
Theories and Empirical Studies of International Institutions

Lisa L. Martin and Beth A. Simmons

The role of international institutions has been central to the study of world politics at least since the conclusion of World War II. Much of this research was, and continues to be, pioneered in the pages of *International Organization*. In this article we take stock of past work on international institutions, trace the evolution of major themes in scholarship over time, and highlight areas for productive new research. Our central argument is that research should increasingly turn to the question of how institutions matter in shaping the behavior of important actors in world politics. New research efforts should emphasize observable implications of alternative theories of institutions. We advocate approaching international institutions as both the object of strategic choice and a constraint on actors' behavior, an idea that is familiar to scholars of domestic institutions but has been neglected in much of the debate between realist and institutionalist scholars of international relations.

The article is organized into three major sections. The first section provides an analytical review of the development of studies of international institutions. From the beginning, the pages of *IO* have been filled with insightful studies of institutions, in some cases asking questions consistent with the research agenda we propose in this essay. But the lack of a disciplinary foundation in the early years meant that many good insights were simply lost, not integrated into other scholars' research. With the professionalization of the discipline since the late 1950s, scholarship on international institutions has become more theoretically informed, and empirical research has begun more often to conform to social-scientific standards of evidence, with results that provide both caution and inspiration for future research. One of the most consequential developments for our understanding of international institutions came in the early 1970s, when a new generation of scholars developed insights that opened up inquiry beyond that of formal organizations, providing intellectual bridgeheads to the study of institutions more generally.

Our thanks for comments on previous versions go to Marc Busch, Peter Katzenstein, Bob Keohane, Steve Krasner, and participants in the *IO* fiftieth anniversary issue conference.

The second section explicitly addresses a theme that arises from the review of scholarship on institutions: whether international politics needs to be treated as *sui generis*, with its own theories and approaches that are distinct from other fields of political science, or whether it fruitfully can draw on theories of domestic politics. As our review shows, developments in studies of American politics, such as studies of voting and coalitional behavior, have often influenced the way that scholars approached international institutions. Most of these efforts did not pay off with major insights. The functionalist approach to institutions adopted in the 1980s owed little to theories of domestic politics, drawing more on economic models. Today, we see the pendulum swinging back, as more scholars turn to modern theories initially developed to study domestic political phenomena (see Helen Milner's article in this issue). Here, we assess whether these new attempts are likely to be any more successful than previous efforts.

The third section turns to the problem of research agendas. Where does scholarship on international institutions go next? Our primary argument in this section is that attention needs to focus on *how*, not just *whether*, international institutions matter for world politics. Too often over the last decade and a half the focal point of debate has been crudely dichotomous: institutions matter, or they do not. This shaping of the agenda has obscured more productive and interesting questions about variation in the types and degree of institutional effects, variations that were in fact well documented in the less theoretical but well-researched case studies of the journal's earliest years. Of course, we do not suggest a return to idiographic institutional analysis. Rather, we suggest a number of lines of theoretically informed analysis that may lead to research that both asks better questions and is more subject to empirical testing. These paths include more serious analysis of the distributional effects of institutions, the relation between international institutions and domestic politics, the problem of unanticipated consequences, and a typology of institutional effects.

The Evolution of an Idea: Institutions in International Politics

Early Studies of the Institutionalization of the Postwar World

The "poles" of realism and idealism—of which much is made in graduate seminars—had little to do with the highly practical organizational analysis that dominated the pages of *IO* in the first decades after the war. The focus of attention was on how well these newly established institutions met the problems that they were designed to solve. On this score, few scholarly accounts were overly optimistic. Overwhelmed by the magnitude of the political and economic reconstruction effort, few judged postwar organizations as up to the task. Central to this debate was a highly realistic understanding that international politics would shape and limit the effectiveness of postwar institutions; virtually no one predicted that these would triumph over poli-

tics. The UN,¹ the General Agreement on Tariffs and Trade (GATT),² the International Monetary Fund³—all were the subject of highly critical review.

A number of important studies grappled explicitly with the impact of these institutions on the policies of the major powers and the outcomes for the central political and military competition between them. The answers, predictably, were derived from little more than informed counterfactual reasoning, but they displayed a sensitivity to the broad range of possible impacts that institutions such as the League and the UN could have on the major powers. In their examination of the ideal of collective security, Howard C. Johnson and Gerhart Niemeyer squarely inquired into the role that norms, backed by organizations such as the UN, play in affecting states' behavior. They asked whether states were "prepared to use force or the threat of force for the sake of public law and order rather than for the sake of their national advantage in relation to that of other states. . . . How has the behavior of states been affected by these standards?"⁴ Though ultimately more confident in the balance of power than in norms embodied in the rule of law, these scholars were correct to push for a mechanism that might explain the effects of institutions on behavior: "We cannot claim to have learned much about the League experiment until we know *how* it has affected the problem of harnessing and controlling the factors of force and their role in the relations of power."⁵

A flurry of studies in the early 1950s suggested possible answers. Pointing to the U.S. role in decolonization and military aid for Korea, collective institutions were said to raise U.S. "consciousness of broader issues" that might affect American interests and thereby make the U.S. more responsive to world opinion.⁶ By subjecting policies to global scrutiny—a mechanism not unlike those of transparency and reputation central to the literature in the 1980s—the UN was viewed as having had an (admittedly marginal) effect on some of the most central issues of world politics.

Though lacking the elaborate theoretical apparatus of current research, early studies of postwar organizations had many of the same insights that have informed "modern" institutionalism. Paralleling much contemporary argument on the form of cooperation,⁷ one study as early as 1949 argued that multilateralism was precluded in cases where there were significant bargaining advantages and discrimination advantages of proceeding bilaterally.⁸ Foreshadowing more theoretically sophisticated treatments of informal versus formal agreements,⁹ studies of GATT as early as 1954 recognized that some agreements gain strength through their informal nature, and

1. See Goodrich 1947, 18; Fox 1951; Hoffmann 1956; Claude 1963; and Malin 1947. But for the optimistic view, see Bloomfield 1960.

2. Gorter 1954.

3. See Knorr 1948; and Kindleberger 1951a.

4. Johnson and Niemeyer 1954, 27.

5. Niemeyer 1952, 558 (italics added).

6. Cohen 1951. For a parallel analysis of institutional effects on Soviet behavior, see Rudzinski 1951.

7. See Oye 1992; and Martin 1992b.

8. Little 1949.

9. Lipson 1991.

prescient of the regimes literature viewed the value of GATT as “a focal point on which many divergent views on appropriate commercial policy converge.”¹⁰ Lacking a theoretical hook on which to hang these observations, and without a professionalized critical mass of scholars to develop these insights, many important findings were only rediscovered and advanced more than two decades later.

Nowhere is this more true than in the rediscovery of the relationship between international institutions and domestic politics. The idea that international institutions can influence state behavior by acting through domestic political channels was recognized by scholars writing in the mid-1950s. Referring to the example of the International Finance Corporation, B. E. Matecki wrote that international organizations could be “idea generating centers” with the ability to set in motion national forces that directly influence the making of national policy.¹¹ Reflecting on the efforts of the Council of Europe to gain acceptance of its vision for Europe in national capitals, an early study by A. Glenn Mowers pointed out the conscious strategy of direct lobbying of national governments through national parliaments.¹² And in a fascinating study of the role of the Security Council in influencing Dutch colonial policy, Whitney Perkins pointed to the crucial interaction between authoritative international decisions and democratic politics: “By defiance of the Security Council the Dutch alerted powerful monitors who allied their strength with domestic forces in requiring them to live up to principles [of decolonization].”¹³ “In this type of interaction between democratic governments and the UN emerge some of the essential elements of a world political process.”¹⁴ Anticipating a mechanism for institutional effects that have recently resurfaced in contemporary studies, he concluded that “The role of the UN is to exert pressures designed to enable the loser in public sentiment to accept the consequences of its loss.”¹⁵ This research approach reflected an effort to flesh out the mechanisms by which the policies and perspectives of international institutions could work through national politics.

In short, the early postwar literature on international institutions, while highly focused on formal organizations, was far less naive and legalistic, more politically sensitive and insightful than it is often given credit for being. Early insights included the recognition that the nature of the international political system provided a context for the effectiveness of international institutions, that institutionaleffectiveness should be subject to empirical investigation, and that elaborate organizational structure is not always the best approach to achieving international cooperation. Moreover, the best of this early literature was concerned not merely with *whether* international institutions had an impact, but *how* one might think about a mechanism for their effects. Transparency, reputation, and legitimacy as well as domestic political pressures were suggested in various strands of thought. But there was no conceptual

10. Gorter 1954, 1, 8.

11. Matecki 1956.

12. Mowers 1964.

13. Perkins 1958, 40.

14. *Ibid.*, 26.

15. *Ibid.*, 42

framework that could tie these insights together; nor was there a systematic comparative enterprise to check for their regularity. Rather, another research agenda, replete with fancy methodological tools imported from American politics, was to demote these questions in favor of an only partially fruitful examination of the internal politics of international organizations.

*The Influence of Behavioralism:
Politics Within International Institutions*

If few thought international organization would liberate the world from politics, it arguably became important to understand who has power in these organizations and how that power was being exercised. Especially since the use of the veto had apparently rendered the Security Council toothless, concern began to focus on the development of rules and norms in the General Assembly. The supposed “specter” of bloc voting in that forum—increasingly of concern to American scholars and policymakers as the Cold War extended its gelid reach—became a central concern.¹⁶

This debate took what appears today to be an odd early direction. Perhaps due to new and exciting work in U.S. legislative behavior, the research program quickly became focused on how to describe patterns of voting in the General Assembly, without a systematic attempt to sort out the usefulness of the voting behavior approach. Despite warnings that the international system was fundamentally different from domestic political systems,¹⁷ this research program easily accepted that voting in the UN was a proxy for power in that institution. Certainly there were skeptics: Rupert Emerson and Inis L. Claude, for example, cautioned that voting in an international body does not have the same function as in a democratically elected parliament; an international conference is a negotiating rather than a legislative body. Voting in such a situation, they noted, was unlikely to play a deliberative role, since such votes were no more than propaganda efforts.¹⁸ Few of these studies explicitly defended their assumption that General Assembly resolutions somehow mattered to the conduct of world politics. But the fascination with the method for analyzing voting behavior overcame fairly readily the caution that the domestic–international logic should be subject to close scrutiny. Moreover, the hope of providing an explicitly political (legislative) model inspired by American politics may have been a reaction against the overly “anarchic” systems analysis of the late 1950s.¹⁹

Much of this work can be traced directly to developments in the study of American politics. Hayward Alker and Bruce Russett’s study *International Politics in the Gen-*

16. For one of the earliest studies of bloc voting, see Ball 1951. For a study focusing primarily on the behavior of the Commonwealth countries, see Carter 1950. Concern with the influence of the Commonwealth grew as former British colonies gained independence and membership in the early 1960s. See Millar 1962.

17. Hoffmann 1960, 1–4.

18. Emerson and Claude 1952. See also Jebb 1952.

19. Alker and Russett 1965, 145, explicitly refer to Liska 1957 and Kaplan 1957. They argue that “[i]t is simply erroneous to think of international politics as anarchic, chaotic, and utterly unlike national politics.” Alker and Russett 1965, 147.

eral Assembly, for example, acknowledged “that studies of the American political process by Robert Dahl, Duncan Macrae,²⁰ and David Truman were theoretically and methodologically suggestive of ways in which roll-call data could be used to test for the existence of a pluralistic political process in a quasi-legislative international organization.”²¹ Influenced by James March²² and Robert Dahl, this study sought to understand various influences on UN voting behavior across issue areas in which the dimensions of power and influence were likely to differ. Certainly, one factor influencing this research agenda was the priority given to reproducible and “objective” forms of social science; the focus on General Assembly voting was acknowledged to be an artifact of the availability of fairly complete voting records.²³

Largely related to the ferment in American voting studies, politics within the UN dominated the research agenda for most of the decade from the mid-1960s. Central was the concern to explain why certain countries had a tendency to vote together, to vote in blocs, or to form “legislative coalitions.”²⁴ Also obviously inspired by American politics, another branch of inquiry focused on the determinants of successfully running for elective UN office.²⁵ Much of this literature was methodologically rather than conceptually driven and highly inductive with respect to its major empirical findings.²⁶ Little effort was made to explore the extent to which the concept of representation or the winning of elections in the domestic setting could travel meaningfully to an international institution. The research program lost steam under heavy fire from scholars who demanded a stronger justification for focusing on the General Assembly as a microcosm for world politics.²⁷

Partially in response to the critique that the General Assembly was hardly the center of world politics, and partially influenced by another trend in American politics growing out of the study of bureaucratic politics and political systems, another research path was taken by Robert Cox and Harold Jacobson’s study of eight specialized agencies within the UN.²⁸ In their edited volume, the focus was on the structure and process of influence associated with these institutions and their outputs, rather than on their formal character. Reflecting once again a major thread in American politics, the underlying assumption was that international organizations could be fruitfully analyzed as distinct political systems in which one could trace out patterns of influence: “The legal and formal character and the content of the decision is less important than the balance of forces that it expresses and the inclination that it gives to the further direction of events.”²⁹

20. MacRae 1958.

21. Alker and Russett 1965, vii.

22. March 1955.

23. On objectivity, see Alker and Russett 1965, 2–3; on availability of data see p. 19.

24. See Riggs 1958; Hovet 1958; Keohane 1967, 1969; Weigert and Riggs 1969; Gareau 1970; Alker 1970; Volgy 1973; and Harbert 1976.

25. See Volgy and Quistgard 1974; and Singer and Sensenig 1963.

26. See, for example, Rieselbach 1960.

27. For two systematic reviews of the quantitative research on the UN and international organizations, see Riggs et al. 1970; and Alger 1970.

28. Cox and Jacobson 1973.

29. *Ibid.*

The work of Cox and Jacobson also encouraged the study of international organizations to consider a more transgovernmental model of their influences. Whereas other research inspired by behavioralism typically assumed a unified model of state interests and actors, this work focused on transgovernmental coalitions involving parts of governments and parts of international organizations. One of the most important insights generated was highly consonant with developments in transgovernmental relations that had come on the intellectual scene in the 1970s:³⁰ the observation that one channel through which international organizations could affect state policies was through the potential alliances that could form between international bureaucracies and domestic pressure groups at the national level.³¹ Although this was an interesting insight, and case studies tended to confirm the importance of such “transnational coalitions” for policy implementation, their effect on policy formulation remains unclear.³² Meanwhile, the issues facing the international community changed drastically in the early 1970s, giving rise to a new approach to the study of international institutions, discussed in the following section.

Finally, a strand of research stimulated by Ernst Haas’s “neofunctional approach” to integration also left a telling mark on the study of empirical effects of international institutions in the 1970s. Neofunctionalism ascribed a dynamic role to individuals and interest groups in the process of integrating pluralist communities.³³ By virtue of their participation in the policymaking process of an integrating community, interest groups and other participants were hypothesized to “learn” about the rewards of such involvement and undergo attitudinal changes inclining them favorably toward the integrative system. According to Haas, “political integration is the process whereby actors shift their loyalties, expectations, and political activities toward a new center, whose institutions possess or demand jurisdiction over preexisting national states.”³⁴ The implications for empirical research on such institutions were readily drawn: those who participate in international organizations should exhibit altered attitudes toward their usefulness and effectiveness.

American politics provided yet another methodological instrument that dovetailed nicely with what was thought to be an empirically testable proposition of Haas’s theory: survey research! From the late 1950s into the early 1980s, a plethora of studies tried to establish whether international organizations could contribute to “learning,” whether cognitive or affective.³⁵ The attitudes of civil servants,³⁶ political appointees, and even national legislators³⁷ were scrutinized for evidence that the length or nature of their association with various kinds of international organizations had induced attitudinal change. The impact of methods from American politics was obvi-

30. Keohane and Nye 1974.

31. See Cox 1969, 225; and Cox and Jacobson 1973, 214.

32. See, for example, Russell 1973; and Keohane 1978.

33. See Haas 1958; and Pentland 1973.

34. Haas 1958, 10.

35. See Kelman 1962; Alger 1965; and Jacobson 1967. See also Wolf 1974, 352–53; and Volgy and Quistgard 1975.

36. See Ernst 1978; and Peck 1979.

37. See Bonham 1970; Kerr 1973; Riggs 1977; and Karns 1977.

ous: in some cases, indicators were used that precisely paralleled the “thermometers” used by the National Opinion Survey Research project.

Three problems bedeviled this research approach for years. First, it failed to produce consensus on the effect of international institutions on attitudes.³⁸ Second, attitudes were never reconnected with outcomes, policies, or actions.³⁹ Third, researchers were never able to overcome the problem of recruitment bias, which itself accounted for most of the positive attitudes held by personnel associated with international institutions. As neofunctionalism as a theoretical orientation lost favor over the course of the 1970s and integrative international organizations such as the European Community and the UN seemed to stagnate in the face of growing world problems beyond their purview, this research program declined, though today a version is pursued primarily in studies that attempt to document mass attitudes toward the European Union.

*Politics Beyond Formal Organizations:
The Rise of International Regimes*

As the study of international institutions progressed over the post–World War II years, the gulf between international politics and formal organization arrangements began to open in ways that were not easy to reconcile. The major international conflict for a rising generation of scholars—the Vietnam War—raged beyond the formal declarations of the UN. Two decades of predictable monetary relations under the Bretton Woods institutions were shattered by a unilateral decision by the United States in 1971 to close the gold window and later to float the dollar. The rise of the Organization of Petroleum Exporting Countries and their apparent power to upset previously understood arrangements with respect to oil pricing and availability took place outside the structure of traditional international organizations, as did consumers’ response later in the decade. For some, the proper normative response seemed to be to strengthen international organizations to deal with rising problems of interdependence.⁴⁰ Others more familiar with the public choice literature argued that a proper extension of property rights, largely underway in areas such as environmental protection, rather than a formal extension of supranational authority per se, was the answer to solving problems of collective action.⁴¹ Overall, few doubted that international life was “organized,” but, increasingly, it became apparent that much of the earlier focus

38. Studies that failed to confirm expectations of attitudinal change include Siverson 1973; and Bonham 1970. A few studies even found negative impacts on attitudes due to association with international organizations: Smith 1973; and Pendergast 1976.

39. To the extent that such associations affected outcomes, the results were generally innocuous. See, for example, Mathiason 1972.

40. Brown and Fabian, for example, modestly call for “a comprehensive ocean authority, an outer space projects agency, a global weather and climate organization, and an international scientific commission on global resources and technologies.” See Brown and Fabian 1975. See also Ruggie 1972, 890, 891; and Gosovic and Ruggie 1976.

41. Conybeare 1980.

on formal structures and multilateral treaty-based agreements, especially the UN, had been overdrawn.⁴²

The events of the early 1970s gave rise to the study of “international regimes,” defined as rules, norms, principles, and procedures that focus expectations regarding international behavior. Clearly, the regimes movement represented an effort to substitute an understanding of international organization with an understanding of international governance more broadly.⁴³ It also demoted the study of international organizations as actors: prior to the study of international regimes an inquiry into the effects of international institutions meant inquiring into how effectively a particular agency performed its job, for example, the efficiency with which the World Health Organization vaccinated the world’s needy children.⁴⁴ When regimes analysts looked for effects, these were understood to be outcomes influenced by a constellation of rules rather than tasks performed by a collective international agency.

But just what effects regimes analysis sought to uncover has changed as the research program has unfolded.⁴⁵ A first collective effort by the scholarly community to address regime effects was primarily interested in the distributive consequences of the norms of the international food regime, arguing that it is important to consider the “ways in which the global food regime affects . . . wealth, power, autonomy, community, nutritional well-being, . . . and sometimes physical survival.”⁴⁶ In this view, regime “effects” were to be reckoned in terms of the distributive consequences of the behavior of a myriad of producers, distributors, and consumers, and, in a minor way, by international organizations and state bureaucracies. Certainly, there was in this early volume little thought that regimes were somehow efficient or efficiency-improving outcomes, as later theorizing would imply; rather, the food regime was characterized by “broad and endemic inadequacies,” which are the result of national policies that are “internationally bargained and coordinated . . . by multilateral agreement or unilateral dictate.”⁴⁷

Further research on international regimes moved thinking in three important directions. First, distributive consequences soon fell from the center of consideration as research began to focus on how international regimes are created and transformed in the first place as well as the behavioral consequences of norms or rules,⁴⁸ rather than the distributive consequences of behavior itself. (We argue later that attention to distributive issues ought to be restored.) Second, in one (though not dominant) strand

42. On skepticism regarding the centrality of the GATT regime, see Strange 1988. On the declining importance of “public international agencies” in general and the FAO in particular, see McLin 1979.

43. See, for example, Hopkins and Puchala 1978, especially 598.

44. Hoole 1977. The focus on international organizations as actors providing collective or redistributive goods has a long history. See Kindleberger 1951a; Ascher 1952; Wood 1952; Loveday 1953; Sharp 1953; and Gregg 1966.

45. We focus here on effects of international regimes because, as argued later, we think this is the question on which future research should concentrate. For a review of theories that purport to explain international regimes, see Haggard and Simmons 1987.

46. Hopkins and Puchala 1978, 598.

47. *Ibid.*, 615–16.

48. Krasner 1983b, introduction and conclusion.

of research, attention to the normative aspects of international regimes led naturally to consideration of the subjective meaning of such norms and to a research paradigm that was in sympathy with developments in constructivist schools of thought.⁴⁹ (See the essay by Martha Finnemore and Kathryn Sikkink in this issue of *IO*.)

Third, by the mid-1980s explanations of international regimes became intertwined with explanations of international cooperation more generally. The work of Robert Keohane especially drew from functionalist approaches that emphasized the efficiency reasons for rules and agreements among regime participants.⁵⁰ Based on rationality assumptions shared by a growing literature in political economy, this research sought to show that international institutions provided a way for states to overcome problems of collective action, high transaction costs, and information deficits or asymmetries. This approach has produced a number of insights, which we will discuss and extend later. But its analytical bite—derived from its focus on states as unified rational actors—was purchased at the expense of earlier insights relating to transnational coalitions and, especially, domestic politics. Furthermore, the strength of this approach has largely been its ability to explain the creation and maintenance of international institutions. It has been weaker in delineating their effects on state behavior and other significant outcomes, an issue to which we will return.

This weakness opened the way for an important realist counterthrust in the late 1980s: the challenge to show that international institutions affect state behavior in any significant way. Some realists, particularly neorealists, raised logical and empirical objections to the institutionalist research agenda. On the logical side, Joseph Grieco⁵¹ and John Mearsheimer argued that relative-gains concerns prevent states from intensive cooperation. The essence of their argument was that since the benefits of cooperation could be translated into military advantages, states would be fearful that such benefits would disproportionately flow to potential adversaries and therefore would be reluctant to cooperate in substantial, sustained ways. Responses by Duncan Snidal and Robert Powell showed that, even if states did put substantial weight on such relative-gains concerns, the circumstances under which they would greatly inhibit cooperation were quite limited. Mearsheimer, in his extensive challenge to institutionalism, also argued that the empirical evidence showing that institutions changed patterns of state behavior was weak, especially in the area of security affairs. While we might dispute the extreme conclusions drawn by Mearsheimer, we take seriously his challenge to provide stronger empirical evidence. In the third section of this article we suggest lines of institutionalist analysis that should lend themselves to rigorous empirical testing, avoiding some of the inferential traps and fallacies that Mearsheimer and other realists have identified.⁵²

49. See Haas 1983; and Ruggie 1972.

50. Keohane 1984.

51. See Grieco 1988; and Mearsheimer 1994.

52. See Snidal 1991; and Powell 1991. See also Baldwin 1993.

Institutions Across the Level-of-Analysis Divide: Insights from Domestic Politics

Early studies of international institutions were often motivated by the attempt to apply new methods used in the study of domestic politics. As just reviewed, studies of voting behavior in the General Assembly, electoral success in the UN governing structure, and surveys regarding attitudinal change as a result of international organization experience are all prime examples. Similar studies continue today, for example, in calculations of power indexes for member states of the European Union.⁵³ These approaches have not, however, been widely influential recently and have been subject to trenchant criticisms.⁵⁴ In spite of this less-than-promising experience, scholars today are turning once again to models of domestic politics to suggest new questions and approaches to the study of international institutions. In this section, we briefly consider whether these new approaches are more likely to bear fruit.

We find reasons to be relatively optimistic about today's attempts to transport models across levels of analysis, as long as such attempts are undertaken with some caution. In particular, we see substantial potential in looking at theories of domestic institutions that are rooted in noncooperative game theory. Rationalist theories of institutions that fall into the category of the "new institutionalism" have applicability at both the domestic and international levels. Virtually all the early attempts to apply techniques and research strategies from domestic politics to the international level were implicitly based on the assumption that agreements among actors are enforceable. Indeed, this was the *only* assumption under which it made sense to look at the politics that underlay voting and decision making in international institutions at all. Models that assume that agreements will be enforced by a neutral third party are especially inappropriate for the international setting; calculating voting power in the General Assembly in a world of unenforceable agreements may have more than a passing resemblance to arranging deck chairs on the Titanic. Thus, it is not surprising that these models have not had great influence when transported to the international level.

However, recent models of domestic institutions as a rule draw, often explicitly, on noncooperative game theory. The basic assumptions of noncooperative game theory are that actors are rational, strategic, and opportunistic, and that no outside actor will step in to enforce agreements. Therefore, agreements that will make a difference must be self-enforcing. These conditions are remarkably similar to the usual characterization of international politics as a situation of anarchy and self-help.⁵⁵ As long as models use the same basic assumptions about the nature of actors and their environment, the potential for learning across the level-of-analysis divide could be enormous.

53. Hosli 1993.

54. Garrett and Tsebelis 1996.

55. Waltz 1979.

As one example, consider what international relations scholars might learn from looking at current debates on the nature of legislative institutions.⁵⁶ Analogously to how realist theory portrays states with a mixture of common and conflicting interests but without supranational enforcement, these models treat legislators as self-interested, individualistic actors in a situation where they must cooperate with one another to achieve mutual benefits.⁵⁷ They ask how legislators under these conditions might construct institutions—such as committees or parties—that will allow them to reach goals such as reelection.⁵⁸ Similarly, international relations scholars are interested in how states or other entities design institutional forms (organizations, procedures, informal cooperative arrangements, treaty arrangements) that assist in the realization of their objectives. The point is *not*, as much of the earlier literature assumed, that “legislative activity” at the international level is interesting *per se*. The power of the analogy rests solely on how actors choose strategies to cope with similar strategic environments. In general, we suggest that more progress can be made by drawing out the aspects of domestic politics that are characterized by attempts to cooperate by actors with mixed motives, who cannot turn easily to external enforcement, and applying them selectively to the study of international relations.

The debate about legislative organization, which we argue may provide insights into international institutions more generally, has been roughly organized into a contrast between informational and distributional models. Informational models concentrate on the ways in which legislative structures allow legislators to learn about the policies they are adopting, thus avoiding inefficient outcomes.⁵⁹ Researchers have argued that properly structured legislative committees can efficiently signal information about the effects of proposed policies to the floor, and that informational concerns can explain both the pattern of appointment of legislators to committees and the decision making rules under which committees operate. All of these claims have stimulated intense empirical investigation, which has been challenged by the distributional perspective discussed later. Informational models can be used to extend and clarify arguments in the international literature that stress the role of institutions in the provision of information, as Keohane has argued, and in the learning process, as Ernst and Peter Haas have emphasized. They can lead to predictions about the conditions under which international institutions can effectively provide policy-relevant information to states, about the kinds of institutions that can provide credible information, and about the effects of such information provision on patterns of state behavior. An example of an issue area where these effects might be prominent is environ-

56. The work on legislative institutions is just one example of the application of noncooperative game theory to domestic institutions. But since it is a particularly well-developed literature, we concentrate on it here, without wishing to imply that this is the only branch of research on domestic institutions that may have interesting analogies to international institutions.

57. Shepsle and Weingast 1995.

58. Although much of the work on legislative organization concentrates on the American context, in recent years creative efforts have been made to develop such models in non-U.S. settings. See Huber 1996b; Tsebelis and Money 1995; Ramseyer and Rosenbluth 1993; G. Cox 1987; and Shugart and Carey 1992.

59. See Gilligan and Krehbiel 1990; and Krehbiel 1991.

mental institutions, where it is highly likely that the ability of organizations to provide reliable, credible information about the effects of human activities on the environment is a key factor in explaining the success or failure of negotiations on environmental treaties. Another possible application might be the creation of international financial institutions, such as the Bank for International Settlements, an original function of which was to provide credible information to markets on German creditworthiness.⁶⁰ Within the European Union, the Commission's role as a relatively independent collector of policy-relevant information is a plausible explanation for its ability to exercise considerable influence over policy outcomes.⁶¹

Distributional models, on the other hand, assume that information is not all that problematic. Instead, they concentrate on the fact that legislators are heterogeneous in their tastes, caring differentially about various issues.⁶² Achieving mutual gains, in this framework, means cutting deals that will stick across different issues. Since exchanges of votes cannot always be simultaneous, legislators have developed structures such as committees and agenda-setting rules that allow them to put together majorities on the issues of most intense particularistic interest to them. This structure provides predictions about the distribution of benefits to individual legislators. Distributional benefits flow through appointment to powerful legislative committees. Like researchers in the informational tradition, those in the distributional tradition have used such models to explain and predict various aspects of legislative organization. For example, they argue that committees will be composed of preference outliers—those legislators who care most intensely about particular issues—and that such committees will be granted agenda-setting power, which is necessary to keep cross-issue deals from unraveling on the floor. Distributional models may be especially useful in exploring in a rigorous fashion the role of international institutions in facilitating or hampering mutually beneficial issue linkages that have been an important research agenda in international relations.⁶³

The debate between informational and distributional models of legislative organization has been highly productive, in both theoretical and empirical terms. It has provided new insights into the types of problems confronted by legislators, the types of solutions available to them, and the role of institutions in democracies. On the empirical side, it has generated a plethora of alternative observable implications, for example, about the composition of congressional committees or the conditions under which actors gain gatekeeping or amendment power. Empirical research on both sides has led to deep insights about how the structure of institutions, such as legislative committees, influences their ability to help individuals overcome collective-action problems, and the conditions under which individuals will be willing to delegate substantial decision-making authority to such institutions. Both types of questions are highly relevant and essential to an understanding of the role of institutions in international politics as well. For example, the informational model suggests

60. Simmons 1993.

61. See Haas 1989; and Bernauer 1995.

62. Weingast and Marshall 1988.

63. On issue linkage, see Stein 1980; and Martin 1992c.

that institutions should be most influential in promoting cooperation when they are relatively independent, “expert” sources of information and when such information is scarce and valuable to states. We should expect this model to be most useful in international issue areas characterized by information asymmetries or in the development of expert knowledge (such as financial and banking regulation). The distributional model predicts that institutions will be most successful in allowing for credible cross-issue deals between states when those with the most intense interest in any particular issue dominate policymaking on that dimension and when institutional mechanisms inhibit states from renegeing on cross-issue deals, even if performance on different dimensions is not simultaneous. Institutions that try to cope with environmental protection and development needs in the same package (such as UNCED and the Agenda 21 program) provide a plausible example. For our interests, another striking analogy between the international arena and the legislative literature is the degree to which the terms of the debate—information versus distribution—reflect the emerging debate about the significance of international institutions.

In many essential respects the problems faced by individual legislators mirror those faced by individual states in the international system. Individual actors face situations in which they must cooperate in order to achieve benefits but also face temptations to defect from cooperative arrangements. No external authority exists to enforce cooperative agreements; they must be self-enforcing. Self-enforcement takes the form of exclusion from the benefits of cooperation, a coercive measure. Given these analogies, there is every reason to expect that some of the methods, insights, and results of these new studies of legislators could usefully inform new studies of international institutions, *in spite of the fact that legislators (usually) operate in a more densely institutionalized environment*.⁶⁴ More generally, rationalist models of institutions that have been developed in domestic settings have the potential to be translated to the international level. As long as we are considering mixed-motive situations in which actors must cooperate in order to pursue their objectives, the incentives to construct institutions to structure and encourage cooperation are similar.

How Institutions Matter

Since the 1980s, work on international institutions has been defined for the most part by the demand that scholars respond to a realist agenda: to prove that institutions have a significant effect on state behavior. While structuring the debate in this manner may have stimulated direct theoretical confrontation, it has also obscured some important and tractable research paths. Allowing realism to set the research agenda has meant that models of international institutions have rarely taken domestic poli-

64. One could make a similar argument about domestic theories of delegation. See Epstein and O'Halloran 1997; Lohmann and O'Halloran 1994; and Lupia and McCubbins 1994. The analogy between politicians deciding to delegate authority to bureaucrats or committees and states delegating authority to international institutions is strong.

tics seriously, treating the state as a unit. The debate has also been reduced to a dichotomy: either institutions matter or they do not. Insufficient attention has been given to the mechanisms through which we might expect institutional effects to work. Institutionalists, in response to realism, have treated institutions largely as independent variables, while playing down earlier insights that international institutions are themselves the objects of strategic state choice. Treating institutions as dependent variables has mistakenly been understood as an implicit admission that they are epiphenomenal, with no independent effect on patterns of behavior.⁶⁵

Although it has been important to go beyond merely explaining the existence of international institutions, productive new lines of research emerge if we accept that institutions are *simultaneously causes and effects*; that is, institutions are both the objects of state choice and consequential. In a rationalist, equilibrium framework, this statement is obvious and unexceptionable: states choose and design institutions. States do so because they face certain problems that can be resolved through institutional mechanisms. They choose institutions because of their intended effects. Once constructed, institutions will constrain and shape behavior, even as they are constantly challenged and reformed by their member states. In this section, we outline a number of lines of research that show promise to take us beyond the “do they matter or don’t they” structure of research on international institutions.

The following research agenda is firmly in the rationalist tradition. Although this approach allows for substantial variation in patterns of preferences over outcomes, and indeed provides predictions about outcomes based on exogenous change in such preferences, it provides relatively little explanatory leverage with respect to the sources of change in such preferences. A few words on how this agenda is related to the constructivist research program may be in order. To the degree that constructivist approaches prove powerful at making changes in actors’ fundamental goals endogenous, providing refutable hypotheses about the conditions for such change, the constructivist and rationalist approaches will be complementary. Although rationalist approaches are generally powerful for explaining how policy preferences change when external constraints or information conditions change, alternative approaches, such as constructivism, are necessary for explaining more fundamental, internal changes in actors’ goals. However, the rationalist research program has much to contribute even without strong theories about the reasons for change in actors’ goals. One of the core insights of theories of strategic interaction is that, regardless of actors’ specific preferences, they will tend to face generic types of cooperation problems over and over again. Many situations give rise to incentives to renege on deals or to behave in time-inconsistent ways that make actors happy in the short run but regretful in the long run. Likewise, many situations of strategic interaction give rise to benefits from cooperation, and conflicts over how to divide up this surplus will plague cooperative efforts. Thus, considerations of how to prevent cheating and how to resolve distributional conflict, to give two prominent examples, are central to theories of cooperation regardless of the specific goals of actors. Rationalist ap-

65. Mearsheimer 1994.

proaches are powerful because they suggest observable implications about patterns of cooperation in the face of such dilemmas, even absent the kind of precise information about preferences that scholars desire. It is to such dilemmas that we now turn our attention.

Collaboration Versus Coordination Problems

The most productive institutionalist research agenda thus far in international relations has been the rationalist–functionalist agenda, originating with Keohane’s *After Hegemony* and Steve Krasner’s edited volume on international regimes.⁶⁶ This work was informed by a fundamentally important insight, inspired by the metaphor of the Prisoners’ Dilemma (PD). Individually rational action by states could impede mutually beneficial cooperation. Institutions would be effective to the degree that they allowed states to avoid short-term temptations to renege, thus realizing available mutual benefits.

Some authors, recognizing that PD was only one type of collective-action problem, drew a distinction between collaboration and coordination problems.⁶⁷ Collaboration problems, like PD, are characterized by individual incentives to defect and the existence of equilibria that are not Pareto optimal. Thus, the problem states face in this situation is finding ways to bind themselves and others in order to reach the Pareto frontier. In contrast, coordination games are characterized by the existence of multiple Pareto-optimal equilibria. The problem states face in this situation is not to avoid temptations to defect, but to choose among these equilibria. Such choice may be relatively simple and resolved by identification of a focal point, if the equilibria are not sharply differentiated from one another in terms of the distribution of benefits.⁶⁸ But some coordination games, like the paradigmatic Battle of the Sexes, involve multiple equilibria over which the actors have strongly divergent preferences. Initially, most authors argued that institutions would have little effect on patterns of state behavior in coordination games, predicting substantial institutional effects only in collaboration situations. Interestingly, these arguments led both to expectations about institutional effects on state behavior and to state incentives to delegate authority to institutions, consistent with the kind of equilibrium analysis we find most promising for future research.

As the logic of modern game theory has become more deeply integrated into international relations theory, and as authors have recognized the limitations of the collaboration–coordination distinction, we have begun to see work that integrates the efficiency concerns associated with collaboration and the distributional concerns associated with coordination. Krasner made a seminal contribution to this line of analysis.⁶⁹ He argued that when states are attempting to cooperate with one another, achieving efficiency gains—reaching the Pareto frontier—is only one of the challenges they

66. See Keohane 1984; and Krasner 1983b.

67. See Snidal 1985a; Stein 1983; and Martin 1992b.

68. Garrett and Weingast 1993.

69. Krasner 1991.

face and often not the most difficult one. Many equilibria may exist along the Pareto frontier, and specifying one of these as the locus of cooperation, through bargaining and the exercise of state power, dominates empirical examples of international cooperation. Krasner's insight is perfectly compatible with the folk theorems of noncooperative game theory that show that repeated play of a PD-type game gives rise to many—in fact, infinite—equilibria. Thus, repetition transforms collaboration problems into coordination problems. In most circumstances, states have simultaneously to worry about reaching efficient outcomes and resolving distributional conflict.

Once we recognize this fact, our approach to international institutions becomes both more complex and more closely related to traditional international relations concerns about power and bargaining. To be effective, institutions cannot merely resolve collaboration problems through monitoring and other informational functions. They must also provide a mechanism for resolving distributional conflict. For example, institutions may construct focal points, identifying one possible equilibrium as the default or “obvious” one, thus reducing state-to-state bargaining about the choice of a particular pattern of outcomes. The role of the European Court of Justice (ECJ), discussed elsewhere in this article, is captured in part by this type of constructed focal-point analysis. The Basle Banking Committee's role in devising international standards for prudential banking practices similarly helped to coordinate national regulations where a number of plausible solutions were available.⁷⁰ Where states fear that the benefits of cooperation are disproportionately flowing to others, institutions can provide reliable information about state behavior and the realized benefits of cooperation to allay such fears. Trade institutions perform many functions; one function that could stand more analytical scrutiny is the provision of such information about the distribution of benefits among members. Another way institutions could mitigate distributional conflict is to “keep account” of deals struck, compromises made, and gains achieved, particularly in complex multi-issue institutions. The networks created within the supranational institutions of the European Union, for example, provide the necessary scope for issue-linkage and institutional memory to perform the function of assuring that all members, over time, achieve a reasonably fair share of the benefits of cooperation.⁷¹ Unless the problem of equilibrium selection is resolved, all the third-party monitoring in the world will not allow for stable international cooperation.

Thus, a promising line of research will involve bringing distributional issues back into the study of international institutions, issues that were in fact the focus of some of the early regimes literature discussed earlier. Institutions may interact with distributional conflict in a number of ways. Most simply, they reflect and solidify settlements of distributional conflict that have been established through more traditional means. These means include the exercise of state power, which Krasner emphasizes, market dominance, and alternative methods of bargaining such as making trades across issues.⁷² In this perspective, institutions can make a difference if they lock in a

70. Simmons 1998.

71. Pollack 1997.

72. Fearon 1994a.

particular equilibrium, providing stability. But rather than merely reflecting power in an epiphenomenal fashion, as realists would have it, institutions in this formulation prevent potential challengers from undermining existing patterns of cooperation, explaining why powerful states may choose to institutionalize these patterns rather than relying solely on ad hoc cooperation.

Institutions may also serve a less controversial signaling function, therefore minimizing bargaining costs. This would be the case if institutions construct focal points or if they primarily keep account of the pattern of benefits over time, as discussed earlier. In either case, they effectively increase path dependence. Once a particular equilibrium is chosen, institutions lock it in. Researching the ways in which institutions do this—how do they enhance path dependence, and under what conditions?—would be intriguing. Normative questions also rise to the top of the agenda once we recognize the lock-in role of institutions. If they do in fact solidify a pattern of cooperation preferred by the most powerful, we should question the ethical status of institutions, turning our attention to equity, as well as efficiency, questions.

In the most traditional, state-centric terms, institutions reflect and enhance state power; in Tony Evans and Peter Wilson's words, they are "arenas for acting out power relations."⁷³ On the other end of the spectrum, we may want to ask about situations in which institutions play a more active role in resolving distributional conflict. Perhaps institutions sometimes do more than lock in equilibria chosen through the exercise of state power, having an independent part in the selection of equilibria. Such an argument has been made most clearly in the case of the ECJ. Here, Geoffrey Garrett and Barry R. Weingast find that there are a number of ways in which the European Community could have realized its goal of completing the internal market.⁷⁴ The ECJ made a big difference in the course of European integration because it was able to construct a focal point by choosing one of these mechanisms, that of mutual recognition. This choice had clear distributional implications but was accepted by member states because it was a Pareto improvement over the reversion point of failing to complete the internal market. A distinct research tradition emphasizes the legitimizing role that international institutions can play in focal-point selection. Some scholars point out that institutionally and legally enshrined focal points can gain a high degree of legitimacy both internationally and domestically.⁷⁵ This legitimacy, in turn, has important political consequences.⁷⁶

To develop a research agenda on how institutions resolve problems of multiple equilibria and distribution, we would have to build on these insights to ask conditional questions. When are states, particularly the powerful, willing to turn the problem of equilibrium selection over to an institution? What kinds of institutions are most likely to perform this function effectively—those that are strategic or those that

73. Evans and Wilson 1992.

74. Garrett and Weingast 1993. They also argue that the multiple equilibria were not sharply distinguished from one another in terms of efficiency and do not concentrate on distributional conflict among equilibria. They have been criticized on these points. See Burley and Mattli 1993.

75. See Franck 1990; and Peck 1996, 237.

76. Claude 1966, 367.

are naive; those that rely on political decision making or those that rely heavily on relatively independent experts and/or judicial processes; those that broadly reflect the membership of the institution or those that are dominated by the powerful? Under what conditions are constructed focal points likely to gain international recognition and acceptance? Overall, bringing the traditional international relations focus on distributional conflict back into the study of international institutions holds the potential for generating researchable questions that are both positive and normative in nature.

International Institutions and Domestic Politics

In allowing their agenda to be defined by responding to the realist challenge, institutionalists have generally neglected the role of domestic politics. States have been treated as rational unitary actors and assigned preferences and beliefs. This framework has been productive in allowing us to outline the broad ways in which institutions can change patterns of behavior. But in privileging the state as an actor, we have neglected the ways in which other actors in international politics might use institutions (a central insight of earlier studies of transgovernmental organization) and the ways in which the nature or interests of the state itself are potentially changed by the actions of institutions (an implication of the early neofunctionalist literature). Here we outline a few lines of analysis that should be fruitful for integrating domestic politics and international institutions in a systematic manner, rather than treating domestic politics as a residual category of explanation. Because the lines of analysis here have foundations in specific analytical frameworks with explicit assumptions, applying them to the problem of international institutions should result in productive research paths, rather than merely the proliferation of possible “explanatory variables” that has characterized many attempts to integrate domestic politics and international relations. We should note that bringing domestic politics back into the study of international institutions is an agenda that should be understood as analytically distinct from that of applying institutionalist models developed in the domestic setting to the international level, an agenda addressed elsewhere in this article.

As we will argue, one of the more fundamental ways in which international institutions can change state behavior is by substituting for domestic practices. If policies formerly made by domestic institutions are now made on the international level, it is reasonable to expect substantial changes in the patterns of world politics. Three related questions are central to understanding the relations between domestic and international institutions. First, under what conditions might domestic actors be willing to substitute international for domestic institutions? Second, are particular domestic actors regularly advantaged by the ability to transfer policymaking authority to the international level? Third, to what extent can international institutional decisions and rules be enforced by domestic institutions, and what are the implications for compliance? These questions are tied together by the assumption that domestic actors intentionally delegate policymaking authority to the international level when this action furthers pursuit of their interests.

Domestic institutions can at times be a barrier to the realization of benefits for society as a whole. Failures of domestic institutions can arise through a number of mechanisms. Perhaps most obviously, domestic institutions can be captured by preference outliers who hold policy hostage to their demands. Recent research suggests that this may be the case with respect to the settlement of territorial disputes between bordering states in some regions: repeated failure to ratify border agreements in the legislature is one of the most important domestic political conditions associated with the willingness of states to submit their disputes to international arbitration.⁷⁷ More generally, this situation is likely to arise when some actors, such as those looking for particularistic benefits, find it easier to organize than do actors more concerned with the welfare of the average citizen. Such is the story often told about trade policy. Import-competing producers and others with an interest in protectionist policies may find it easier to organize than those who favor free trade, a coalition of exporters and consumers. This differential ability to organize will bias policy in favor of protection, decreasing overall welfare. Transferring the policymaking process to the international level, where exporters can see that they have a stake in organization in order to gain the opening of foreign markets, can facilitate a more evenhanded representation of interests. Those actors who have the most to gain from pursuit of general welfare—such as executives elected by a national constituency—will show the most interest in turning to international institutions under such circumstances. Judith Goldstein provides an analysis along these lines when she explains the paradox of the U.S. president agreeing to bilateral dispute-resolution panels in the U.S.–Canada Free Trade Act (FTA), in spite of the fact that these panels predictably decide cases in a way that tends to deny protection to U.S. producers.⁷⁸

We can identify other incentives for domestic actors to transfer policymaking to the international level. One common problem with institutions that are under the control of political actors is that of time-inconsistent preferences. Although running an unexpectedly high level of inflation today may bring immediate benefits to politicians up for reelection, for example, allowing monetary policy to be made by politicians will introduce a welfare-decreasing inflationary bias to the economy. Putting additional constraints on policy, for example, by joining a system of fixed exchange rates or a common currency area, can provide a mechanism to overcome this time-inconsistency problem, as argued by proponents of a single European currency. In general, if pursuit of gains over time involves short-term sacrifices, turning to international institutions can be an attractive option for domestic policymakers.

A second and related question about domestic politics is whether particular kinds of actors will regularly see an advantage in turning to the international level. At the simplest level, it seems likely that “internationalist” actors—those heavily engaged in international transactions,⁷⁹ those who share the norms of international society,⁸⁰

77. Simmons 1998.

78. See Goldstein 1996; and Gilligan 1997.

79. Frieden 1991.

80. Sikkink 1993a.

or those who have a stake in a transnational or global resource⁸¹ —will have an interest in turning to the international level. This may especially be the case when such groups or parties are consistently in a minority position in domestic politics. Drawing on these ideas, we could begin to develop hypotheses about the kinds of domestic interest groups that will most favor transferring some authority to the international level.

Certain domestic institutional actors may also have a tendency to benefit from international-level policymaking. One such actor, which is just beginning to enter political scientists' analysis of international institutions, is the judiciary. Increasingly, international agreements are legal in form. This means that they often are interpreted by domestic courts, and that judges can use international law as a basis on which to make judgments.⁸² Because international law provides this particular actor with an additional resource by which to pursue agendas, whether bureaucratic or ideological, we might expect that the judiciary in general tends to be sympathetic to international institutions.

Overall, as we work toward more sophisticated specification of the causal mechanisms through which institutions can influence behavior, we will have to pay much more attention to domestic politics than studies of international institutions have thus far. The development of general theories of domestic politics provides an opening for systematic development of propositions about domestic actors. We no longer need to treat the domestic level as merely the source of state preferences, nor as a residual category to explain anomalies or patterns of variation that cannot be explained by international factors. Instead, we can move toward genuinely interactive theories of domestic politics and international institutions, specifying the conditions under which certain actors are likely to prefer that policy be made on the international level. This focus allows us to specify conditions likely to lead to the delegation of policymaking authority to the international level, some of which we have outlined here.

Unanticipated Consequences

In a rationalist framework, institutions are both the object of state choice and consequential. The link that ties these two aspects of institutions together, and allows the analyst to develop refutable propositions about institutions within an equilibrium framework, is the ability of actors to anticipate the consequences of particular types of institutions. For example, in the preceding discussion of domestic politics, we assumed consistently that domestic actors were able to gauge with some degree of accuracy the ways in which working within international institutions would affect their ability to pursue their material or ideational goals.

The rationalist approach stands in distinction to a historical or sociological approach to institutions.⁸³ These approaches see institutions as more deeply rooted and

81. Young 1979.

82. See Alter 1996; and Conforti 1993.

83. See Steinmo, Thelen, and Longstreth 1992; and Pierson 1996b. Historical institutionalism stresses the path-dependent nature of institutions, explaining why apparently inefficient institutions persist. Socio-

draw attention to their unanticipated consequences. Although we may question whether many international institutions reach the same degree of “taken-for-grantedness” that we see in domestic politics or smaller-scale social relations, it seems undeniable that they sometimes have effects that surprise their member states. It is important to differentiate between unintended and unanticipated effects. Effects may be anticipated but unintended. For example, it is generally expected that arrangements to lower the rate of inflation will lead to somewhat higher levels of unemployment. Thus, higher unemployment is an anticipated, although unintended, consequence of stringent monetary policies. It is best understood as a price actors are sometimes willing to bear to gain the benefits of low inflation. Such unintended but anticipated consequences of institutions present little challenge to a rationalist approach, since they fit neatly into a typical cost-benefit analysis. Genuinely unanticipated effects, however, present a larger challenge.

Specific examples of apparently unanticipated consequences of international institutions are not difficult to find. States that believed that human-rights accords were nothing but meaningless scraps of paper found themselves surprised by the ability of transnational actors to use these commitments to force governments to change their policies.⁸⁴ In the European Community, few anticipated that the ECJ would have the widespread influence on policy that it has.⁸⁵ Prime Minister Margaret Thatcher was apparently quite surprised at the results of agreeing to change voting rules within the European Community, such as the adoption of qualified-majority voting, which she accepted in the Single European Act.⁸⁶

How might a rationalist approach deal with these events? One productive approach might be to attempt to specify the conditions under which unanticipated consequences are most likely. This specification would at least allow us to suggest when a simple rationalist model will provide substantial explanatory leverage and when it might become necessary to integrate the insights of other schools of thought. If unanticipated consequences dominate political outcomes, we would have to draw on alternatives to rationalist models in a way that goes far beyond using them as a way to specify preferences and goals. Here, we begin specifying when unanticipated consequences are most likely to confound patterns of international cooperation.

Inductively, it appears that changes in secondary rules—that is, rules about rules—are the changes most likely to work in unexpected ways. Changes in voting rules within an institution, for example, can give rise to new coalitions and previously suppressed expressions of interest, leading to unpredicted policy outcomes. Changes in decision-making procedures can have even more widespread and unexpected effects if they open the policy process to input from new actors. Many examples of unanticipated consequences arise from decision-making procedures that provide access to nongovernmental and transnational actors, as, for example, Kathryn Sik-

logical institutionalism emphasizes the social nature of institutions, stressing their role in defining individuals’ identities and the fact that many important institutions come to be taken for granted and therefore not seen as susceptible to reform.

84. Sikkink 1993a.

85. Burley and Mattli 1993.

86. Moravcsik 1991.

kink's work has shown.⁸⁷ Both as sources of new information and as strategic actors in their own right, such groups are often able to use new points of access to gain unexpected leverage over policy. Changes in decision-making rules will have widespread effects on a variety of substantive rules and are thus more likely to have unanticipated effects on outcomes than changes in substantive rules themselves. If this observation is correct, we should see more unanticipated consequences in situations that have relatively complex and permutable secondary rules, such as legalized institutions. Traditional state-to-state bargaining with a unit veto, which has little secondary rule structure, should provide less opportunity for nonstate actors or coalitions of the weak to influence outcomes unexpectedly.

One question that often arises, especially in the international arena, is why governments are willing to live with unanticipated outcomes. After all, participation in international institutions is voluntary. If unpleasant and unexpected outcomes frequently occur, states as sovereign actors retain the right to pull out of institutions. Why might they choose to remain in? The trivial answer is that the benefits of remaining in are greater than the costs. But we can turn this answer into something non-trivial by thinking about the conditions when institutional membership is likely to provide the greatest benefits. Some of these have been spelled out in functionalist theory. Keohane argues that the demand for international institutions will be greatest under conditions of interdependence, when states face a dense network of relations with one another and where information is somewhat scarce.⁸⁸ We could generalize that states are least likely to be willing to withdraw from an institution in the face of unanticipated consequences when they are dealing with issues that exhibit increasing returns to scale, which, in turn, create conditions of path dependence. Consider the creation of regional trading arrangements in the 1990s. These arrangements provide their members with economic benefits, and those on the outside of the arrangements find themselves losing investment and trading opportunities. We therefore see eastern European, Caribbean, and other states clamoring to become members of the relevant regional trading arrangements. This is a good example of how increasing returns to scale create a high demand for institutional membership. Under these conditions, it seems likely that these states will be willing to put up with a high level of unexpected outcomes before they would seriously consider withdrawing from an institution. However, this example begs the question of whether trade agreements are likely to have substantial unanticipated effects. They are only likely to do so in the case of rapid technological change or large international economic shocks, such as the oil shocks of the 1970s.

Typology of Institutional Effects

As we turn our attention to the problem of how, not just whether, international institutions matter, it becomes essential to understand alternative mechanisms through which institutions might exert their effects. To prod our thinking in this direction, we

87. Sikkink 1993a.

88. Keohane 1983a.

introduce a preliminary typology of institutional effects. The reasoning behind this typology is that different institutions, or perhaps similar institutions in different settings, will have different types of effects. Specifying these effects will not only allow us to develop better insights into the causal mechanisms underlying the interaction between institutions and states or societies. It will also provide for more testable propositions about how and when we should expect institutions to exert substantial effects on behavior.

The typology we suggest is analytically informed but aims first to provide a language for describing patterns of change in state behavior after creation of an international institution. Here we spell out the typology and present some illustrative examples. The next step will be to link the typology to causal processes, and we suggest some preliminary ideas along these lines. We begin by suggesting two types of institutional effects: *convergence* and *divergence* effects. Of course, the null hypothesis is that institutions have no effect. Development of a clearer analytical framework may force us to consider situations in which we combine effects: for example, perhaps some types of states are subject to convergence effects and others to divergence effects.

We begin with convergence effects, since the logic of most rationalist, economic, and functionalist theories of international institutions leads us to expect such effects. These models posit goals that states find it difficult to achieve on their own, whether for reasons of time-inconsistent preferences, collective-action problems, old-fashioned domestic political stalemate, or other failures of unilateral state action. In this functionalist logic, states turn to international institutions to resolve such problems; institutions allow them to achieve benefits unavailable through unilateral action of existing state structures. Functionalist analysis sees international institutions as important because they help states to solve problems. Many of these problems have their roots in the failures of domestic institutions, and their resolution involves turning some types of authority over to the international level. Once policy is delegated to an international institution, state behavior will converge: members will tend to adopt similar monetary, trade, or defense policies.

What has been missing from functionalist accounts of institutionalization is the systematic connection between domestic political conditions and incentives to construct and comply with international institutions. But once we recognize that international institutions may make a difference because they effectively substitute for domestic practices (making policy decisions, setting policy goals, or undertaking monitoring activities), our attention turns to the domestic political conditions that make such substitution a reasonable policy alternative. If domestic institutions are the source of persistent policy failure, if they somehow prevent the realization of societal preferences, or if they interfere with the pursuit of mutual benefits with other states, turning functions over to the international level can enhance national welfare.⁸⁹ Monetary policy is a prime example of this logic. Other examples might

89. Some would argue that this process is antidemocratic. See Vaubel 1986. However, such an argument rests on weak foundations. First, it assumes that domestic institutions are necessarily responsive to

include trade policy, if domestic trade policy institutions are captured by protectionists; or environmental policy, if domestic institutions encourage a short-term rather than a long-term perspective on the problem. Thinking about the logic of substitution requires much more attention to inefficient domestic politics than most functional theories have provided to date.

A classic example of international institutions acting as substitutes for domestic institutions and therefore having convergence effects lies in arguments about why high-inflation states such as Italy might choose to enter the European Monetary Union (EMU).⁹⁰ High inflation is a public bad, leading to lower overall welfare than low inflation. However, the short-term benefits to politicians from allowing spurts of unanticipated inflation make it difficult to achieve low rates of inflation unless institutions that set monetary policy are independent of political influence.⁹¹ Thus, transferring authority to an institution that is relatively insulated from political influence, and that itself has a preference for low inflation, can provide overall welfare benefits for the country. This is the logic that leads a state like Italy to take the unusual step (for a relatively rich, developed country) of transferring a core aspect of sovereignty—control over the currency—to a European Central Bank.

Given this logic of delegation, states that become members of the EMU should see a convergence in their rates of inflation.⁹² Although the debate rages among economists about whether the European Monetary System has in fact worked in this manner,⁹³ there is little doubt that one of the major motivations for monetary union is for high-inflation states to “import” low German rates of inflation, leading to similar inflation rates in all member states. If we looked at the variation in inflation rates prior to entry into monetary union (or into a monetary system more generally), and compared it to inflation rates after entry, we should see a decline in the level of variation.

Although monetary union is a prominent and intriguing example of convergence effects, we can imagine a similar dynamic in other issue areas as well. Environmental institutions should lead to convergence of environmental indicators, such as carbon dioxide emissions.⁹⁴ Human-rights institutions acting as substitutes should lead members to adopt increasingly similar human-rights practices. Even if full convergence does not occur, the major effect of an institution that is acting as a substitute will be to bring state practices more closely in line with one another.

A convergence effect could be measured and identified by decreased variation in relevant indicators of state practices, whether inflation rates, pollution, or human-

national preferences. For the kinds of reasons just discussed, such as time-inconsistent preferences, or institutional capture, this assumption is often false. Second, the argument assumes that international institutions are necessarily more difficult to monitor, constrain, and influence than domestic institutions. Although this may be a reasonable assumption for some kinds of societal actors and some states, it is not universally true.

90. For a contrasting argument on the logic of EMU, see Gruber 1996.

91. Rogoff 1985.

92. Fratianni and von Hagen 1992.

93. See Giavazzi and Giovannini 1989; and Weber 1991.

94. Levy 1993.

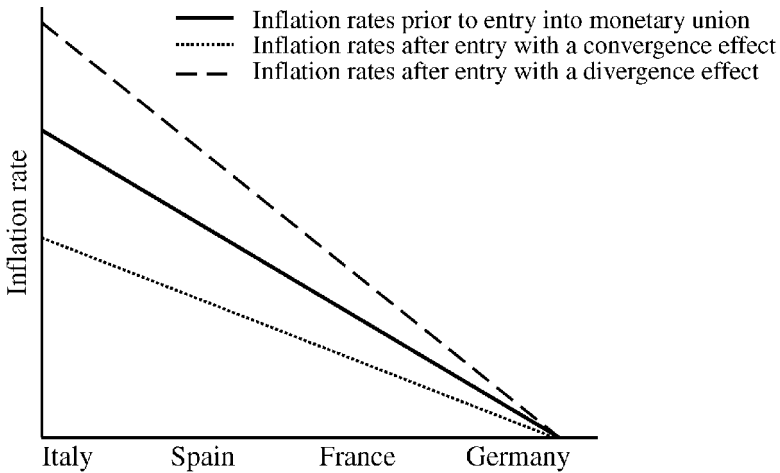


FIGURE 1. *International institutions with convergence or divergence effects*

rights abuses. The existence of a convergence effect could also be identified through graphical means. Figure 1 gives an example. On the *x*-axis, states are arrayed in order of their performance on the outcome dimension, say inflation rates. These rates are indicated on the *y*-axis. The solid line indicates inflation rates prior to entry into monetary union. Its steep slope indicates that the states exhibit substantial variation in inflation rates. The dotted line represents the outcome of monetary union acting as substitute, causing convergence in inflation rates. The more shallow slope indicates less variation than observed before entry into monetary union.

The notion that international institutions might substitute for domestic ones underlies functionalist theories of institutions. However, some empirical work on the effects of institutions has found a pattern quite different from the convergence of outcomes predicted by such a mechanism. Instead, some authors have found that the primary effect of institutions is to exaggerate preexisting patterns of behavior. For example, Andrew Moravcsik has found, in a regional comparison of human-rights institutions, that these institutions only led to an improvement in practices in those states that already exhibited a high level of respect for human rights.⁹⁵ Thus, West European states, through participation in institutions, have improved their already very good human-rights records, whereas Latin American states, according to his evidence, show little impact of institutional participation.

This pattern suggests that international institutions sometimes lead to divergence of state practices, in effect complementing and magnifying preexisting tendencies rather than overriding them. In this case, institutions will have a divergence effect. This effect results when states whose initial practice falls far from institutional guidelines will show little change from behavior, whereas those near the guidelines move

95. Moravcsik 1995.

even closer to them. In contrast, a convergence effect appears when institutions exert their greatest influence on precisely those states whose behavior deviates substantially from institutional norms. Divergence is likely to emerge when institutions exaggerate domestically generated tendencies of state behavior or when they primarily mimic domestic institutions. Anne-Marie Slaughter has argued something along these lines in pointing out that liberal states are the ones most likely to create and abide by relatively liberal international institutions.⁹⁶ According to this logic, liberal institutions will change the behavior of liberal states but not illiberal ones, leading to divergence of state behavior.

A divergence effect means that those states that already come close to institutional norms will move further toward them, whereas the behavior of those that deviate from such norms will remain unchanged. If we were to develop a measure of state behavior, we would see a divergence effect in increased variation of state behavior after institutional creation. We can also illustrate divergence effects graphically, as in Figure 1. Here, institutional effects result in a steeper line, indicating greater divergence in the relevant outcome variable. For ease of comparison, we continue to use the EMU-inflation example. Although such an outcome seems unlikely in practice, for the sake of argument we could imagine that monetary union that allowed for decentralized, unconstrained fiscal policymaking while providing additional resources to cover national debts could lead to such a perverse outcome. Another, perhaps more plausible, example of a divergence effect is in the area of overseas development aid. In the 1970s, OECD countries agreed to devote a set percentage of their GDP, 0.7 percent, to development assistance. Although some countries have come close to providing this level of aid and use the target figure as a tool in domestic debates, others have wholly neglected this target and instead decreased the percentage of their national income that they devote to foreign aid.

If this typology provides a useful way to describe alternative institutional effects, the next challenge is to begin to link up these patterns of behavior to alternative causal mechanisms. This project appears promising, and we outline preliminary ideas here. As suggested earlier, institutions that lead to convergence of state behavior link up nicely to the functionalist approach that has dominated studies of international institutions, regimes, and organizations over the last fifteen years. In this situation, the failure of domestic institutions or of unilateral state action creates incentives to rely on international mechanisms. The kinds of problems that would prompt states to use international institutions that lead to convergence of behavior are relatively well understood. They include time-inconsistency problems that create incentives for states to bind themselves and collective-action problems among states or within polities. When states turn to international institutions as the result of such problems, and when these institutions are operating as intended, we would expect to see convergence of state behavior.

The conditions that would prompt states to use institutions that lead to divergence of behavior are not as well understood. We can begin by noting that states facing

96. Slaughter 1995.

collective-action problems, such as a PD or a coordination game, would be unlikely to rely on an institution that exaggerated differences in state behavior. The fundamental problem in such cases is to create incentives for states to adopt similar policies: free trade, stringent fiscal policies, arms control, and so on.⁹⁷ In such a situation, an institution that led to increased divergence of state practice would quickly become irrelevant as states ignored its constraints. Thus, one initial expectation is that institutions should not lead to divergence in situations where incentives exist to adopt similar policies, as when strong externalities to divergent or unilateral state behavior exist. Perhaps this helps us understand why we appear to see some divergence effects in the human-rights issue area. Although human rights are a matter of concern around the globe, human-rights practices usually do not involve the kinds of externalities and incentives for strategic interaction that exist in issue areas such as the environment or monetary policy.⁹⁸

However, lack of externalities does not provide a direct answer to why divergence would occur. To understand this effect, it is likely that we need to consider domestic politics, returning us to an argument made earlier in this article. International agreements, even those without enforcement mechanisms such as the OECD aid target, can provide “hooks” by which interest groups that favor the international agreement can increase their influence on the domestic agenda. For example, in Scandinavian countries the OECD target has become a potent arguing point in parliamentary debates. In states without a well-organized group to grab onto this hook, or in those with a more closed political process, agreements without enforcement mechanisms or substantial pressure from other states to comply are unlikely to have any effect. These contrasting domestic political dynamics are likely to give rise to divergence of state behavior among members of the institution.

A rationalist research agenda for the study of international institutions is rich and promising. This agenda begins by recognizing that, in equilibrium, institutions are both causes and effects, and that empirical researchers must begin to consider the question of how institutions matter, not just whether they do. Thinking in these terms turns our attention to the problem of how institutions might resolve bargaining and distributional conflict as well as the more recognized problems of cheating. It forces us to differentiate anticipated from unanticipated effects of institutions and to ask about the conditions under which unanticipated effects are most likely. Rationalist theories provide a mechanism for bringing domestic politics more systematically into the study of international institutions, an area of research that has been slighted by the development of the field thus far. Finally, a rationalist approach allows us to

97. There may be some coordination situations, for example, some discussed by Simmons, in which the solution to the coordination problem does not involve adoption of similar policies by all states but clear division of responsibilities among states. See Simmons 1994. The Bretton Woods systems, for example, coordinated state behavior by creating expectations that the United States would behave differently from other members of the system.

98. Donnelly surveys the landscape of human rights regimes. See Donnelly 1986. There may be exceptions to the generalization that international strategic interaction on human rights is minimal, for example, when severe human-rights abuses lead to massive refugee flows. This kind of logic could lead to testable propositions within the issue area, for example, that institutions should function differently when such externalities exist than under “normal” circumstances.

begin differentiating between different types of institutional effects and developing refutable propositions about the conditions under which we are most likely to observe such effects.

Conclusions

Studies of international institutions have varied in their theoretical sophistication and frequency over time but have remained a staple of international relations research and the pages of *IO* over the last fifty years. In this article we have examined the development of these studies and outlined some promising directions for future research on international institutions. Early studies of institutions were very much problem-driven, focusing on the problems of the postwar world that some hoped international organizations could solve. Although on balance realistic and insightful, the results of these studies failed to cumulate, likely due to the lack of a disciplinary or theoretical framework in which to situate the studies. A more scientific approach showed itself in a newer wave of work on institutions, drawing on methods and models of American politics. But because these models were in general poorly suited to the realities of international politics, they failed to generate substantial new insights. It was not until the 1980s, with the development of work on international regimes and functionalist theories, that a more progressive research program on institutions arose.

One failing of the current research program, however, has been its intense focus on proving that institutions matter, without sufficient attention to constructing well-delineated causal mechanisms or explaining variation in institutional effects. We consider two approaches that might move research beyond this impasse. First, we ask whether applying recent models of domestic politics might be more successful than have past attempts. We find scope for optimism here, since modern theories of domestic institutions typically draw on similar assumptions of unenforceable agreements and opportunistic behavior by individuals that characterize most work in international relations. Finally, we turn to some more specific research directions that are likely to give rise to important and testable propositions. These include more careful consideration of distributional issues, the role of domestic politics, unanticipated consequences, and a typology of institutional effects.

As we consider international institutions as both objects of strategic choice and consequential, allowing them to serve as both dependent and independent variables in our models, the potential for increasing our understanding of institutions and of international politics in general is substantial, as preliminary empirical work has begun to show. The earliest work on international institutions produced insights that failed to add up to much because of the lack of an analytical framework in which to situate these insights; the next generation of work had the benefit of such a framework, but one that was poorly suited to the task at hand. In this article, we hope to have identified lines of research that will combine the best of both worlds: theoretically grounded research on institutions that draws on assumptions that are appropriate for the persistent problems of international relations.