Managing the Evolution of Multilateralism
George W. Downs, David M. Rocke, and Peter N. Barsoom

In the past five years a relatively extensive literature has emerged that explores the demand for multilateral cooperation, that is, those factors that motivate states to develop (or resist developing) formal institutions that operate to increase interstate economic, military, or environmental cooperation. These factors include the transaction costs associated with ad hoc and multiple, bilateral arrangements; increased interstate trade; the diffusion of liberal trade theory; information about the costs of environmental degradation; and trends in elite ideology. To date, however, our understanding of the supply side of multilateralism—the standards that are set for admission, the order and speed with which candidate states are admitted, and the impact of expansion on cooperation and future evolution—has been relatively undeveloped. Partly as a result, theorists have had little to say about how multilateral institutions are likely to evolve or what the policy consequences of their expansion will be.

Supply-side issues are important because multilateral organizations usually do not “spring forth full blown”—they grow. Instead of forming an “inclusive” agreement—that is, one that covers nearly all of the states that its designers eventually hope to include—many multilateral organizations start out with substantially smaller memberships and generally expand over time. We argue that, among the many possible explanations for the choice of this design strategy, there is a rational choice argument.

We thank Bruce Bueno de Mesquita, Mark Fey, Robert Keohane, Walter Mattli, Ken Oye, Robert Padre, Adam Przeworski, and the editors of International Organization, Peter Gourevitch and David A. Lake, for comments on earlier drafts. Tam Bayoumi and Jeff Gable generously provided some of the data used in this paper.

1. On the demand for integration and multilateral trade agreements, see Bond and Syropoulos 1995; Bond, Syropoulos, and Winters 1996; Anderson and Blackhurst 1993; Bagwell and Staiger 1997a,b; de Melo and Panagariya 1993; Leidy and Hoekman 1992; Schneider and Weitsman 1994; Schneider 1994; Snidal 1994; Frankel 1997; and Snidal 1996.
2. For a review of the literature on multilateralism, see Keohane 1990.

International Organization 52, 2, Spring 1998, pp. 397–419
© 1998 by The IO Foundation and the Massachusetts Institute of Technology
that contains behavioral implications not found in other theories. We show how the strategy of admitting potential members sequentially over time based on their preferences for cooperation is able to generate endogenously a series of structure-induced equilibria. The resulting path-dependent process produces a multilateral organization that will often be deeper at every stage of its development than would be obtained by an inclusive strategy; and it mitigates, even if it does not fully eliminate, the breadth-depth trade-off so prominent in the existing literature. As a result, large multilaterals that start out small will tend to become considerably “deeper” in a cooperative sense than those that start out with many members. This outcome holds whether or not sequential growth has been pursued for strategic reasons. As long as those states favoring deeper cooperation tend to be admitted before more conservative states (that is, those desiring less cooperation), the efficiency advantages of this design strategy should hold.

In what follows we first demonstrate the evolutionary advantages of the sequential strategy over an inclusive approach using a simple and generic formal model. We then discuss the source of the sequential model’s power and the robustness of its advantage. Finally, we evaluate our claims on data drawn from the historical experience of two different types of multilaterals. We first examine the extent to which cooperation in the European Union (EU), comprising the deepest trade multilateral institutions, has evolved in the way that the theory predicts. Next, we examine the relative performance of environmental multilaterals that have been organized in a manner consistent with the precepts of the sequential construction model compared to those organized more inclusively.

**Sequential Versus Inclusive Treaty Construction**

Throughout this article, we follow the common practice of focusing on membership issues surrounding the evolution of a multilateral subsequent to its creators choosing some well-defined way of reaching collective decisions. The creation decision in its purest form is an unstructured bargaining problem of enormous complexity. For example, in a world of fifty states there are over two million combinations by which a five-state multilateral might form. The evolution of a multilateral committed to following a set of (relatively) well-defined decision rules that function as its constitution is, in contrast, an “easier” if still analytically difficult problem that involves \( n - k \) possible states, where \( n \) is the total number of states in the system, and \( k \) is the number of states presently in the agreement. In recent years political economists

4. There is an extensive literature on the potentially negative consequences of membership expansion on depth of cooperation. For some of the more prominent contributions in this debate, see Baldwin, Haaparanta, and Kiander 1995; Kahler 1992; Lawrence 1996; Nugent 1992; Phare 1995; Pinder 1992; Taylor 1983; and Ungerer 1993.

5. Alberto Alesina and Vittoria Grilli and Walter Mattli also sidestep the formation of multilateral agreements and focus instead on the evolution of an established agreement; see Alesina and Grilli 1994; and Mattli 1995.

6. For an interesting attempt to cope with some of these problems, see Snidal 1996.
have made far more progress in modeling such decisions than in modeling the outcome of unstructured bargaining problems. The only assumption that we make about the circumstances surrounding the creation of a multilateral is that the initial group of members is relatively more "liberal" (in the sense of desiring larger rather than smaller reductions in trade barriers or more rather than less environmental regulation) than are the remainder of states in the pool of potential members. Since these states have the greatest incentive to overcome the transaction costs associated with the formation of the multilateral, this assumption seems reasonable.7

To make the following discussion as general as possible, we have omitted any full-scale specification of a trade, collective defense, or environmental model and focus instead on the narrow issue of the public choice implications of sequential versus inclusive formation. Inclusive formation is basically a one-step process in which invitations are issued to every potential member.8 Sequential construction consists of a multistep process in which the multilateral expands only gradually as either potential members or the multilateral itself attains some property not originally possessed at the time the institution was created. This property might emerge as the result of economic development, the diffusion of more liberal trade ideology or greener environmental attitudes, the installation of an administration more disposed to regulation, or a change in relative prices. The only requirement is that there be a predictable trend in the change and that it affects states’ preferences for cooperation or regulation.

For scholars who typically assume that preferences are fixed, expecting that preferences will change and will do so in a basically predictable manner over the short run may seem eccentric. This expectation, however, is associated with the creation of virtually every trade and environmental agreement. From the very beginning, the German and French officials who worked to create the Coal and Steel Agreement expected it to grow and evolve into something much more ambitious as other European states (as well as their own) increasingly recognized the benefits of regional trade integration. Many U.S. officials behind the creation of the North American Free Trade Agreement (NAFTA) had comparable expectations.9 By every account, expectations about how preferences will change as the result of the diffusion of information about the cost of environmental degradation, increases in per capita income, and collective deliberation have played an even greater role in the design of environmental agreements such as the Montreal Protocol and the Climate Change Agreement.10 As we demonstrate, the existence of such expectations has profound, if rarely acknowledged, implications for regime design.

7. There are a handful of instances where relatively more conservative states appear to have created a multilateral in an effort to set the regulatory agenda before more liberal states could do so (for example, The International Convention for the Regulation of Whaling).
8. Inclusive formation is obviously an ideal type. Most institutions that start out with even large memberships still admit additional members over time. However, a qualitative difference exists between institutions that begin with a membership believed to include most potential members (for example, the UN) and one that does not (for example, the EU or NAFTA).
In the next two sections we compare the relative performance of each of the design strategies by examining the difference in the treaty level they can be expected to achieve.

**Inclusive Formation**

We first consider how the inclusive formation approach performs when there is no preference change. Suppose that a small group of “liberal” states (that is, those that prefer the most cooperation) wish to form a larger multilateral entity. These states share a roughly common vision about which other states they eventually hope to include as members (for example, the states in Western Europe or North and South America) and have extended invitations to nearly all of them. The initial meeting of the full membership will therefore contain a group of \( n \) participating states who will vote on the initial treaty level. The initial states have decided among themselves that any proposed treaty level must obtain a majority level \( \zeta \), which here is two-thirds, and that any subsequent change in the voting rule \( \zeta \) will require the same supermajority.\(^{11}\) Each state has an ideal point \( \theta_i \) with quadratic loss, so that state \( i \) will always prefer a treaty level of \( \bar{x} \) that is nearer to \( \theta_i \) than to one that is farther away. We assume that the states have been ordered so that \( \theta_1 \leq \theta_2 \leq \cdots \leq \theta_n \), with \( \theta_1 \) being the most liberal member. An initial treaty level \( \bar{x} \) will be set in period 1 (with lower levels of \( \bar{x} \) corresponding to deeper levels of cooperation) and updated in each period if the required majority to do so exists, and a discount rate \( \delta \) applies in calculating a strategy’s total benefits.

To approach the problem of determining what treaty level the full membership will agree to we can use classic cooperative game theory to identify the “core” of the game. This consists of the set of feasible outcomes in the interval \([\theta_L, \theta_U]\) that cannot be overturned. In the context of our two-thirds rule, the upper bound of the core is \( U = \lfloor (2/3)n \rfloor \) or the smallest integer greater than or equal to \((2/3)n\), and the lower bound is \( L = n - U + 1 \) (see proposition 1 in Appendix A). To understand how this works in a concrete example, assume there are thirty-one states in an agreement, and a two-thirds majority is required to make or change the treaty level. Then \( U = \lfloor (2/3)(31) \rfloor = 21 \), and \( L = 11 \). Cooperative game theory tells us that once established, any outcome in the interval \([\theta_{11}, \theta_{21}]\) cannot be defeated since a blocking coalition of the first eleven states will refuse to increase \( \bar{x} \), and a blocking coalition of the last eleven states will refuse to decrease it (see the first line of Figure 1, discussed later).

The limitation of cooperative game theory is that it tells us nothing about the critical question of where inside the core the treaty will first be established. Fortunately, noncooperative game theory can sometimes provide more specific predictions about the nature of a structure-induced equilibrium—if the decision process is suffi-

\(^{11}\) We have chosen a specific value for the voting majority to make the exposition clearer, in particular two-thirds, because this is a voting rule commonly used by multilaterals.
ciently constrained or well structured. In one prominent example, David Baron and John Ferejohn are able to employ noncooperative game theory to predict the outcome of legislative decision making by adding assumptions about the process by which a legislative proposal is adopted, whether closed or open rules are employed, and the number of proposals that are voted on during each period. Here we will obtain the needed structure by assuming that a state chosen at random will propose a treaty level, subject to substitute amendment, and then have it voted on, all using a two-thirds majority. Under this simple rule the equilibrium strategy is for the randomly chosen member to propose either its own ideal point (if it lies in the core) or the closest point in the core to its ideal point otherwise. This, in turn, leads us to anticipate a set of possible treaty levels that is distributed more or less uniformly on the core, with spikes at the two boundaries of the core.

Noncooperative game theory answers the question we asked at the beginning of the section. It tells us that although any particular treaty within the core is a possible outcome of an inclusive strategy, the expected treaty level of the inclusive formation strategy is near the middle of the core. If the number of states is fairly large and the ideal points are distributed symmetrically, this is near the median voter’s ideal point. In the earlier example with thirty-one states the inclusive strategy would produce a treaty level near $\theta_{16}$, the median state.

Now we are in a position to see what happens to an inclusive multilateral with a given voting rule if the preferences of the states continue to liberalize. Under collective liberalization, all of the $\theta_i$ decrease. If the treaty level $\bar{x}$ remains unchanged while this process is occurring, its relative position in the sequence of states will shift to the right. The ideal point of the state, which formerly corresponded to the treaty level, is now to the left of the treaty level, whereas a state whose ideal point had formerly been to the right is now the pivotal state. So long as $\bar{x}$ is in the core (which itself shifts as the $\theta_i$’s shift), no vote can defeat it, and eventually the treaty will lie at the most conservative (right) end of the core. From this point on, the future evolution of the treaty will be tied to the preferences of the most conservative state. This is so because whenever the states liberalize sufficiently that the treaty is no longer in the core, the most conservative bloc of states has enough votes to prevent any outcome more liberal than moving the treaty level back to the nearest point of the core. Needless to say, this is an outcome that the original states (who lie at the left end of the preference ordering) would prefer to have avoided. Although in an absolute sense the treaty level has remained the same, the original states would have preferred that the treaty had evolved and liberalized at the same rate as they had.

Figure 1 illustrates this phenomenon. The first line shows the thirty-one ideal points; the core is the box, and the expected treaty level is denoted by $\bar{x}$. In the second line the states have liberalized sufficiently that the treaty level (which has not changed in absolute terms) is now at the most conservative edge of the core. As

12. Early examples of structure-induced equilibria in the political science literature are Shepsle 1979; and Shepsle and Weingast 1981.
liberalization proceeds further, the treaty level, which can never lie outside the core, persists in its relative location on the conservative boundary of the core. The treaty level has now liberalized in absolute terms but is far more conservative than the original liberal group of states would like.

**Sequential Treaty Construction**

States wanting to create an agreement could deal with the liberalization problem by waiting until the liberalization process has stopped before forming the multilateral. The treaty level would then reflect (in expectation) the preferences of the median voter rather than those of the most conservative voter in the core, and (at least in the absence of some unexpected event) it would remain there. In the previous example that would result in a treaty level at $\theta_{16}$ instead of $\theta_{21}$. The problem with this scheme is that the opportunity costs associated with waiting until liberalization comes to a halt before starting to cooperate will often be high because the liberalization process does not unfold overnight. The costs of having no international trade regime until every state in the world embraced free trade and eliminated all nontariff barriers or of having no EU until every state in Europe embraced the idea of total economic integration would clearly have been considerable. What the core states would like instead is a design strategy that will *both* yield a treaty level more conducive to their preferences and allow them to capture a substantial portion of the cooperative benefits that are available while the liberalization process runs its course. The question is, what properties should this design strategy possess?

The first, rather obvious point to make is that simple majority rule provides the best protection against the future treaty level of the multilateral falling under the control of the relatively conservative states. Since under liberalization the treaty level invariably drifts to the right, or conservative boundary of the core of whatever decision rule is used, the initiator states have an incentive to choose the voting rule where the relative position of this boundary is as liberal as possible. This turns out to be majority rule because (1) its core, the median voter, is initially to the left of the most
conservative state in the core of every other voting rule; and (2) the fact that the median voter is only one or at most two states ensures that virtually no aggregate liberalization will take place without some commensurate liberalization of the treaty level also occurring.

Now consider the issue of the timing of admissions decisions in the context of active liberalization. Suppose we have an existing agreement of \( k \) states and a group of \( k_1 \geq 1 \) states that are candidates for admission. A plausible model of liberalization is that the ideal point \( \theta_{it} \) of state \( i \) at time \( t \) is given by \( \theta_{it} = \theta_{it0} e^{-\lambda t} + (\theta_{it1} - \theta_{it0}) t \), so that the ideal point proceeds smoothly from \( \theta_{it0} \) to \( \theta_{it1} \). This predicts that just as per capita incomes can be expected to grow (on average) at a faster rate in developing states than in developed states, relatively more conservative states will liberalize at a faster rate than those that are quite liberal to begin with. The process is partially converging in the sense that at the end of the process the states will be relatively closer to each other than they were at the beginning, but they will still vary with respect to their preferences.

If the liberalization rate of each state in or out of the agreement can be estimated, at least approximately, and the majority rule is used for voting, we get the following proposition (which is proved in Appendix A).

**Proposition 2.** Suppose that a treaty group of \( k \) states with majority rule is considering whether to admit a group of \( k_1 \) additional states. Suppose that \( U_i(t) \) is the utility that the original group of states derives from the current treaty group, with associated treaty level as chosen by the group, and \( V_i(t) \) is the utility for the original states corresponding to the larger group. Then admission of the new states to the group can occur only at \( t = 0 \), or at discrete intervals when the value of \( U_i(t) \) and \( V_i(t) \) are equal for some state \( i \) in the treaty. In some cases it may never be optimal to admit the states.

This somewhat surprising result is equivalent to saying that admissions decisions will depend only on whether the admission of prospective member state(s) will yield immediate benefits; states will not be admitted on the expectation of future benefits. This means that if there is considerable variation in state preferences, the admission of states will generally be strung out over time. Some states will be admitted almost immediately, whereas others will be admitted gradually as each state liberalizes sufficiently so that it passes a threshold where its admission costs less in terms of decreasing the rate of treaty liberalization than the benefits the state provides by being in the treaty. If the preferences of the potential members remain too far to the right of the pivotal state, the group may never be admitted.

Sequential admissions, the requirement that states yield immediate benefits before admission and the existence of expected liberalization also have implications for discussions of the breadth-depth trade-off, because it tells us that states will not be willing to trade the benefits that depth currently creates for the future benefits provided by breadth. This biases the decision in favor of depth—at least compared to the trade-off that would be justified in expected value terms. Breadth increases will still take place, but they will take place gradually; and when a new member is eventually
admitted, its impact on depth will be modest because it will have become more liberal as the process has unfolded.

Figure 2 illustrates the admissions decision in the context of active liberalization. In the first line we see nine states—three are at the liberal end, and the other six are more conservative and vary widely. The group of three states can form an agreement (defined by the dotted box) with treaty level $\bar{x}$, which is better than the treaty level $\bar{y}$ that would result from all states being admitted and given a vote. In the second line of the figure all state have liberalized, but the conservative states have liberalized more, thus reducing the variation. The three states can now contemplate admitting two more states with only a small degradation of the treaty level from what it would be in the original three-state agreement. The treaty level $\bar{x}$ is more conservative than if it had been kept at $\theta_2$, but it is still more liberal than in period 1. Note that the agreement level from the median of the full group, $\bar{y}$, continues to differ significantly from the treaty level of the smaller group. The third line shows further liberalization and the expansion of the treaty from five to seven states. In the fourth line all states have been admitted, and $\bar{x}$ and $\bar{y}$ are finally the same; what was gained by sequential admission is better treaty levels during the process of liberalization.

Consider the implications of the preceding for the EU and for economic integration more generally. The conventional wisdom, supported by the British foot-dragging experience of the 1970s and the debates surrounding the accession of eastern European states in the 1990s, is that a negative relationship exists between expansion and depth. The arguments used to support this thesis are varied. For example, many analysts argue that expansion magnifies the complexity of the bargaining process and increases transaction costs by introducing a host of new concerns that must

FIGURE 2. Sequential admission under liberalization

14. For a more extensive discussion and analysis of the EU debates surrounding enlargement, see Baldwin 1995b; Flam 1995; Granell 1995; Nugent 1992; Redmond 1990; Taylor 1983; and Winters 1993a.
be addressed with each successive step toward deeper integration. In addition Miles Kahler suggests that an increase in numbers is also likely to increase the amount of preference heterogeneity on any decision. Thus every increase in depth creates a constraint that an applicant state may not be able to or want to satisfy. From this perspective, the only possibility for deeper integration following significant expansion lies in a variable-speed geometry.

Support for the conventional view is not, however, unanimous. Richard E. Baldwin argues that deep integration increases the prospects for expansion by decreasing the competitiveness of alternative arrangements. Robert Pahre goes still further to argue that not only does deep integration increase the prospects for expansion, but expansion increases the prospects for deep integration. This occurs because the demand of member states for deeper integration increases with expansion “because adding members worsens the non-integration outcome and raises the optimal level of integration.” Moreover, the supply of integration also increases with expansion because the ability and willingness to punish treaty violators increases with size.

Greater attention to the potential impact of the sequential design strategy suggests that both the conventional view and the Baldwin–Pahre view contain a good deal of truth. With regard to the conventional view, the preference heterogeneity that is likely to increase with growth, particularly rapid growth, can potentially pose a hazard for deeper cooperation whenever late entrants are more conservative in their preference for cooperation. Fortunately, more liberal members can manage a multilateral’s evolution to reduce considerably the adverse consequences of expansion by devoting most of its early expansion efforts to admitting other relatively liberal states while refusing to admit more conservative states until they have liberalized to the point where their negative impact on the rate at which the multilateral is growing deeper will be small at worst. Even better, the impact of relatively conservative states on the absolute depth of cooperation in the multilateral is likely to be nonexistent as long as they continue to liberalize.

More generally, multilaterals that have employed sequential construction along the way should be more deeply cooperative than those of comparable size that have been constructed inclusively. Interestingly, the assumption that there is a conscious strategy being employed is not critical. As long as more liberal states tend to be admitted before more conservative states, which presumably will often be the case

15. Werner Ungerer, for instance, describes the weaknesses of the EU decision-making process and concludes that enlargement of the Community would make decision making more cumbersome, see Ungerer 1993. For a sophisticated analysis of the consequences of enlargement on EC decision making, see Widgrén 1994; and Hosli 1993.
17. For recent expositions of a variable-speed Europe, see Alesina and Grilli 1994; Martin 1993; Michalski and Wallace 1992; Pinder 1992; and Martin 1995. The relationship between variable-speed agreements and sequential strategies is too complicated to be dealt with here; however, some sort of sequential strategy for admission and transfer between the outer group and the inner group is needed for variable speed to work properly.
20. Ibid., 112.
since they are by definition more in favor of cooperative regulation, and as long as the process of adding additional states takes place slowly enough that when these states are admitted they are more liberal than they were initially, the prediction should hold. Its impact should be most pronounced in large multilaterals that have been constructed over a number of years. This is true because the distance between the median and the right edge of the core of a supermajoritarian voting rule is greatest in large agreements, and the amount of liberalization that the most conservative states will have undergone prior to admission will be greatest in multilaterals that expand slowly.

The Endgame

In the previous section we concluded that a small group of states faced with a treaty design decision in a climate of rapid liberalization should favor majority rule for setting the treaty level and should admit other states sequentially once they liberalize sufficiently. However, liberalization does not go on forever; eventually, it must level off, putting us back in a relatively static situation where some variance remains in state preferences with respect to the ideal treaty level. This expectation has its own design implications.

To understand these implications, suppose that our liberalization process ended abruptly after our three states had let in two more states and agreed to set a treaty level \( \bar{x} \) that corresponds to the ideal point of the median voter. States that the original membership would have admitted had they become liberal enough are still left outside the agreement, and now there is no prospect that their (still) conservative preferences will liberalize further. One option that the five states in the multilateral have is to simply halt expansion at this point and remain a five-member multilateral forever. However, let us further assume that the current members of the multilateral meet and conclude that by admitting ten of the remaining states they would still be collectively better off. Their problem is how to do this while minimizing the negative consequences of expansion on the liberality of the treaty level.

If they maintain their simple majoritarian voting rule and admit all of the members simultaneously, the treaty level will move to the right from the ideal point of the third state to the ideal point of the eighth state. From the point of view of the most liberal states in the original group this represents a notable cost. If they admit them sequentially, the same thing will happen.

Yet there is a strategy that can be used to reach a more liberal outcome. It consists of the simultaneous use of sequential admission and a supermajoritarian voting rule. To see this, consider Figure 3. As before, the states’ ideal points are indicated by \( \theta \), the treaty level is shown as \( \bar{x} \), and the core is shown as a solid box, with the treaty membership shown by a dotted box. When liberalization stopped, the treaty level was at the median or third state. With progressively more conservative states added one at a time to the right of the five-member multilateral, we have the time track of events shown in Figure 3.
Note that the original treaty level does not change at all as the first three new states are added. This is because the self-interest of the three most liberal states will lead them to vote against any change in the status quo that would move the treaty to the right, and this is sufficient to block any new proposal under the two-thirds rule. Only when a ninth state is added to the original membership is the treaty level and the state whose ideal point it represents forced outside the core. Such a situation is by definition unstable and creates the opportunity for members to vote to move the treaty level somewhere back in the core. The question is, where inside the core will it be moved? It turns out that except in rare cases the only level that will be able to defeat the existing status quo ($\theta_3$) under a two-thirds majority vote consists of those at or near $\theta_4$. This is because the cooperation of state 4 is needed before any new proposal can attain the two-thirds majority to overturn the status quo treaty level, and only a treaty level close to $\theta_4$ will make it better off than the current treaty at $\theta_3$. As a consequence the treaty level periodically moves to the right with the addition of new, more conservative states, but never very far. The last two lines of Figure 3 show the progression of the agreement to full membership.

The analysis given in this section can be formalized as the following proposition:

**Proposition 3.** Suppose that there are $n$ states in an international system, with ideal treaty levels $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_n$, that the most liberal $k$ states (states $i$ with $1 \leq i \leq k$) enter into an agreement to set a treaty level $\bar{x}_0$, and that a $\zeta$-fraction majority is required to adopt a treaty level. Suppose also that the remaining $n - k$ states are added in sequence from smallest remaining value of $\theta_i$ to the largest, and that a new treaty level is adopted by voting after each admission decision, with the status quo ante holding if no alternative can achieve a $\zeta$-fraction majority. Any state may propose any treaty level, and the order in which proposals are made is random. After a sufficient number of new states are admitted, the only treaty value that can be adopted will lie at the most liberal end of the core. The number $n^k$ of states to be added before
this occurs depends on the starting treaty level, $\bar{x}_0$, as well as $k$ and $\zeta$, and satisfies $n^* < (\zeta k + 1)/(1 - \zeta)$. The final location of the treaty level within the core does not depend on the original treaty level chosen.

Proof. See Appendix A. In this context where liberalization is no longer taking place, and where the states’ ideal treaty levels still vary, we have seen that larger supermajorities are better. The treaty level will lie at the liberal end of the core in the expanded group, and the larger the supermajority, the more liberal the liberal end of the core will be. The result is an impressive amount of de facto control over the agreement’s destiny; this control becomes more important in absolute terms as the number of treaty states that are added after liberalization increases and hence the size of the core increases.

More generally, in the initial liberalization period, simple majority rule is preferred, whereas, in the later stages, high supermajorities are better. This suggests that in terms of simple efficiency multilateral should switch decision rules when they believe liberalization has come to an end. However, we think that such a complicated strategy will be rare, largely because it would necessitate the multilateral taking an action that would conspicuously disenfranchise the poorest, least developed states who will often be the last states to join. What seems more likely is that a single decision rule will be chosen at the outset that is a compromise between these extremes. This may help to explain why the two-thirds and three-quarters voting rules are so common.

The Robustness of the Sequential Strategy

In order to assess the generalizability of our conclusions regarding the sequential strategy, we have to look at two partially overlapping dimensions of robustness. The first dimension relates to the degree of relative advantage it offers within a given regulatory arena. The second dimension concerns the stability of this advantage as we move from one regulatory arena (such as trade) to another (such as environment).

The two major sources of difference between the sequential and inclusive strategies lie in (1) the distribution of preferences (θ) across states and (2) the speed of liberalization (λ). Obviously, if all states had the same value of θ and the same liberalization rates, the decision rule and the method of treaty formation would be irrelevant. As the variance of θ grows and especially as the distribution becomes more “skewed” with a disproportionate number of very conservative states, the advantage of the sequential construction strategy grows because it “dampens” the negative impact on the treaty level of adding conservative states. This is especially

21. Theoretically, the original states could employ a Stackelberg strategy by adopting a treaty level that could be changed only by unanimous consent. With sufficient knowledge, one could choose this level to maximize the gains. However, this undemocratic strategy is in most cases politically infeasible.

22. Here skewed means that the right or conservative side of the set of preferences is highly variable, whereas the left of liberal side is more closely spaced together. Otherwise said, Germany and Sweden are likely to be more similar than Uganda and India.
important near the end of the liberalization process or if that process is cut short for some unexpected reason (for example, the impact of severe worldwide recession on the diffusion of liberal trade ideology). If the variance of $\theta$ is large and liberalization is taking place at a moderate pace, the handicap of admitting all the states at once is heaviest, at least from the perspective of the liberal states that often create agreements in the first place. This is because control soon rests in the hands of a relatively conservative state and remains there. With respect to the impact of the speed of liberalization ($\lambda$), in those rare cases when liberalization is taking place extremely rapidly, strategy is again irrelevant because the pace of liberalization ensures that any advantage that might be gained by keeping control in the hands of the median state rather than a more conservative state is quickly rendered insignificant.

A third determinant of the relative performance of different strategies is the discount rate. Since the choice of majority is a compromise between majority rule, favored in the early stages, and a higher supermajority favored in the later stages, the particular compromise chosen will be more heavily weighted toward majority rule if the future is discounted more heavily. Note that a high discount rate does not particularly favor the inclusive process; if it is immediately advantageous for a given state to be admitted at first, the sequential admission strategy also admits the state; the only states not admitted are those that are of no immediate help.

Evaluating the utility of the sequential strategy across different regulatory arenas is difficult because its relative utility may be affected by a host of incentive compatibility issues connected with the nature of the regulated good (for example, whether exclusion is possible) that we have not dealt with here. In terms of the dimensions that the simple model does capture, trade and the environment appear to offer relatively fertile contexts for the application of the sequential construction strategy, though for somewhat different reasons. With regard to trade multilaterals, a sequential approach is more likely to pay dividends than an inclusive approach because considerable variation exists in the amount of liberalization states prefer, and the ideology of trade liberalization is diffusing at a moderate pace. With regard to environmental multilaterals where the liberalization rate often appears slower than it does in trade, a sequential approach offers a means of coping with a degree of preference heterogeneity that is greater than that commonly connected with multilaterals in trade. Security multilateralism appears to provide the least hospitable context for sequential construction. Although the security preferences of prospective alliance members vary, this variation is usually more modest than that connected with multilaterals in other areas. More importantly, rarely does the same sort of predictable liberalization process unfold.23

Empirical Corroboration

To evaluate the cooperative benefits generated by the sequential strategy, we conducted two very different tests on data drawn from the historical experience of the

23. Beyond this, voting is less frequently employed for decision making in security multilaterals, and they are often dominated by a few hegemonic security providers.
two types of multilaterals where the sequential strategy appears most likely to be useful. First, we consider whether the sequential strategy employed by the EU has had the effect on the breadth-depth trade-off that we predict. Second, we consider the relative performance of environmental multilaterals designed in a manner consistent with the sequential model and of those designed more inclusively. On the design side, substantial consistency appears to exist between the strategy employed by the EC and then the EU and the sequential strategy, but this is hardly a “discovery.” Obviously, before we began this project we knew that the EU was formed sequentially and employed supermajorities during a substantial portion of its history. The prediction that the ability of a multilateral like the EU to manipulate the order and timing of entries allows it to reduce the negative consequences of increasing size or breadth on the depth of cooperation is more interesting because the extent to which this is true is not so obvious.

Given the absence of any direct measure of the degree of cooperation embodied in every trade regulation created during the lifetime of the EU, we rely on a gravity model framework to measure depth of cooperation. We consider the consequences of EU enlargements on the rate of change in trade flows, where such changes are used to measure changes in the depth of cooperation. According to the standard gravity model, the volume of trade between any two countries is proportional to income and population and inversely proportional to the distance between them. In addition to these economic variables, we include a dummy variable indicating whether or not countries are members of the EU. Since the gravity model allows us to control for the direct trade effects of geographic and economic variables, the coefficient on the dummy variable should isolate the effect of EU membership (and the associated trade policy of the EU) on the rate of trade increase. To the extent that the standard variables do not account for the observed trade, a positive EU coefficient is taken as evidence of cooperation among EU members. The resultant model is:

\[
d\log(\text{TRADE}_{it}) = \alpha + \beta_1 d\log(Y_i Y_p) \\
+ \beta_2 d\log(P_i P_p) + \beta_3 d\log(R_i P_p) + \beta_3 d\log(\text{EU}_i)
\]

Changes in the value of the EU coefficient provide an estimate of the rate of growth in trade that has taken place as a result of membership in the EU while controlling for state size, income, and other economic factors. Ideally, in order to measure the impact of breadth on depth we need a reliable counterfactual estimate of

24. The analysis here is similar to Tamim Bayoumi and Barry Eichengreen’s, although they are interested in estimating the welfare effects of the EU and EFTA, not the depth of cooperation within the EU; see Bayoumi and Eichengreen 1995.

25. The data sample consists of annual bilateral trade flows among twenty-one industrial countries from the International Monetary Fund’s Direction of Trade Statistics for 1953–92. The data were converted to constant dollars using the U.S. gross domestic product (GDP) deflator. Real GDP and exchange rates were drawn from Summers and Heston 1991. To reduce business cycle effects, we used three-year averages of successive, nonoverlapping annual figures.

26. \(Y_i\) is real GDP of country \(i\), \(P_i\) is population, and \(R_{it}\) is the real exchange rate at time \(t\).

how much depth would have existed if a given state or set of states had not been admitted. The difference between this counterfactual level of cooperation and that which was actually realized would provide us with our measure of the impact of breadth on depth. Unfortunately, no reliable method of generating this counterfactual level of cooperation presently exists. What we can do, however, is to look at the time series created by our estimate of the impact of the EU on trade growth and see whether the increase in EU size or breadth has had a consistently negative (or positive) impact on the rate of growth in our measure of depth. If our theory is correct that multilaterals can manage their evolution in such a way as to mitigate the potentially negative effects of increases in breadth on the depth of cooperation, we expect the effect of EU growth to range mostly from neutral to varying degrees of a positive relationship.

Figure 4 plots the regression coefficient of the EU from 1958 to 1992. We ran a separate regression for each enlargement of the EU: 1958–73 (six original EC members), 1972–83 (accession of Britain, Ireland, and Denmark), and 1980–92 (twelve EU members), and we transformed the coefficient to indicate the proportional increase in the predicted value of trade between two member states ($e^{EU} - 1$). These findings indicate that contrary to the expectations embodied in much of the literature, increases in EU breadth did not have a negative impact on the depth of cooperation. After the first enlargement, the rate of growth in cooperation declined relative to the initial years of the EU, but cooperation grew more deep in all but four of the thirty-four years for which we have data. This is what the model has led us to expect through its prediction that states would not be admitted if their immediate contribution were not positive.
Our second test involves environmental regimes. Instead of focusing on the development of a single trade regime, we conduct a comparative evaluation of the cooperative depth attained by environmental multilaterals that evolved incrementally as the sequential strategy prescribes versus that attained by those multilaterals that were designed inclusively. Apart from giving us the opportunity to evaluate the robustness of our claims, this test has at least two important advantages over the previous test. First, the advantage that we have claimed for the sequential model is explicitly comparative. It can only be calculated when the counterfactual of the inclusive strategy can be estimated reliably. No such reliable estimate is available for the EU case. Although the theory suggests that an inclusively organized EU would have achieved less cooperation more slowly, direct empirical evidence for the claim is scant. Second, the test gives us a chance to see whether the benefits of sequential strategy extend to multilaterals that may have adopted it for nonstrategic reasons. Although we have good direct evidence that the French and Germans went out of their way to exclude Britain from the Coal and Steel Agreement, we possess no such evidence regarding the intentions of the states in the environmental agreements. Indeed, we find it difficult to believe that such calculations were always prominent. Nonetheless, if the states that made up the membership of the multilateral at the beginning of its history before it expanded are usually more liberal than those admitted subsequently, the model suggests that the multilateral will likely obtain an advantage with regard to the level of regulatory cooperation it will eventually achieve.

The data were drawn from the set of one hundred existing international agreements and international legal instruments recorded in the UN Environmental Program’s Register of International Treaties and Other Agreements in the Field of the Environment. From this set we eliminated agreements that were no longer in force, that had two or fewer members, or for which no data were available. We identified thirteen inclusive agreements that contained at least twenty states when they were first created and these states constitute at least 75 percent of their current membership. We also identified seven sequentially constructed agreements; these are agreements that initially comprised only a subset of the states that the agreements’ designers eventually hoped would become members. Intent was determined by the treaty documents and narratives describing the treaty’s goals. The issue of intent is critical because it makes no sense to include small agreements that clearly intend to stay small or to limit the sample only to those small agreements that become larger. The first type of agreement must be excluded because it is not pursuing a sequential strategy. An exclusive focus on the second type of agreement would likely create a selection effect, because only successes would remain in the sample.

The dependent variable that we focus on is the depth of cooperation achieved today. This is a measure of how much more cooperation exists today than existed at the time the treaty was signed. Agreements in which the behavior of states within the

27. Different definitions of large (for example, agreements with at least twenty-five members and 85 percent of their current membership) do not yield significantly different results.
treaty differs little from the status quo at the signing of the treaty are coded 1. Agreements in which the behavior of almost all member states is significantly different from their pre-treaty behavior are coded 5 (see Appendix B for further information about coding rules). Obviously, such coding has its limitations. We have no accurate way of estimating what portion of the change in behavior was “caused” by the treaty and what portion of the change would have occurred even in the absence of the treaty. This limitation may not be as serious as it might first appear, however, because we are primarily interested in the relative performance of the two types of design approaches, not in their absolute performance. The difficulties in accurately estimating the counterfactual will only distort performance comparisons between inclusive and sequential agreements if one type of agreement is more affected by some extra-treaty force than the other. If both are affected, the impact of such extra-treaty forces should cancel out each other.

The results of the comparative analysis are dramatic. The average depth of the inclusive agreements is 1.62, whereas that of the average sequentially constructed agreement is 3.43, more than twice as deep. Six of the inclusive agreements have yielded no behavioral change at all (states’ behavior remains at the status quo level or at the same level as when the treaty was signed), and only one has led to moderate or large depth. All the sequentially constructed agreements produced behavioral change, and six of the seven have achieved moderate or large depth.28

None of this would matter very much if the sequentially constructed agreements were dramatically smaller than the inclusive agreements. It would do nothing more than reaffirm the breadth-depth trade-off that so many suspect. Fortunately, this does not appear to be the case. The average inclusive agreement contains eighty-two states today, whereas the average sequential agreement contains ninety-four states, growing from an average size of twenty-three states.

The empirical analysis not only reveals that sequentially constructed agreements behave as the theory predicts in trade regimes but suggests that the relative superiority of sequential construction as compared to inclusive formation extends to other arenas of international cooperation as well, including global environmental regulation.

**Conclusion**

We have argued that the pattern of evolutionary growth that so frequently characterizes successful multilateral organizations possesses an important advantage over more “one shot” development. To demonstrate this we first constructed a formal model that focuses on the choice of whether to build a multilateral sequentially by admitting different states at different times or inclusively by admitting all of the potential members simultaneously. We found that the sequential construction process provides the more regulation-oriented or liberal states that usually play an important role in creat-

28. This difference cannot be attributed to a difference in the average age of the agreement. If we regress depth on age and agreement type, the coefficient of age is insignificant, and the type coefficient remains significant.
ing multilaterals with an ability to expand an institution’s size while maintaining more control over the evolution of its policies than would otherwise be possible. The result is a multilateral that is far more liberal than if the desires of the “median state” of the potential population of members carried the day. The sequential strategy yields a calculated path-dependency reminiscent of state constitutions, which in a very real sense is exactly what it is.

It is important to keep in mind that the significance of our argument does not lie in the claim that multilaterals comprising a small number of states will often be able to achieve a deeper level of cooperation than those comprising many states. In itself this is hardly remarkable; a subset of the most liberal states will always create a more liberal agreement than will the entire population of states. Rather its significance lies in the prediction that, all else being equal, large multilaterals that start out small will be able to achieve considerably more depth than those that start out relatively large. This result appears robust. It does not depend on an assumption that the model we have described is the only source of inspiration for a multilateral’s expansion strategy; we would be the first to acknowledge that a host of considerations, ranging from transaction costs to geography to membership in critical military alliances, will influence a multilateral’s expansion strategy. Nonetheless, as long as relatively liberal states (that is, those favoring deeper cooperation) tend to be admitted before relatively conservative states, the evolutionary advantages of sequential construction should hold.

The evidence corroborates the results of the formal analysis. A gravity model-based estimate of growth in trade cooperation within the EU reveals that the policy of sequential growth has enabled the EU to overcome the breadth-depth trade-off that would have almost inevitably been present if a more aggressive growth strategy had been employed and would have certainly been present if the EU had been formed inclusively. Data from twenty environmental multilaterals further substantiates the theory. Multilaterals created sequentially have achieved a level of cooperation as deep as those formed inclusively. Even if we take into account the obvious shortcomings of any Likert-scale estimate of depth and the possibility of selection problems, the difference is provocative.

Appendix A: Proofs of Propositions

Proof of Proposition 1

**Proposition 1.** Suppose that *n* states with ideal treaty levels \( \theta_1 \leq \theta_2 \leq \ldots \leq \theta_n \) enter into an agreement to set a treaty level \( \bar{x} \). Suppose that a \( \zeta \)-fraction majority is required to adopt a treaty level. The core then consists of all values of \( \bar{x} \) satisfying \( \theta_L \leq \bar{x} \leq \theta_U \), where \( U \) is the least integer satisfying \( U \geq \zeta n \), and \( L = n - U + 1 \).

**Proof.** The core consists of all values that could attract a winning coalition that could stand against all alternatives. Suppose that \( \bar{x} \) satisfies the given inequality. Suppose a different value \( \bar{x}^* \) is proposed and tries to attract voters from \( \bar{x} \). If \( \bar{x}^* > \bar{x} \), the only voters who could be
attracted are those whose ideal points lie above \( \bar{x} \), which is less than \( U \) and therefore does not comprise a \( \zeta \)-fraction majority. Similar results hold if \( x^* \) is a critical value. Suppose now that \( \bar{x} \) is in the core, but \( \bar{x} > \theta_B \). Then there exists a proposed value of \( x^* \) strictly better for \( n \geq \zeta n \) voters, and thus the coalition supporting \( \bar{x} \) is not stable.

**Proof of Proposition 2**

**Proposition 2.** Suppose that a treaty group of \( k \) states with majority rule is considering whether to admit a group of \( k_j \) additional states. Suppose that \( U_i(t) \) is the utility that the original group of states derives from the current treaty group with associated treaty level as chosen by the group, and \( V_i(t) \) is the utility for the original states corresponding to the larger group. Then admission of the new states to the group can only occur at \( t = 0 \), or at discrete intervals when the value of \( U_i(t) \) and \( V_i(t) \) are equal for some state \( i \) in the treaty. In some cases it may never be optimal to admit the states.

**Proof.** Consider this decision from the point of view of one of the original \( k \) member states, such as state \( i \). If the treaty were to stay at the original size, the utility to state \( i \) at time \( t \) would be \( U_i(t) \). If, at time \( T \), the other \( k_j \) states are admitted, the regime shifts to one with a larger group of states, so that the utility of state \( i \) becomes \( V_i(t) \). The value of \( V_i(t) \) does not depend on the time \( T \) that the new states were admitted, since the median voter’s ideal point will always be the treaty level. In this case the value to state \( i \) of the utility flows discounted at rate \( \delta \) is

\[
\int_0^T U_i(t) \delta^t dt + \int_T^\infty V_i(t) \delta^t dt.
\]

The admission decision problem is to choose \( T \) to maximize this quantity. From first principles, the maximum must be at \( T = 0 \), \( T = \infty \), or where the derivative is 0. Now consider the specific context. The states already in the treaty have preferences for tighter controls than states not in the treaty. This means that the new states will, at time \( t = 0 \), vote for loosening the restrictions if possible. A few such states can be admitted without damaging the treaty, so long as they would be outvoted. Let us then assume that the maximum number of these dangerous states has been admitted as well as all the states with liberal preferences (these form the group of \( k \) original states). In this case, clearly \( U_i(0) > V_i(0) \), so \( T = 0 \) cannot be optimal. \( T = \infty \) cannot be immediately eliminated as a possible optimal choice, since a group of states that is very opposed to controls and also liberalizes at a very slow rate may never be admissible. Subject to that possibility, the optimal value occurs when the first derivative reaches 0; that is, when

\[
0 = U(T) \delta^T - V(T) \delta^T
\]

\[
= U(T) - V(T).
\]

Thus, the decision to admit, considering all the effects of discounted utility flow, will occur at the instant that the benefits with the new group equal the benefits of the treaty with only the original group.

**Proof of Proposition 3**

**Proposition 3.** Suppose that there are \( n \) states in an international system, with ideal treaty levels \( \theta_1 \leq \theta_2 \leq \ldots \leq \theta_n \); that the most liberal \( k \) states (states \( i \) with \( I \leq i \leq k \) enter into an
agreement to set a treaty level $\bar{x}_0$ and that a $\zeta$-fraction majority is required to adopt a treaty level. Suppose also that the remaining $n - k$ states are added in sequence from the smallest remaining value of $\theta_0$ to the largest, and that a new treaty level is adopted by voting after each admission decision, with the status quo ante holding if no alternative can achieve a $\zeta$-fraction majority. Any state may propose any treaty level, and the order in which proposals are made is random. After a sufficient number of new states are admitted, the only treaty value that can be adopted will lie at the most liberal end of the core. The number $n^*$ of states to be added before this occurs depends on the starting treaty level, $\bar{x}_0$, as well as $k$ and $\zeta$, and satisfies $n^* < (\zeta k + 1)/(1 - \zeta)$. The final location of the treaty level within the core does not depend on the original treaty level chosen.

**Proof.** The initial treaty level $\bar{x}_0$ is by assumption in the core of the $k$-state initial group. As states are added, so long as the initial treaty level is still in the core, it cannot be defeated. Suppose that $m$ is the first treaty size at which it is not in the core. Let $U$ and $L$ be as in proposition 1, and let $L_1$ be the number of states whose ideal points are less than or equal to $\bar{x}_0$. By hypothesis $L_1 = L - 1$. The $L_1$ states that have ideal points less than or equal to $\bar{x}_0$ will vote against any proposal to raise $\bar{x}$, so the critical vote will be state $L_1$, which can essentially choose the result, and will therefore choose $\bar{x}_0 = \theta_{L_1}$. This then lies at the lower end of the core. The same logic takes hold whenever a sufficient number of new states are added. Clearly, at first the treaty level may or may not be at the most liberal end of the core (of the initial $k$-state treaty). A treaty revision only occurs when the initial treaty level $\theta_0$ no longer lies in the core, and the revision then takes the treaty level to the liberal end of the core, which properly continues to hold as states are added. The worst case from this point of view is when the initial treaty is at the most conservative point in the initial core, which will happen when $\bar{x}_0 = \theta_{U_0}$, where $U_0 = \lceil \zeta k \rceil$. A sufficient value of $n$ is one in which $n - U = U_0 - 1$, where $U = \lfloor \zeta n \rfloor$. This condition implies that the initial treaty level is itself now at the liberal end of the core. Manipulation of these equations is eased by setting $U_0 = \zeta k + a$ and $U = \zeta n + b$, with $0 \leq a < 1$ and $0 \leq b < 1$. Then

\[ n = U + U_0 - 1 = \zeta n + a + \zeta k + b - 1 \]

so

\[ n = \frac{\zeta k + a + b - 1}{1 - \zeta} < \frac{\zeta k + 1}{1 - \zeta} \]

as was to be shown.\(^{29}\)

---

29. We have assumed for simplicity that the treaty level is one of the ideal points. If that assumption is removed, the treaty level still remains the same as states are added until it falls outside the core. Then the set of feasible treaties (those that could defeat the status quo in a $\zeta$-fraction vote) consists of a small interval around the ideal point of the voter whose ideal point is just above the status quo. The simplifying assumption that the value chosen is that voter’s ideal point, rather than a value near the ideal point, introduces no substantial distortion into the model.
TABLE B.1. Depth today variable

<table>
<thead>
<tr>
<th>Depth measure today</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status quo</td>
<td>1</td>
<td>Behavior of no more than 5% is significantly different from the pre-treaty status quo, and no member state has substantially different behavior.</td>
</tr>
<tr>
<td>Little depth</td>
<td>2</td>
<td>Behavior of minority of member states (i.e., 6–20%) is significantly but not substantially different from the pre-treaty status quo, or the behavior of a very few member states (i.e., 1–5%) is substantially different from the pre-treaty status quo.</td>
</tr>
<tr>
<td>Moderate depth</td>
<td>3</td>
<td>Behavior of many member states (i.e., 21–50%) is significantly but not substantially different from the pre-treaty status quo, or the behavior of a minority of member states (i.e., 6–20%) is substantially different from the pre-treaty status quo.</td>
</tr>
<tr>
<td>High depth</td>
<td>4</td>
<td>Behavior of most member states (i.e., 51–80%) is significantly but not substantially different from the pre-treaty status quo, or the behavior of many member states (i.e., 21–50%) is substantially different from the pre-treaty status quo.</td>
</tr>
<tr>
<td>Very high depth</td>
<td>5</td>
<td>Behavior of almost all member states (i.e., 81–100%) is significantly but not substantially different from the pre-treaty status quo, or the behavior of most member states (i.e., 51–80%) is substantially different from the pre-treaty status quo.</td>
</tr>
</tbody>
</table>

The same depth coding scheme expressed in another way:

<table>
<thead>
<tr>
<th>Percentage of member states</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–5%</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

Note: Significant change is defined as a 10–50 percent change in the behavior being regulated, normalized to the pre-treaty status quo (for example, for emissions regulation, one would normalize emissions output to the pre-treaty level and measure change in emissions from that point). Substantial change is defined as a 51–100 percent change in the behavior being regulated, normalized to the pre-treaty status quo.

Appendix B: Data Coding Rules

The process of identifying and coding the treaties was complex. The full set of citations and methods of coding is too extensive to include here but can be obtained from the authors, c/o Peter N. Barsoom, Mitchell Madison Group, 520 Madison Avenue, New York, NY 10022. Table B.1 gives the guidelines that were used in coding the depth variable.

References


Curry, Robert L. 1993. AFTA and NAFTA and the Need for Open Regionalism. Southeast Asian Affairs 17:54–70.


Silva, Miguel. 1996. *Implications of Incorporating Chile into the NAFTA Agreement*. Unpublished manuscript, Strategic Research Department, Center for Naval Warfare Studies, Naval War College, Newport, R.I.


