Theories of international regimes
Stephan Haggard and Beth A. Simmons

Over the last ten years, international regimes emerged as a major focus of empirical research and theoretical debate within international relations. The interest in regimes sprang from a dissatisfaction with dominant conceptions of international order, authority, and organization. The sharp contrast between the competitive, zero-sum "anarchy" of interstate relations and the "authority" of domestic politics seemed overdrawn in explaining cooperative behavior among the advanced industrial states. The policy dilemmas created by the growth of interdependence since World War II generated new forms of coordination and organization that fit uneasily in a realist framework.

Intellectual traditions emphasizing the "societal" dimension of international politics suffered, however, from a lingering taint of idealism. Realism questioned the importance of international law as a constraint on state behavior and by the 1970s, its positive study by political scientists was virtually moribund. The subfield of international organization, and particularly the study of regional integration, generated rich theoretical debates during the 1960s. Yet the field remained closely tied to the study of formal organi-

We are indebted to Don Babai, David Baldwin, Esther Barbe, Jack Donnelly, Jeff Frieden, Barbara Geddes, Joe Grieco, Ernst Haas, Peter Haas, Stanley Hoffmann, Peter Katzenstein, Robert Keohane, Stephen Krasner, Fritz Kratochwil, Charles Kupchan, David Lake, Doug Nelson, Eric Nordlinger, Helen Milner, Andy Moravcsik, Craig Murphy, Joseph Nye, Robert Putnam, John Ruggie, Raymond Vernon, and two anonymous reviewers for their comments.

organizations, missing a range of state behavior that nonetheless appeared regulated or organized in a broader sense. Few strong theories started from the assumption that, as John Ruggie put it in a seminal article, "international behavior is institutionalized."  

Regime analysis attempted to fill this lacuna by defining a focus that was neither as broad as international structure, nor as narrow as the study of formal organizations. Regime analysts assumed that patterns of state action are influenced by norms, but that such norm-governed behavior was wholly consistent with the pursuit of national interests. Hence, the regimes literature can be viewed as an experiment in reconciling the idealist and realist traditions.

After a decade of development, it is time to submit this research program to critical review. A plethora of contending theories have explained regime creation, maintenance, and transformation, but the relationship among them is unclear and empirical research has yet to determine which are the more plausible. While earlier work on regimes focused on interdependence, the widening variety of state goals, and the importance of non-state actors and international organizations, recent work on regimes and international cooperation unfortunately reverts to an approach which treats states as unified, rational actors. In addition, little research has addressed whether, and in what ways, regimes "matter."  

Do regimes have independent influence on state behavior and, if so, how?

We begin by briefly surveying the contending definitions of international regimes, which range from patterned behavior, to convergent norms and expectations, to explicit injunctions. We then suggest a number of dimensions along which regimes vary over time or across cases; these dimensions have been or might be used to operationalize "regime change." They include strength, organizational form, scope, and allocational mode. The third section examines four theoretical approaches to regimes—structural, game-theoretic, functional, and cognitive—and attempts to clarify what each theory can and cannot tell us about regimes.

In the conclusion, we ask how and whether regimes "matter." So far, little research addresses this problem. Testing for the significance of regimes—even to verify international systemic theories—demands careful tracing of national level processes, structures, and values. We outline a research program that focuses greater attention on issues raised by theorists of complex interdependence; these issues have been neglected in the revival of game theory, and include the erasure of the boundaries between domestic and foreign policies, the importance of transnational coalitions, and, above

all, the way in which domestic political forces determine patterns of international cooperation.

1. Regimes: Definitions

How do we know a regime when we see one? Three answers have been offered to this question. The most comprehensive equates regimes with patterned behavior. Donald Puchala and Raymond Hopkins argue that “a regime exists in every substantive issue-area in international relations . . . Wherever there is regularity in behavior, some kinds of principles, norms or rules must exist to account for it.” 4 But the existence of patterned behavior alone should not lead one to suspect that a regime lurks below the surface. A broad definition runs the risk of conflating regularized patterns of behavior with rules, and almost certainly overestimates the level of normative consensus in international politics. 5 Deducing regimes from patterned behavior makes it difficult to decide how they mediate, constrain, or influence behavior. The term “regime” is sometimes used in a purely descriptive way to group a range of state behaviors in a particular issue-area, but since the potential for tautology is high, this approach has largely been abandoned. 6

Stephen Krasner’s influential definition seeks a middle ground between “order” and explicit commitments; it stresses the normative dimension of international politics. Krasner defines a regime as “implicit or explicit principles, norms, rules and decision-making procedures around which actors’ expectations converge in a given area of international relations.” 7 Despite the care with which this complex hierarchy of components is defined, “principles” (which include not only beliefs of fact and causation, but also of “rectitude”) shade off into norms, “standards of behavior defined in terms of rights and obligations.” Norms, in turn, are difficult to distinguish from rules, “specific prescriptions or proscriptions for action.” 8

Not surprisingly, analysts have had serious disagreements about how the "norms" and "principles" of a given regime should be defined; their differences have strong implications for interpretations of regime stability and change. Susan Strange, John Zysman, and Stephen Cohen argue that state practices such as selective safeguards and industrial targeting have undermined the fundamental premises of the trade regime.9 John Ruggie, on the other hand, contends that these derogations are consistent with a broader ideological framework, which he calls "embedded liberalism," which will endure as long as there is some accommodation of economic efficiency to social stability.10 It is exceedingly difficult to imagine any state abandoning this "norm"; the result is a bias in favor of seeing continuity and a stable order.

The explicitness of commitment required before a regime can be said to exist creates a related problem. Focusing on "implicit regimes" captures the convergence of actor expectations and may help us summarize a complex pattern of behavior. But it begs the question of the extent to which state behavior is, in fact, rule-governed. Robert Keohane's effort to trace postwar cooperation in oil and Charles Lipson's study of changes in the treatment of foreign investment exemplify unnecessary, if not misleading, uses of the regime concept.11 In both cases, the rights and duties of different actors were open to significant question and conflict. Violations of the norms governing the treatment of foreign investment in the 19th century, for example, really violated European (and primarily British) expectations. Do we gain explanatory leverage in analyzing 20th-century investment disputes when we point out that "'national treatment' [became] an international norm'?12 The core of Lipson's argument is not normative; national policies towards

11. Keohane, in After Hegemony, notes that a formal international regime was never established in oil (p. 140), but later refers to the postwar oil regime (p. 190). Charles Lipson, Standing Guard: Protecting Foreign Capital in the Nineteenth and Twentieth Centuries (Berkeley: University of California Press, 1985), p. 32, includes in his object of analysis both patterns of behavior and normative structures; the concept of regime is not central to his analysis, and is used primarily as a summary of state behaviors. This is true also of Robert E. Wood, From Marshall Plan to Debt Crisis (Berkeley: University of California Press, 1986). For a defense of this usage, see Jack Donnelly, "International Human Rights Regimes," International Organization 40 (Summer 1986), pp. 599–642.
12. Lipson, Standing Guard, p. 262, emphasis in original.
foreign investment are neatly explained in terms of the investing countries’ relative power and the hosts’ regulatory capabilities.

The “oil regime” between 1945 and 1970 consisted of the market activities of oligopolistically interdependent firms, the national rules of the producers, and ad hoc interventions by the United States. The Seven Sisters no doubt followed certain “rules” concerning production, exploration, and marketing. But their behavior must be understood in terms of supply conditions, a particular market structure, and national regulatory environments. The behavior of the major powers seems even less constrained by international rules. How “convergent” were British, French, and American interests during the Suez crisis, in which the United States’ diplomatic use of the oil weapon played a key role? Clear American interests and the power to back them adequately explain the outcome of the crisis. By endorsing the idea of “tacit” or “implicit” regimes, Keohane overemphasizes the importance of convergent expectations, particularly among governments.

A third, more restricted definition treats regimes as multilateral agreements among states which aim to regulate national actions within an issue-area. Regimes define the range of permissible state action by outlining explicit injunctions. Regimes often contain rules which govern or specify their own transformation, but to explain “regime change” per se is to explain why states would agree to modify the codified rights and rules that regulate their behavior. This approach risks the charge of formalism—a charge which has plagued the study of international law. On the other hand, it focuses attention on the evolution of the texts constituting international agreements; it also clearly separates normative consensus from the definition of regimes, treating it rather as a causal or constitutive variable that may be useful in explaining cooperation.

This definition also allows a sharper distinction between the concept of regime and several cognates, such as cooperation. Regimes are examples of cooperative behavior, and facilitate cooperation, but cooperation can take place in the absence of established regimes. A recent example was the package of measures adopted by the advanced industrial states at the 1978 Bonn summit. Regimes must also be distinguished from the broader concept of “institutions,” the essential feature of which is “the conjunction of


14. See particularly Aggarwal, Liberal Protectionism, p. 16.

15. On cooperation, see Keohane, After Hegemony, pp. 51–52.

convergent expectations and patterns of behavior or practice." 17 Regimes aid the "institutionalization" of portions of international life by regularizing expectations, but some international institutions such as the balance of power are not bound to explicit rights and rules. "Convergent expectations" may or may not be tied to explicit agreements—they might, in fact, arise in a milieu characterized by substantial conflict. Finally, we should distinguish between regimes and "order" or "stability." Regimes may facilitate order and stability but are not coterminous with them. 18 In some instances, regimes may unintentionally contribute to instability, as when commitments to maintain parities under the Bretton Woods regime in the late 1960s produced chaotic exchange markets.

2. Regimes: Dimensions of variance and change

Many studies fail to specify what they mean by regime transformation or treat it in a unidimensional way. Regimes may change over time or vary across cases in at least four ways: strength, organizational form, scope, and allocational mode. As we shall argue, different theoretical approaches address one or more of these variables, but are less useful in explaining others.

a. Strength

The majority of "regime change" studies try to explain why regimes eventually weaken or decay. 19 Strength is measured by the degree of compliance with regime injunctions, particularly in instances where short-term or "myopic" self-interests collide with regime rules.

b. Organizational form

In its quest to move beyond the study of concrete international organizations, recent regimes literature has largely ignored problems of organizational design and operation. Some issues are conducive to decentralized regulation: regime injunctions may only call on states to share information, or to refrain from certain actions, such as polluting, over-fishing, nuclear testing, or raising tariffs. Other regimes, such as a fixed-exchange-rate re-

17. Oran Young, Compliance and Public Authority, p. 16; and his "International Regimes: Toward a New Theory."


19. Explaining regime strength is central to the analysis of Lipson, Standing Guard; Aggarwal, Liberal Protectionism; Keohane, After Hegemony; and Stephen Krasner, Structural Conflict: The Third World Against Global Liberalism (Berkeley: University of California, 1985).
regimes, demand positive interventions by states, but remain largely decentralized.

Most regimes, however, are likely to have at least some minimal administrative apparatus for the purpose of dispute settlement, the collection and sharing of information, or surveillance. Complex cooperative tasks require more elaborate, and potentially autonomous, organizational structures. If cooperation is already highly institutionalized, theories resting on assumptions of anarchy are highly misleading; blackboxing organizational structure and processes will lead to simplistic predictions.20

The principles governing representation are another dimension of organizational variance. Most universalist regimes are structured either on the "one nation, one vote" principle or, as in the International Monetary Fund (IMF) and World Bank, on weighted voting. Alternative principles of membership, however, are based on discrimination along functional or sectoral lines (the Tokyo Round codes) or regional ones (the Caribbean Basin Initiative). Principles of membership have important distributional consequences, since they affect international agendas and organizational resource allocation.

c. Scope

Scope refers to the range of issues the regime covers. Though changes in regime scope have attracted little theoretical attention, its neglect can cause misleading characterizations. The failure to comply with certain GATT (General Agreement on Tariffs and Trade) provisions signaled a weakening of the trade regime in the 1970s. Yet at the same time, the regime’s scope expanded through the negotiation of the Tokyo Round codes.21 The most contentious questions on the current trade agenda concern the regime’s scope—namely, how the GATT will address new issues such as trade in services, industrial policy, and national rules governing foreign direct investment.

The jurisdictional scope of a regime is not incidental to its success. Overly broad jurisdiction raises administrative costs and complexity, but overly narrow agreements may allow little room for bargaining and issue-linkage. One important cause of regime change is the “externalities” associated with


inadequate scope. GATT negotiations in the 1950s and 1960s virtually eliminated tariff barriers as an important impediment to trade. The result, however, was to expose, and even encourage, non-tariff barriers. These externalities drove the reform efforts which culminated in the Tokyo Round.

d. Allocational mode

Regimes can endorse different social mechanisms for resource allocation. A market-oriented regime supports the private allocation of resources, discourages national controls, guarantees property rights, and facilitates private contracting. As Oran Young states, “free enterprise systems . . . are not institutional arrangements operating outside or in the absence of any regime. Such systems clearly require explicit structures of property or use rights.” At the other extreme, authoritative allocation involves the direct control of resources by regime authorities, and will demand more extensive, and potentially autonomous, organizational structures. The IMF’s role in the balance-of-payments financing regime provides an example.

The nature of the issue-area and the extent of cooperation sought will partly determine the preference for market-oriented versus authoritative modes of allocation. Many issue-areas could be organized either way, however, with sharply different distributional consequences. The Group of 77’s proposal for a New International Economic Order (NIEO) provides the clearest example. Virtually all of the NIEO debates centered on allocation mechanisms, with the South generally favoring authoritative ones.

3. Theoretical approaches to regime change and variance

The literature on regime development and change can be grouped into four families: structural, game-theoretic, functional, and cognitive. These categories are not mutually exclusive, and the most persuasive interpretations are likely to draw from more than one theoretical tradition. Yet these ap-

22. Stephen Krasner makes this point in Structural Conflict.
23. Oran Young, Compliance and Public Authority, p. 55.
25. Structural models have been given game-theoretic treatment by David Lake, “Beneath the Commerce of Nations: A Theory of International Economic Structures,” International
approaches often speak past one another—in part because of fundamental differences in their underlying assumptions, in part because they address different dimensions of regime change and variance. We can gain a great deal by being explicit about underlying assumptions and what each approach can and cannot explain.

Most structural, game-theoretic, and functional theories of regimes are state-centered, presuming unified rational actors, even if the assumption is relaxed to gain explanatory leverage. Structural explanations, particularly including the theory of hegemonic stability, attempt to show how international conditions define the possibilities for cooperation. Structuralists argue that we cannot infer national policies from intentions because structures tend to mold state behavior "toward a common quality of outcomes even though the efforts and aims of agents and agencies vary." Game-theoretic approaches incorporate exogenously determined preference orderings into the analysis. While these preference rankings, in principle, include all domestic factors that may impinge on a state’s overall preferences, most research emphasizes that actors are primarily constrained by the structure of the interstate game. Functional theories, of which Robert Keohane’s After Hegemony is a leading example, also assume rational actors, but introduce market imperfections, transactions, and information costs and uncertainty.

We argue that all three approaches downplay the central insight of interdependence theorists: foreign policy is integrally related to domestic structures and processes.

A group of theorists we have labeled (for want of a better term) "cognitivists" have suggested a radically different research program. Focusing on the intersubjective meaning structures that bind actors together, they necessarily see a looser fit between structural constraints, interests, and choices. Where functional theories see regimes as more or less efficient responses to fixed needs, cognitive theories see them as conditioned by ideology and consensual knowledge and evolving as actors learn. Cognitivists argue that "there is no fixed ‘national interest’ and no ‘optimal regime.’" Presumably determinative "structural" constraints must always be understood in terms of historically conditioned, interpretive frameworks. This requirement promises some corrective to the spare assumptions of other theoretical approaches, particularly by reminding us of historical context and the substan-
tive issues over which conflict and cooperation occur. Yet, to date, the cognitivists have not adequately resolved their relationship with the structuralist research program. The relationship between power, ideology, and knowledge is one of the most exciting areas of theoretical debate, but one strewn with basic methodological problems.29

a. Structuralism: The theory of hegemonic stability

The theory of hegemonic stability offers the most parsimonious and widely employed explanation of regime dynamics; it links regime creation and maintenance to a dominant power’s existence and the weakening of regimes to a waning hegemon. Ironically, the theory’s early proponents were interested primarily in the world economy’s stability and openness and did not mention the role of rules or regimes.30 In Power and Interdependence, Keohane and Nye first linked regime dynamics to waxing and waning state power. They argued that, since power is not wholly fungible under conditions of complex interdependence, linkage across issue-areas is not likely to be effective. Keohane and Nye proposed a more disaggregated “issue-structural” model, predicting that “stronger states in the issue system will dominate the weaker ones and determine the rules of the game.”31

The theory of hegemonic stability has subsequently been challenged on both empirical and theoretical grounds.32 First, the hazards of power analy-

29. The social theorist most interested in the relationship between knowledge and power is Michel Foucault. See particularly Colin Gordon, ed., Power/Knowledge: Selected Interviews and Other Writings by Michel Foucault, 1972–1977 (New York: Pantheon, 1980).


31. Keohane and Nye, Power and Interdependence, pp. 50–51.

sis plague efforts to explain concisely the structure of regime dynamics. Inferring interests from capabilities implies that there is some unambiguous way to assess the distribution of capabilities. From this structure flows a greatly restricted set of possible international outcomes. The delineation of the relevant structure, however, is bound to be arbitrary. David Lake offers the most careful conceptualization of international structure within the trade area, including a measure of relative productivity as well as relative size. These measures seek to explain both the capacity to underwrite or support free trade and the willingness to do so. But as Jeff Frieden points out (using Lake’s data), the shift in the United States’ share of world trade is only minuscule between 1929, when the United States refused to lead, and 1960, the high point of American hegemony. Can a move from 13.9 percent of world trade in the earlier period to 15.3 percent in the latter account for the sharp contrast in trade practices?

Second, since structure alone is a poor predictor of regime characteristics and national policies, “structural” theories must continually revert in an ad hoc way to domestic political variables. As Kindleberger recognized, America’s refusal to lead in the interwar period had a domestic foundation. Similarly, Krasner found that structure predicted the openness of the trading system in only half the cases he examined, with domestic institutional “lags” introduced to explain the rest. If such “slippage” is common, or if similarly situated states tend to respond differently to international con-

mon’s Dilemma: Great Britain, the United States, and the International Economic Order,” International Organization 38 (Spring 1984), pp. 355–86. More positive assessments of the theory’s utility can be found in William P. Avery and David P. Rapkin, eds., America in a Changing World Political Economy (New York: Longman, 1982), though most of the pieces address the question of U.S. policy, as opposed to regime outcomes.


Finally some have suggested that the theory is simply misspecified. Bruce Russett and Susan Strange argue that American decline has been exaggerated; see Russett, “The Mysterious Case of Vanishing Hegemony, or Is Mark Twain Really Dead?” International Organization 39 (Spring 1985) pp. 202–31; and Strange, “Cave! Hic Dragones.” Snidal (in “Hegemonic Stability Theory”) and Russett both argue that the theory has been misapplied to cases where provision of a public good is not at stake.

straints, than the primacy of structural theory is called into question. We need theories of domestic processes or theories linking the international and domestic levels.

Hegemonic interpretations of regimes are not always clear about what hegemons actually do to promulgate and maintain a given set of rules; structural theories are, by their nature, less useful for understanding processes than for establishing correlations. The two major schools of thought on the mechanics of hegemonic leadership stress different solutions to the provision of public goods.36 The "malign" view, associated with Gilpin's work, sees the hegemon providing coercive leadership. The hegemon enforces regime rules with positive and negative sanctions; it even extracts payment from smaller states to maintain the regime. Thucydides describes, for example, how Athens's "leadership" of the Delian League, based on consensus, was gradually transformed into coercive empire.37 This interpretation of hegemony would cast doubt on the assumption that regimes are "cooperative" institutions, even though benefits might accrue to the coerced actors.38 Empirical studies suggest that we cannot assume the hegemon will enforce regime compliance, however. Timothy McKeown's study of 19th-century trade concludes that European liberalization had little to do with pressure from England, but rather was the result of national policies towards industrialization.39 Similarly, the United States tolerated derogations from "liberal" norms in the first two postwar decades. Only when its relative power eroded did the United States insist on "reciprocity."

The "benign" view of hegemony turns realism on its head.40 Because they benefit greatly from a well-ordered system, hegemons are willing to provide

36. Snidal, "Hegemonic Stability Theory."

international public goods. More precisely, the hegemon constitutes a "privileged group," for which the cost of supplying the public good is less than the benefit. But since the hegemon has an independent interest in supplying public goods regardless of the contributions of others, beneficiaries of the system will have an incentive to free-ride. Rather than the strong exploiting the weak, it is the weak who exploit the strong.

Two criticisms have been raised against the collective goods approach to the hegemonic theory of regimes; both point to the inappropriate application of otherwise sound theory. First, most regimes do not provide pure public goods, which are characterized not only by jointness of supply but by non-excludability. Most regimes have developed mechanisms that enhance compliance through exclusion—for example, trade reciprocity, reciprocal rights of innocent passage, and technological assistance in exchange for adherence to non-proliferation agreements. Keohane and Snidal also point out that nothing in the theory of collective action suggests that a single hegemon is required to provide public goods. If a privileged subgroup exists such that each member benefits from providing the good even without cooperation from other members, the public good will be provided. Individual incentives not to contribute will increase as the size of the privileged subgroup increases, but certain characteristics of the regimes, such as those which increase the transparency of state action, may mitigate this effect.

A broader criticism may be leveled against the theory of hegemonic stability, however. The relevant "structure" is usually defined as the distribution of power within the international capitalist system rather than within the world political system as a whole; regimes are seen primarily as responses to the problems of collective action among the advanced capitalist countries rather than as an integral part of high politics and alliance solidarity. Despite all the attention lavished on "international structure" as an explanation of regime creation and maintenance, bipolarity has been all but overlooked. This hardly conforms with the views of the postwar planners who constructed the Bretton Woods order. The World Bank and IMF can be seen, and indeed were seen by the Soviet Union, as an American effort to close capitalist ranks at the Eastern bloc’s expense; contrast America’s insouciant response to negotiations for the Bank for International Settlements in 1929

41. Trade reciprocity is not an example of an exclusionary device as long as there is a most-favored nation provision. Most reductions in tariffs are negotiated by principle suppliers, however, and discriminatory treatment couched in the language of "reciprocity" has clearly grown. For a discussion of impure public goods and so-called "club goods," which are excludable at a cost, see Richard Carnes and Todd Sandler, The Theory of Externalities, Public Goods, and Club Goods (New York: Cambridge University Press), chaps. 10–12.

42. Keohane, After Hegemony; Snidal, “Hegemonic Stability Theory.”


44. See, for example, David Calleo and Benjamin M. Rowland, America and the World Political Economy (Bloomington, Ind.: Indiana University Press, 1973).
and its largely uncooperative attitude towards outstanding war debts throughout the 1920s and 1930s. The link between economic and security relations is often ignored by American students of international political economy; it is, however, a recurrent theme in the writing of Europeans. It has been highlighted in new interpretations of the emergence of the postwar Pacific order, and it is an underlying theme in the work of contemporary diplomatic historians such as Paul Kennedy.\textsuperscript{45} By ignoring grand strategy, hegemonic interpretations may have missed this century’s most important “structural” cause of cooperation among the advanced industrial states.

\textit{b. Strategic and game-theoretic approaches}

Game theory has recently been used to explain how cooperation can evolve under anarchic conditions which lack supranational authority to enforce compliance.\textsuperscript{46} Again, we must distinguish regimes from cooperation, which is clearly possible in the absence of regimes. Game theory can readily explain the conditions under which regimes might arise as an instance of cooperative behavior, and it can also suggest the conditions conducive to stable compliance, but it has difficulty explaining organizational form, scope, or change.

The attraction of game theory is its spare elucidation of a strategic interaction’s structure. The attendant risk is oversimplification, particularly given the daunting problems which surround the application and testing of its insights. Where multilateral interactions cannot be disaggregated into actor dyads, the $2 \times 2$ game may mislead more than it clarifies. The alternative—the n-person game—does not yield determinate outcomes.\textsuperscript{47} The $2 \times 2$ game also radically simplifies the choice set into unambiguous categories—cooperate and defect—though choices in most areas, such as macroeconomic coordination, are continuous.\textsuperscript{48} Games are also frequently depicted in their


\textsuperscript{47} Duncan R. Luce and Howard Raiffa, \textit{Games and Decisions} (New York: Wiley, 1957), chap. 7.

normal rather than extensive form, though game theorists have shown that this can lead one to overlook how the dynamics of bargaining and sequencing of moves can determine outcomes. The game may change in the course of negotiations; indeed, bargaining is precisely an effort to restructure preference orderings.

The greatest difficulty is in specifying preferences at the outset, which includes not only their ordering but their intensity. This effort must consider all relevant factors of the game environment: economic and technological conditions, domestic politics, transnational relations, and the rules under which the game is played. Deciding what the game is may be as difficult as solving it. One is tempted to construct the payoff in light of the outcomes, which throws into question the predictive value of the exercise.

Game theory can cope in principle with a high degree of complexity. Games may be "graduated" by introducing subgames, extended out into meta- or super-games, "linked" to games in other issue-areas or modeled in an evolutionary fashion. The result, however, is to detract from the selling point, parsimony, and to question how the theory actually explains particular cases. Empirical studies too often lapse into descriptive history followed by elaborate, and perhaps redundant, translation into game theoretical argot.

Most game theoretic studies of international cooperation and regimes have focused on the Prisoner's Dilemma (PD). PD is attractive since it can produce cooperative behavior under "realist" conditions. If play is repeated, the costs of defecting on any single move must be calculated not only with reference to the immediate payoff, but with reference to the opportunity costs associated with future interactions. Yet under assumptions of complex interdependence, the "dilemma" of PD diminishes. The very existence of a network of regimes and transnational relations among the ad-


50. The intensity of players' preferences has generally been given less attention in recent applications of game theory than their ordering. As Robert Jervis notes, however, "cooperation is more probable when mutual cooperation is only slightly less attractive than exploiting the other, when being exploited is only slightly worse than mutual competition and when the latter outcome is much worse than mutual cooperation." "From Balance to Concert: A Study of International Security Cooperation," in World Politics 38 (October 1985), p. 64.

51. For a review of this problem, see Duncan Snidal, "The Game Theory of International Politics," ibid., pp. 25–57.

vanced industrial states facilitates communication, enhances the importance of reputation, and lengthens the "shadow of the future." In its heuristic use, PD indicates why these institutions deter suboptimal outcomes; more compelling theories of how these institutions actually evolve are still in their infancy.

Most applications of game theory to the study of international cooperation have assumed unified state actors and the dominance of the international game. These presumptions are strongly challenged by earlier literature on transnational relations and bureaucratic politics. As with structural approaches, the domestic processes affecting payoff structures are frequently blackboxed as "exogenous." As we shall argue in the conclusion, one must analyze the tradeoffs a state, or more accurately, a particular political leadership, is willing and able to make between domestic and foreign games. This certainly doesn't rule out the application of game theory, but it demands that we reorient our effort towards the analysis of domestic/international linkages.

Game-theoretic approaches are strongest when they reveal the conditions which enable cooperation and stability; they say far less about whether regimes will actually arise, how they will be institutionalized, and, above all, the rules and norms which will comprise them. Functionalist logic provides a useful supplement.

c. Functional theories

Functional theory explains behaviors or institutions in terms of their effects. If regimes serve to reduce information and transaction costs among their adherents, for example, the rewards of compliance will reinforce the regime. Thus, anticipated consequences explain the persistence of the regime and compliance with its injunctions. Similarly, the modification of regimes or their weakening is likely to occur when they become "dysfunctional." Functional theories explain regime strength, particularly the puzzle of why compliance with regimes tends to persist even when the structural conditions that initially gave rise to them changes.

Functional theories have important limitations. They are not causal in a strong sense. They are better at specifying when regimes will be demanded rather than suggesting how or when they will be supplied. Regimes and

55. "Functional" here refers to a particular form of explanation, and should be distinguished from earlier functional and neofunctional theories of international organization. For a discussion of earlier work, see Ernst Haas, Beyond the Nation State, chaps. 1–5, 13–14. For a general discussion of functional explanatory logic, see Arthur L. Stinchcombe, Constructing Social Theories (New York: Harcourt Brace, 1968), pp. 80–98.
cooperation in one issue-area may arise as an unintended consequence of cooperation in some other area. For example, three countries may respond to a security threat from a fourth with a cooperative venture, yet the regime may persist because of the positive externalities generated by military cooperation. Unless functional theories suggest how a selection mechanism operates, they can easily become teleological, a problem suffered by earlier functionalist writing, which viewed international cooperation as a response to imputed "system-maintenance" or "equilibrating" functions. Collective action may not occur because of the free-rider problem or if the "need" goes unperceived. This criticism need not deter one from formulating a functionalism that operates through individual motivations, particularly where existing institutions mitigate the collective action dilemma. The regime's benefits are simply that it provides incentives to certain forms of cooperative action, for example, by offering forums in which reputation comes into play or where games can be iterated and linked. Regimes may be "supplied" when there is sufficient "demand" for the functions they perform, but the market analogy has obvious limitations. In any case, this approach seems little different from an intentional theory that begins from old-fashioned interests.

The specific functions that regimes perform naturally vary from issue-area to issue-area. Most functionalist theorizing suggests that some generalized functions are underprovided, given conditions of anarchy or market failure. Regime analysts working in this vein have drawn heavily from the economic literature on transaction costs. In this literature, organizations, or "hierarchies," evolve as solutions to the opportunism, uncertainty, information costs, measurement problems, and difficulties of contract enforcement which plague arms-length market transactions. Drawing an analogy between the market and the uncoordinated actions of states, Keohane, Oye, and Aggarwal have shown that regimes reduce transactions costs and facilitate decentralized rule-making. Aggarwal notes that "construction of a multilateral mechanism is organizationally less expensive than is the development of many bilateral contracts." Keohane notes that the marginal cost

of dealing with an additional issue will be lower with a regime, an insight that casts light on the important question of why regimes often expand in scope.\(^{61}\)

Once these functions are enumerated, the remaining positive research program which follows is unclear. Crude functionalism has been criticized for simply noting that some behavior or institution—in this case, a regime—does, in fact, perform an anticipated function. Even if we knew that every regime performed some specified set of functions, this knowledge would not explain why regimes emerge in some issue-areas and not in others.

Nor would it explain why some regimes develop impressive formal organizations, while others do not. Unlike the transactions cost literature, the institutional solution resulting from failed interstate coordination is not a “hierarchy” in Williamson’s sense. Numerous institutional solutions to a given problem of cooperation are possible, and little work has been done to show how particular contracting problems may yield particular governance solutions. In general, the “new functionalism” does not distinguish clearly between institutions and organizations, nor indicate the conditions that lead to the international development of the latter.\(^{62}\)

In addition, we are interested not only in the fact that regimes perform certain tasks, but the importance or weight regimes have in motivating and explaining state behavior. The proper test of a functional theory is not the mere existence of a regime, but the demonstration that actors’ behavior was motivated by benefits provided uniquely, or at least more efficiently, through the regime, or by reputational concerns connected to the existence of rules.

Finally, a strong liberal bias operates in the “new functionalism” just as it did in the old. Though Keohane admits that cooperation is not, in itself, a good, he tends to emphasize the functions which enhance global welfare, or at least the collective welfare of the regime’s adherents. The bias is suggested by Keohane’s premise of the “suboptimality” of market failure. But the institutions “provided” may or may not be Pareto superior, let alone optimal, and can certainly result in different distributional outcomes. The institutions that emerge in world politics are certainly more likely to reflect the interests of the powerful than the interests of the weak, a complaint repeatedly raised by commodity producers against the operation of the GATT. Aggarwal argues that regimes control large states’ behavior toward the small by reducing the need to exercise power directly. He demonstrates that regimes have occasionally kept domestic protectionist interests from


achieving their preferred solutions. Yet his analysis of the textile regime also demonstrates that most of the functions which regimes perform can be used to control states when the initial discrepancies of power are large, and that stronger players frequently ignore the "restraints" placed on them. The regime changed when it could not adequately serve the interests of the developed country importers.63 How do we decide between Cheryl Payer's Marxist functionalism, which asserts that the World Bank operates as an instrument of Northern control, and a benign, liberal functionalism, which sees it as a more-or-less optimal institutional response to "failures" in international capital markets?64 Douglass North has outlined the problem in a trenchant review of the transaction costs literature:

One cannot have the productivity of an industrial society with political anarchy. But while . . . a state is a necessary condition for realizing the gains from trade, it obviously is not sufficient. A state becomes the inevitable source of struggle to take control of it in the interests of one of the parties. The state then becomes the vehicle by which the costs of transacting are raised to capture the gains that will accrue to any interested party that can control the specification and enforcement of property rights.65

Substituting "regime" for "state" in the foregoing passage raises a problem for functionalist analysis. Functional theories emphasize how the facilitating role of regimes helps them realize common interests. But regimes are also arenas for conflict and the exercise of power. Because functional theories assume highly convergent interests and downplay divergent ones, they do not explore how regimes may institutionalize inequalities.

d. Cognitive theories: Knowledge, ideology, and regimes

"Cognitive" theories explore what structural, game-theoretic, and functional approaches bracket.66 The core cognitive insight is that cooperation

63. Aggarwal, Liberal Protectionism, p. 28.
66. "Cognitive" would best be left in quotation marks throughout, since we use the term to refer to a quite disparate group of approaches. We focus on the commonalities among those writers emphasizing ideology, belief systems, and knowledge as explanations of regime change. For other cuts at "cognitive" approaches that focus more squarely on the individual decision-maker see Deborah Welch Larson, Origins of Containment: A Psychological Approach (Princeton: Princeton University Press, 1985); Robert Jervis, Perception and Misperception in International Politics (Princeton: Princeton University Press, 1976).

cannot be completely explained without reference to ideology, the values of actors, the beliefs they hold about the interdependence of issues, and the knowledge available to them about how they can realize specific goals. Cooperation is affected by perception and misperception, the capacity to process information, and learning. While structural, game-theoretic, and functional theories assume that cooperation operates within an issue-area which is relatively unambiguous, cognitivists point out that issue-areas are never simply given. Cognitive approaches are therefore particularly important in explaining the substantive content of regime rules and why they evolve. By elevating the importance of actor learning, cognitive theories have a dynamic other theoretical approaches lack.

Cognitive approaches cannot predict at what point consensual values or knowledge will produce cooperation. Consensus still may not overcome problems of collective action. More importantly, "ends, or purposes, of action are not self-evidently derivable from the scientific understanding of relationships among variables."67 The generation of new knowledge just as easily might render a game less cooperative by exposing new incentives to defect.68 Nor can cognitive approaches argue that a particular regime is uniquely suited to realize some common values. Criticizing the hegemonic stability theory, Ruggie argues that regimes reflect not only a configuration of power, but also a configuration of dominant social purpose.69 Ruggie's concept of "embedded liberalism" tries to describe the common social purpose that arose after the Great Depression across the advanced capitalist states—the need to reconcile the advantages of liberalism and the costs of an unfettered market system. Yet this "purpose" is elastic enough to subsume a fairly wide range of "norm-governed changes," including the move from fixed to flexible exchange rates and the rise of the new protectionism during the 1970s. For this very reason, it cannot fully explain these changes.

Cognitivists argue that learning and, in a somewhat different fashion, ideology, affect international rules and cooperation by showing the merit (or futility) of certain lines of action. Knowledge and ideology, including the knowledge provided by regimes, can alter actor interests. But the cognitive

67. Ernst Haas, "Is There a Hole in the Whole?" p. 848.
69. John Ruggie, "International Regimes."
critique reaches deeper. Even this causal language—knowledge "affects" interests—is inappropriate, since the dichotomy between ideology and knowledge on the one hand, and interests on the other, is wholly artificial and misleading. "Interests" only emerge within particular normative and epistemic contexts and cannot be understood outside them. Cognitivists argue that the assumption of rational utility maximization, for example, is too spare to be of explanatory value, since it ignores the way that historically situated actors interpret their constraints. Nor is the problem solved simply by showing that rationality is "bounded" in various ways. Such a modification predicts deviations from "rational" decision-making in environments characterized by incomplete information, uncertainty, complexity, but still divorces the concept of interest from shared meanings. The recurrent structuralist fallacy is to expect different states or individuals to respond similarly to the same structural constraints and opportunities; much depends on past history, knowledge, and purpose.

It is very difficult to generalize from the central cognitive insight. The predictive value of cognitive theories is problematic, particularly when they emphasize the importance of consensual knowledge. Historical episodes of cooperation may be inexplicable without reference to shared knowledge and meanings, but since future knowledge is, by definition, impossible to foresee, prediction about the substantive content of cooperation is ruled out. Nonetheless, the degree of ideological consensus and agreement over causal relationships, regardless of the nature of the issue, is an important variable in explaining cooperation.

Studies taking a cognitive perspective frequently have an evolutionary, historicist, or, less kindly, post hoc flavor that highlights contingency and path-dependence. Ernst Haas summarizes the methodological biases of one cognitivist:

The type of systems theory I find useful features the inductive method in the construction of reality and uses the perceptions and actions of concrete human beings in grappling with reality as its main data. Such systems are assumed "open" in the sense that they do not tend, by definition, toward a given state, such as equilibrium. They are "constructed" in the sense that the theorist considers them as heuristic approximations rather than networks of determinative "laws" constraining choice.  

Ideaology and knowledge can be construed and combined in a number of ways, creating a wide range of possible research strategies. Writers emphasizing the role of consensual knowledge have naturally focused on areas of technical cooperation. Richard Cooper has shown how a slow and painful evolution of consensus on the causes of disease transmission was critical for

70. Ernst Haas, "Is There a Hole in the Whole?" p. 839.
international cooperation on public health in the 19th century.\textsuperscript{71} Don Babai argues that structuralist models are ill-equipped to explain the types of changes which took place at the World Bank under Robert McNamara. The evolution of knowledge, as well as intellectual fashion concerning development, shaped the Bank's lending strategy.\textsuperscript{72} Cognitive theories which explain international cooperation primarily with ideological variables are more rare. Judith Goldstein's work on U.S. trade policy, however, argues that a liberal ideology has become institutionalized in a way that systematically conditions the propensity to cooperate, even where defection might be more rational.\textsuperscript{73}

Sorting out the autonomous influence of knowledge and ideology can prove extremely difficult in practice, particularly where there is a congruence between ideology and structural position. If structural theories are weak on cognitive variables, most cognitive theories cannot describe clearly how power and ideas interact. The same set of objections may even extend to knowledge-oriented "cognitivism." While few would claim that scientific knowledge is reducible to social interests, scientific evidence can be resisted. The range of scientific investigation, if not its content, is also shaped by social and political purposes. Cognitive theory needs to specify more clearly the types of issues and conditions under which consensual knowledge is likely to drive cooperation.\textsuperscript{74} Are they many or few? No one would disagree that purposes, values, and knowledge matter, but when? What is the relationship between the cognitivist research program and that outlined by various forms of structuralism?

Cognitivists pose a simple, yet profound, question: can interests in an issue area be unambiguously deduced from power and situational con-


strains? Frequently they cannot. Without shifts in power position, interests change as a result of learning, persuasion, and divine revelation. Knowledge and ideology may then become an important explanation of regime change, but when posed in this fashion, the question is an empirical as much as a theoretical one. The resolution of the debate between structuralists and cognitivists will depend on tests that allow a confrontation between the two approaches without violating the epistemological tenets of either.

4. Conclusion: Recapturing politics

Current theories of international regimes have ignored domestic political processes, in part because of the lure of parsimonious systemic theory. This neglect has extended to the issue of how regimes actually influence national policy choices, a question closely related to the issue of compliance and regime strength. More broadly, there have been few studies of the domestic political determinants of international cooperation. There are both methodological and theoretical reasons to open the black-box of domestic politics. Even if one adopts a structural explanation of compliance and defection, validating such claims demands careful reconstruction of decision-making at the national level. But the neglect of the domestic political and economic realm has had deeper costs, including a neglect of the substantive issues over which states are likely to seek cooperation and the basic forces leading to regime change. To address these difficulties, we suggest a research program that views international cooperation not only as the outcome of relations among states, but of the interaction between domestic and international games and coalitions that span national boundaries.

Do regimes "matter"?

Regimes are said to affect state behavior in two ways. One, emphasized in functionalist and game-theoretic approaches, is that regimes have altered the situation or setting in which states interact so that cooperation is more likely. The literature on iterated Prisoner's Dilemma claims that altering the institutional environment—by lengthening the shadow of the future, limiting the number of players, increasing the transparency of state action, and altering the payoff structure—can increase the incentives to cooperate. Functionalist theories emphasize that regimes reduce the transactions costs associated with bilateral contracting.

How do we know if these claims are correct? More importantly, how do we know that these explanations of cooperative behavior are superior to alternative explanations? Two methods could substantiate these claims. Large-n studies could determine whether or not regimes are, in fact, associated with the institutional factors specified ex ante by the theory. Far more
convincing, however, would be evidence that domestic policymakers were actually concerned with reputation, reducing transactions costs, the need for transparency, and so forth, when facing decisions about regime creation and compliance. Even if one advances a structural explanation, the most convincing evidence must be found in the calculations of national decision-makers. Drawing on the work of Alexander George, Vinod Aggarwal employs such a method of “process tracing” in analyzing the influence of international structure on national textile policies. \(^{75}\) Surprisingly little work of this kind has been done.

An even stronger claim for regimes is that they can alter actors’ interests or preferences, which are generally held constant in functional and game-theoretic formulations. Cognitivists claim that regimes may change basic definitions of reality. Over the 1970s, for example, the United Nations Conference on Trade and Development (UNCTAD) provided the institutional locus for a new Third World critique of the existing international economic order. UNCTAD virtually served as a secretariat for the Group of 77 developing countries; it provided studies and documentation that were crucial in defining individual LDC’s positions in international forums. \(^{76}\)

Where regimes are actually suspected of altering state preferences, the proposed research program should focus on domestic decision-making for verification, since other forces may also be at work simultaneously. This can be seen by examining the problems of compliance and defection.

A fit between regime rules and national behavior may not occur for three reasons. The first is that the norms characterizing the regime may not be *formulated* to be authoritatively binding. Some regimes, such as those governing human rights, allow self-selected national exemptions or represent only broad collective aspirations; they are, in effect, born weak. \(^{77}\) Opportunism presents a second reason for disjuncture between regime norms and state behavior. States may negotiate regimes with the intention of breaking them or knowingly exploit others’ compliance in order to extract higher payoffs.

A final possibility, and a more common and politically interesting one, is what Robert Putnam has called “involuntary defection.” \(^{78}\) This defection happens when a party reaching or supporting an international agreement is unable to sustain commitments because of domestic political constraints. Three examples suggest the pervasiveness of this phenomenon. The first is

---

77. Jack Donnelly shows that the international human rights regime is composed of widely accepted substantive norms and internationalized standard-setting procedures, but a very limited degree of international implementation and no international enforcement. As he concludes, “such normative strength and procedural weakness, however, is the result of conscious political decisions,” “International Human Rights,” p. 614.
the rise of the "new protectionism" in the advanced industrial states. Taking the form of quantitative restrictions in particular products aimed at particular exporting countries, these "orderly marketing agreements" and "voluntary export restraints" clearly violate GATT norms, especially the non-discrimination norm. Yet they are driven by efforts to diffuse political pressures, while simultaneously limiting the extent of protection, both by product and by trading partner. This practice has, in turn, affected the nature of the regime. Not surprisingly, there is little agreement on a code governing the use of selective safeguards, but the consensus seems to be moving towards a tolerance of their use.79 Recent research on the implementation of IMF stabilization programs suggests a similar phenomenon.80 Some programs are no doubt negotiated in bad faith, with little intention of fulfillment. More commonly, those at the negotiating table are simply unable to control those in the palace or in the streets. A final example may be drawn from the interwar period. Great Britain's strong commitment to maintaining the gold standard during the 1920s stimulated strong industrial opposition to the Bank of England's policies, contributing to the final decision to go off gold in 1931.81

"Defection" in such instances is not the result of calculating unified actors, but the outcome of domestic political conflicts which no single actor can control. How international agreements play into these domestic political fights is still poorly understood. Reputation has appeared in arguments against "rash" protectionist moves. Occasionally, the IMF has been used to implement programs that would have been impossible without outside pressure. The Bank of England argued that defending parity was key to British leadership in international financing affairs. On the other hand, opposition forces have used the government's compliance with regime injunctions to their own political benefit, arguing that they could have extracted a superior bargain by defecting, or that bowing to outside pressure is intolerable. Joanne Gowa shows how those defending the closing of the gold window portrayed regime rules as overly restrictive of national autonomy.82

Towards a new theory of international cooperation

Building a theory of cooperation and regime change demands that we return to the central insight of the interdependence literature: growing interdependence means the erasure of the boundaries separating international

and domestic politics. "Domestic" political issues spill over into international politics and "foreign policy" has domestic roots and consequences. Governments, when making choices about regime creation and compliance, try to preserve the benefits of cooperation while minimizing the costs that may fall on politically important groups. This insight appears to have been lost in much recent writing about regimes. One approach would be to start with the domestic-level game, and consider the conditions favorable to cooperation. This approach assumes that a state's decision to enter into and abide by international rules can be treated in the same way as any other public policy choice. An institutionalist approach, for example, would suggest that the arena in which decisions are made conditions the ability of state elites to sustain commitments. Decisions made in insulated arenas may not be pulled as much by "uncooperative" forces. On the other hand, the decentralization of certain policies, such as banking regulation and tax policy in the United States, may inhibit the formation of international regimes. A coalitional approach would emphasize the constellation of domestic actors standing to benefit from cooperation. The Reagan administration's 1985 decision to seek international cooperation on the dollar and exchange rates was prodded in part by export interests and firms facing import competition.

Starting with the "unit" level is risky, since the interactive effects of international structures, bargaining, and rules on domestic politics can get lost. Peter Katzenstein's work on small European democracies attempts to capture such feedback, at least over a long historical period. The position of small states in the international system influences the development of domestic institutions in predictable ways; these institutions, in turn, affect these states' ability to adjust to the dictates of a liberal international and regional order.

The next step is to develop interactive models that link domestic and international politics more closely. Robert Putnam has suggested that episodes of international cooperation must be viewed as "two-level games." At one level, representatives of different countries seek to reach or sustain international agreements; at a second level, those same representatives must build the political support required to sustain commitments and establish credibility. These processes often occur simultaneously and expectations about the likelihood of gaining acceptance in one arena influence the bargaining process in the other.

Another approach focuses on the transnational coalitions that span countries. This method is adopted by Ernst Haas in his classic study The Uniting

83. For an insightful review of different approaches to public policy, see Herbert Kitschelt, "Four Theories of Public Policy Making and Fast Breeder Reactor Development," International Organization 40 (Winter 1986).
85. Putnam, "The Logic of Two-Level Games."
of Europe, but this line of research has not been adequately pursued. For example, does some synchronization of coalitional patterns or electoral cycles across states make cooperation more likely? Where a nation may lack a winning coalition domestically, can pressure or support from external actors tip policy in the direction favored by the cooperative minority? Robert Putnam and Nicholas Bayne suggest this may have occurred in the German decision to reflate in 1978.86

The central point is that growing interdependence means that groups at the domestic level increasingly have "regime interests." Welfare is tied not only to particular policy decisions, but to other states' compliance with regime norms, the way in which international cooperation is institutionalized, and the access regimes provide for private actors. Structural theory alerted us to the myopia of the second image. It would be a misfortune if the quest for parsimonious systemic theory displaced other research and theoretical traditions that study domestic histories and, above all, politics.