APPLYING RATIONAL CHOICE THEORY TO INTERNATIONAL LAW: THE PROMISE AND PITFALLS

ALEXANDER THOMPSON*

I. INTRODUCTION

There is good reason to believe that rational choice approaches can be usefully applied to the study of international law (IL). The law and economics tradition has brought rational choice insights and analysis to bear on domestic legal affairs, and political scientists in the international relations (IR) field have applied rational choice approaches—especially game theory, expected utility theory, and transaction cost economics—with great success. Indeed, given these movements in IL’s sister fields of domestic law and IR, one puzzle is why it has taken so long for international lawyers to turn to rational choice theory.¹ Most IL scholarship has focused on describing the nature and sources of law and defending its legitimacy. However, with the growth of international law and economics,² and increasing cross-fertilization between IL and IR,³ the scholarly landscape is beginning to change in ways that will benefit the study of international legal behavior and institutions. Rational choice offers a promising foundation—including a coherent set of

* Department of Political Science, Ohio State University. I would like to thank Ken Abbott, Eric Posner, Susan Pratt, and Duncan Snidal for comments on earlier drafts.

¹ As Jack Goldsmith and Eric Posner point out in their introductory paper, this puzzle may be explained partly by the normative commitment to law felt by many IL scholars. After all, historically they have not been mere observers and researchers: their writings have constituted a source of law and legal interpretation.


[Journal of Legal Studies, vol. XXXI (January 2002)]
© 2002 by The University of Chicago. All rights reserved. 0047-2530/2002/3101-0020$01.50

S285
common assumptions and various analytical tools—on which to base positive theorizing regarding the creation and influence of IL.

Its most zealous proponents tout rational choice as a universal framework with which to understand all realms of human behavior,\(^4\) and indeed the notion of rational utility maximization has been used to illuminate a wide range of activities, from macrohistorical events to the selection of marriage partners.\(^5\) This had led to critiques of rational choice as excessively ambitious.\(^6\) Nevertheless, its ability to generate insights in yet another domain—international legal relations—is a testimony to its flexibility and explanatory power. Notwithstanding its limitations, some of which are outlined below, the assertion of George Stigler and Gary Becker that “no other approach of remotely comparable generality and power is available”\(^7\) remains true a quarter century later.

I argue that the power and versatility of rational choice is at once its greatest virtue and its potential weakness, and I use the papers in this volume to illustrate the tension. With versatility come certain dangers, and this trade-off is the central theme of this essay. The authors benefit from the scope and flexibility of a rational choice approach but also suffer from various pathologies derived from these same characteristics.

The next section briefly outlines the range of rational choice explanations and methods employed by the authors to illustrate its expansiveness as a social scientific paradigm. The sections that follow point to three aspects of rational choice theory’s versatility and the attendant dangers—or potential dangers—of each. First, while rational choice has the virtue of accommodating a wide range of actors as the locus of analysis, its use is limited when the preferences of individual actors must be aggregated. And yet corporate actors are often treated unproblematically as unitary actors. Second, while rational choice has demonstrated flexibility in explaining behavior in a variety of empirical settings, various analytical problems occur when existing models and theories are transferred from one empirical domain to another. This may result in explanations that are stretched too thin to be useful. Third, precisely because a rationalist explanation can be constructed (often post hoc) to explain virtually any behavior, research design and empirics are of heightened importance for rational choice practitioners. Theorists must be careful not just to present confirmatory evidence but also to construct arguments that are precise and falsifiable and to link their arguments to testable implications.

---


\(^7\) Stigler & Becker, supra note 4, at 77.
These are requisites for rational choice contributions to IL to become a cumulative research program. The concluding section discusses how the contributors to this symposium issue have pushed the bounds of rational choice to overlap in complementary ways with social constructivist explanations involving distinct motivations (for example, values and emotions) and constraints (for example, norms).

II. THE VERSATILITY OF RATIONAL CHOICE THEORY

Rational choice theory—which is more properly thought of as a paradigm or family of theories—has few inherent limitations in terms of the problems to which it might be applied. Its underlying assumption is that individuals engage in purposive, means-ends calculation in order to attain their goals—that is, they select actions so as to maximize their utility. Rational choice does not tell us what our goals should be; it focuses on the means, not the ends. Nevertheless, the theory of rational choice, and the notion of rationality more generally, has both a normative and a positive dimension. Rationality is a normative concept insofar as “it points to what we should do in order to attain a given end or objective.” Its positive application comes when it is used to describe, explain, and predict human behavior. From the premise of optimizing behavior, a positive rational choice argument deduces the action that an actor did or will choose on the basis of the actor’s goals and beliefs and a specification of relevant environmental constraints (such as the availability of information or, in a strategic setting, the actions of other actors).

Though we could draw normative conclusions or policy lessons from some of the papers in this issue—for example, maybe trade dispute settlement that allows efficient breach is desirable and maybe “interest actors” should mobilize norms if they hope to be successful—the authors engage almost exclusively in positive analysis. Julie Roin offers policy proposals for achieving tax base harmonization at the end of her paper, but she stresses repeatedly that she is less concerned with the merits of this harmonization than with explaining why it will be difficult to achieve. However, this focus on positive

---


10 For more on rational choice as both a normative and a positive theory, see Jon Elster, Introduction, in id.; and Cristina Bicchieri, Rationality and Coordination 3 (1993).


12 See Kenneth W. Abbott & Duncan Snidal, Values and Interests: International Legalization in the Fight against Corruption, in this issue, at S141.

13 Julie Roin, Taxation without Coordination, in this issue, at S61.
rather than normative theorizing is the only respect in which the volume as a whole delimits the scope of rational choice theory.

The papers employ an inclusive conception of what constitutes rational choice (a reasonable approach at the early stages of this research program). To the extent the authors discuss what “rational” behavior entails, the requirements vary considerably. Most seem to implicitly assume substantive rationality on the part of actors but do not address thornier issues of procedural rationality. The major exception is Peter Huang, who critically reflects on the assumptions that underlie rational choice theory and incorporates the extrarational forces of emotion and psychology. The authors also do not explicitly impose stricter requirements such as the completeness and transitivity of preferences, that is, that all outcomes can be consistently compared and ranked. Setear, for example, points to no more than a rough functionalism guiding political actors. For the most part, rationalism is thus defined in its most basic form as purposive utility maximization. As with Huang, some of the authors seek to push the limits of the rationality assumption, a topic I return to in the concluding section.

The authors also exploit the versatility of rational choice theory by treating legal institutions—including customary law, treaties, and international organizations—as both endogenous and exogenous. By doing so, this volume joins the rationalist strain of the “new institutionalism” in the social sciences, especially economics and political science, by examining both actor preferences and choices in the context of institutional constraints (institutions as exogenous) and how institutions are created and designed as a function of actor preferences and other exogenous constraints (institutions as endogenous).

Three of the papers are principally concerned with explaining the existence and nature of law and institutions (the endogenous perspective). Kenneth Abbott and Duncan Snidal show how a combination of values and interests shape the creation and design of legal rules and institutions. Thus the Organization for Economic Cooperation and Development (OECD) Anti-bribery Convention is treated as a piece of law whose origins must be explained. And Warren Schwartz and Alan Sykes explain the structure and rules of the World Trade Organization (WTO) as an institutional outcome.

17 Although Abbott and Snidal have shown us how both values and interests can shape IL, their argument is indeterminate with regard to what a particular piece of law will look like. Their framework cannot explain which perspective will tend to prevail—interests or values—and how this affects the nature of resulting law.
Roin’s paper is more complicated because it is somewhat speculative; she attempts to explain why the cooperation needed for international tax base harmonization will not be realized. This is a nice complement to the other papers written from an endogenous perspective since she is predicting an institutional outcome and since she analyzes a nonevent; such papers are usually selected out of studies of international cooperation and institution building.

The other papers take law largely as given and seek to explain state behavior (the exogenous perspective). Guzman and Simmons take the WTO and its dispute settlement mechanism (DSU) as given in trying to explain negotiating and paneling behavior. It should be noted, however, that they do not in fact use characteristics of the DSU to do much explaining; most of their independent variables come from other sources, such as the nature of the issue being disputed (its “lumpiness”) or of the disputants themselves (for example, regime type, relative power). Even institutional characteristics that are typically used to explain outcomes—such as the transaction costs of litigation in the WTO and the mutual availability of information—are assumed to be unimportant.

James Morrow takes the set of law of war treaties as given and attempts to explain the compliance behavior that results. However, the paper does contain implicit suggestions about how we might understand law endogenously. For example, he points to some advantages of formal treaties over custom and treats “shared understandings”—an ingredient for the development of formal law—as endogenous when he writes that “formal negotiation can sharpen the understanding among states about custom.” This implies that certain strategies, by fostering shared understandings, may produce formal treaties rather than custom. Jack Goldsmith and Eric Posner treat law as exogenous and seek to explain the existence of moral and legal rhetoric. They do devote some effort to turning their argument around: they explain the existence of treaties as “attempts to clarify actions in order to facilitate

18 See Andrew Guzman & Beth A. Simmons, To Settle or Empanel? An Empirical Analysis of Litigation and Settlement at the World Trade Organization, in this issue, at S205; Jack L. Goldsmith & Eric A. Posner, Moral and Legal Rhetoric in International Relations: A Rational Choice Perspective, in this issue, at S115; and George W. Downs & Michael A. Jones, Reputation, Compliance, and International Law, in this issue, at S95. In fact, the Downs and Jones and Setear papers do not fit neatly into the endogenous-exogenous dichotomy. For Downs and Jones, law is exogenous insofar as they take it as a given and focus on explaining compliance behavior by states. However, their independent variable is the degree to which states are concerned about their reputations, which is in turn a function of state characteristics and the importance of the substantive issue at stake. The nature of legal institutions is not an important explanatory variable in their analysis. Setear discusses the treaty ratification process in the United States and does not analyze international law per se, except insofar as a failed attempt at legislative approval may scuttle a treaty or render it nonbinding.

coordination.” This is a good example of how an exogenous argument about institutions can be logically extended to have endogenous implications, a common practice among rationalist institutionalists in other fields. Similarly, Huang treats law as largely exogenous to focus on compliance behavior but then extrapolates from the model to explain the demand for nonbinding environmental law, thereby endogenizing legal institutions.

The most comprehensive effort to treat international law as both exogenous and endogenous is offered by Barbara Koremenos. She provides both a theoretical argument for how states design flexibility provisions into treaties when they face uncertainties and an explanation for cooperation and compliance behavior once the regime is established. This dual approach is also reflected in her empirics, where a description of the negotiations leading to the International Coffee Agreement (law as endogenous) is followed by a discussion of reneging behavior under the agreement (law as exogenous).

Rational choice is also versatile insofar as it can be deployed with a wide range of analytic tools and subtheories. The authors reveal this diversity in their contributions. They borrow insights and methods from game theory (Downs and Jones, Koremenos, Guzman and Simmons, and Huang), transaction cost economics (Guzman and Simmons, and Roin), public choice theory (Roin, and Schwartz and Sykes), the economic theories of contracting and bargaining (Schwartz and Sykes, Guzman and Simmons, and Koremenos), and signaling models (Goldsmith and Posner, and Setear). These tools allow the authors to capture the strategic interdependence of actors engaged in cost-benefit decision making and facilitate a rationalist approach to studying legal and political institutions.

By relying on a very general conception of what constitutes “rational” behavior, by treating law as both an explanatory variable and an outcome, and by employing a range of analytical tools and subtheories to operationalize their rationalist theorizing, the authors embrace the versatility of rational choice theory. Indeed, Goldsmith and Posner point to the demonstrated flexibility of rational choice as a “general lesson of this symposium.” The next three sections examine additional sources of versatility in the rational choice approach, each of which is also exhibited in this volume: it can accommodate almost any actor, it can be applied to virtually any empirical domain, and it can be used to explain a wide range of behaviors. Each of these apparent virtues of rational choice theory entails a potential pitfall, as I illustrate below with examples from the papers.

20 Goldsmith & Posner, supra note 18, at S137.
III. THE AGGREGATION OF PREFERENCES

In a rational choice model, outcomes are the product of individual decisions, where "individual" is defined as a unitary actor. In theory, rational choice explanations can be applied to the behavior of any unitary actor (whether those explanations are accurate in practice is, of course, an empirical matter). As a practical matter, however, analysts—especially those studying international affairs and foreign policy—often deal with aggregate actors like international organizations, interest groups, publics, and especially states. Those who treat states as unitary assume either that the state aggregates all domestic preferences—of individuals, interest groups, and various intragovernmental actors—and acts as if it were a single actor or that state decision making is in fact channeled through a single or small group of crucial individuals who make important decisions. In practice, the unitary-actor assumption usually combines an element of methodological convenience with some belief that it is an empirically accurate portrayal of state behavior.

So what are the potential problems of treating states as unitary? To begin with, the very notion of the post-Westphalian sovereign state has been called into question by recent analyses. Stephen Krasner critiques state-centered theories of IR in light of the historical frailty of national sovereignty, arguing that they "cannot offer much guidance in situations where . . . [their] basic ontological assumption that states are autonomous actors is violated." Sovereignty in thought and practice has been fluid and contingent on various historical and cultural factors.

A more profound challenge to the state as unitary actor comes from within the rational choice tradition and concerns the difficulties of endowing any corporate actor with coherent preferences or agency. Strictly speaking, aggregate actors such as societies and nations—and even more intimate groups like families—cannot be treated as the basis for rational choice theorizing, which relies on decisions by individuals. Building on Kenneth Arrow's impossibility theorem (showing that under majority-rule voting no single,
collectively preferred policy exists), Kenneth Shepsle examines legislatures as corporate actors. He argues that individual legislators’ intents, “even if they are unambiguous, do not add up like vectors.” Legislation does not reflect a simple aggregation of the preferences of legislators, and therefore “legislative intent is an internally inconsistent, self-contradictory expression.” The more general point is that collective actors do not behave according to rationalist principles in the same way unitary actors do—they do not have coherent beliefs, goals, and preferences. This casts further doubt on the validity of applying rational choice models to the behavior of states.

Turning to the articles in this symposium, we see that several of the authors decompose the state and look at domestic and transnational actors to varying degrees. Roin uses public choice theory, which was developed to capture the insight that societal preferences are not always aggregated effectively, and considers a variety of governmental (the president, the Treasury Department, Congress, and even individual legislators) and subnational (taxpayers, especially corporations) actors. Setear focuses mostly on the incentives of the president as an individual and further disaggregates the state by treating Congress as an actor and by discussing the role of the courts. Thus the incentives of individuals do make their way into these analyses, even if fairly complex corporate actors also play key roles.

Abbott and Snidal move furthest away from the assumption of the state as unitary actor—indeed, they embrace the diversity of relevant actors, whom they distinguish according to social roles. Thus we have interest actors, motivated by selfish and usually material interests, and value actors, whose interests are based on principles and norms. Both types of actors may include government agencies, corporations, international nongovernmental organizations, and intergovernmental organizations, although the authors adduce evidence on the impact of key individuals within these larger entities (such as Dan Tarullo of the State Department and Mark Peith of the OECD). Nevertheless, because it is ultimately the state (or its representative) that must negotiate international agreements, Abbott and Snidal cannot escape treating states as aggregators of domestic actors’ interests when it comes to the final stages of IL formation. However, they recognize that ultimately states as unitary actors are fictitious entities and that their preferences fluctuate over time as domestic values and interests change. In other words, preference aggregation is an ongoing political process.

Other authors clearly treat the state as a unitary actor. Some do so by

---


28 Shepsle, supra note 27, at 239.

29 Setear, supra note 16, at S35–S36.

30 Abbott & Snidal, supra note 12, at S155.
making the assumption that the state acts as if it were an individual. George Downs and Michael Jones, Koremenos, and Huang make this analytical move, and so do Guzman and Simmons, in effect. Guzman and Simmons do consider the potential constraining effects of domestic politics and institutions on international negotiations. See Guzman & Simmons, supra note 18, at S209, S218–S219.

Others make an implicit or explicit empirical claim that one person does in fact represent the state. Schwartz and Sykes focus on signatory officials to international agreements, and, although they mention that domestic constituencies might influence the behavior of such officials, they do not explore in any detail how the preferences of subnational actors map into the choices made by treaty framers (see the section below for further discussion of this point). Goldsmith and Posner and Morrow refer most often to ‘national leaders’ and ‘government leaders,’ respectively, however, Goldsmith and Posner argue that publics (both domestic and foreign) may play a role in some cheap-talk games. By treating the public as an undifferentiated group, they arguably do not encounter an aggregation-of-preferences problem insofar as the public is treated as a passive actor to whom politicians appeal with their rhetoric, not as an active agent that is making choices and driving their explanation of cheap talk.

The bottom line for studies of international affairs is simply to be sensitive to the potential pitfalls of models that assume states are unitary actors. As Jon Elster notes, ‘To be taken seriously, . . . any account that imputes goal-oriented behavior to aggregate entities has to explain why we should expect consistency in their behavior.’ The unitary-actor assumption is sometimes justified on empirical grounds (a case made by Setear) and sometimes justified as a reasonable simplification of reality (a case made by Abbott and Snidal); on the other hand, it may confound explanations and predictions when actors and institutions within the state play prominent roles.

The evolution of the literature on institutions in IR is instructive. Early efforts by ‘neoliberal institutionalists’ treated states as unitary actors and showed little sensitivity to the potential problems of preference aggregation. International institutions were assumed to be ‘like all social institutions’ in terms of their fundamental properties, even though they are made up of states rather than individuals. As the analysis of institutions became more sophisticated and realistic, scholars increasingly incorporated substate actors even at the cost of more complex models. Nevertheless, state centrism was
a reasonable assumption for launching a new institutionalist research program in IR—after all, most rational choice models must be complemented by other theories to apply in the real world anyway.\footnote{This is true when the logic of utility maximization does not apply and when rational choice models produce multiple equilibria.} As rational choice makes inroads into the study of IL, treating the state as a unitary actor will lead to productive theorizing but should eventually be replaced by a more nuanced vision of precisely who creates and is influenced by international legal institutions and how these actors interact.

IV. BRIDGING EMPIRICAL DOMAINS

The explanatory power and elegance of many rational choice explanations make them attractive as models to be applied across empirical domains. The law and economics movement is premised on the notion that models used to explain individual behavior in the market can also shed light on legal rules and behavior: "[M]any legal decisions are indeed market-like choices. They may be said to be so on the ground that legal rules create implicit prices on different behaviors and that legal decisions makers conform their behavior to those prices in much the same way as they conform their behavior to the relative prices there."\footnote{Robert Keohane, After Hegemony (1984), for example, builds on economic models borrowed from the work of economists such as Oliver Williamson, Ronald Coase, and Mancur Olson. Neorealism, another leading approach to IR, rests on the assumption that structural models from microeconomics can be applied across empirical domains if the basic structure remains the same. See Kenneth Waltz, Theory of International Politics 80 (1979).} Similarly, theories of international institutions have borrowed extensively from rational choice theorizing on local-level institutions and collective action, such as theories of the firm, bargaining theory, and collective action theory.\footnote{On the notion of external validity, see Thomas Cook & Donald Campbell, Quasi-Experimentation: Design and Analysis Issues for Field Settings 39 (1979) (defining it as the "validity with which conclusions are drawn about the generalizability of a causal relationship to and across populations of persons, settings, and times").} It is often useful to borrow from fields that are more theoretically developed; however, the practice of applying theories across empirical domains can also be dangerous. When concepts and models from one area are transplanted to another, their explanatory power may be stretched too thin to be illuminating. At worst, this approach to theorizing can lead to misleading conclusions.

We can think of the problem of bridging empirical domains as a special case of generalizability or external validity: it requires an argument to be robust across realms of human behavior, not just across cases within the realm for which it was created.\footnote{Thomas Ulen, Rational Choice Theory in Law and Economics, in The Encyclopedia of Law and Economics 790, 797 (Boudewijn Bouckaert & Gerrit De Geest eds. 1993); Helen Milner, Interests, Institutions, and Information (1997); and Daniel Drezner, Locating the Proper Authorities: The Interaction of Domestic and International Institutions (2002).} This question of validity is always relevant...
when we take assumptions and arguments out of their original setting—where, after all, they are already simplifications of reality—and apply them to a different area of social behavior. Theorists of international affairs face an especially acute version of this problem since they often borrow theories from lower levels of activity, which raises additional issues of scale. This has been referred to as the problem of theoretical “scaling up.”

The premise of this symposium is that rational choice models can improve our understanding of IL. Since every rational choice model carries with it an interpretive structure, including assumptions about actors and their contexts, we must be sensitive to potential problems of bridging empirical domains if we apply existing models to a different—in this case, higher-level—realm. There are at least two sources of such problems. First, they arise from differences in social setting, including legal and political institutions as well as less formal norms and rules. The context or structure of a situation—including both social and material factors—can shape actor behavior in important ways and can even “induce” a certain equilibrium outcome. Second, problems arise when shifts in domain entail differences in the substantive issue being studied. When theory A is developed to explain empirical subject X, the subsequent application of A to subject Y may not be appropriate if the analytical nature of X and Y vary in important ways.

The contributions to this volume reveal considerable use of theories originally designed to explain behavior in different domains, including, at the domestic and individual levels, contract theory (Schwartz and Sykes, and Koremenos), explanations of litigation and settlement (Guzman and Simmons), and reputational models that are based on individual interactions (Downs and Jones), to name a few examples. At times the authors show explicit sensitivity to the potential problems of transplanting theories across domains, and some of the empirical findings reveal that explanations developed for phenomena in other areas do not always translate well to the international legal arena. Roin, for example, demonstrates that tax base harmonization, easily explained in economic efficiency terms within the domestic polity, has proved out of reach for national states, even among members of the European Union. She shows how differences in social context, including legal and political institutions, obstruct the transfer of theoretical

40 On the practice of scaling up lower-level theories on institutions and collective action to IR, including the potential dangers of doing so, see Oran Young, The Problem of Scale in Human/Environment Relationships, in Local Commons and Global Interdependence: Heterogeneity and Cooperation in Two Domains 27 (Robert Keohane & Elinor Ostrom eds. 1995); and Alexander Thompson, Scaling Up: Applying Lower-Level Theories to International Relations and Law (paper presented at the annual meeting of the Midwest Political Science Association, Chicago, April 25, 2002).

predictions from the domestic to the international level. For example, while the federal income tax predated most state tax systems within the United States, providing a baseline, this background institutional condition does not exist for nation-states at the international level, meaning any collective understanding of tax bases would have to develop “from scratch.” There is no central government to provide a coordination point for standardization.

Downs and Jones’s argument also reflects differences in social setting as we move across empirical domains. Many of the influences of reputation imputed to state behavior, such as adherence to commitments, have been overstated, they argue, and reputation may not operate at the international level as it does within domestic society. The sparse social atmosphere of the international system, captured by Downs and Jones with the deductive framework of repeated games and the chain store literature in economics, dampens the development of reputations and of reputation-based incentives. Huang’s analysis is much less sensitive to the move across empirical domains. His states are concerned about saving face and gaining approval from others, though the theories of cognitive and social psychology that constitute the microfoundations of his argument were designed to apply to individuals within society, not to corporate actors at the international level (a point with obvious implications for the aggregation-of-preferences issue discussed above). This is not to say that we cannot ascribe emotions to states, only that the underlying assumptions should be justified as we move into a radically different social setting.42

In their study of WTO dispute settlement, Guzman and Simmons scale up bargaining theory—specifically, theories of litigation and settlement in the domestic context. Generally speaking, the metaphor of bargaining applies very well to the interaction of states. Many of the specific assumptions and implications of bargaining models, on the other hand, travel less well to the complexity of issues in international affairs. State power and influence may come from a variety of sources in addition to market power, for example. Moreover, if the decision to grant concessions or stand firm depends on domestic political structures and interests, bargaining may take a very different form.43 Guzman and Simmons, however, are quite sensitive to issues that arise from shifting to a new issue area; rather than take assumptions from lower-level theories as given, they test them empirically. For example, they recognize that in international trade disputes it may be transaction costs more than information asymmetries that obstruct settlement. This is because states typically negotiate over specific issues, many of which are “lumpy” (that is, not easily divisible), and do not offer cash on the side as domestic

42 Huang does address this issue briefly, asserting that a concern over face is “universal” and that it motivates countries as well as people. Huang, supra note 15, at S253.

litigants might. Bargaining is thus of a different nature. They also control for many political variables, such as the relative power and regime type of states. While they mine the law and economics literature on litigation and settlement for basic insights—for example, that litigants should prefer settlement over going to court, subject to transaction costs and the symmetry of information—they also show that actors negotiate under different constraints in the area of international trade politics.

Schwartz and Sykes argue that "the lessons developed by law and economics scholars regarding the way that private contracting parties structure their bargains accordingly have much to teach us about the structure of treaties." They are explicitly transplanting ideas from one substantive issue, domestic contracting, to another, international treaty formation (specifically, provisions on violations, remedies, and dispute settlement in the General Agreement on Tariffs and Trade (GATT)/WTO). Even if these two subjects share certain strategic or analytic properties, the substantive differences between the two issues corrupt the application of contracting theory as conceived in economics to a complex political phenomenon like the WTO's dispute settlement mechanism, and the authors recognize this by making significant modifications.

Nevertheless, they may overstate the extent to which international institutions represent rationally efficient equilibria. Economists have various tools for defining efficiency that capture important aspects of market exchange and aggregate welfare. However, the notion of efficiency becomes problematic when applied to political issues, especially at the international level. Efficient outcomes are not easy to define. Most trade officials would not accept compensation (either in the form of a payment or the withdrawal of substantially equivalent concession) as a "solution" to a trade dispute even if it left both parties better off in the aggregate. The reasons lie in the institutions and dynamics of domestic politics. Domestic interest groups and their congressional representatives seeking to remove a foreign unfair trade practice are typically not satisfied with anything short of compliance (that is, removal of the offending practice), especially since those harmed by the practice do not necessarily benefit from compensation or sanctions. This may explain why the compensation remedy is rarely used and why alternative rational choice explanations centered more explicitly on domestic politics explain the features of the GATT/WTO quite well. International institutions are politically

---

44 Schwartz & Sykes, supra note 11, at S204.
45 For example, they argue that governments seek a treaty that will maximize their political rather than their economic welfare, although this political welfare seems to reduce to the economic benefits attained by certain producer groups (which are presumed to translate into votes, contributions, and so on, to politicians).
efficient rather than economically efficient,\textsuperscript{47} and the bargaining that leads to their creation is of a two-level nature. The theory of efficient breach was not designed with either of these possibilities in mind.

In general, studies that are open-minded about the motivations of actors are best suited to accommodating changes in the social setting and the nature of substantive issues. This may include theoretical approaches that combine rationalist and sociological logics, such as Abbott and Snidal's effort to integrate "interest actors" and "value actors" into the same framework or Morrow's discussion of "shared understandings" as common conjecture in game theory. The problems associated with transferring rational choice models across empirical domains illustrate the necessity of modifying explanations originally designed for other settings. Of course, analysts will always face a trade-off between parsimony and contextual accuracy when they transplant theories across domains, a choice that should be driven largely by the empirical accuracy of the resulting model.

V. Research Design and Empirical Evidence

Virtually any behavior can be explained within a rational choice framework—that is, a rationalist explanation can be constructed post hoc to fit events. And nonrationalist explanations of behavior can often be reinterpreted in rationalist terms.\textsuperscript{48} This versatility is seen as a virtue of rational choice by its proponents, although Amartya Sen makes the point in the form of a criticism: "It is possible to define a person's interests in such a way that no matter what he does he can be seen to be furthering his own interests in every isolated act of choice."\textsuperscript{49} Precisely because of this flexibility, it is important for rational choice theorists not only to seek empirical support for their arguments but also to create a research design that allows for falsification and for the gathering of evidence that rules out alternative explanations.

As rational choice theory moves into the IL field, its success will be measured in large part by the empirical success of such studies. Can they explain and predict more effectively than other approaches? Empirics—and by implication, research design—are especially important since rational choice theorizing has been criticized generally and in the context of international affairs for not being backed up by significant and novel empirical


\textsuperscript{48} For example, see George Tsebelis, Nested Games 22-23 (1990).

\textsuperscript{49} Amartya Sen, Rational Fools, in Scientific Models and Man 1, 5 (Henry Harris ed. 1979).
And since law scholarship more generally is accused of too often ignoring the "rules of inference." This section outlines very general principles for sound research design and discusses the papers in this volume in light of these guidelines and in light of their empirical analyses.

To maximize the consequence of a theoretical argument, the researcher should present it in its most vulnerable form. The key to achieving this is falsifiability, which has several components. First, a theory must be stated clearly enough so specific, noncontradictory hypotheses can be generated. Otherwise the theory cannot be shown to be right or wrong. Second, the theorist should produce observable implications from these hypotheses that point to both confirming and disconfirming evidence. We need to know what evidence would prove the theory wrong, or at the very least that such evidence could be marshaled in principle. Finally, research projects should outline potential rival explanations and related alternative hypotheses. These research design criteria are important apart from the actual empirical analysis, which is not likely to be effective absent the logical and practical possibility of an empirical challenge. Rational choice theories that are not falsifiable are usually tautological, hence the importance of these issues.

Of course, not all scholarly works—especially article-length ones—need to conduct comprehensive empirical tests of their arguments. Single case studies, for example, may be used for various purposes other than theory testing: they may be used to illustrate the logic of an argument, to generate more refined hypotheses, and as plausibility probes into the validity of a theory. Other studies are entirely deductive and theoretical, leaving empirics for future analysis. Nevertheless, theories should be constructed to allow for empirical testing and researchers should at least establish a foundation for such testing if they hope to generate knowledge about a topic. In the discussion of the papers that follows, I am not necessarily pointing out intrinsic weaknesses of the works; my goal, rather, is to provide comments with the goal of turning "rational choice and international law" into a productive and cumulative research program.

Goldsmith and Posner's article on moral and legal rhetoric, for example, illustrates how scholarship can usefully contribute to the development of a research program without itself containing systematic empirics. Relying on insights from game theory, they propose various rationalist logics for the

---


52 For a more thorough discussion of falsifiability as a component of good theory, see Gary King, Robert Keohane, & Sidney Verba, Designing Social Inquiry (1994); Epstein & King, supra note 51; and Spiro Latsis, A Research Programme in Economics, in Method and Appraisal in Economics 1, 7–9 (Spiro Latsis ed. 1976). A general discussion of the falsification principle in the context of philosophy of science is Karl Popper, The Logic of Scientific Discovery (1968).
existence of legal rhetoric and use well-chosen examples from history to illustrate the logic. Most important, they end the paper with specific hypotheses that could drive empirical research. The main weakness, perhaps, is that they do not present the most plausible competing arguments and a strategy for adjudicating among alternative explanations.\textsuperscript{53} Perhaps the authors felt that doing so was unnecessary since they are arguing against conventional ILS scholarship, which holds that rhetoric is inconsistent with rational choice. Even if the alternatives seem self-evident, however, they should contrast these predictions with their own.

Beyond illustrative examples, Downs and Jones also do not offer systematic evidence for their claims about the role of reputation in shaping international cooperation and compliance. Their research design, however, is quite effective. They begin by identifying the principal competing argument, drawn from mainstream IR and game theory, that reputation is central to explaining why actors keep commitments in a decentralized setting. They then argue that the relationship between reputation and compliance is weaker and more variable than the literature suggests; the bulk of the paper fleshes out the logic behind this assertion with a theoretical discussion. Like Goldsmith and Posner, they end the paper with several specific propositions—in effect, hypotheses for future research—that follow logically from the discussion: reputation promotes compliance most in trade and security affairs and least in environmental and human rights law, reputation has more influence on large states than on small ones, and new states (or new regimes) will be more sensitive on average to reputational considerations, for example. Downs and Jones do not, however, provide strategies for teasing out directly observable implications of these predictions. To be sure, reputation is often difficult to observe empirically, and they point out that it is very difficult to gather information on reputation.\textsuperscript{54} It nevertheless would have been helpful for them to begin to rectify this problem by identifying some strategies for observing reputational effects (and thus for testing reputation-based arguments).

Abbott and Snidal and Guzman and Simmons take the role of empirics most seriously. Though Abbott and Snidal are primarily concerned with establishing a framework for thinking about how interests and values interact in the development of law, they use qualitative case-study work to illustrate and develop their logic and as a sort of plausibility probe of their argument. Thus they are not “testing” their argument per se, they are simply showing how it might manifest itself in the real world and are careful not to overstate the implications of their empirics. Their case-study approach is the most appropriate for their project since qualitative description is needed to separate

\textsuperscript{53} They do briefly present one partial alternative explanation, that leaders engage in cheap talk for domestic political reasons, and argue effectively against it as an important independent variable.

\textsuperscript{54} Downs & Jones, \textit{supra} note 18, at S113–S114.
value actors from interest actors and to uncover the role of individual actors in the law-formation process. With falsifiability in mind, however, Abbott and Snidal could have been more specific about what the alternative arguments are and indicated specific strategies for ruling them out. This is especially important since few would agree with the apparent alternative—that either interests or values alone explain international legalization—in its simplest form. Guzman and Simmons present hypotheses that are precise enough to generate observable implications and use statistical analysis to control for a number of plausible competing variables. In terms of research design and empirics, the paper is thoughtful and effective. Nevertheless, even a modest amount of qualitative evidence (for example, interviews with government officials) would enhance the reliability of their finding that transaction costs actually influence the decision making of those pursuing trade disputes.

Roin asks whether tax base harmonization among the economies of the world will succeed and argues that it probably will not. She uses evidence from related situations, such as the behavior of states within the United States with respect to tax law harmonization, to illustrate her logic, but she provides no evidence of whether important actors are motivated by the factors she proposes. For example, she makes a strong deductive and intuitive case for why national legislatures might resist tax base harmonization, but she provides no evidence that legislators are in fact worried about losing control over tax policy in this area, despite her stated desire to focus on “attitudinal impediments.”

She also does not present any counterarguments for the potential failure of international tax base harmonization and evidence that might adjudicate between hers and other accounts. The paper implicitly rejects another hypothesis, that government leaders simply choose policies that maximize aggregate welfare and efficiency, but Roin does not explicitly address alternative explanations. Her mistake is to infer this conclusion after seeking evidence for her theory but not against it or for a rival theory.

Similarly, both Setear and Koremenos only briefly consider and dismiss counterarguments offered in the extant literature.

Schwartz and Sykes make a clear and provocative argument regarding the GATT/WTO provisions on violations and remedies: “We believe that the explanation for these provisions lies in the desire of signatories to facilitate efficient breach and in the relative superiority of a liability rule [rather than a property rule] approach to that task.”

The assumptions underlying their empirical evidence are worth noting. They are concerned with the intentions

55 Roin, supra note 13, at S62.
56 For more on this common violation of the rules of inference, see Epstein & King, supra note 51, at 7.
57 Setear, supra note 16, at S27 n.38; Koremenos, supra note 21, at S273–S274.
58 Schwartz & Sykes, supra note 11, at S187.
and "desires" of the framers, but, like Roin, they offer no evidence about the actual motivations or decisions of these individuals. They use the design of the WTO itself as their evidence. In other words, their economic rationality assumption is strong enough for them to assume that certain institutional features would not exist if their argument were not correct; they treat revealed preferences and outcomes as adequate evidence. Consider the following empirical claim in favor of their efficient breach argument: "We simply note that the provisions of the DSU, taken as a whole, allow a violator to continue a violation in perpetuity, as long as it compensates or is willing to bear the costs of the retaliatory suspension of concessions." This constitutes evidence for their arguments because "we see no other purpose to the provisions... Such a provision can only represent an institutional means for setting an appropriate price for violating commitments."

Though not unconventional in rationalist research in economics and political science, empirical arguments of this sort can be problematic and do not overcome the "correlation equals causation" problem. Ideally, confirmatory evidence of a rational choice explanation should show not only that the resulting action (in this case, an institutional design choice) is consistent with the stipulated goals, it should show that the action is caused by them. The institutional features adduced as evidence in the paper may reflect the preferences imputed by Schwartz and Sykes to the treaty creators, but they also may not. In fact, the authors cite arguments made by the United States to the effect that the purpose of sanctions in the WTO is to induce compliance and to be punitive, not to foster efficient breach, and they cite several provisions of the WTO that could be interpreted as obligating violators to comply rather than simply to compensate. So we are left wondering precisely what WTO members' preferences really were as they framed the new institution, that is, as they made rational choices.

Koremenos offers a nice counterpoint in terms of empirical strategies. The style of her argument is similar to that of Schwartz and Sykes—explaining an international institution as an equilibrium outcome—and yet her supportive evidence includes not only the resulting features of the International Coffee Agreement but also an overview of the negotiations that led to its creation. In other words, her empirics get more directly at the rational choices being made.

The Morrow paper offers an intriguing theoretical discussion, but its logic is not always precise enough to generate specific and noncontradictory hy-

59 Id.
60 Id. at S191.
61 Id. at S192; emphasis added.
62 See Elster, supra note 10, at 16.
63 Schwartz & Sykes, supra note 11, at S189–S190.
64 Koremenos, supra note 21, at S273–S276.
potheses. He asks why some law-of-war treaties succeed while others fail and argues that treaties codify a "shared understanding" of what behavior is unacceptable and what consequences might follow and that this shared understanding operates as a common conjecture to sustain an equilibrium of military competition. He develops a deductive framework for explaining when shared understandings might constitute common conjectures but provides no evidence that this is in fact driving a treaty equilibrium—manifested in higher compliance rates—in any given case. Giving some detail for even a single example would have helped clarify the causal logic, which is unclear at times. Consider the "screening effect" he attributes to treaties: if states must choose to sign or not sign a formal treaty before a war, this provides information to other states because it divides them into those who intend to comply and those who intend not to comply, thereby establishing an "in-group" that shares a common understanding. He gives the example of the United States's failure to support the International Criminal Court. But can we really infer from this that the United States will be less likely to abide by the standards of such a court? Is the United States not a member of the in-group when it comes to the commission of international war crimes? Morrow surely has a response to these questions (perhaps that the relevant standard is the actual procedures of the court, not the commission of war crimes more generally), but taking us through a case, or at least an example, would help flesh out his logic.

In addition, Morrow argues that formal law (that is, treaties) is able to serve as common conjecture, whereas customary law is not since it lacks sufficient precision. At the same time, however, he notes that even formal treaties have "a mixed record of compliance" and gives examples of customary laws of war that have produced compliance. These facts imply that treaties do not always embody shared understandings that can support an equilibrium and that customary law sometimes does. In any case, the logic and implications of the argument are not outlined clearly enough for one to generate specific hypotheses, let alone observable implications, concerning the role of treaty versus custom—tasks Morrow himself does not perform.

This section suggests that we should ask two questions of these and other contributions to the rational choice and IL research program. Is the theorizing designed to be falsifiable? Are empirics used effectively? Of course, different empirical methods have different virtues and limitations, and not every scholarly work hangs its hat on powerful empirical findings. The evidence needed to support some of the arguments might be difficult to collect, such as that on the effects of reputation in Downs and Jones's paper and evidence on the

---

65 In game theory, a common conjecture is an understanding among all players of each player's strategy. It is a necessary condition for an equilibrium to be attained.

66 Morrow, supra note 19, at S49–S50.

67 Id. at S41.
actual motivations behind actor behavior. The above critiques on empirics are meant to apply only insofar as empirics are deemed important and in order to compare the papers in terms of their empirical findings. On the theory construction side, however, the stakes are much greater. Without falsifiable arguments—based on clear logic, precise hypotheses, and strategies for disconfirmation—explanations based on rational choice will be inherently unconvincing and future research will not build effectively on extant theorizing.

VI. CONCLUSION: PUSHING THE BOUNDS OF RATIONAL CHOICE

The rational choice paradigm provides a powerful set of theories and tools for understanding international legal behavior and rules. It provides a rationalist and positivist counterweight to a field that sometimes relies on what Setear has called “mystical appeals to such concepts as ‘consent’ or ‘legitimacy.’”68 With the versatility of rational choice, however, comes potential pitfalls, and this paper points to three analytical and research design issues to which rational choice practitioners in IL and IR must be sensitive. First, I discussed the preference aggregation problem involved in treating corporate actors as unitary. Second, I noted various barriers to transferring models and propositions across empirical domains. Third, I offered guidelines for research design and empirical analysis that should foster falsifiable theorizing and contribute to a cumulative research program. Examples from the contributions to this volume were used to illustrate how these issues manifest themselves in practice (and, in many cases, how they can be dealt with effectively).

A notable feature of this special issue is that, while it helps introduce rational choice to the study of IL, it has already begun to take the next step: to show how a rationalist approach might complement, and be complemented by, other approaches. Rational choice theory almost always provides insight for understanding human behavior; on the other hand, it almost always needs to be complemented by other perspectives on what motivates action. After all, human decision making in the real world routinely violates the assumption of rationalist utility maximization,69 and we know that social norms and a desire to behave appropriately may also condition behavior.

Goldsmith and Posner, Morrow, and Abbott and Snidal push the bounds of rationalist assumptions to include various factors that are more typically associated with constructivism or norm-based analysis. They show that the logic of appropriateness often complements the logic of consequences. In their discussion of the content of cheap talk, Goldsmith and Posner avowedly


69 A seminal work in the vast experimental literature on decision making is Daniel Kahneman, Paul Slovic, & Amos Tversky, Judgment under Uncertainty (1981).
make a “concession to the limits of rationality” when they argue that states may form in-groups that are based on ideals and values. However, this sociological or identity factor is meant only to establish a focal point for rhetoric when there are multiple equilibria. Goldsmith and Posner also assume that domestic publics are less rational than states, both because they cannot process information as well and because they care more about emotions like “fear” and “hatred” that are not attributed to states.

The theoretical frameworks of two papers explicitly and fundamentally combine rationalism and constructivism. Morrow uses the game theory notion of common conjecture, a necessary component of an equilibrium, to describe the importance of “shared understandings,” a concept that lies at the heart of social constructivism. An equilibrium, in the form of a law-of-war treaty, thus reflects both calculated self-interest and a shared understanding among actors of acceptable conduct. He goes so far as to argue that states maintain moral stances: “states may comply with codes of conduct because their leaders believe that humane treatment is moral and appropriate.” This combination of interests based on consequences and interests based on appropriateness is central to Morrow’s understanding of why international treaties and law more generally work some times and not others; compliance is a function of both “reciprocal enforcement” and “state values.” The central aim of Abbott and Snidal’s project is to show the interrelation of the logic of consequences and the logic of appropriateness. Thus they divide actors into interest actors and value actors and show deductively and empirically how the two influence each other. They also do not impose fixed preferences on their actors, in contrast to the conventional rational choice approach (and consistent with social constructivism).

Huang pushes rationalism in a somewhat different direction. He presents a model of emotional rational choice: “This model extends unemotional rational choice theory by considering actors that are in part motivated by a desire to avoid losing face in the international community.” By folding such motivations into actor preferences and a game-theoretic framework, Huang’s paper has implications and inspires future research that go well beyond the issue of compliance with environmental agreements. It suggests that political psychology as well as rational choice theory might be usefully applied to the study of international law.

While sometimes pushing the boundaries of rationalism, all of the contributors to this volume assume that actors behave instrumentally and strategically. This suggests that even if rational choice theory as developed by economists, involving strict utility maximization, does not become the con-

---

70 Goldsmith & Posner, supra note 18, at S133.
71 Id. at S129.
72 Morrow, supra note 19, at S51.
73 Huang, supra note 15, at S240.
ventional approach to the study of IL, perhaps a near consensus will develop around the assumption that actors behave instrumentally and in systematic ways. This will facilitate positive theorizing based on insights from rational choice and will allow the IL field to continue to borrow from and contribute to the social sciences under a unified framework.