be used to combat incentives to defect, such linkage is usually unnecessary. Asymmetric externalities, however, present more severe distribution and enforcement problems.<sup>4</sup> An institution limited to the single issue of an asymmetric externality would provide benefits only to victims and impose costs only on perpetrators. To create incentive-compatible institutions in the face of such distributional problems, states dissatisfied with the status quo must broaden institutional scope, using the linkage of incentives or coercion to convince perpetrators to join the institution (Rational Design conjecture S2, SCOPE increases with DISTRIBUTION problems).<sup>5</sup> Victims also realize that membership restrictions and the retaliatory noncompliance of reciprocity—which can support enforcement in symmetric settings (conjecture M1, restrictive MEMBERSHIP increases with ENFORCEMENT problems)—are either unavailable to them or ineffective at inducing upstream states to join and comply. Thus enforcement problems reinforce the tendency in asymmetric externalities to broaden institutional scope, as victims incorporate linkages to ensure perpetrators comply once they join (conjecture S3, SCOPE increases with ENFORCEMENT problems). Whether scope is broadened by coercion or exchange depends on the power of the downstream state. For weak downstream states, the exchange of side payments for cooperation is the only available means of engaging stronger perpetrators in resolving the problem. Downstream states that are stronger than the perpetrators may employ such positive linkage but also can use negative linkage to coerce perpetrators into mitigating an externality and can do so without the aid of an institution. The distribution of costs and benefits in asymmetric externalities makes actors reluctant to join an institution and encourages them to violate institutional rules if they do, suggesting that distribution problems are not always separable from, and indeed sometimes drive, enforcement problems.<sup>6</sup>

While our analysis strongly validates two of the three Rational Design conjectures regarding scope (conjecture S2, SCOPE increases with DISTRIBUTION problems; and conjecture S3, SCOPE increases with ENFORCEMENT problems), it directs attention to the design interactions mentioned in Barbara Koremenos, Charles Lipson, and Duncan Snidal's introduction to this volume, suggesting in particular that restricting membership and increasing institutional scope (conjecture M1, restrictive MEMBERSHIP increases with ENFORCEMENT problems; and S3, SCOPE increases with ENFORCEMENT problems) can serve as substitutes for enforcement, with the former more

3. Milner 1997, 8–9.

4. Asymmetries are discussed as part of the NUMBER variable in Koremenos, Lipson, and Snidal, this volume, 778.

5. Koremenos, Lipson, and Snidal, this volume.

6. On interactions among variables, see the comments by Koremenos, Lipson, and Snidal, this volume, 779–80.

likely in symmetric settings and the latter more likely in asymmetric ones.<sup>7</sup> Our analysis also sheds light on two other Rational Design conjectures. We provide some support for conjecture C4, CENTRALIZATION increases with ENFORCEMENT problems. Victims of an asymmetric externality realize that perpetrators have no incentives to cooperate unless compensated. Centralized and explicit coordination of compensation reassures each victim that (and how much) other victims will contribute and, by pooling resources, increases compensation to the perpetrators. We also refine the role of flexibility in institutional design. In many cases, increasing flexibility allows states to agree on some institutional form that provides benefits to all parties despite the remaining bargaining problems resulting from distributional problems or uncertainty. In asymmetric externalities, however, fostering institutional creation depends on limiting flexibility to reassure both sides that the carefully negotiated terms of agreement will not be subject to later renegotiation or reinterpretation. In such contexts both victims and perpetrators want to limit the flexibility of the other side to prevent it from separating the two distributional problems that, when linked, make the institution worthwhile.

## Definitions and Clarifications

Although we recognize that states can regulate behavior through informal social practices, we follow Koremenos, Lipson, and Snidal's narrower definition of international institutions: explicit arrangements, such as treaties and conventions, that regulate behavior. We also use the term *regime* interchangeably with *institution*. We refer to states that initiate regime formation as either "dissatisfied" or "victim" states and to those whose behavior they seek to influence as "externality-generating" or "perpetrating" states to capture the fact that all the regimes we analyze attempt to mitigate externalities resulting from at least one state's behavior. These terms refer to a state's role in a problem rather than to inherent characteristics of a state. Indeed, symmetric situations are defined as circumstances in which states that generate an externality are also dissatisfied with the status quo. But these terms let us capture situations in which victims are distinct from perpetrators.

To simplify the situations states face, we use a model composed of two groups of states addressing a single issue. The increased clarity gained by this approach must be balanced against the decreased accuracy in depicting the problems states face and the ways states resolve them.<sup>8</sup> Focusing on a single issue constrains our ability to generalize, since states sometimes try to create institutions to address multiple, logically linked issues. In those cases the interplay of power, interests, and other factors that vary across issue-areas can produce complex dynamics not captured in the model. Yet frequently issues are not connected by "objective" necessity but by

7. *Ibid.*, 796.

8. For a similar simplification of international economic interactions, see Oye 1992.

states finding it compelling or in their interests to link them. Indeed, it is precisely the endogeneity of institutional scope—that states choose to make (or not make) certain linkages rather than take linkages as givens—that interests us here. Even when states seek to create institutions to resolve problems within a single area of behavior,<sup>9</sup> as seems particularly common in environmental affairs,<sup>10</sup> they still face choices of institutional scope, choices we seek to explain through our model.

Limiting the model to two actors also has virtues and costs. In single-issue contexts, it seems possible and appropriate to categorize all states as falling into one of two groups: those who prefer to reduce the externality and those who do not. If no states prefer a reduction, then no “problem” exists and no institution will be created. When some prefer a reduction, costs and benefits will produce differences in the strength of preferences, and some may prefer to free-ride on others’ contributions without contributing themselves. A two-actor or two-group view does not, however, remove collective-action problems even among those who dislike the externality, and we consciously examine how conflicts both between groups and within groups influence institutional design.

Finally, we seek to explain institutional design, not compliance or effectiveness. The incentives to defect from an institution’s rules influence the rules states adopt, even as those rules, once adopted, influence the propensity to defect and hence compliance and effectiveness. We hope that by highlighting how incentives to defect influence institutional design, and not just compliance, we will encourage scholars concerned with effectiveness to control more explicitly for, rather than assume that, different regime designs were adopted in equivalent circumstances.<sup>11</sup> In short we argue that differences in design reflect differences in strategic structure.

## Enforcement and Distribution Problems Across Situation Structure

States have incentives to establish international institutions whenever doing so offers improvements over the status quo. Often these incentives are the response of states seeking to mitigate some other state’s externality.<sup>12</sup> Variation in these externalities affects regime design.<sup>13</sup> This point is commonly illustrated by noting that regimes addressing collaboration situations have carefully designed compliance

9. Limiting negotiations to a single issue helps states avoid the potentially debilitating complexity of linkages to other issues. See Sebenius 1983; and McGinnis 1986.

10. Thus states have created separate regimes for acid precipitation, climate change, and stratospheric ozone loss and for biodiversity, endangered species, deforestation, and desertification rather than broader ones covering atmospheric or wildlife issues, respectively. Many environmentalists criticize this tendency to compartmentalize as ineffective in resolving environmental problems that are fundamentally integrated. Esty 1994.

11. This insight has not been fully incorporated in the debate between the “managerial” and “enforcement” schools. See Chayes and Chayes 1995; and Downs, Rocke, and Barsboom 1996.

12. See Oye 1992, 17; and Milner 1997, 8–9.

13. Mitchell 1999.

mechanisms while those addressing coordination situations do not because the former pose more severe enforcement problems.<sup>14</sup> Yet the enforcement problems in collaboration situations are still not as severe as those in the asymmetric externalities addressed by many environmental regimes. Although international institutions addressing symmetric externalities exhibit some design variation, more fundamental differences exist between these and institutions designed to address asymmetric externalities.

Much international relations literature focuses on the symmetric and reciprocal externalities analogized as Prisoners' Dilemmas or Tragedies of the Commons. Consider states sharing a lake, each polluting it while also using it for drinking water.<sup>15</sup> These states are symmetric both in their capacities to cause the problem—each generates an externality that harms others—and in their preferences for alleviation of the problem—each feels victimized by the corresponding externalities of others. Each believes its benefits from others' halting those activities are greater than its costs of halting its own such activities. Cooperation nevertheless remains elusive because each state prefers unilateral defection even more than collective mitigation of the externality. Although the payoffs of mutual cooperation are rarely evenly distributed, we consider these situations "symmetric" in the sense that all believe they would benefit from mitigation of the externality, and all can either exacerbate or mitigate the problem by engaging or not engaging in the behavior generating the externality.

Consider states sharing a river, each polluting it but also using it for drinking water, with some situated upstream from the rest. These states are asymmetric in two respects. They are asymmetric in their capacities to cause the problem (upstream states generate externalities and downstream states do not) and in their preferences for solving the problem (downstream states prefer alleviation and upstream states do not). Although often assumed otherwise, states do create regimes to address such asymmetric externalities in which "some actors obtain their most preferred outcome while others are left aggrieved."<sup>16</sup> With symmetric externalities, actors differ in their preferences for alternative institutions, but all prefer some institution to none. With asymmetric externalities, perpetrating actors prefer no institution because, absent compensation, they would bear the institutional costs but receive no institutional benefits.<sup>17</sup> Such distributional asymmetries can arise from material conditions or from states having different values, with some preferring that all undertake externality-mitigating acts and others preferring the status quo.<sup>18</sup> States can have downstream-type preferences if they do not engage in the externality-generating behavior, mitigate the externalities of such behaviors for indepen-

14. See Stein 1983; Krasner 1991; and Martin 1992a.

15. Waltz uses a similar analogy. Waltz 1979, 196.

16. The quote is from Stein 1983, 120. On such externalities, see Coase 1960; Conybear 1980; Oye 1992; and Bernauer and Ruloff 1999.

17. Koremenos, Lipson, and Snidal, this volume.

18. Asymmetric externalities have received scholarly attention only recently but do not appear particularly rare empirically. See Rittberger and Zürn 1990, 31–32; and Martin 1992a,b.

dent reasons, or engage in externality-generating behaviors that do not materially affect other states or have material effects to which other states are indifferent. Even when all states actually prefer mutual cooperation to the status quo, some may dissemble if, by doing so, they can extract concessions from other states.<sup>19</sup> Whether reflecting true or strategically manipulated preferences, asymmetric situations reflect the willingness of some states to act as if they were pure perpetrators and not victims, indifferent to whether others reduce some externality. Asymmetric situations create greater enforcement problems precisely because they involve unidirectional dependence rather than reciprocal interdependence.

Power has a role in symmetric settings (as we elaborate later), but it particularly influences institutional design in asymmetric settings. Although notoriously difficult to assess, the distribution of power among states nonetheless influences the likelihood and shape of the institutions states create. Downstream states, whether weak or strong, have incentives to try to induce an upstream polluter, or other externality generator, to desist. We define strong states as those that possess resources (such as military might or a strong trade relationship in a crucial good) that can be used to impose costs on others for undesirable behavior. Although weak states lack resources to coerce or compel by imposing costs, they may be able to persuade or induce behavioral changes by using other resources as rewards (such as technological or financial aid).<sup>20</sup> A state's power is thus relational and issue-specific, dependent on how committed it is to achieving its own goals and on how vulnerable and sensitive other countries are to the resources it controls.<sup>21</sup> Weak states cannot get other states to "do something against their will" but may be able to get them to be "willing to do something."<sup>22</sup> Weak states may simply suffer the harms imposed by upstream states. Indeed, if an externality imposes costs on several states but the costs of inducing the perpetrator to desist are greater than the benefits to any single state of getting it to do so, collective-action problems will almost ensure the externality continues. However, weak states sometimes find ways to overcome these problems or suffer harms sufficiently large that unilateral action becomes worthwhile. They may then seek to design institutions that look much different from those initiated by strong victims.

## Institutional Design Options

Both symmetric and asymmetric situations give at least some states ongoing incentives to defect. To counter these incentives, states design institutional mech-

19. We are indebted to James Morrow for clarifying this point.

20. See Knorr 1975, 310–19; Baldwin 1979, 184; and Morgenthau 1993, 31.

21. Keohane and Nye 1989.

22. This distinction coincides with the "paradox of unrealized power," that whether a state can convert its control over resources into influence over outcomes depends on how it deploys those resources. Baldwin 1979.

anisms that often rely on altering the relative costs and benefits of cooperation and defection.<sup>23</sup> We view these institutions as varying in the scope of their “fundamental bargains,” that is, in the behaviors externality-generating states agree to limit and the threats and/or promises other states make in response. These responses are designed to restructure the incentives for engaging in externality-generating behaviors so that externality-generating states deem it worthwhile to participate. We use the term *bargain* loosely since these institutions can involve coerced as well as voluntary participation. In the context of the Rational Design framework, we see institutional scope as the core trait that leads a state to join rather than remain outside an institution. Scope can be manipulated for many purposes, but we focus on how states limit or extend it to balance the twin goals of participation and compliance.

Unlike Andrew Kydd’s analysis of NATO expansion in which “who to include” is endogenous to institutional design,<sup>24</sup> in the externalities we analyze membership is not a design choice reflecting enforcement problems (conjecture M1, restrictive MEMBERSHIP increases with ENFORCEMENT problems) but a parameter dictated by the number of actors (NUMBER) who must be included to resolve the problem.<sup>25</sup> Our argument suggests that the Rational Design conjectures on membership and scope (conjecture M1; and conjecture S3, SCOPE increases with ENFORCEMENT problems) are variants of an overarching conjecture: the more severe the enforcement problems, the more institutional design reflects efforts to target benefits at contributors and sanctions at noncontributors. Restricting membership is one way to prevent non-contributors from receiving institutional benefits (conjecture M1). Expanding scope is a substitute strategy—the benefits of positive linkage and the costs of negative linkage can be targeted and controlled in ways that the effects of issue-specific reciprocity often cannot (conjecture S3).

In manipulating scope to create an incentive-compatible institution, states choose among three ideal-type bargains: issue-specific reciprocity, coercion (negative linkage), and exchange (positive linkage).<sup>26</sup> Although states can combine two or more of these mechanisms, these ideal-types capture essential design differences. All three involve attempts by dissatisfied states (1) to get perpetrators to take some action they would not otherwise take, (2) to do so by adopting contingent behaviors that present perpetrators with only two possible outcomes, and (3) to structure those outcomes so that perpetrators’ resultant preferences over alternatives match dissatisfied states’ preferences.<sup>27</sup> Coercion and exchange differ from issue-specific reciprocity, however, in that they increase institutional scope.

States must also choose the degree of centralization appropriate for solving the collective-action and informational problems that arise with all three ideal-types.

23. In contrast to Oye, we seek to explain only cases in which the institution changes, rather than clarifies, the contingent response of the dissatisfied state, thus excluding his category of “explanation.” Oye 1992.

24. Kydd, this volume.

25. Koremenos, Lipson, and Snidal, this volume, 777–78.

26. For other views on coercion, contracts, and extortion, see Oye 1992, 35; and Krasner 1999, 26.

27. See Hasenclever, Mayer, and Rittberger 1997, 106; Krasner 1991, 340f; and Amini 1997, 7.



Centralizing information can facilitate implementation of all three strategies by improving each state's knowledge about what other states are doing. Centralizing "enforcement," in the broad sense of responding to compliance and violation when they occur, can help states overcome collective-action problems that plague all three strategies. Centralization constitutes a not-always-successful attempt to help states coordinate retaliatory noncompliance in reciprocity, to coordinate sanctioning in coercion, and to pool resources in exchange.

Regimes adopting issue-specific reciprocity rely exclusively on intertemporal linkage within an issue-area, with dissatisfied states mitigating their externality if others also do so. The fundamental bargain involves a contingent offer quite narrow in scope: "if you do *X*, I'll do *X*, but if you don't do *X*, I won't do *X*," where *X* is some externality-reducing action.<sup>28</sup> Such reciprocity combines the promise of whatever benefits accrue to perpetrators from sustained mutual cooperation with the threat of reversion to the no-agreement status quo of mutual defection. Although states can estimate the benefits of mutual cooperation, reciprocity-based institutions rarely delineate these benefits explicitly. We use *reciprocity* to refer to the issue-specific contingent behaviors of Tit-for-Tat (as opposed to more diffuse reciprocity),<sup>29</sup> a strategy not always available and, when available, perhaps not effective or Pareto improving.<sup>30</sup> Equally important, in many multilateral settings, states discover that retaliatory noncompliance itself involves a collective-action problem among regime supporters, who can find it difficult to target the effects of such noncompliance so that only the original defector is punished, thereby undermining rather than reinforcing cooperation.<sup>31</sup>

Expanding institutional scope helps remedy these problems. Devising contingent responses other than mitigating their own externality allows dissatisfied states access to outcomes other than those dictated by the power distribution within the issue-area. Although issue linkage often involves combining complementary Prisoners' Dilemma conflicts, we use *issue linkage* in the very limited, tactical sense of a state using resources other than engaging in the externality or not.<sup>32</sup> Negative linkage involves dissatisfied states threatening some sanction (*S*) unless the perpetrator does *X*: "if you do *X*, I won't do *S*, but if you don't do *X*, I'll do *S*." Unlike reciprocity or exchange, the success of coercion requires that victims preclude perpetrators from continuing to receive the status quo payoff and instead choose

28. *X* can be either a positive or negative action, that is, *A* can be attempting to get *B* to commence a new activity or to halt or change an existing activity.

29. Although many regimes replace or reinforce issue-specific reciprocity with reciprocity based on positive and negative linkage to other issue-areas, the core of Axelrod's argument lay in demonstrating that Tit-for-Tat fosters cooperation without such linkage. Axelrod 1984.

30. Michael Zürn and Thomas Bernauer helped us clarify this point. Thus states can threaten to punish their political dissidents unless other states stop punishing theirs but do not do so because the nature of the situation makes it obviously ineffective.

31. See Axelrod and Keohane 1986; and Oye 1992, 17–33.

32. See McGinnis 1986; Sebenius 1983; and Haas 1980, 371–72.



between the costs of doing *X* or the costs of the sanctions.<sup>33</sup> Leaving the exact nature of the sanctions unstated can strengthen the threat's deterrent effect. Such sanctions can reduce the "targeting" problem of retaliatory noncompliance but may leave the collective-action obstacles of sanctioning unresolved.

States can also offer positive linkage. In such an "exchange" regime, dissatisfied states offer some side payment to perpetrators that prefer it to continuing current policies. The fundamental bargain involves an exchange promising some reward (*R*) if the perpetrator does *X* and the weak threat of remaining at the no-agreement status quo position otherwise: "if you do *X*, I'll do *R*, but if you don't do *X*, I won't do *R*." Exchange and reciprocity share a basic model of contingent behavior but differ in the rewards: they allow careful modulation of the benefits to the perpetrator, but the benefits are dictated by the payoffs of mutual cooperation when using issue-specific reciprocity. Unlike reciprocity and coercion, exchange usually requires specifying both the magnitude of the reward and the terms for granting it.

Using these theoretical distinctions to differentiate empirical cases requires careful consideration of the status quo and existing expectations. Consider two states (or groups of states), *A* and *B*, with a given level of trade. State *A* is not engaged in an influence attempt if it simply promotes increased trade with *B* without making this increase contingent on some policy of *B*'s.<sup>34</sup> State *A* is engaged in issue-specific reciprocity if it makes its future level of trade contingent on *B*'s level of trade. State *A* is engaged in negative linkage if it threatens to reduce levels of trade from those that *B* had previously expected would continue (or to block improvements in trade that *B* had previously expected would occur) and promises to maintain previously expected trade levels or improvements only if *B* improves its human rights policies. State *A* is engaged in positive linkage if it offers to increase trade with *B* beyond the level that *B* had previously expected only if *B* improves its human rights policies. Table 1 summarizes these distinctions.

## How Situation Structure Influences Institutional Design

Situation structure strongly influences, without dictating, whether states create a narrow institution based on reciprocity or a broader one based on exchange or coercion. Asymmetric and symmetric externalities produce different incentives to defect. Differences in how institutional designs address these incentives to defect give rise, in turn, to different incentives for membership. Thus mechanisms adopted to restructure incentives to defect also restructure the incentives to join the institution. The bargaining or distributional problem (creating an agreement states will join) and the enforcement problem (inducing those that have joined to comply)

33. Krasner 1999, 26.

34. Baldwin 1971, 24.

TABLE 1. *Distinguishing reciprocity, exchange, and coercion*

	<i>Alternatives offered by A (dissatisfied state) to B (externality-generating state) [and B's payoff]</i>	
	<i>If B does X</i>	<i>If B does not do X</i>
Issue-specific reciprocity (no linkage)	A does X* [SQ - Cx + Bx]	A continues status quo behavior** [SQ]
Exchange (positive linkage)	A provides Reward [SQ - Cx + R]	A continues status quo behavior [SQ]
Coercion (negative linkage)	A continues status quo behavior [SQ - Cx]	A imposes Sanction [SQ - S]

- \* = mutual cooperation
- \*\* = mutual defection
- SQ = Status quo payoff to B
- Cx = B's costs of doing X
- Bx = Benefits to B from A doing X
- R = Value of A's reward to B
- S = Cost of A's sanction to B

thus become linked, interacting in ways that reinforce each other, as suggested in the volume's introduction.<sup>35</sup>

Faced with a symmetric externality, states most dissatisfied with the status quo will tend to prefer reciprocity to coercion or exchange because it helps resolve the bargaining problem simply and the enforcement problem adequately. For most international negotiations, a wide range of outcomes are possible. Issue-specific reciprocity provides a strong and simple criteria for constraining bargaining to institutional rules that apply universally. Without creating a single focal point for discussion or eliminating bargaining conflict, imposing nominally equal requirements on all externality generators—for instance, a universal ban, equal reductions, or common technology requirements—significantly narrows the range of outcomes to be considered. Reciprocity rarely imposes equal burdens or provides equal benefits, but it avoids making those distributional differences explicit and salient aspects of the negotiation. In a symmetric setting the prospect of creating a Pareto-improving institution will lead those who could benefit from mutual cooperation to join the negotiations, hoping to define requirements that maximize their benefits while minimizing their burdens and risks. Their ability to use exit or voice will tend to produce agreements that make all better off, even if not making all equally better off.<sup>36</sup>

35. See Fearon 1998; and Morrow 1994c.

36. Hirschman 1970.

Of course, the shared preference for mutual cooperation that defines symmetric situations does not make it stable. States have ongoing incentives to violate. Direct reciprocity can support mutual cooperation if, as in trade agreements, the effects of violation are sufficiently concentrated that actors have incentives to carry out retaliatory noncompliance and can target the effects of such noncompliance onto the initial violator. But if large harms of violations fall on a diffuse set of actors, as often occurs in environmental affairs, retaliatory noncompliance will be unlikely. The individual costs of retaliating will exceed the individual benefits, creating collective-action problems, and even victims with incentives to retaliate will worry that their retaliation will undermine cooperation without succeeding at altering the initial violator's behavior. States can centralize the process of responding to violators in different ways in their effort to overcome these collective-action problems. They can promote information exchange, information dissemination, and alternative sanction mechanisms to increase the ability and incentives to retaliate and the ability to target initial violators, though these may not eliminate the collective-action problem. We view even regimes that authorize such sanctions as fundamentally reciprocal if sanctioning states engage in the behaviors they are seeking to induce in the targets and targets consider themselves as benefiting from those states' behaviors.<sup>37</sup>

States can, however, design truly coercive regimes in symmetric contexts. The Pareto-deficient outcomes of a Tragedy of the Commons do not require Pareto-efficient solutions. Malevolent hegemonic states can coerce weaker states to join regimes in which the weaker states provide collective goods while the hegemon free-rides.<sup>38</sup> They will do so if the benefits from the cooperation they can coerce (less the costs of coercing it) exceed those from the cooperation they could induce by their own cooperation (less the costs of that cooperation). Yet such coercion seems unlikely to be formalized since doing so requires coercing acceptance of inequitable institutional terms in the face of international norms that "stress social and economic equity as well as the equality of states [and make] opposition look more harshly self-interested and less defensible."<sup>39</sup> Institutions imposed by powerful states are unlikely to have highly centralized information and enforcement provisions: the powerful state can induce considerable compliance on its own, making the institution's enforcement problem less severe than when power is distributed more equally (Rational Design conjecture C4, CENTRALIZATION increases with ENFORCEMENT problems), and weaker states will resist efforts to get them to contribute to their own coercion.

Positive linkage is also possible but less likely as a response to symmetric externalities. Offering side payments would require identifying a distribution of benefits that all consider more equitable than reciprocal cooperation, but such

37. Axelrod and Keohane 1986.

38. See Snidal 1985; and Kindleberger 1981.

39. Keohane and Nye 1989, 36, 235–36. It remains unclear how often hegemonic states view imposed agreements as cheaper ways to coerce cooperation than traditional threats exercised outside of an international institution. Young 1989, 84–89.

alternatives will be difficult to identify. In Tragedies of the Commons, even participating in the negotiations reveals a preference for mutual cooperation over the status quo, making subsequent demands for compensation suspect. States may cheat on their obligations but will avoid public threats designed to extort compensation because doing so risks undermining the cooperation of others from which they benefit. At the same time, the biggest beneficiaries of cooperation can point to norms that “all who benefit should contribute” to resist demands for compensation. For these beneficiaries, such compensation is unattractive in itself and because it sets a precedent for future compensation, creating a reputation for “caving in” or getting “the short end of the stick.” Thus, although institutionalizing reciprocity in symmetric contexts may be difficult, institutionalizing exchange or coercion is likely to be even more difficult.

These institutional choices appear quite different in cases of asymmetric externalities. In line with two of the Rational Design conjectures, both distributional and enforcement problems lead to institutions broader in scope than in symmetric situations (conjecture S2, SCOPE increases with DISTRIBUTION problems; and conjecture S3, SCOPE increases with ENFORCEMENT problems). Although institutions that provide absolute gains for all states but greater relative gains for some must address distributional difficulties (see Koremenos, Lipson, and Snidal’s analysis of conjecture M3, MEMBERSHIP increases with DISTRIBUTION problems, in the introduction), those distributional problems become even more severe with asymmetric externalities. Reciprocity limited to the issue-area poses two distinct problems. First, perpetrators receive no benefits if dissatisfied states cooperate. This distribution, or bargaining, problem means perpetrators have no reason to join. Second, perpetrators are not harmed if dissatisfied states defect. This enforcement problem means perpetrators who nonetheless join have no reason to comply. Because reciprocity is not Pareto improving, dissatisfied states must expand institutional scope in ways that induce perpetrators to join while reassuring dissatisfied states that the perpetrators will comply. Whether that increased scope will entail coercion or reward depends on the relative power of the perpetrator.

When the victims of an asymmetric externality are stronger than the perpetrators, the former may simply threaten the latter to compel them to mitigate an externality at their own expense.<sup>40</sup> They may create international institutions to do so, but even states that have coercive resources at their disposal may have difficulty making credible threats because of the costs involved, the constraints of domestic political opinion, or the difficulties of getting other states to cooperate in imposing sanctions.<sup>41</sup> Nor are the benefits of formalizing coercive relations clear. Although the strong state still must expend resources to induce participation and compliance, formalization makes the coercion more explicit, allowing weaker states to use legal norms in appealing to other states when resisting such coercion. A formal agreement

40. See Young 1989, 84–89; Martin 1992a,b; and Rittberger and Zürn 1990.

41. Drezner 2000.

may enhance the stronger state's reputation for coercive strength, but reputation effects also make weak states more likely to resist.

Strong states may turn to rewards because making threats credible and potent can be difficult.<sup>42</sup> If strong states can choose between rewards and sanctions, weak states must choose between rewards and simply "suffering what they must."<sup>43</sup> Realizing that they lack the resources to coerce and that offering reciprocity will not alter perpetrators' incentives, weak states facing asymmetric externalities correctly see positive linkage as the only viable remedy.<sup>44</sup> For both weak and strong victims, the benefits of inducing the perpetrator's cooperation are evident in the damage from, or costs of mitigating, the externality. If this benefit to the victim exceeds the value the perpetrator places on continuing its current behavior, a Pareto-improving agreement is possible.

And institutionalizing an exchange offers advantages that do not exist with coercion. Formalizing the terms of exchange enhances the credibility of the exchange to both sides by clarifying what was promised and defining iterative bargains that reduce both sides' fear of being suckered.<sup>45</sup> Given that a perpetrator of an asymmetric externality will not willingly participate without side payments, centralizing those side payments through the institution pools resources and spreads the costs while helping each victim know how much all other victims will contribute. This information reduces the perpetrator's ability to extort additional compensation. In a form of weak centralization, making offers of rewards both collective and public engages reputational effects that help overcome the greater enforcement problems of asymmetric externalities (conjecture C4, CENTRALIZATION increases with ENFORCEMENT problems). Victims fear that perpetrators will "take the money and run"; perpetrators fear that victims will renege on compensation. Therefore, public, explicit, and formal obligations benefit each side by increasing normative and social pressures on the other to carry out its part of the bargain. Those offering rewards want clear terms of exchange to avoid moral-hazard problems that often plague offers of rewards, as evident in the postwar trade and payments cases described by Thomas Oatley.<sup>46</sup>

This discussion regarding formalizing compensation clarifies the role of flexibility in institutional design. When states agree that any of several possible institutions would be better than the status quo, institutional creation can still be held hostage by disagreements over which institution to create. In these cases flexibility can allow institutional creation to move forward without fully resolving distribution problems or uncertainty about the future state of the world (conjecture F1, FLEXIBILITY increases with UNCERTAINTY about the state of the world; and conjecture F2,

42. See Baldwin 1971; and Schelling 1960, 177.

43. Thucydides 1954.

44. See Hasenclever, Mayer, and Rittberger 1997, 106; Keohane and Nye 1989, 52–53, 122; and Krasner 1991, 340f.

45. This is evident, for example, in the agreement to induce North Korea to forgo its nuclear weapons ambitions. Dorn and Fulton 1997.

46. Oatley, this volume.

FLEXIBILITY increases with DISTRIBUTION problems). Allowing for redesign and reinterpretation ensures the agreement remains beneficial to all parties as circumstances change. In contrast, the benefits of an exchange institution in resolving an asymmetric externality depend on, and are sensitive to, the agreement's exact terms. Terms are carefully crafted to avoid the twin obsolescing bargains of behavioral change without compensation and compensation without behavioral change. Each side sees the institution as beneficial only if it is implemented as agreed. Here, limiting flexibility and precluding renegotiation make institutional creation possible. Thus examining asymmetric externalities affirms the Rational Design contention that states manipulate flexibility to facilitate institutional creation but suggests that this manipulation may involve either increasing it or decreasing it depending on the nature of the problem being addressed.<sup>47</sup>

Finally, a caveat is in order. As Koremenos, Lipson, and Snidal note, rational design does not require that policymakers immediately identify the situation structure they face and propose reciprocity, coercion, or exchange accordingly.<sup>48</sup> Informational and perceptual obstacles may prevent dissatisfied states from initially recognizing the structure of the situation they face. States make mistakes, viewing asymmetric problems as symmetric, projecting their own preferences onto others, or hoping that simple reciprocity can produce their desired outcome. Even weak states may try coercion first, because it may work and, if it does, will be less costly than offering rewards. Thus reciprocity, coercion, or exchange may result not only from a rational calculus by policymakers but also from a process in which the refusal of other states to join reciprocal or coercive institutions clarifies preferences and hence situation structure. Our argument requires only that dissatisfied states respond to such insights by moving from reciprocity to coercion to exchange until perpetrating states accept.<sup>49</sup> Such a trial-and-error process of design, though taking longer, is no less rational or purposive.

### **Examining the Empirical Evidence**

Three predictions on institutional scope stem from our argument. Although we expect these to apply in many issue-areas, we examine them here only in the environmental realm.

1. Issue-specific reciprocity should be the most common institutionalized response to symmetric externalities.
2. Coercion should be a more common institutionalized response to asymmetric externalities with strong victims than exchange, but exchange will still be possible.

47. See John Richards' example of airline regulation and Milner and Rosendorff's discussion of escape clauses, in this volume.

48. Koremenos, Lipson, and Snidal, this volume, 766–67 and fn. 19.

49. We are indebted to Michael Zürn for this insight.



3. Exchange should be the only institutionalized response to asymmetric externalities with weak victims.

Although the third prediction is nonprobabilistic, a full evaluation of the other two would require contingency tables of situation structures and institutions for a large, representative sample of cases. In the absence of such a database,<sup>50</sup> we assess the initial plausibility of these predictions using five separate cases presented by three environmental institutions. Environmental problems complement the other issues analyzed in the Rational Design project and present many examples of both symmetric and asymmetric externalities.

We evaluate whaling, stratospheric ozone loss, and Rhine River chloride pollution because they allow us to observe how situation structure, and hence distribution and enforcement problems, influence the institutional bargains in five distinct cases. Ozone loss provides two cases distinguished by different situation structures among the states involved, with a symmetric externality among states concerned about the problem intertwined with an asymmetric externality between concerned and unconcerned states. International whaling provides two similar cases distinguished by time period: the initial symmetric externality among whaling states gained an asymmetric element as underlying state preferences changed and nonwhaling "victims," led by a strong United States, sought to induce whaling states to end commercial whaling. The Rhine River case presents a clear asymmetric externality with a weak victim.

Besides allowing our primary independent variables, distribution and enforcement problems, to vary, these cases share other important traits. Each involved the "low politics" of transboundary environmental problems, attracting less policy attention and concern than security and economic issues. All five fit the scope of our argument, involving cooperation on a single issue that actors could either support or oppose. Some parties in all the cases had ongoing incentives to defect. We could categorize each case by situation structure and fundamental bargain. Finally, despite these many similarities, the resultant institutions vary in the fundamental bargains they devised to deal with states' incentives to defect. These five cases allow us to evaluate Rational Design conjecture S2, SCOPE increases with DISTRIBUTION problems; conjecture S3, SCOPE increases with ENFORCEMENT problems; and conjecture C4, CENTRALIZATION increases with ENFORCEMENT problems.

### *International Whaling*

After World War II, whaling states faced a quintessential Tragedy of the Commons.<sup>51</sup> Interest in recommencing whaling to bolster postwar supplies of food and oil prompted anxieties about repeating the overexploitation of whale stocks and

50. Mitchell 2001.

51. Peterson 1992, 158.

overcapitalization of the whaling fleet that had plagued the prewar period.<sup>52</sup> Whaling states were symmetrically situated—each state’s overappropriation contributed to the externality of a declining whale stock, and each preferred that others reduce their overappropriation. In 1946 these states negotiated the International Convention for the Regulation of Whaling (ICRW) and created an International Whaling Commission (IWC) composed of one representative from each member state. The IWC became a “whaling club” in which all whaling states, but primarily the five major whaling states (Japan, Norway, the Netherlands, the Soviet Union, and the United Kingdom) sought to avert the declining whale populations, the increasing effort per whale killed, and the low profits of the prewar period.<sup>53</sup> Each state’s economic incentives to support “rational management of a renewable common pool resource” were counterbalanced by predictable incentives for (and fears of) free riding as well as discount rates that often exceeded stock-replenishment rates.<sup>54</sup> Whaling states desired a healthy stock but preferred not to contribute to its production. For almost twenty years, major whaling states rejected any meaningful restraints within the IWC, and some smaller whaling states (Chile, North Korea, and Peru) refused to join. The smaller whaling states knew that refusing to join the regime would not prevent cooperation among major whaling states, cooperation from which they would benefit.

To the extent that the ICRW influenced state behavior at all, it did so almost exclusively through implicit, but nonetheless clear, issue-specific reciprocity. States accepted three-quarters-majority voting rules to set collective annual quotas on total whales killed and on some specific species. Governments agreed to promptly report catch statistics to the International Bureau of Whaling Statistics, which would close the whaling season on the date it estimated the quota would be reached.<sup>55</sup> Negotiated quotas and reported catches, even with misreporting, improved each state’s ability to predict and respond to the behavior of others. States could “opt out” of any quota found objectionable (or that other states had opted out of), as well as invoke the standard withdrawal clause, facilitating the reciprocity of reversion to the status quo.<sup>56</sup> States used these mechanisms, often simply to escape from the agreement but at other times in reciprocal but decentralized Tit-for-Tat behavior, either to avoid being taken advantage of or to enforce the agreement.<sup>57</sup> The ICRW required governments to punish infractions committed by individuals, but did not specify sanctions for particular states’ fleets exceeding the collective quota.<sup>58</sup> Reciprocity

52. Levy 1988, 5.

53. See Stoett 1997, 57; and Tonnessen and Johnsen 1982, 509.

54. See Stoett 1997, 57; Peterson 1992, 158, 160; and Levy 1988, 17.

55. See ICRW 1946, Art. IX; and Walsh 1999.

56. On escape clauses, see Milner and Rosendorff 2001.

57. Walsh 1999, 313. In 1959 Norway and the Netherlands withdrew altogether in protest of quotas they considered too restrictive. In subsequent cases opting out created the awkward but unsurprising situation of species-specific quotas that bound only states that did not hunt those species.

58. Even development of a system of independent inspectors to verify compliance failed to address the question of “after detection, what?” Ikke 1961.

was nonetheless clear in the implicit threat that no fleet could keep whaling after the International Bureau of Whaling Statistics closed the season without provoking continued whaling by other fleets, a strategy that appears to have prevented postseason whaling (although not misreporting of whaling during the season). Although several cases of egregious, particularly Soviet, violations have come to light, most states generally abided by regime quotas.<sup>59</sup>

During the 1970s, declining economic interests in whaling and increasing environmental animosity toward whaling transformed this symmetric externality among whaling states into an asymmetric externality between whaling states and antiwhaling states. Whaling was no longer solely the concern of those hunting them.<sup>60</sup> Although antiwhaling states did not contribute to the problem by hunting whales, they considered the whaling states' behaviors as threatening various whale species and conflicting with a growing moral sentiment against any killing of whales.<sup>61</sup> Taking advantage of a provision allowing universal membership, "non-whaling nations and conservation [nongovernmental organizations] attempted to persuade other non-whaling nations to join the IWC in an effort to obtain the three-quarters majority needed to establish a moratorium on [commercial] whaling."<sup>62</sup> Between 1978 and 1982, IWC membership grew from eight nonwhaling and eleven whaling states to twenty-seven nonwhaling and twelve whaling states. Nonwhaling states joined the regime precisely because they were not engaged in the activity and did not share the preferences of the whaling states, but considered their interests as harmed by whalers, even if whaling imposed no material impacts on them.<sup>63</sup> The new membership adopted a commercial moratorium in 1982, over the opposition of all but one whaling state. The major whaling states (Japan, Norway, and the Soviet Union) filed objections to it and threatened to leave the regime.<sup>64</sup>

The moratorium involved a new fundamental bargain. Had it been negotiated only among whaling states, it would have constituted a simple continuation of reciprocity. *But* the moratorium was adopted by nonwhaling states who viewed whalers as perpetrating a negative externality. Although nonwhaling states could have offered positive inducements to those who ceased whaling, they instead used coercion to address this asymmetric externality. Rather than relying on centralized enforcement, the United States, as a powerful dissatisfied state, became the "self-appointed 'policeman' of the IWC."<sup>65</sup> The United States made or carried out threats to reduce fishing rights or restrict fish imports both to induce member states to stay

59. Earlier Soviet violations seem quite distinct from more recent Japanese violations that have been smaller in magnitude and coupled with quite public denunciations of the IWC. See Yablokov 1994; Baker and Palumbi 1994; and Peterson 1992.

60. In terms of problem structure, a "conflict over means" had become a "conflict over values." Rittberger and Zürn 1990.

61. D'Amato and Chopra 1991.

62. Stedman 1990, 168.

63. See Levy 1988, 29; and Andresen 1989, 109, 116.

64. See Birnie 1985, 616; Sigvaldsson 1996, 330 citing Holt 1985, 192; and Stedman 1990, 168.

65. See Andresen 1989, 111; Wilkinson 1989; and Martin and Brennan 1989.

in the IWC and comply with the moratorium and to induce nonmember whaling states to join the IWC.<sup>66</sup> A U.S. threat of issue-specific reciprocity (that is, to recommence commercial whaling) would have lacked credibility because of its domestic unpopularity and would have been ineffective if carried out because its effects could not be targeted on whaling states who, in any event, increasingly considered some species as no longer threatened. In contrast, threats of economic sanctions that favored domestic fishing interests were both credible and targetable, soon leading major whaling states to discontinue their commercial whaling. Nongovernmental organizations also used direct action and publicity campaigns to sanction whaling, whether conducted within or outside the regime's rules.

Over time, whaling states have responded as one would expect of perpetrators of an asymmetric externality. They have switched from clandestine to public rejections of institutional norms, rejecting even the illusion of voluntary participation and regularly denouncing the IWC. Iceland has withdrawn, Norway has recommenced commercial whaling, and Japan and Russia have threatened to do both. Whaling states have granted numerous scientific permits over IWC objections and formed the North Atlantic Marine Mammal Commission as an alternative institution with membership restricted to commercial whaling interests.<sup>67</sup> This decreasing cooperation reflects, as our model predicts, the shift in underlying situation structure. Initially, symmetric interests among whaling states led them to accept and often, if reluctantly, to comply with a reciprocity-based regime. The emergence of strong antiwhaling sentiment created an asymmetric setting, but one with a strong "victim," the United States. The United States could have pressed for an exchange regime, but it chose coercion through economic sanctions, not only because this strategy appeared cheaper than offering rewards but also because of domestic resistance to paying states to cease a behavior that many viewed as morally wrong in the first place. Predictably, whaling states have resisted the institution's new form, increasingly participating and complying only under duress.<sup>68</sup>

### *Stratospheric Ozone Depletion*

In the 1980s stratospheric ozone depletion resulting from chemicals released into the atmosphere by human activities presented states concerned about the problem (mainly industrialized states) with two interconnected strategic situations. Dissatisfied states faced a symmetric Tragedy of the Commons among themselves, since most industrialized states were both perpetrators and victims of ozone loss. They were perpetrators because their firms produced and their publics consumed most of the world's chlorofluorocarbons (CFCs) and other ozone-depleting substances. They

66. DeSombre 2000.

67. See Hoel 1993; and Caron 1995.

68. If the strategy is eventually judged ineffective, the United States, and even nongovernmental organizations, may yet decide that paying states not to whale is a more effective strategy, however morally repugnant.

were victims because growing awareness of the problem had mobilized their publics to demand protection of the ozone layer, for both material and symbolic reasons. Although reducing CFC use proved not particularly costly once cheap alternatives became available, the benefits of incurring those costs depended on the extent of similar action by others.<sup>69</sup> As in any environmental Tragedy of the Commons, the costs of cooperation were worth incurring and promised environmental benefits only so long as familiar obstacles to inducing cooperation by others could be overcome.

Concerned states faced an asymmetric externality, however, in relation to many developing states. Many developing states, with more pressing policy priorities, high discount rates, and weak domestic concern, considered the immediate benefits of using CFCs—to improve food refrigeration, for example—as outweighing any small future benefits of protecting the ozone layer.<sup>70</sup> Industrialized states feared that China, India, and other developing states would increase their use of ozone-depleting substances and hasten depletion of the ozone layer. Publics in industrialized states worried that their efforts to protect the ozone layer would make them “downstream” victims of developing states relatively unconcerned about damage to the ozone layer. Although the use of ozone-depleting substances by concerned states would constitute free riding, their use by many developing states would reflect a more deep-seated asymmetry of interests: “they were more concerned with accelerating industrial development than with saving the ozone layer, no matter what actions other states took.”<sup>71</sup> Yet resolving the problem required their involvement.

As in the early whaling regime, issue-specific reciprocity was adequate for agreement among industrialized states. The 1987 Montreal Protocol to the 1985 Vienna Convention for the Protection of the Ozone Layer established deadlines for phasing out CFC use that applied to all industrialized states. The relatively mild distributional differences among industrialized states meant negotiators did not discuss positive incentives among industrialized states and expressed little concern about sanctions for violations, leaving their development to an executive body.<sup>72</sup> The issue-specific reciprocity of clear phase-out deadlines applicable to all industrialized states was sufficient to garner signatures by every Organization for Economic Cooperation and Development (OECD) state except Turkey by the end of 1989.

Developing states, however, were considered to be in a “special situation.” The 1987 Montreal Protocol granted them ten years to meet phase-out deadlines and required industrialized states to “facilitate access to” aid and environmentally safe technology to foster their use of alternatives to CFCs. Yet this grace period and the vague offers of rewards, coupled with the threat that member states would halt all

69. Sell 1996, 100.

70. *Ibid.*, 99–100.

71. *Ibid.*, 102.

72. Three years later, that body adopted a weak “list of measures that might be taken” (including “issuing cautions” and suspending treaty rights and privileges) if industrialized states failed to meet deadlines for phasing out ozone-depleting substances or to fund the financial mechanism. UNEP 1991.

trade in ozone-depleting substances with nonmember states, failed to convince most developing states to sign on.<sup>73</sup> By June 1990, only three of the thirteen developing countries predicted to be the biggest CFC consumers—and only thirty-one of over one hundred developing countries—had signed the Montreal Protocol.<sup>74</sup> In the London Amendments of that month, industrialized states, recognizing this, established a Financial Mechanism involving centralized disbursement of pooled financial resources to cover developing states' compliance costs. The amendments carefully centralized and specified the exchange process. This reassured developing states by requiring that financial aid be "additional" and making their phase-out obligations explicitly contingent on receiving aid. It also reassured industrialized states by establishing specific criteria for developing states to receive aid and requiring the monitoring of performance.<sup>75</sup> Only this unambiguous codification of side payments convinced most developing states to join the regime.<sup>76</sup> Within three years fifty more developing countries, including all major prospective CFC users, had joined the regime. Sanctions played little role in this increase: the amendments added no sanctions, and countries known to be smuggling CFCs have yet to be sanctioned.<sup>77</sup>

States concerned about ozone depletion devised different fundamental bargains to deal with different problems. Facing a symmetric externality among themselves, concerned states promptly accepted noncontingent obligations that involved neither rewards nor sanctions, but the reciprocity of mutual CFC phase-outs. These same states faced an asymmetric externality in which less concerned developing states joined the institution only when offered explicit, well-codified side payments. Here, positive linkage was not driven by necessity but apparently was chosen by strong victims as less costly and more effective than using negative linkage to compel developing states to reduce their use of ozone-depleting substances.

### *Rhine River Chloride Pollution*

Pollution of the Rhine River by chlorides involves a classic asymmetric externality that dramatically illustrates the distribution and enforcement problems they can pose.<sup>78</sup> Since the 1930s, French, German, and Swiss enterprises had dumped steadily increasing amounts of chlorides, among other chemicals, into the Rhine.<sup>79</sup> French and German enterprises contributed 90 percent of the chloride load, the

73. Sell 1996, 100. Restrictions on trade in ozone-depleting substances "would have no inhibiting effect on China and India because of their huge potential domestic markets." Benedick 1991, 100.

74. Benedick 1991, 151; and ratification list compiled by authors.

75. Montreal Protocol 1987/1990, Art. 5(5–7).

76. Weiss and Jacobson 1998.

77. Clapp 1997. Victor considers the threat of cutting off multilateral funds as a sanction, but it is one that would have been unavailable without initial adoption of the reward-based strategy. Victor 1998, 165–66.

78. The section that follows builds extensively on the excellent analyses of the Rhine River case by Bernauer 1995 and 1996.

79. See LeMarquand 1977, 125; and Mingst 1981, 164.



Swiss a small percentage, and the Dutch practically none.<sup>80</sup> France, Germany, and Switzerland are exclusively perpetrators since the nature of their water supplies “leaves them unaffected by chloride pollution.”<sup>81</sup> The Dutch, in contrast, are exclusively victims, with high salt levels imposing significant costs on their water works and agricultural interests.

Within the larger regime based on the International Commission for the Protection of the Rhine Against Pollution (or Rhine Commission), the Dutch eventually succeeded in addressing this asymmetric externality through explicit side payments. As early as the 1930s, the Dutch had protested France’s policy of allowing the Alsace Potassium Mine (MdPA; the single largest source, contributing almost 40 percent of the Rhine’s chloride load) to discharge chlorides into the Rhine to avoid contaminating Alsatian groundwater. Although the Dutch government and individuals might have turned to international law to coerce a change in French policy, international legal norms did not lend much support to their position until the mid-1970s.<sup>82</sup> A 1963 Dutch proposal for issue-specific reciprocity with all states “freez[ing] the Rhine’s chloride load at the 1954 level” was, not surprisingly, flatly rejected by France and Germany.<sup>83</sup> Dutch efforts only began to succeed in 1972 with a proposal that France reduce MdPA’s chloride discharges by 60 kilograms per stere in exchange for the Dutch, Germans, and Swiss covering 34 percent, 30 percent, and 6 percent, respectively, of the costs.<sup>84</sup> These cost shares reflected an ad hoc balancing of each state’s contribution to the pollution problem and the “intensity of their demand for chloride reductions.”<sup>85</sup> Once cost estimates grew, however, France refused to implement the agreement.<sup>86</sup> Yet this ad hoc, historically contingent, and rejected proposal provided a surprisingly robust foundation for subsequent institutionalized cooperation.

Revisions to the 1972 proposal produced the 1976 Convention on the Protection of the Rhine Against Pollution by Chlorides. France agreed to reduce discharges by 60 kilograms per stere in three phases, with phase 1 requiring the French to reduce chloride discharges by 20 kilograms per stere by installing a system to inject salts underground.<sup>87</sup> The Dutch, Germans, and Swiss agreed to prevent any net increase in their own discharges. French cooperation was not a response to this reciprocity but to the application of the 1972 cost-sharing formula to the Fr 132 million costs of the injection system and to the costs of the deeper phase 2 and 3 reductions.<sup>88</sup> A 1991 Protocol applied this same cost-sharing formula to two additional projects.

80. See Kamminga 1978, 66; and LeMarquand 1977, 119.

81. Bernauer 1995, 372.

82. Bernauer 1996, 220–21.

83. *Ibid.*, 209.

84. Although not specified in the convention and often misinterpreted as *kilograms per second*, the unit of measurement is *kilograms per stere*, with 1 stere equal to 1 cubic meter of water.

85. Bernauer 1996, 210.

86. LeMarquand 1977, 118.

87. Mingst 1981, 168.

88. Bernauer 1995, 377.

One involved eight years of carefully parsed Dutch, German, and Swiss contributions, totaling Fr 400 million, toward MdPA's costs of stockpiling salt on land during periods of high chloride concentrations in the river. Far more surprising, however, was an agreement by the three upstream states (France, Germany, and Switzerland) to contribute 32 million florins on the same cost-sharing basis for a Dutch project to reduce chlorides entering the IJsselmeer, a major source of water for Dutch waterworks.<sup>89</sup>

French acceptance of the 1976 convention illustrates how the situation structure shapes the fundamental bargain of an international institution. The distribution problem explains the absence of reciprocity. Dutch attempts to gain acceptance of a joint cap on discharges was, in the absence of positive linkage, not compelling to France, Germany, or Switzerland who, as upstream states, would not benefit from any changes in Dutch discharges, which were low in any event. The absence of any Dutch resources to coerce the more powerful upstream perpetrators, particularly France and Germany, explains the absence of sanctions. The Dutch surely would have preferred liability arrangements or adherence to a polluter pays principle, but they lacked legal or material means to force France or Germany to accept such arrangements.<sup>90</sup> The Dutch had to either accept the negative externality or identify positive linkages attractive to the French. They took advantage of the fact that the French mines were a single source that would soon be exhausted for economic reasons anyway.<sup>91</sup> By targeting MdPA reductions in particular, the Dutch minimized the costs of, and French resistance to, proposed reductions while simultaneously gaining German willingness to subsidize French reductions to avoid demands from the Dutch and from German domestic environmental groups for reductions in Germany.<sup>92</sup> The Swiss appear to have agreed more "in the name of basin-wide solidarity."<sup>93</sup> Although Swiss and German contributions mitigated the Dutch cost burden, the Dutch "had little choice except to contribute to the costs," and at a level that was slightly higher (34 instead of 30 percent) than that of the major polluters and much higher than its own contribution to the problem.<sup>94</sup> Here, a victim state compensated a state that was unambiguously capable but simply unwilling to halt its externality-generating behavior.

Events since 1976 illustrate how important positive linkage was to initial French participation, how positive linkage is not free of implementation problems, and how influential initial institutional bargains are to subsequent ones. Implementation of the 1976 agreement was anything but smooth. Although the Netherlands, Germany, and Switzerland paid their cost-shares in 1976, the French withdrew the agreement from parliamentary consideration in 1979, leading the Swiss to reclaim their

89. Bernauer 1996, 216.

90. See LeMarquand 1977, 119; and Bernauer 1996, 205.

91. Indeed, Bernauer argues that the agreement had only a small effect on French behavior. Bernauer 1996. As noted earlier, however, we are explaining regime design, not regime effectiveness.

92. Bernauer 1996, 209–10.

93. See LeMarquand 1977, 124; and Bernauer 1995, 372.

94. See LeMarquand 1977, 119; and Bernauer 1996, 221.

payment in 1981 and the Dutch to recall their ambassador to France.<sup>95</sup> The French finally ratified the agreement in 1985 and began implementing discharge reductions in 1987, but at a rate less than originally agreed (15 instead of 20 kilograms per stere). Nor were the reductions achieved by injecting chloride into the subsoil, as agreed, but by stockpiling it on land, leaving the prospect that the French subsequently would dispose of the chloride into the river.<sup>96</sup> That said, target levels for Rhine chloride concentrations have been achieved through discharge reductions that were due at least in part to agreements based on positive incentives, agreements that would not have been accepted by the French had they been based on sanctions or reciprocity.<sup>97</sup>

The Rhine River case highlights the power of formalized exchange. Initially, one might wonder whether this case involves simply an ad hoc arrangement rather than an institution.<sup>98</sup> Indeed, the underlying asymmetry does well at explaining Dutch side payments to France in 1976 but less well at explaining German and Swiss contributions, which require more context-contingent explanations. Institutional influence, however, is suggested in the design of the 1991 protocol. One anomaly is the application of the 1970s-era cost-sharing formula to the 1991 French project. Rather than renegotiate cost-shares to reflect current levels of pollution contribution and political concern, the states simply applied the institutionalized formula. This formula proved a particularly “sticky” focal point, as evident in Switzerland’s being granted a Fr 12 million “credit” toward its share (for having closed a Swiss chloride-discharging enterprise) rather than recalculating cost shares. More surprising, and more indicative of the power of institutionalization, is the French, German, and Swiss agreement to contribute to the Dutch cleanup of the IJsselmeer, a project offering them neither environmental nor economic benefits. And this project also went through with no renegotiation of cost shares, cost shares that were based on a political and environmental reality almost twenty years old. Absent the Rhine Chloride Convention, it is difficult to explain why France and Germany contributed 30 percent each to a project in which the Netherlands was to stop polluting its own IJsselmeer. These outcomes seem explicable only in institutional terms and illustrate how institutional structures and forms, once created, can wield considerable influence over subsequent outcomes. Table 2 summarizes the evidence from the whaling, ozone-depletion, and Rhine River cases.

### *Comparing Cases and Alternative Hypotheses*

These cases confirm that variation in situation structure, and corresponding variation in distribution and enforcement problems, influences institutional scope (conjecture

95. Bernauer 1995, 378.

96. *Ibid.*, 378–79. See also the failure of efforts to induce Eastern European states to decommission unsafe nuclear reactors detailed in Connolly and List 1996.

97. Bernauer 1996, 225.

98. See Hasenclever, Mayer, and Rittberger 1997, 33, 42–43; and Keohane 1983, 153.

TABLE 2. *Summary of cases*

<i>Case</i>	<i>Situation structure</i>	<i>Fundamental bargain</i>
Whaling among whaling states	Symmetric externality	Issue-specific reciprocity (no linkage)
Whaling between whaling and nonwhaling states	Asymmetric externality, with strong victim	Coercion (negative linkage)
Ozone depletion among industrialized states	Symmetric externality	Issue-specific reciprocity (no linkage)
Ozone depletion between industrialized and developing states	Asymmetric externality, with strong victim	Exchange (positive linkage)
Rhine River chloride	Asymmetric externality, with weak victim	Exchange (positive linkage)

S2, SCOPE increases with DISTRIBUTION problems; and conjecture S3, SCOPE increases with ENFORCEMENT problems). Although the timing of institutional formation surely depended on concern about the problem reaching certain levels, institutional shape depended more on underlying situation structure. Distribution problems among industrialized states concerned about ozone depletion and initially among whaling states were sufficiently mild that reciprocity was a readily negotiated and adequate solution. States joined institutions that had neither significant sanctions for non-compliance nor rewards for compliance. They joined simply because of the unequal and unspecified benefits each believed would arise from mutual cooperation and the desire to avoid decentralized retaliatory noncompliance and corresponding reversion to the status quo.<sup>99</sup> The more severe enforcement problems of asymmetric externalities produced greater centralization, evident in the Financial Mechanism of the Montreal Protocol and the cost-sharing formulas of the Rhine Convention (conjecture C4, CENTRALIZATION increases with ENFORCEMENT problems).

Although reciprocity proves a nonstarter in asymmetric externalities, both the asymmetry and the futility of reciprocity may not be initially obvious. As Koremenos, Lipson, and Snidal observe in the volume's introduction, states use negotiations to collect information about others' preferences.<sup>100</sup> The reluctance of developing states to join a reciprocity-based regime to restrict CFC use showed that they did not share the industrialized states' concerns about stratospheric ozone loss. Industrialized states hoped, incorrectly, that banning trade in CFCs with nonmember states would coerce developing states to join. Likewise, the Dutch saw their reciprocity-based proposal that all Rhine River states reduce chloride discharges fall flat, and the failure of early lawsuits confirmed that they had few effective threats.

99. Hardin 1968.

100. Koremenos, Lipson, and Snidal, this volume, 782.

Such institutional false starts support the intuition that states extend institutional scope only when enforcement problems dictate that narrower institutions will be ineffective. The declining commitment of whaling states to the IWC despite coercive threats by the United States suggests that institutional survival may depend on states and/or nongovernmental organizations offering side payments to those still interested in whaling.

To confirm our analysis, consider alternative explanations for the variation we have documented. Certainly the observed differences in institutional design do not reflect variation in the incentives to defect. In all our cases at least some states had, and acted on, incentives to defect, through violating IWC rules and refusing to join the IWC, smuggling CFCs and missing phase-out deadlines, and failing to reduce chloride discharges on schedule.<sup>101</sup>

The choice of reciprocity, coercion, or exchange might reflect variation in dissatisfied states' costs of making them effective, rather than the difficulty of doing so.<sup>102</sup> The evidence refutes such an interpretation. Reciprocity was not considered and rejected as too expensive by victim states (the Dutch in the Rhine River case or the industrialized states in the ozone case) but rather was proposed by those states and rejected by perpetrator states (the upstream Rhine states and the developing states). Making offers of reciprocity more credible would have been easier and cheaper than devising sanctions or rewards, but doing so would not have made reciprocity more attractive to the perpetrators. In the ozone case, industrialized states had incentives to carry out their threats to end CFC trade with nonmember states since doing so would increase demand for the CFC alternatives they had to offer. They dropped this strategy because it failed to induce developing states to join the regime, and because offering rewards appeared a more effective means of achieving that end than finding more effective sanctioning tools.

Finally, the choice of positive linkage may reflect normative and domestic political constraints on the use of coercion. Policymakers, perhaps pressed by their publics, may reject coercion as inappropriate, even if effective, especially when dealing with developing states.<sup>103</sup> Such concerns may well have influenced the decision to frame side payments in the ozone agreement as targeting "lack of capacity" rather than "lack of will." Yet such a norms-based argument fails to conform to most of the evidence. First, negotiators did ban CFC trade with nonmembers, a ban that effectively applied only to developing states, since all OECD states planned to join. Second, norms were indeterminate. In the ozone case the norm against paying polluters contradicted the norm against sanctioning poorer states. In the Rhine River case, the norm that the polluter should pay was reinforced by the fact that France could clearly afford to stop polluting and merely lacked the will. The choice by states in both cases to pay polluters in the face of a counter-

101. See Yablokov 1994; Clapp 1997; Biermann 1997; Victor 1998; and Bernauer and Moser 1996.

102. We are indebted to an anonymous reviewer for clarifying these points.

103. On the distinction between a logic of appropriateness and a logic of consequences, see March and Olsen 1989; and Finnemore 1996.

vailing norm lends force to the argument that the situation structure made normatively more appropriate mechanisms unavailable or patently ineffective.

## Conclusion

The choices states make in designing international institutions reflect rational efforts to create mechanisms compatible with the incentives in the strategic situations they face. Negative externalities create incentives for victims to induce perpetrators to change their behavior. But such externalities may be either symmetric or asymmetric. The latter pose more severe distributional and enforcement problems, which lead states to create institutions that involve linkages that broaden their scope, strongly confirming two of the Rational Design conjectures (conjecture S2, SCOPE increases with DISTRIBUTION problems; and conjecture S3, SCOPE increases with ENFORCEMENT problems). In symmetric, Tragedy of the Commons, externalities, all perpetrators are also victims, so those most dissatisfied can devise acceptable institutions through simple issue-specific reciprocity. Reciprocity is Pareto improving yet sufficiently attractive to induce participation without the complications of linkage. In contrast, in asymmetric externalities upstream states prefer the status quo to any agreement limited to the issue that concerns downstream states. Whenever issue-specific reciprocity is not Pareto improving, large distribution and enforcement problems arise that can only be addressed if dissatisfied states increase institutional scope through linkage. Linkage may involve those who would benefit by changes in behavior compensating those who must change their behavior. Indeed, weak victims that want an externality to stop must design institutions involving side payments to attract perpetrator participation. However, dissatisfied or victim states, if they are stronger than the perpetrators, may also choose the negative linkage of coercion, exacting “obedience” without institutions or imposing a regime.<sup>104</sup>

In the context of the Rational Design project, our findings demonstrate how the more severe distributional and enforcement problems of asymmetric situations lead states to expand institutional scope. Restricting membership will not induce participation by perpetrators in such situations (see conjecture M1, restrictive MEMBERSHIP increases with ENFORCEMENT problems), so states must use the substitute strategies of offering positive or negative linkage (conjecture S3, SCOPE increases with ENFORCEMENT problems). Situations with symmetric externalities tend to produce narrow, reciprocity-based institutions. Situations with asymmetric externalities and strong victims tend to produce broader coercion-based institutions, and those with weak victims tend to produce exchange-based institutions. Distribution and enforcement problems are tightly intertwined and mutually influential parts of international

104. Young 1979.



cooperation.<sup>105</sup> The argument provides some support for the conjecture that centralization increases with the severity of enforcement problems (conjecture C4, CENTRALIZATION increases with ENFORCEMENT problems), as evident in the careful attention paid to compensation schemes in the Rhine River and ozone-depletion cases. It also suggests, however, that the role of flexibility in institutional design is more complex than suggested by the Rational Design project's framers (conjecture F1, FLEXIBILITY increases with UNCERTAINTY about the state of the world; and conjecture F2, FLEXIBILITY increases with DISTRIBUTION problems). States will embrace flexibility if it allows them to reap near-term institutional benefits while reducing longer-term risks; they will eschew such flexibility and accept more binding, specific rules if, as in asymmetric externalities, each side's institutional benefits depend critically on the other side carrying out the exact terms of the agreed-upon exchange.

The argument reminds us that, along with symmetric Prisoners' Dilemmas and Tragedies of the Commons, asymmetric or unidirectional externalities are important features of the international landscape. These less symmetric contexts also create pressure for institutional formation. Variation in the symmetry and power of the underlying structure influences not only whether states will create institutions but also the mechanisms they design into those institutions they do create.<sup>106</sup> The issue-specific reciprocity common to symmetric externalities has received considerable study, as have coercive regimes imposed by strong states.<sup>107</sup> Exchange regimes have received far less attention. As Ronald Coase would have predicted,<sup>108</sup> states that lack the resources to force others to internalize a negative externality can, for a price, devise institutions that provide an alternative to simply accepting it. Finally, our argument sharpens the debate over whether sanctions are always (or never) the source of compliance with international regimes,<sup>109</sup> demonstrating that reciprocity, sanctions, and rewards tend to be adopted in circumstances that vary systematically, a source of variation that must be considered before the relative effectiveness of different strategies can be properly evaluated.

105. See Fearon 1998; and Morrow 1994c.

106. See Martin 1992b; and Rittberger and Zürn 1990.

107. On imposed regimes, see Young 1989; Martin 1992a; and Gruber 2000.

108. Coase 1960.

109. See Chayes and Chayes 1995; and Downs, Rocke, and Barsboom 1996.