



Discover Archive

Vanderbilt University's Institutional Repository

This article was originally published as:

Timothy Meyer

**Power, Exit Costs, and
Renegotiation in International Law**

51 Harvard International Law Journal 379 (2010)

HEINONLINE

Citation: 51 Harv. Int'l L.J. 379 2010



Content downloaded/printed from
HeinOnline (<http://heinonline.org>)
Fri Aug 28 15:49:26 2015

- Your use of this HeinOnline PDF indicates your acceptance of HeinOnline's Terms and Conditions of the license agreement available at <http://heinonline.org/HOL/License>
- The search text of this PDF is generated from uncorrected OCR text.
- To obtain permission to use this article beyond the scope of your HeinOnline license, please use:

[https://www.copyright.com/ccc/basicSearch.do?
&operation=go&searchType=0
&lastSearch=simple&all=on&titleOrStdNo=0017-8063](https://www.copyright.com/ccc/basicSearch.do?&operation=go&searchType=0&lastSearch=simple&all=on&titleOrStdNo=0017-8063)

Power, Exit Costs, and Renegotiation in International Law

Timothy Meyer*

Scholars have long understood that the instability of power has ramifications for compliance with international law. Scholars have not, however, focused on how states' expectations about shifting power affect the initial design of international agreements. In this paper, I integrate shifting power into an analysis of the initial design of both the formal and substantive aspects of agreements. I argue that a state expecting to become more powerful over time incurs an opportunity cost by agreeing to formal provisions that raise the cost of exiting an agreement. Exit costs—which promote the stability of legal rules—have distributional implications. Before joining an agreement, an “ascendant” state will therefore often require either a greater share of the benefits from cooperation, or a reduction in exit costs through the use of formal provisions such as withdrawal clauses, sunset clauses, and provisions affecting the legality of an agreement. I analyze how states determine which concessions to make in order to reach agreement with an ascendant state. This analysis helps explain a number of puzzles in the international legal literature, such as why states with poor compliance rates are sometimes observed to join international agreements at the same or higher rates than states with good compliance rates; why weak agreements often evolve into more constraining agreements; and why multilateral agreements are more likely to have low exit costs than bilateral agreements.

INTRODUCTION

It is an inescapable fact of international relations that states rise and fall. The eighteenth century was dominated by France, the nineteenth by Britain, and the twentieth by the emergence of superpowers like Germany, the Soviet Union, and the United States, only to see the demise of the first two. Although only in its early years, the twenty-first century has witnessed the ongoing consolidation of the European Union, the rise of China and India as major economic powers, and the resurgence of Russia. Many of these changes in national fortunes occurred over historically brief periods. Germany became a major world power in the first fifteen years of the last century, was defeated in World War I and remained relatively weak during most of the 1920s, and then rearmed and reemerged as a major world power in the 1930s, only to be defeated again in World War II. The Soviet Union

* Assistant Professor, University of Georgia School of Law. I would like to thank Karen Alter, Brian Broughman, Bob Coorer, Dan Drezner, Mike Gilbert, Andrew Guzman, Larry Helfer, Ian Hurd, Mike Meurer, Jide Nzelibe, Bob Powell, Cesare Romano, Alvaro Santos, Jeff Staton, and the participants in the International Law/International Relations Roundtable at Vanderbilt University Law School and the International Organization/International Law Workshop at Northwestern University. All errors remain my own.

broke up in the early 1990s, but less than two decades later Russia is reasserting its regional authority and reviving historic east-west tensions.

These shifts in the international political landscape invite the question: what effect do the rise and fall of nations have on international law? More specifically, how do states bargain over the design of international agreements when they believe that they or their cooperative partners might become stronger (or weaker) down the road? Although power may be defined in a variety of different ways, this question deals with a specific type of power: the bargaining power that flows from being able to credibly threaten to withhold one's participation in an institution, such as an international agreement.¹ It is in this sense that I use the terms "power" and "bargaining power" throughout this Article.² The ability to credibly threaten not to participate in an institution depends on one's alternatives to cooperation. A state only has a credible threat to withhold its participation if its payoff from noncooperation—what I refer to as a state's outside option—is greater than its payoff from cooperation. Thus, a state's bargaining power depends on its outside option. As the value of a state's outside option increases relative to the value of cooperation, so too does a state's bargaining power.

In acting as international lawmakers, states must surely take into account potential shifts in bargaining power. A state might, for example, delay making an agreement because it expects that it will be more powerful relative to its potential cooperative partners later, and hence able to extract beneficial concessions. Alternatively, a state might be willing to make concessions now to avoid the possibility of bargaining from a weaker position in the future. And, of course, if a state finds itself having become more powerful since it entered into an agreement, it may seek to renegotiate or revise that agreement to reflect the new underlying power distribution.

Yet despite both their pervasiveness and importance to the stability of international law, shifts in bargaining power have largely been omitted from the legal scholarship on how states design international agreements. To be sure, international relations scholars have long understood the importance of shifts in power for conflicts between states,³ but until recently they had only

1. To some extent, this definition of "power" resembles the concept of "go-it-alone power" as set forth by LLOYD GRUBER, *RULING THE WORLD: POWER POLITICS AND THE RISE OF SUPRANATIONAL INSTITUTIONS* 6–7 (2000). Like Gruber's notion of "go-it-alone power," states here have leverage to renegotiate the terms of cooperation when they can make a credible threat to abandon the existing terms of cooperation. However, the term "go-it-alone power" suggests that a state's alternative to cooperation is always noncooperation. As I explain later, I have a more expansive notion of what constitutes alternatives to cooperation that includes cooperating in other fora or with other partners.

2. This is, of course, not to deny that other conceptions of power may be appropriate for answering other questions.

3. See, e.g., HANS J. MORGENTHAU, *POLITICS AMONG NATIONS* (1967); ROBERT POWELL, *IN THE SHADOW OF POWER: STATES AND STRATEGIES IN INTERNATIONAL POLITICS* (1999) (analyzing how power configurations affect the probability of war).

rarely ventured to apply those insights to the field of international law.⁴ For their part, international legal scholars studying the design of international agreements have focused their attention on describing the tradeoffs between formal legal provisions—those that affect the credibility of a commitment—and substantive legal provisions—those obligations related directly to the subject matter of an agreement.⁵ These scholars have not, however, examined how states' expectations about shifting bargaining power—and thus their expectations about the possibility of renegotiation—affect the design of international agreements. And while legal scholars have by no means turned a blind eye to power's relationship to international law more generally, they have tended to focus on the ways in which power can shape observance of the law, rather than exploring the antecedent question of how power shapes the law.⁶

The seeds of future actions are, however, sown in the soil of contemporary behavior. One cannot understand why shifts in power cause states to abandon international rules to which they have previously agreed without first understanding how states thought about the possibility of shifts in power when creating those same rules. In short, one must integrate shifting power into one's theory of how states *make* law before one can fully comprehend how shifting power affects states' *observance* of law.

This Article evaluates how the fragility of power can affect the design of international agreements, and specifically how it can affect the tradeoff between form and substance. In so doing, this Article adds to our understanding of how states bargain over the substantive and formal aspects of international agreements and the creation of international law more generally. Recognizing wide variation in the types of formal provisions used in international agreements—including escape clauses, withdrawal clauses, sunset provisions, dispute resolution provisions, and provisions governing the legality of an agreement—legal scholars have sought to explain what motivates states to include particular provisions but not others. What has emerged is a healthy literature that explains much of the variation we observe in international agreements. Risk and uncertainty manifest themselves in different ways, and different types of provisions add value, or fail to when one might expect they would, for different reasons. Moreover, by pairing different formal provisions with different substantive commitments, states are able to generate agreement and cooperation in areas where it might not

4. For examples of political science research exploring the design of international agreements, see Kenneth W. Abbott & Duncan Snidal, *Hard and Soft Law in International Governance*, 54 INT'L ORG. 421, 446–47 (2000); Barbara Koremenos, *Loosening the Ties that Bind: A Learning Model of Agreement Flexibility*, 55 INT'L ORG. 289 (2001).

5. See generally Andrew T. Guzman, *The Design of International Agreements*, 16 EUR. J. INT'L L. 579 (2005) [hereinafter Guzman, *Design*]; Kal Raustiala, *Form and Substance in International Agreements*, 99 AM. J. INT'L L. 581 (2005).

6. See, e.g., JACK L. GOLDSMITH & ERIC A. POSNER, *THE LIMITS OF INTERNATIONAL LAW* 11–13 (2005).

otherwise have been possible. At the end of the day though, states make these tradeoffs between form and substance based on political and economic realities that may not be the same from day to day or year to year. States' expectations about how their political fortunes will change should, then, play a major role in the design of international agreements.

The key to understanding how shifting power affects the tradeoff between form and substance is recognizing that such shifts can allow a state to credibly threaten to exit an agreement. A credible threat to exit an international agreement confers power on a state by allowing the state to demand a greater share of the gains from cooperation in exchange for participating. Drawing on this notion of credible threats to exit, I conceptualize bargaining power as the value of a state's outside option—its next-best alternative to the existing or proposed agreement. If a state's outside option becomes more valuable than the current agreement, a state can demand to renegotiate the agreement in its favor. And if renegotiation is unsuccessful, the state can withdraw from the agreement. Although it is perhaps counterintuitive, such renegotiations and withdrawals can make other states worse off relative to the status quo agreement and yet still be rational for all parties.⁷

Thus, when bargaining over the initial creation of an agreement states must be cognizant of their expectations about how bargaining power may shift in the future. In designing international agreements, states have some control over the likelihood that a state will develop a credible threat to exit. That is, within certain limits, states can mitigate the effects of shifting power and therefore the likelihood of renegotiation. States have at their disposal a variety of different formal provisions that can raise or lower the cost of exiting an agreement, including withdrawal provisions, sunset or duration provisions, and provisions affecting the legality of an agreement. Because a state must incur the costs of exiting an existing agreement to avail itself of its outside option, raising exit costs reduces the likelihood of a state later developing a credible threat to exit—and therefore reduces the likelihood of renegotiations that are not Pareto-improving.

As even the casual observer can tell, however, states routinely employ low exit costs and thus raise the possibility of future redistributive amendments. The possibility of non-Pareto-improving amendments is a puzzle, because we would generally think states would want their agreements to be renegotiation-proof (or as close to renegotiation-proof as possible), both to avoid the possibility of losing the gains tomorrow that they bargained for today and to induce reliance on the promises exchanged. In this Article, I explore how expected shifts in bargaining power can explain this oddity. A state that expects to see its outside option improve tomorrow (an "ascendant state") incurs an opportunity cost by entering into an agreement today. Specifically, because it will have to pay the exit costs associated with the agree-

7. See GRUBER, *supra* note 1, at 38–40.

ment to obtain its outside option in the future, the ascendant state reduces its expected future leverage. The ascendant state will therefore demand a greater share of the benefits from cooperation—it will, in other words, demand substantive rules that are more favorable. High exit costs therefore narrow the set of substantive rules that make an agreement today worth foregoing tomorrow's bargaining power. In extreme situations, high exit costs may preclude an agreement altogether.

Alternatively, the ascendant state may demand, or a "declining state" may offer, low exit costs. These low exit costs allow the ascendant state to retain the ability to renegotiate when it becomes more powerful. By allowing the ascendant state greater freedom to exit its obligations, the declining state may not have to make as many concessions on the substantive rules in order to get an agreement.

Shifting bargaining power, in other words, creates incentives for both ascendant and declining states to concede on the issue of exit costs. Declining states may actually wish to have low exit costs because they prefer a greater share of the benefits from cooperation today, and conceding on the issue of exit costs may allow them to extract preferable substantive rules. Low exit costs also facilitate entrance by ascendant states, making more agreements possible. In contrast, ascendant states may in some cases be willing to concede high exit costs, precisely because of the bargaining power it confers on them in negotiating substantive rules.

The fact that states may capture expected shifts in power in either the formal or substantive aspects of an agreement, and that in doing so they often opt to reduce exit costs, complicates our understanding of how states behave in authoring international law. My analysis shows first that in the presence of shifting power, stable legal rules—those with high exit costs—are not distributionally neutral.⁸ Intuitively, high exit costs favor declining states by locking in the terms of cooperation despite future shifts in bargaining power. Perhaps less intuitively, stable and enduring international legal rules will tend to substantively favor ascending states in a way that may undermine their attractiveness. Put differently, all else equal, we would expect to see less stable and durable legal rules in the presence of shifting bargaining power. Second, states with little to gain from existing rules are "powerful" in the sense in which I use that term, precisely because their outside options likely present them with credible threats to walk away from existing rules of cooperation. This means that states on the outskirts of international cooperation—rogue states, or developing nations that have not historically participated in the formation of international legal rules—will have a disproportionate influence on the shape of legal rules when those rules are important to broader efforts at cooperation.⁹ Third, states have an

8. See *infra* Part III.C.2.

9. See *infra* Part IV.

incentive to undermine cooperation by investing in their outside options. A state might, for example, develop alternative legal institutions or technological capabilities that make existing legal rules relatively less attractive to it. A state that benefits from contributing to a public bad might withhold its participation in schemes aimed at solving that public bad when its activities that contribute to the public bad are on the rise. Consider a developing nation that industrializes despite the negative effect of industrialization on climate change: as a growing industrial economy becomes more important to a state, limitations on industrial growth—such as those that might be imposed as part of a climate change regime—become costlier for that state. And as legal rules limiting industrial growth become costlier, a state is in a position to withhold its participation if the rules are not favorable to it. Improved outside options, in other words, may allow a state to negotiate or renegotiate legal rules in its favor.¹⁰

This Article proceeds in five parts. In Part I, I briefly review the existing literature on the design of international agreements and the tradeoff between form and substance. Although this literature explains much of the variation we observe, shifts in power—a commonplace occurrence in the real world—are absent from discussions on form and substance. In Part II, I explore the doctrinal and functional issues related to exit costs: what provisions affect exit costs and how? I also discuss how, from the standpoint of inducing renegotiation, the ability to exit an agreement differs from the ability to violate an agreement. Part III develops the theory of how states will design international agreements in light of changes in bargaining power. The theory shows that where power is shifting, states will often face a tradeoff between pursuing high exit costs, which reduce the likelihood of future renegotiation, and pursuing favorable substantive rules. The paradoxical result is that states that expect to become weaker over time may sometimes prefer an agreement that leaves them vulnerable to having to give up their bargained-for gains, while states that expect to become more powerful will be willing to give away their future bargaining leverage if they can obtain sufficiently favorable substantive rules now.

Part IV explores the implications of this theory. I argue that the theory of bargaining against the backdrop of expected shifts in power helps explain a number of puzzles in international law, including why states with poor compliance rates sometimes join international agreements at higher rates than states with good compliance rates, why weak agreements often evolve into more constraining agreements, and why multilateral agreements are more likely to have low exit costs than bilateral agreements. Part V concludes.

10. See *infra* Part IV.

I. EXISTING THEORIES OF AGREEMENT DESIGN

Shifting power has been an omitted variable in certain parts of the international law literature. Scholars have, to be sure, paid heed to the relationship between power and law.¹¹ But legal scholars have sometimes ignored the instability of power. Moreover, legal scholarship that does incorporate power has generally not distinguished between power's effect on *compliance* with international law and power's effect on the *creation* of international law itself, emphasizing the former over the latter.¹² Yet one cannot understand how states behave as subjects of international law without first understanding how they behave as creators of international law.

Over the past several years, scholars have begun to rise to this challenge, putting forward a number of theories as to how international agreements are designed. To varying degrees, these theories shed light on how states attempt to cope with uncertainty generally and the likelihood or desirability of future renegotiations. This section discusses and evaluates some of these existing theories.

A. Form and Substance

Before discussing theories, it is useful to clarify terminology. Scholars working on the design of international agreements generally agree that different "design elements" or "dimensions"¹³ of international agreements interact with each other to produce the wide variation we see in international agreements. More specifically, scholars have generally found it useful to categorize provisions of international agreements as relating to matters of "form" or matters of "substance."¹⁴ Broadly speaking, "substance" refers to those obligations directly related to the subject matter of an agreement. For example, in an agreement governing trade in nuclear technology, an obligation not to sell specific types of nuclear materials or equipment to states that have not accepted international safeguards is a substantive commitment.¹⁵ Substantive commitments can be viewed through a variety of lenses. Many scholars studying the form-substance tradeoff focus their analysis of substance on the depth of cooperation, or the extent to which an agreement's

11. The effects of power—but not shifts in power—on international law generally have, of course, been a subject of great discussion in international law for decades. These discussions have generally drawn on the study of international institutions by political scientists and scholars of international relations. *See, e.g.,* Richard H. Steinberg & Jonathan M. Zasloff, *Power and International Law*, 100 AM. J. INT'L L. 64 (2006).

12. Notable studies that do deal with the creation and evolution of international agreements include GRUBER, *supra* note 1; Richard H. Steinberg, *In the Shadow of Law or Power? Consensus-Based Bargaining and Outcomes in the GATT/WTO*, 56 INT'L ORG. 339 (2002).

13. Raustiala, *supra* note 5, at 583.

14. *See generally id.*; Guzman, *Design*, *supra* note 5.

15. Timothy Meyer, *Soft Law as Delegation*, 32 FORDHAM INT'L L.J. 888, 926 (2009) (discussing the safeguard requirements for the transfer of certain nuclear technology) [hereinafter Meyer, *Delegation*].

obligations require deviation from status quo behavior.¹⁶ I focus below on the distributional implications of substantive commitments and ask a particular question: are the subject-matter specific obligations contained in an agreement more favorable to one state or another?

“Form” refers to those aspects of an agreement that relate to the credibility of the substantive obligations created, or which prescribe procedures for monitoring compliance and shaping the content of the substantive obligations. So, for example, the decision whether to include dispute resolution provisions is a matter of form. Tribunals provide both monitoring functions and expositional functions; that is, tribunals create information about whether a state is complying with its obligations, and they also clarify the content of those obligations.¹⁷ But the obligation to submit to a tribunal’s jurisdiction is not a substantive one; it is separate and apart from the subject-matter of the agreement. Provisions of an agreement affecting exit costs, such as sunset provisions, withdrawal provisions, and legality are classic examples of an agreement’s formal aspects.

The categories of form and substance are useful because they allow scholars to think about how states trade off different design elements in creating international agreements. If, for example, an agreement is to include invasive monitoring procedures which increase the risk of violations being detected, states might wish to reduce the risk of violation by making their substantive obligations easier to comply with.¹⁸ Another area in which the tradeoff between form and substance is obvious is in the difference between hard law, or legally binding obligations, and soft law, or non-binding obligations that nevertheless have legal consequences. All else equal, a violation of a hard obligation carries with it a larger reputational sanction for non-compliance than violation of a soft law obligation. The reason for this increased sanction is that agreeing to make an obligation hard law is a more serious signal about a state’s intention to comply with the obligation.¹⁹ If a state sends a signal about its intention to comply and then fails to follow through, its reputation for compliance will suffer, hurting its ability to extract concessions in future negotiations.²⁰ States may therefore wish to control their risk of sanctions by making hard law obligations easier to comply with than soft law obligations. Consider, for example, export controls on nuclear materials. Article III of the Nuclear Nonproliferation Treaty (“NPT”) provides that states may only transfer nuclear equipment or material that can be used to produce or process fissionable material if the transfer

16. Guzman, *Design*, *supra* note 5, at 602–03; Raustiala, *supra* note 5, at 584.

17. Andrew T. Guzman & Timothy L. Meyer, *Customary International Law in the 21st Century*, in *PROGRESS IN INTERNATIONAL LAW* 214–15 (Rebecca Bratspies & Russell Miller eds., 2008).

18. Guzman, *Design*, *supra* note 5, at 603.

19. Andrew T. Guzman & Timothy Meyer, *International Soft Law*, 2 J. LEG. ANALYSIS (forthcoming 2010) (manuscript at 20–21), available at <http://ssrn.com/abstract=1353444> [hereinafter Guzman & Meyer, *International Soft Law*].

20. *Id.*

occurs under safeguards.²¹ The exact types of technology or equipment that trigger Article III's obligations and the safeguard measures necessary when transferring such technology are, however, nowhere specified in the NPT. Instead, soft law rules such as those produced by the Zangger Committee or the Nuclear Suppliers Group are the source of regulations governing precisely what technology may be transferred and how.²²

B. Uncertainty and Risk

Having clarified terms, the first group of theories worth mentioning is that based on uncertainty and risk. Following the form-substance distinction, this group of theories deals directly with the question of how states, in bargaining over international agreements, trade off the substantive provisions of an agreement (distributional concerns) and the formal aspects of agreements that are designed to control uncertainty and risk. These theories approach that general question from a variety of angles. Some studies have hypothesized that where states are uncertain about whether conditions affecting the size or distribution of the cooperative benefits created by an agreement may change, they may actually wish to mandate future negotiations, rather than try to avoid them.²³ Planned renegotiation, such as may be provided by sunset provisions²⁴ or framework conventions that mandate negotiations on future protocols,²⁵ address this type of uncertainty by forcing states back to the drawing board to consider changed or evolved circumstances.

This view makes sense when applied to situations in which power is expected to be relatively stable. If bargaining power is expected to remain roughly the same over time, planned renegotiation provides states with a mechanism to deal with uncertainty or unforeseen changes in conditions.²⁶ But the explanatory power of uncertainty can be overstated. In particular, this account makes considerably less sense in situations in which states expect power to shift over time.

Changing conditions can, to be sure, result in shifts in the distribution of benefits under an international agreement. For example, the strong intellectual property protections contained in the Trade-Related Aspects of Intellectual Property Rights Agreement ("TRIPS") have had the effect of raising

21. Treaty on the Non-Proliferation of Nuclear Weapons art. III, July 1, 1968, 21 U.S.T. 483, 729 U.N.T.S. 161 [hereinafter NPT].

22. Meyer, *Delegation*, *supra* note 15, at 924–25.

23. See JOSEPH GRIECO, COOPERATION AMONG NATIONS: EUROPE, AMERICA, AND NON-TARIFF BARRIERS TO TRADE 228 (1990) ("If two states are worried or uncertain about relative achievement of gains, each will prefer a less durable cooperative arrangement, for each will want to more readily be able to exit from the arrangement in the event that gains favor the other."); Koremenos, *supra* note 4, at 291–92.

24. Koremenos, *supra* note 4, at 309.

25. John K. Setear, *Ozone, Iteration, and International Law*, 40 VA. J. INT'L L. 193, 213 (1999).

26. Koremenos, *supra* note 4, at 292–93.

the cost of medicine above what developing countries can pay to fight the HIV/AIDS epidemic those countries face.²⁷ As a result of pressure by developing nations, TRIPS was effectively amended in 2001, through the Declaration on the TRIPS Agreement and Public Health, to give the least developed WTO members an additional ten years before they must protect pharmaceuticals.²⁸

However, the fact that conditions may change and disrupt the anticipated benefits of the existing agreement does not mean that states will be willing or able to negotiate a return to the initial distribution of benefits, as planned renegotiation assumes they would. If bargaining power shifts between the initial negotiation and a renegotiation, one would expect the outcome of renegotiation to reflect the new underlying distribution of power. Thus, at the initial negotiation of an agreement, one would predict that states expecting to lose bargaining power in the future should be opposed to mandating renegotiation through sunset provisions or framework agreements.²⁹ Such provisions, by dissolving the existing agreement just as some state or group of states is getting stronger, confers bargaining power on those rising states. Planned renegotiation should only emerge in areas in which major shifts in bargaining power are not expected, or where uncertainty about an agreement's implementation is so severe that it swamps concerns about shifting power.

A second theory based on uncertainty holds that because states cannot write perfectly enforceable contracts, they will balance the marginal expected costs associated with steeper penalties for undeterred violations against the benefits from deterred violations.³⁰ Because international law lacks a system of compensatory damages, sanctions assessed against a violating state are a net loss to the parties.³¹ While not specifically concerned with renegotiation, this approach sheds some light on how states may think about controlling the likelihood of renegotiation.³² Like exit, violations or the threat thereof can be used to withhold cooperative behavior and thereby inflict costs on one's cooperative partners. Public violations are thus a bit like temporary (and illegal) exits; they can be used to spur renegotiation.³³

27. Laurence R. Helfer, *Regime Shifting: The TRIPS Agreement and New Dynamics of International Intellectual Property Lawmaking*, 29 YALE J. INT'L. L. 1, 5 (2004) [hereinafter Helfer, *Regime Shifting*].

28. World Trade Organization, Ministerial Declaration of 14 Nov. 2001 on the TRIPS Agreement and Public Health, ¶ 7, WT/MIN(01)/DEC/W/2, 41 I.L.M. 755 (2002).

29. For high levels of uncertainty about an agreement's effects, states might still prefer renegotiation in the presence of shifting power. Shifting power, however, reduces the appeal of renegotiation. Koremenos, *supra* note 4, at 296 n.26.

30. Guzman & Meyer, *International Soft Law*, *supra* note 19, at 23; cf. Laurence R. Helfer, *Overlegalizing Human Rights: International Relations Theory and the Commonwealth Caribbean Backlash Against Human Rights Regimes*, 102 COLUM. L. REV. 1832 (2002) [hereinafter Helfer, *Overlegalizing*].

31. Guzman & Meyer, *International Soft Law*, *supra* note 19, at 20.

32. Meyer, *Delegation*, *supra* note 15.

33. *Id.*

Just as raising exit costs reduces the threat of renegotiation, so too does raising the sanctions associated with violation.

Perhaps the theories with which this Article has the most in common are those that argue that states may use withdrawal clauses or escape clauses as a form of insurance.³⁴ Such insurance, while suboptimal from an *ex post* standpoint, can be optimal from an *ex ante* standpoint because it helps to facilitate agreement where agreement might not otherwise have been possible. Withdrawal and escape clauses facilitate *entrance* at the expense of also facilitating *exit*. These studies thus highlight the relationship between form and agreement membership over the life of an agreement. In some situations, broad membership at an agreement's conception can come with an increased likelihood of declining membership over time.³⁵ Moreover, these studies have provided preliminary empirical support for the idea that the legal right to exit is meaningful to states. Helfer has shown that treaty denunciation, conventionally understood to be an aberration, is relatively commonplace, if not frequent.³⁶ He finds 1,547 treaty denunciations between 1945 and 2004, or roughly one denunciation for every twenty-one ratifications during that period.³⁷ This figure, although seemingly small, highlights the importance of exit as an empirical phenomenon, particularly when one considers that a) we should only observe an exit when the threat of exit has failed to spur renegotiation;³⁸ and b) Helfer's figure includes only denunciations and withdrawals from hard law agreements, but not exits from soft law agreements or exits that occur through other mechanisms, such as a failure to renew an agreement.³⁹

Like those theories that explicitly contemplate renegotiation, however, these latter theories generally do not take into account bargaining power, let alone changes in bargaining power. Moreover, while these theories offer valuable insights into how an individual agreement is designed, they do not offer a window into how the *next* agreement is designed or how renegotiation occurs. In short, they do not sufficiently explain the evolution of international law. An analysis of shifting power, renegotiation, and provisions affecting exit costs more broadly—including not only withdrawal provisions

34. Laurence R. Helfer, *Exiting Treaties*, 91 VA. L. REV. 1579, 1599–1601 (2005) [hereinafter Helfer, *Exiting Treaties*]; Alan O. Sykes, *Protection as a 'Safeguard': A Positive Analysis of the GATT 'Escape Clause' with Normative Speculations*, 58 U. CHI. L. REV. 255, 279 (1991).

35. Indeed, some treaties explicitly contemplate declines in membership that may render their continuation in force impractical. *See, e.g.*, Convention Concerning the Unification of Road Signals art. 15, Mar. 30, 1931, 150 L.N.T.S. 247 (stating that the treaty will terminate if and when fewer than five states remain as parties).

36. Helfer, *Exiting Treaties*, *supra* note 34, at 1601–02.

37. *Id.* at 1604.

38. *See infra* Part III.

39. Aggregate data on soft law agreements are extraordinarily difficult to obtain or develop. Unlike hard law agreements, which are generally reported to the United Nations or tracked at the national level through mechanisms such as the Case Act, 1 U.S.C. § 112(b)(a) (2004), soft law agreements are not subject to any centralized reporting system.

but also sunset provisions and provisions affecting legality—offers a dynamic theory of international law that current theories do not.

C. Institutionalism and Neorealism

Other studies of the design of international institutions more explicitly take into account the role of bargaining power but do so at the expense of legal nuance. Political scientists working in the institutionalist tradition, for example, have explored how the power to exclude parties from cooperation can shape the cooperative regime that ultimately emerges.⁴⁰ The essential idea here is that groups of like-minded states can form “enacting coalitions.”⁴¹ Rather than compromise on the terms of cooperation with those outside the coalition, the coalition is willing and able to “go it alone” if other states do not adopt rules preferred by the enacting coalition.⁴² By excluding or threatening to exclude other cooperation-minded states, the enacting coalition can ensure that cooperation is on terms at least as good as those it could get by “going it alone.”⁴³ These studies are foundational, highlighting the relationship between credible threats to “go it alone” and the shaping over time of substantive rules of cooperation. Like much of the political science literature on international institutions, though, law is largely omitted from these studies. Instead, these studies focus almost entirely on the substantive terms of cooperation, while neglecting the formal aspects of international institutions that can shape substantive outcomes.

Other scholars have posited a sharp tension between power and the rule of law. Powerful states are sometimes assumed to prefer “shallow” agreements, or agreements that require less of a deviation from status quo behavior.⁴⁴ On this view, powerful states resist paying the costs of changing their behavior because they can—in effect, being powerful means being exempt from the costs of cooperation. A similar hypothesis comes from the neorealist view that international institutions simply reflect the underlying distribution of power.⁴⁵ It has been argued that powerful states exploit their bargaining power in the design of international agreements to ensure that voting rules allow them to control substantive outcomes.⁴⁶ On this view, the formal aspects of international agreements fail to deter renegotiation because the

40. See George W. Downs, David M. Rocke & Peter N. Barsoom, *Managing the Evolution of Multilateralism*, 52 INT'L ORG. 397 (1998); GRUBER, *supra* note 1, at 38–40.

41. GRUBER, *supra* note 1, at 6–7.

42. *Id.* at 7.

43. *Id.* at 38–41.

44. See, e.g., Raustiala, *supra* note 5, at 604 (“States with veto power may demand a shallow agreement or side payments to cooperate (or both).”).

45. As Steinberg and Zasloff point out, this view has sometimes been attributed to all realists. It is more accurately attributed to neo- or structural realists only. See Steinberg & Zasloff, *supra* note 11, at 74–75.

46. See, e.g., GOLDSMITH & POSNER, *supra* note 6, at 13 (“[U]nder our theory, international law does not pull states toward compliance contrary to their interests”); see generally STEPHEN KRASNER, SOVEREIGNTY: ORGANIZED HYPOCRISY (1999); MORGENTHAU, *supra* note 3; Stephen Krasner, *Structural*

powerful ensure that the formal aspects of international agreements simply enshrine their bargaining power. This approach thus recognizes the distinction between formal and substantive provisions, while at the same time maintaining that power relationships render the former generally obsolete.

Neither of these accounts fully explains the relationship between power, form, and substance. With respect to the first, the fact that the powerful sometimes resist cooperation is self-evident; that the effect of power on bargaining is a one-way ratchet toward shallower cooperation is far from obvious. An individual state's net benefit from participating in an international agreement can be thought of as the sum of the benefits it receives from each complying member state's behavior, less its own individual costs of complying.⁴⁷ Formally, this expression can be written as:

$$U_j = (\sum^n b_{i,i \neq j}) + (b_j - c_j).$$

The first term represents the sum of the benefits the j th state receives from every other state's compliance (where b_i is the benefit to the j th state of the i th state's decision to comply) and the second term represents the benefits less the costs to the j th state of its own decision to comply.⁴⁸ Regardless of power, we would expect individual states negotiating an international agreement to attempt to maximize this formula. In other words, we would expect states acting strategically to advocate rules that maximize the net benefits they individually receive. The assumption that powerful states prefer shallow cooperation is really an assumption that states seek to minimize c_j . But an individual state should be willing to increase the marginal cost it pays to comply with an obligation—the marginal depth of its cooperation—up to the point at which that marginal cost equals the marginal benefit it receives from a marginal increase in the cumulative compliance of all states. In terms of the formula above, if an increase in c_j is accompanied by an equal or greater increase in $(\sum^n b_{i,i \neq j}) + b_j$ then state j should be willing to agree to an increase in the depth of cooperation.

Thus, whether powerful states want shallow cooperation should depend on how an obligation that increases their marginal costs of complying translates into marginal benefits. Depending on circumstances, a powerful state might prefer to dictate deep cooperation. The legal regime governing ozone-depletion is a good example. Although the good produced—a reduction in chlorofluorocarbons and halon that deplete the ozone layer—is a public good (and thus the benefits of cooperation were not internalized fully by the

Causes and Regime Consequences: Regimes as Intervening Variables, in *INTERNATIONAL REGIMES* (Stephen Krasner ed., 1983).

47. SCOTT BARRETT, *ENVIRONMENT AND STATECRAFT: THE STRATEGY OF ENVIRONMENTAL TREATY-MAKING* 202 (2003). Put differently, the second term is the net benefit a state gets from acting unilaterally, while the first term is the benefit a state gets from the actions of other states. So, for example, in an arms control agreement, the second term would capture those benefits the j th state receives simply from reducing its own arms: costs savings, reduced risk of accidents, etc. The first term would capture the bulk of the security gains which result from the other state's reductions in arms.

48. *Id.*

parties paying the costs of compliance), the costs of even deep cooperation were sufficiently low in proportion to the overall benefits that leading states such as the United States preferred deep cooperation.⁴⁹

The neorealist view, for its part, overstates its case. Even supposing for the sake of argument that states that are powerful at the moment of negotiation are able to dictate freely those formal provisions of an agreement that allocate future bargaining power among the parties, the assumption that they would always prefer rules that translate contemporary bargaining power into outcomes overlooks the same real world phenomenon as those hypotheses relying on uncertainty and risk: shifts in power. Because low exit costs privilege states with attractive alternatives to cooperation—by allowing them to exploit a credible threat to exit in order to demand favorable amendments—increases in the value of *any state's* outside option can reallocate bargaining power. In other words, writing into an agreement formal provisions that make control of substantive rules depend on exogenous shifts in bargaining power is not likely to be favored by those using an agreement to capture the benefits of transitory bargaining power.⁵⁰ Therefore, as with the uncertainty-risk hypotheses, we would only expect the powerful to dictate the procedural rules neorealists predict in situations in which they expect power to be stable over time.

The puzzle presented by the relationship between form, substance, and shifting power thus remains. Existing theories of agreement design have greatly expanded our understanding of how states approach the tradeoff between form and substance, as well as how power can affect the substantive terms states negotiate. But no theory has yet integrated shifting power into a theory of how states design international law. Absent such a theory, we cannot understand how states use *both* the formal and substantive aspects of an agreement to shape the ways in which international law evolves over time. Simply put, we do not know why states leave their agreements vulnerable to redistributive renegotiation forced by states that have become more powerful.

49. Setear, *supra* note 25, at 196. This was in part because there were low-cost alternatives available, and also because a few countries contributed disproportionately to the emissions of chlorofluorocarbons and halon, thus reducing fears about cheating. *Id.*

50. Of course, formal provisions that allocate bargaining power through the prescription of decision rules will often not be as directly vulnerable to exogenous shifts in power. For example, if an agreement gives a specific state a veto, as the U.N. Charter gives the five permanent members of the Security Council, shifts in power may not as directly reallocate bargaining power within the agreement. With such an arrangement, however, the danger is that if *de facto* power does not translate into some measure of *de jure* bargaining power, disenfranchised powerful states will undermine the agreement's effectiveness by, for example, establishing a rival agreement or engaging in forum shopping. Kal Raustiala & David G. Victor, *The Regime Complex for Plant Genetic Resources*, 58 INT'L ORG. 277, 279–80 (2004).

II. SHIFTING POWER, RENEGOTIATION, AND EXIT COSTS

Facilitating exit from international agreements leads to renegotiations if power shifts among the parties. These renegotiations can make states worse off than under the status quo agreement. Yet states nevertheless routinely employ provisions in agreements that reduce exit costs and thereby make renegotiation more likely. This Part defines exit costs and describes the different types of provisions that affect the cost of exiting an international agreement. It also explains why exit is distinct from violation in terms of the ability to drive a renegotiation.

A. What Are Exit Costs?

States are, in a practical sense, always free to exit an agreement,⁵¹ so before analyzing exit one first needs a theory of what makes some exits more costly than others. What are exit costs and what is the mechanism through which they are imposed?

Exit is, in essence, a legal right. No state is obliged to enter an agreement in the first place, and so as an initial matter all states have the right to control the legal obligations that bind them.⁵² But like most legal rights, the right to exit is neither absolute nor mandatory. It can be given away or circumscribed upon entering an agreement. Indeed, the default rule of international law, as codified in Article 56 of the Vienna Convention on the Law of Treaties (“VCLT”), is that an agreement that is silent on the right to withdraw generally does not contain such a right.⁵³

As discussed below, there are a variety of different provisions that affect a state’s legal ability to exit an agreement. These provisions detail circumstances under which exit is lawful, restrict or make conditional the right to exit, or impose potentially costly bargaining procedures on would-be exiters. The obligation to refrain from exiting an agreement except in accordance with these terms, conditions, and procedures is a legal obligation, and there-

51. See J.H.H. Weiler, *The Transformation of Europe*, 100 *YALE L. J.* 2403, 2412 (1991) (noting that “[i]t takes no particular insight to suggest that should a Member State consider withdrawing from the [European Coal and Steel] Community, the legal argument will not be the critical or determining consideration”).

52. See, e.g., Duncan B. Hollis, *Why State Consent Still Matters—Non-State Actors, Treaties, and the Changing Sources of International Law*, 23 *BERKELEY J. INT’L L.* 137, 138 (2005).

53. Vienna Convention on the Law of Treaties art. 56, May 23, 1969, 1155 *U.N.T.S.* 331 [hereinafter *VCLT*].

1. A treaty which contains no provision regarding its termination and which does not provide for denunciation or withdrawal is not subject to denunciation or withdrawal unless:
 - (a) it is established that the parties intended to admit the possibility of denunciation or withdrawal; or
 - (b) a right of denunciation or withdrawal may be implied by the nature of the treaty.
2. A party shall give not less than twelve months’ notice of its intention to denounce or withdraw from a treaty under paragraph 1.

Id.

fore compliance with that obligation is subject to the same mechanisms that impose costs for violations of international law generally. These mechanisms are reciprocity, retaliation, and reputation.⁵⁴ Reciprocity is not really available to sanction unauthorized exit. Unlike a violation, a state would expect that if it publicly exits an agreement the remaining parties would withdraw any benefits conferred specifically on that state.⁵⁵ Retaliation and reputational sanctions, though, remain available to curb unauthorized exit. In particular, unauthorized exit is a violation of a legal obligation that can result in a reduction of a state's reputation for complying with legal rules.⁵⁶ As with all sanctions, in any given instance the sanction may be insufficient to deter illegal exit. But by raising the sanctions for exiting an agreement, states can make exit less likely.

B. Provisions Affecting Exit Costs

There are a variety of legal provisions that affect the costs associated with withdrawal from a legal agreement. These include exit provisions, sunset provisions, and clauses designed to ensure that an obligation is or is not legally binding. These three types of provisions each reduce exit costs in different ways.

Withdrawal or denunciation clauses generally allow a state the unilateral option to legally withdraw from or denounce an agreement, sometimes when certain conditions are met or sometimes simply at its discretion.⁵⁷ For example, the Marrakesh Agreement establishing the World Trade Organization ("WTO") permits withdrawal unconditionally with six months notice.⁵⁸ By contrast, the NPT only permits withdrawal if a nation determines "that extraordinary events, related to the subject matter of this Treaty, have jeopardized the supreme interests of its country," and requires that the withdrawing state explain those circumstances.⁵⁹

The purpose of withdrawal clauses is to reduce or eliminate the legal consequences of treaty withdrawal.⁶⁰ A state that withdraws in accordance with

54. See ANDREW T. GUZMAN, *HOW INTERNATIONAL LAW WORKS* 33 (2008).

55. Of course, if the good produced by cooperation is a public good, reciprocity is not available as a sanction for a very different reason. To withdraw the benefits of the agreement from the now-exited state would mean ceasing cooperation entirely.

56. Where retaliation is concerned, it is important to distinguish between retaliation that occurs *because* the exit is illegal, and retaliation that would have occurred regardless of whether the exit was illegal. The latter is really a political sanction; it does not depend on the terms of the agreement being exited.

57. Helfer, *Exiting Treaties*, *supra* note 34, at 1582.

58. Marrakesh Agreement Establishing the World Trade Organization, Apr. 15, 1994, 1867 U.N.T.S. 154, 33 I.L.M. 1144.

59. NPT, *supra* note 21, art. X(1).

60. Helfer, *Exiting Treaties*, *supra* note 34, at 1590. Treaty withdrawal, of course, can have non-legal consequences, such as political or information consequences. See *id.* at 1608. My study is in this respect both narrower and broader than prior work on treaty exit. This paper is broader in the sense that it focuses on all provisions that affect the legal cost of permanently exiting an agreement, but it is narrower in the sense that it only focuses on "legal" costs, such as reputational sanctions that are based on the

an exit provision does so lawfully and thus does not, in theory, suffer a reduction in its reputation for complying with international agreements. For example, in recent years Bolivia, Ecuador, Nicaragua, and Venezuela have all threatened to withdraw from the United Nations Convention on the Settlement of Investment Disputes Between States and Nationals of Other States ("ICSID Convention").⁶¹ In 2007, Bolivia followed through, exercising its rights under Article 71 of the ICSID Convention to withdraw with six months notice.⁶² Ecuador, for its part, issued a notification under Article 25(4) that withdrew its consent to ICSID jurisdiction over disputes involving oil and minerals.⁶³ Ecuador also withdrew from nine bilateral investment treaties and used the threat of exit to begin renegotiating others.⁶⁴

Soft law agreements work in a similar fashion to withdrawal clauses: making an agreement soft reduces the cost of exit. Doctrinal distinctions between legally binding agreements are functional signals about the future likelihood of compliance.⁶⁵ From a compliance perspective, then, the chief difference between hard law and soft law agreements is the likelihood of future compliance signaled by the designation. Hard legal obligations, by virtue of being legally binding, represent the most significant pledge of a state's reputation for compliance with international law.⁶⁶ Soft legal obligations, by contrast, are those obligations that are not legally binding, but which nevertheless produce legal consequences.⁶⁷ By designating an agreement as legally binding, states signal a high likelihood of future compliance and so failing to comply, all else equal, results in a larger reputational sanction.⁶⁸ This is so because going forward other states will devalue the signal sent by signing a hard law agreement with the violating state, which can cost the violating state in terms of being able to extract concessions from its cooperative partners or induce their reliance.⁶⁹

The distinction between hard law and soft law is relevant to exit costs in the same way that it is to run of the mine violations of international law.⁷⁰

legality of an agreement or whether certain actions are deemed "legal" as judged against an agreement's rules.

61. Anne van Aaken, *International Investment Law Between Commitment and Flexibility: A Contract Theory Analysis*, 12 J. INT'L ECON. L. 507, 533 & nn.113–14 (2009).

62. See News Release, Int'l Ctr. for Settlement of Inv. Disputes, Bolivia Submits a Notice Under Article 71 of the ICSID Convention (May 16, 2007), available at <http://icsid.worldbank.org/ICSID/StaticFiles/Announcement3.html>.

63. See News Release, Int'l Ctr. for Settlement of Inv. Disputes, Ecuador's Notification under Article 25(4) of the ICSID Convention (Dec. 5, 2007), available at <http://icsid.worldbank.org/ICSID/StaticFiles/Announcement9.html>.

64. Asha Kaushal, *Revisiting History: How the Past Matters for the Present Backlash Against the Foreign Investment Regime*, 50 HARV. INT'L L. J. 491, 493 (2009).

65. Andrew T. Guzman & Timothy Meyer, *International Common Law: The Soft Law of International Tribunals*, 9 CHI. J. INT'L L. 515, 519 (2009).

66. Guzman & Meyer, *International Soft Law*, *supra* note 19, at 23.

67. Meyer, *Delegation*, *supra* note 15, at 890.

68. Guzman & Meyer, *International Soft Law*, *supra* note 19, at 23.

69. *Id.*

70. Meyer, *Delegation*, *supra* note 15, at 915.

“Exiting” a treaty that does not permit withdrawal is a violation of the implied obligation not to exit the agreement and carries with it a reputational sanction that is in part determined by the legal status of the obligation exited. For example, in 2005 the United States announced its “withdrawal” from the Optional Protocol to the Vienna Convention on Consular Relations (“VCCR”), which conferred jurisdiction on the International Court of Justice to hear disputes arising under the VCCR. Notably, however, the VCCR does not contain a provision authorizing such a withdrawal. And though the VCLT recognizes that in some instances it may be legally permissible to withdraw from a treaty that lacks an explicit withdrawal mechanism,⁷¹ as a practical matter states are, in the absence of express authorization, at best likely to have conflicting views about the legality of a specific withdrawal. It thus seems safe to infer that the United States suffered a loss of reputation for exiting the VCCR, and to infer that the loss of reputation was larger than it would have been had the VCCR been a soft law agreement.⁷²

Finally, in contrast to exit provisions and the legality of an agreement, which place the onus of action on states seeking to exit, sunset provisions (and framework agreements, for that matter) usually require collective action to extend the agreement.⁷³ The default rule established in the agreement will determine precisely what the collective action is necessary for, and what decision rule the collective action is taken under. In general, though, the costs of using a sunset provision to exit an agreement will be higher than for using an exit clause, particularly in the multilateral context. This is because sunset provisions typically authorize a majority to make a decision for all. A state wishing to use a sunset provision to exit an agreement must therefore assemble a coalition of states to defeat renewal. The classic example of such a sunset provision is Article X(2) of the NPT. The provision required the parties to the treaty to vote, twenty-five years after the treaty came into force, on whether to extend the treaty indefinitely or for further fixed periods.⁷⁴ Such a provision requires a majority vote to extend the life of the treaty.⁷⁵

71. VCLT, *supra* note 53.

72. Soft law agreements are also not “treaties” and are thus not governed by the rules in the VCLT, including the default rule that withdrawal is generally not permitted unless explicitly authorized. It is possible, therefore, that exiting soft law agreements are governed by a different default rule or no default rule at all.

73. Automatic renewals unless a state objects are another possibility.

74. NPT, *supra* note 21, art. X(2).

75. Another less obvious type of sunset provision is one in which a state commits to completing a certain action by a certain date but does not covenant to continue the action after the date has passed. For example, the Moscow Treaty of 2002 commits the United States and Russia to reduce their arsenal of strategic nuclear warheads to between 1700 and 2200 by December 31, 2012. Regardless of whether that obligation is met, though, the parties have no further obligations under the Moscow Treaty after 2012; indeed, to remove any doubt on this score article IV provides that the treaty shall only remain in force until December 31, 2012—the last day the parties have to complete their obligations under the treaty.

C. Exit versus Violation

Renegotiation challenges the cherished notion of international law as a means of regulating state behavior because renegotiation carries with it a suggestion of willful and opportunistic disobedience. One might therefore wonder why states would bargain over exit costs at all, given that states may always violate their legal commitments, and in so doing, try to prompt a renegotiation. Exit and violation are not equivalent from the standpoint of inducing renegotiation, however. To see why exit is of particular importance to renegotiation, and therefore to agreement design, it is important to understand the various motivations states have for not complying with international law.

I use the term “noncompliance” differently than it is used in most literature on the subject. Noncompliance is often employed as an umbrella term to refer to any deviation from established substantive rules of conduct, regardless of whether those deviations are justified under prevailing rules. Exit and violation are both forms of noncompliance, and the international relations literature has therefore tended to conflate the two.⁷⁶ Although exit can be a violation of an agreement—as it arguably was when the United States “withdrew” from the VCCR—exit can also be perfectly lawful. For example, a state can denounce a treaty in accordance with a provision authorizing such denunciation.⁷⁷

All noncompliance—exit and violation—withholds the benefits of cooperation from other states. All noncompliance can therefore be used to impose costs on cooperative partners in an effort to spur renegotiation. However, noncompliance or the threat thereof undertaken with renegotiation in mind is fundamentally different from the noncompliance that occupies much of the attention of international legal scholars in several respects. First, and perhaps most obviously, the two are not doctrinally equivalent.

After exit, a set of rules no longer applies to the exiting state. The sanctions for breach authorized by the exited agreement, the VCLT, or the customary international rules of state responsibility are only implicated by a violation, not by a lawful exit.⁷⁸

Second, renegotiation-driven noncompliance tends to be public, whereas states frequently try to hide violations of international law. Hiding a violation confers two benefits on a state: 1) it reduces or eliminates the sanctions associated with violation, and 2) it may induce a state’s partners to continue to cooperate. The other states’ continued cooperation means that the violating state reaps the benefits of its partners’ adherence to the law without having to pay the costs of such adherence itself. Provisions in agreements

See Treaty on Strategic Offensive Reductions, U.S.-Russ., arts. 1, 4, May 24, 2002, S. TREATY DOC. NO. 107-8.

76. Helfer, *Exiting Treaties*, *supra* note 34, at 1613.

77. *Id.*

78. *Id.* at 1616–17.

focusing on deterring violations are thus usually aimed at monitoring compliance and producing information about violation upon which sanctions can be based. These information-producing mechanisms are not needed to detect exit, as exit is made public by design.

Third, exit is superior to violation in creating leverage for renegotiation. Even when violations are public, they often fail to demonstrate a state's resolve in seeking a renegotiation. Unlike most violations, exit allows a state to credibly signal that it is serious about renegotiating the existing regime. Exit therefore provides states with a distinct tool as they seek to renegotiate their legal obligations.

Once a state completely exits a regime, it loses access to any of the regime's selective benefits, such as participation in dispute resolution processes, monitoring, rule-making, and information-sharing bodies.⁷⁹ A violating state, on the other hand, leaves itself the possibility of coming back into compliance with its obligations.⁸⁰ And while a state that has exited an agreement may try to reaccede to the agreement, rejoining may be subject to review by other members or by the institution itself. For example, in 1998 Trinidad & Tobago withdrew from both the American Convention on Human Rights and the First Optional Protocol to the International Convention on Civil and Political Rights ("First Optional Protocol"), which confers jurisdiction on the Human Rights Committee to review petitions from individuals.⁸¹ Trinidad & Tobago then tried to reaccede to the First Optional Protocol with a reservation that would have eliminated the jurisdiction of the Human Rights Committee to review petitions from capital defendants.⁸² The Human Rights Committee essentially struck the reservation down, holding it incompatible with the object and purpose of the First Optional Protocol. Consequently, Trinidad & Tobago chose to remain outside of the First Optional Protocol entirely.⁸³

79. *Id.* at 1613–15. In some instances, withdrawing does not mean a total loss of voice in rule-making procedures. Some international institutions, such as the International Criminal Court, still allow non-parties to participate as "observers." See David Scheffer, *Keynote Address: The Future U.S. Relationship with the International Criminal Court*, 17 *PAGE INT'L L. REV.* 161, 175 (2005).

80. At some point the costs of violation—be they reputational or retaliatory—will be large enough to demonstrate the violating state's resolve to hold out for renegotiation. In most cases, however, the costs associated with violation will be small enough to leave states uncertain as to the violating state's resolve.

81. See Helfer, *Overlegalizing*, *supra* note 30, at 1881.

82. *Id.*

83. *Id.* In another example, when Iceland sought to rejoin the International Whaling Commission ("IWC") it did so with a contentious reservation to the IWC's moratorium on whaling. This reservation, and thus Iceland's reaccession to the IWC, was twice rejected before ultimately being approved in a convoluted series of procedural votes. Alexander Gillespie, *Iceland's Reservation at the International Whaling Commission*, 14 *EUR. J. INT'L L.* 977, 977–78 (2003).

III. BARGAINING IN THE PRESENCE OF SHIFTING POWER

The central puzzle of agreement design in relation to renegotiation is why negotiators would agree to give future bargaining power to potentially dissatisfied states. Of course, future bargaining power is good for the potentially dissatisfied state, and in some situations that bargaining power may be good for states generally.⁸⁴ As I have argued elsewhere, where a majority of states are primarily interested in coordinating their behavior regardless of the specific rules chosen, employing soft law to reduce exit costs confers power on a minority of states for whom the specific rules chosen are more important than coordination. Soft law, in effect, acts as a delegation to that minority to drive the evolution of legal rules.⁸⁵ In such situations, the gains from a more efficient system of implementing welfare-enhancing amendments outweigh the costs of opportunistic revisionism. But oftentimes tension in the interests of the parties will be high and distributional concerns will be paramount. How do we make sense of conferring future bargaining power on potentially dissatisfied states in those situations?

Below, I present a theory of bargaining in the shadow of shifting power that answers this question. In so doing, I make standard assumptions about state behavior: states are unitary, behave rationally, and act in their own self-interest (that is, they have no inherent preference for cooperation).⁸⁶ These assumptions ignore important determinants of state behavior, such as constructivist processes of preference formation, interest group interactions, and the interplay of domestic political and legal forces. Nevertheless, they are useful because they allow us to focus our attention on the dynamics of interstate negotiation, rather than on how states determine what objectives to pursue in those negotiations.

A. *The Logic of Renegotiation*

To understand how states bargain in the shadow of shifting power, one first has to understand what happens when power has actually shifted. How does renegotiation occur?

84. See Meyer, *Delegation*, *supra* note 15, at 892.

85. *Id.* at 892–93.

86. These assumptions are widely used to study international legal behavior. See, e.g., GOLDSMITH & POSNER, *supra* note 6, at 4; GUZMAN, *supra* note 54, at 17; ROBERT E. SCOTT & PAUL B. STEPHAN, *THE LIMITS OF LEVIATHAN: CONTRACT THEORY AND THE ENFORCEMENT OF INTERNATIONAL LAW* 9 (2006); Kenneth W. Abbott, *Modern International Relations Theory: A Prospectus for International Lawyers*, 14 *YALE J. INT'L L.* 335, 348–51 (1989); Jeffrey L. Dunoff & Joel P. Trachtman, *Economic Analysis of International Law*, 24 *YALE J. INT'L L.* 1, 2 (1999); Tom Ginsburg & Richard H. McAdams, *Adjudicating in Anarchy: An Expressive Theory of International Dispute Resolution*, 45 *WM. & MARY L. REV.* 1229, 1330 (2004); John K. Setear, *Responses to Breach of a Treaty and Rationalist International Relations Theory: The Rules of Release and Remediation in the Law of Treaties and the Law of State Responsibility*, 83 *VA. L. REV.* 1, 3 (1997); Edward T. Swaine, *Rational Custom*, 52 *DUKE L.J.* 559, 560–62 (2002); Koremenos, *supra* note 4, at 293–94.

1. *Outside Options and Credible Threats*

For states to successfully engage in renegotiation, it must be the case that all states involved do better under the renegotiated terms than they do under the alternative. Critically, however, the alternative is not necessarily the existing agreement.⁸⁷ Consider two different scenarios. In the first, conditions have changed such that the current agreement no longer maximizes the collective welfare of the parties. Put differently, the agreement is no longer on the Pareto frontier. There is thus an amendment or a set of substantive rules that would make at least one party better off without making any other party worse off, relative to the existing deal. In the absence of transaction costs, we would expect renegotiation to occur here regardless of the level of exit costs.

But an amendment that makes everyone better off is not necessary for renegotiation to be rational for all parties. In the second scenario, changed conditions have only affected a single state. That state now has a privately available alternative—an outside option⁸⁸—that is more valuable than cooperation under the existing agreement. Note that the outside option is issue-specific. A state may have a good outside option on one issue but not on others. For example, a state may have a high payoff from noncooperation on climate change because it is a developing nation that benefits from industrialization but poor outside options on security or trade issues. This outside option scenario, in game theoretic terms, allows the ascendant state to make a credible threat to exit the existing agreement.⁸⁹ Either the ascendant state will exit and take its outside option, leaving other states with the value from their next-best alternative to the status quo,⁹⁰ or there will be a renegotiation that makes the ascendant state at least as well off as it would be taking its outside option. Significantly, this renegotiation can leave other states worse off than they would have been under the existing agreement.⁹¹ A declining state will be willing to make concessions to the ascendant state up to the point at which the declining state's outside option is better for it than conceding enough to keep the ascendant state in the existing agreement.

87. GRUBER, *supra* note 1, at 7.

88. David M. Andrews, *Monetary Policy Coordination and Hierarchy*, in INTERNATIONAL MONETARY POWER 91, 102 (David M. Andrews ed., 2006); *see also* Avner Shaked & John Sutton, *Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model*, 52 ECONOMETRICA 1351, 1355 (1984).

89. In the language of game theory, a "threat" is a "promise to carry out a certain action if another player deviates from his equilibrium actions." ERIC RASMUSEN, GAMES AND INFORMATION: AN INTRODUCTION TO GAME THEORY 83 (1989). A threat is "credible" only if the strategy is an equilibrium on all possible paths in a game. Ian Ayres, *Playing Games with the Law*, 42 STAN. L. REV. 1291, 1306 (1990).

90. A declining state's next best alternative could be to carry out the agreement without the ascendant state. This is not really an "outside option" because it is not outside of the existing agreement, but rather is a forced modification of it.

91. GRUBER, *supra* note 1, at 6–7.

2. *How Is the Value of a State's Outside Option Determined?*

Renegotiation is therefore driven by the value of a state's outside option. Put differently, an outside option determines to a large extent a state's bargaining power. It is worth noting that a state's outside option may not be the alternative that makes the state best-off, as measured by the overall welfare of the state's citizens. Rather, the outside option refers to the payoff received by the state's decisionmakers after taking into account the domestic political pressures that come to bear on them. In any form of government, agency problems may emerge in which a state's government's incentives differ from the interests of the population at large. For example, a domestic interest group with veto power may develop a vested interest in continuing to comply with the prevailing terms of cooperation. In such a situation, the interest group may block efforts to use a state's outside option to renegotiate, even when the outside option is in some sense better for the nation as a whole. An emergent middle class in a developing country that has joined the WTO (such as China) may block serious efforts to renegotiate, even if renegotiation might increase citizens' overall welfare.⁹² The reverse might also be true. A domestic political constituency, such as the military in a country building an illicit nuclear weapon, may be vested in using an outside option to drive renegotiation, even if the status quo is arguably better for the state's citizens as a whole.

Outside options are affected by many factors. The existence of competing legal regimes can affect a state's outside option, as when an alternative regime offers rules that are more favorable to a particular state.⁹³ For example, Raustiala and Victor count five different international agreements that set rules governing property rights in plant genetic resources.⁹⁴ Similarly, the International Convention for the Regulation of Whaling and the associated regulatory regime presided over by the International Whaling Commission ("IWC") offer a competing approach to regulating international whaling from the North Atlantic Marine Mammal Commission ("NAMMCO").⁹⁵ Indeed, the IWC and NAMMCO have competed for members, with states such as Iceland moving from the more restrictive IWC to the more permissive NAMMCO (and back again).⁹⁶

A change in how states view issue linkages can also affect the value of a state's outside option. In perhaps the best known case of exit-driven renegotiation, the United States and the European Union withdrew from the General Agreement on Tariffs and Trade ("GATT") (1947) in 1994 and acceded to the WTO, essentially forcing developing nations to follow suit or lose

92. I am grateful to Jide Nzelibe for this example.

93. Helfer, *Exiting Treaties*, *supra* note 34, at 1583; Raustiala & Victor, *supra* note 50, at 295.

94. Raustiala & Victor, *supra* note 50, at 283–84.

95. See David D. Caron, *The International Whaling Commission and the North Atlantic Marine Mammal Commission: The Institutional Risks of Coercion in Consensual Structures*, 89 AM. J. INT'L L. 154, 155 (1995).

96. Helfer, *Exiting Treaties*, *supra* note 34, at 1583.

their most-favored nation trading status with both the United States and the European Union.⁹⁷ This maneuver was conceived of in 1990 and executed to end the Uruguay Round of trade negotiations.⁹⁸ Steinberg has argued that the maneuver was made possible by the end of the Cold War.⁹⁹ While the United States had considered using the possibility of exit to generate favorable bargaining results at the end of the Tokyo Round negotiations in 1979, it ultimately decided not to because of the cost in security terms of pushing developing nations toward the Soviet Union.¹⁰⁰ When the end of the Cold War broke the linkage between security policies and trade policies, the United States and the European Union's bargaining power in the GATT increased considerably.

Relatedly, a shift in a state's preferences can change how it perceives the value of its various options. Such a shift can occur in at least two different ways. First, the preferences of actors within a state can shift. Such shifts can happen through the work of epistemic communities that imbue actors with new patterns of reasoning, leading them to pursue new interests.¹⁰¹ Similarly, transnational networks of activists can use international norms to galvanize public support for positions that constrain or shape the actions of domestic political actors.¹⁰² Second, shifts can happen through changes in government. When a different set of domestic actors with a different set of preferences gains control of the apparatus of state, the state's preferences necessarily change.¹⁰³ The change from the Clinton administration to the Bush administration in 2001 illustrates how a change in government can change a state's preferences. While the Clinton administration signed both the Kyoto Protocol and the Rome Statute of the International Criminal Court, the Bush administration declined to pursue ratification of either and went so far as to "unsign" the latter.¹⁰⁴

A state's outside option can also improve with the development of a particular technological capacity. Consider, for example, states that are interested in developing and testing a nuclear deterrent. Before they have acquired such a deterrent, states such as North Korea or Iran may be happy to remain within a legal framework that denies them the right to possess

97. Steinberg, *supra* note 12, at 359–60.

98. *Id.*

99. *Id.* at 359.

100. *Id.* at 358–59.

101. Peter M. Haas, *Introduction: Epistemic Communities and International Policy Coordination*, in *KNOWLEDGE, POWER, AND INTERNATIONAL POLICY COORDINATION* 1–3 (Peter M. Haas ed., 1992); see also Ryan Goodman & Derek Jinks, *How to Influence States: Socialization and International Human Rights Law*, 54 *DUKE L. J.* 621, 626 (2004).

102. Margaret E. Keck & Kathryn Sikkink, *ACTIVISTS BEYOND BORDERS: ADVOCACY NETWORKS IN INTERNATIONAL POLITICS* 16–25 (1998).

103. See Rachel Brewster, *Unpacking the State's Reputation*, 50 *HARV. INT'L L. J.* 231, 257 (2009) (discussing the reputational consequences of changes in government).

104. See Edward T. Swaine, *Unsigning*, 55 *STAN. L. REV.* 2061, 2061–62, 2064 (2003).

nuclear weapons.¹⁰⁵ Once they develop a nuclear weapon, however, exiting the legal regime becomes a more attractive option and creates leverage to renegotiate the terms of their status in the broader nuclear nonproliferation regime. This is exactly the tactic North Korea employed in withdrawing from the NPT. Following its 2003 withdrawal, North Korea engaged in a prolonged series of negotiations in an attempt to obtain concessions on security and economic aid from the United States, China, Japan, South Korea, and Russia in exchange for nuclear disarmament.¹⁰⁶

Finally, because having valuable alternatives increases bargaining power, states may actively invest in developing outside options. States, in other words, have an incentive to undermine cooperation on existing terms by developing alternatives to be used as leverage.¹⁰⁷ The United States pursued exactly this strategy when it established the Asia-Pacific Partnership on Clean Development and Climate (“AP-6”). The AP-6 is a nonbinding organization, consisting of the United States, Japan, South Korea, India, China, and Australia, that is somewhat of an alternative to the United Nations Framework Convention on Climate Change (“UNFCCC”) and the Kyoto Protocol thereto.¹⁰⁸ Similarly, in the intellectual property area, developed and developing nations have sought out or established competing fora for international lawmaking. The United States, for example, initially pursued the TRIPS Agreement through GATT because of a stalemate on patent protection standards at the World Intellectual Property Organization.¹⁰⁹ The result of developing valuable alternatives is that an agreement that is stable today—one in which no state can force a renegotiation by credibly threatening to exit—may be rendered unstable tomorrow. The likelihood of future renegotiation depends on the likelihood that a state’s option will become more valuable to it than cooperation under existing rules—a probability that states can increase through their own efforts.

3. *How Do Exit Costs Affect a State’s Outside Option?*

States are not without the tools to combat shifting power, but they are also not able to completely contract out of the threat of renegotiation. Bargaining power is, as I have already suggested, affected by the formal terms of an agreement. It is, in other words, affected by law. States can at the time of drafting an agreement raise the costs of exit. They may choose not to include a denunciation clause or to limit the circumstances under which parties may denounce an agreement. They may lengthen the duration of an agreement or

105. Of course, in attempting to develop a nuclear weapons capability while in the nonproliferation regime, states such as Iran and North Korea will typically be violating their nonproliferation obligations.

106. See generally *U.S. and Other Powers Reach Tentative Understanding on North Korea’s Nuclear Program*, 99 AM. J. INT’L. L. 889, 914–16 (2005).

107. Rausiala & Victor, *supra* note 50, at 295–97; Helfer, *Exiting Treaties*, *supra* note 34, at 1583.

108. See *US agrees climate deal with Asia*, BBC NEWS, July 28, 2005, available at <http://news.bbc.co.uk/1/hi/sci/tech/4723305.stm>.

109. Helfer, *Regime Shifting*, *supra* note 27, at 20.

make the agreement of indefinite duration. By drafting exit provisions so as to make exit more costly, states make it less likely that a shift in the value of a state's outside option will allow a state to force a renegotiation. Thus, even if an ascendant state's outside option is more valuable to it than cooperation under current rules, a state may not have a credible threat to exit that will permit it to demand a renegotiation. Instead, an ascendant state's outside option *less exit costs* must exceed the ascendant state's existing share of the gains from cooperation. The effect of exit costs, then, is to shift bargaining power in favor of declining states. Exit costs and the stability they bring to the international legal system therefore have distributional consequences.

That states can reduce the likelihood of renegotiation by curbing exit invites the question why they do not entirely eliminate the possibility. In economic contracting literature, for example, scholars have spent a great deal of effort studying mechanisms that eliminate such future bargaining power and make a contract renegotiation-proof.¹¹⁰ I explain below why states often will not avail themselves of the highest level of exit costs available. However, there at least three independent reasons why states, even if they wanted to, would not be able to entirely eliminate the possibility of renegotiation.

First, as is familiar to scholars of international law, the absence of a third-party enforcer of legal rules capable of awarding damages will generally mean that states cannot create efficient incentives for performance.¹¹¹ Instead, states must balance the benefits of reducing violations and renegotiations with the costs that come with more stringent penalties for exit.¹¹² Second, even where domestic contracts are concerned, incentives to breach are often associated with renegotiation.¹¹³ Many contract remedies are not perfectly efficient—i.e., not perfectly compensatory—and thus renegotiation is often necessary to reach the efficient allocation of resources.¹¹⁴ Third, one might think that states should engage in contingent or complete contracting as a way to eliminate renegotiation. That is to say, states should be able to draft an agreement that adjusts the substantive terms based on shifts in the values of states' outside options. A complete contract that envisioned all possible shifts in power would not require any renegotiation as power shifted. Of course, complete contracting is rarely, if ever, observed in any field of contracting, let alone in international agreements.¹¹⁵ This is in part because contracting is costly, and in most situations the number of possible

110. See, e.g., Oliver D. Hart & Jean Tirole, *Contract Renegotiation and Coasian Dynamics*, 55 REV. ECON. STUD. 509 (1988).

111. See Guzman & Meyer, *International Soft Law*, *supra* note 19, at 20–21.

112. *Id.* at 22–24.

113. See, e.g., Richard Craswell, *Contract Remedies, Renegotiation, and the Theory of Efficient Breach*, 61 S. CAL. L. REV. 629, 633–42 (1988).

114. *Id.*

115. See, e.g., Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, 113 YALE L. J. 541, 594–95 (2003) (“These considerations imply that contracts will inevitably be incomplete.”); Richard Craswell, *The ‘Incomplete Contracts’ Literature and Efficient Precautions*, 56 CASE W. RES. L. REV. 151, 156 n.12 (2005) (noting that “‘complete’ contracts are rarely if ever observed”).

future scenarios is so vast as to prohibit complete contracts.¹¹⁶ Moreover, deferring the cost of negotiating over any given contingency to a future (re)negotiation will often make sense because at the time of initial contracting, such future costs are discounted both by time and by the probability that the particular contingency in question will come to pass.¹¹⁷

The possibility of a future shift in a state's outside option that leads to a credible threat of exit, then, must be a paramount consideration in drafting provisions affecting exit costs. By lowering the cost of exit, an agreement's negotiators confer some increased measure of future bargaining power on states likely to be dissatisfied down the road with existing legal rules. To be sure, explicit decision-making rules also affect the possibility and outcomes of future renegotiations. Examples of decision-making rules include one-state, one-vote majoritarian voting; one-state, one-vote supermajoritarian voting; weighted voting (with either a simple majority or a super-majority necessary to pass an amendment); or unanimity, which, predicated on the doctrine of sovereign equality, is by far the most common type of amendment procedure.¹¹⁸ And bargaining power will also depend on a number of political and economic factors not directly accounted for in the formal provisions of an agreement. But the costs of exit that derive from the agreement are an important source of bargaining power. Reducing these costs makes it more likely that states will be able to demand that the gains from cooperation be redistributed in their favor, regardless of the specific decision rule under which they are operating. Low exit costs confer potential future bargaining power on states whose cooperation in an agreement is most valuable to other agreement members.

B. Bargaining Over Exit Costs in the Shadow of Shifting Power

Renegotiation is not a new problem. What is new to international legal literature is thinking about how states bargain in light of the possibility of future renegotiation occasioned by shifts in bargaining power. Bargaining power is simply not static. If a state expects its bargaining power to increase in the future, it will be reluctant to agree to terms today unless it is guaranteed that those terms will not prejudice its expected future position. At the same time, states that are powerful today may wish to lock in their transitory power through legal mechanisms, such as high exit costs, that reduce the threat of future renegotiations that might occur when they are in a weaker position. In other words, declining states may wish to use formal

116. See Schwartz & Scott, *supra* note 115, at 595 ("There is an infinite number of possible future states and a very large set of possible partner types. When the sum of possible states and partner types is infinite and contracting is costly, contracts must contain gaps. Parties cannot write contracts about everything.").

117. There is also an agency problem. The initial negotiators may not be in power during the renegotiation, and so by deferring negotiations over contingencies into the future, they may entirely avoid the transaction costs of trying to write a complete contract.

118. See Steinberg, *supra* note 12, at 339–40.

mechanisms to convert their present bargaining power into long-term substantive gains, while ascendant states will be hesitant to commit themselves to agreements that lock in substantive terms based on current bargaining power. How will states resolve this conflict?

Consider the following hypothetical. Two states negotiate over an agreement. They have to divide the gains from cooperation. The choice of substantive rules determines how those gains are allocated once the agreement comes into force.¹¹⁹ In negotiating over the substance of an agreement, no state will be willing to accept any rules that leave it worse off than its outside option. Thus, each state's outside option restricts the range of possible agreements. Moreover, the outside option of one state, the ascendant state, is increasing over time as the state grows in power, while the outside option of the declining state remains constant or loses value.

Now imagine that exiting the agreement is costless.¹²⁰ Each time a state's outside option increased so that it exceeded the value to that state of cooperating under the existing agreement, it would have a credible threat to exit the agreement. Faced with such a credible threat, the declining state would concede just enough to make cooperation worth the ascendant state's while.¹²¹ If at any point the sum of the two states' outside options ever exceeded one hundred percent of the benefits from cooperation, the agreement would collapse because neither side would be willing to make the necessary concessions to keep cooperation viable. Under such an agreement, the transaction costs associated with renegotiation would be the only brake on the renegotiation of the substantive rules. If transaction costs were minimal, legal rules would evolve to track the underlying distribution of bargaining power, just as the realist theory discussed above predicts.

It follows that in a world without the ability to commit through exit costs, expected shifts in power are largely irrelevant in negotiating an agreement. Because an ascendant state cannot credibly commit not to exploit its bargaining power tomorrow, the declining state will not make any substantive concessions in order to prevail on the formal aspects of the agreement. An agreement would be based purely on the current configuration of power and would last only so long as that configuration lasted. This situation might disadvantage declining states, which would prefer to negotiate an agreement that locks in their gains and does not leave them susceptible to

119. I assume throughout most of this discussion that all possible agreements are on the Pareto frontier. This assumption allows me to highlight distributional issues.

120. For purposes of this thought experiment, I assume that sanctioning opportunistic noncompliance is a separate and distinct mechanism from sanctioning exit. Thus, one can assume that levels of compliance are the same regardless of the chosen level of exit costs. In reality, of course, factors such as the legality of an agreement affect the sanctions applied both for permanently and publicly exiting an agreement and for opportunistically violating an agreement. However, other mechanisms that influence exit costs, such as withdrawal provisions and sunset provisions, do not affect sanctions for violating an agreement.

121. See Powell, *supra* note 3, at 93-97.

future forced renegotiations. Moreover, the inability to commit can reduce the overall gains from cooperation by removing the potential for agreements that add value over time through investment in cooperation.¹²²

As discussed above, adding exit costs changes this dynamic by allowing states to commit to not take advantage of any changes in their relative fortunes. If an ascendant state must pay a cost to exit an agreement, it can no longer freely obtain its outside option. This reallocation of future bargaining power means that agreeing to high exit costs is a substantial concession for a state that expects to become more powerful over time. A state giving up that power at the time of contracting will expect to be compensated for it. More specifically, if by agreeing to high exit costs a state is giving away its ability to renegotiate better substantive terms in the future, the state will demand better substantive terms now as compensation. Declining states are thus put in the position of having to choose to yield on the level of exit costs, and thereby leave themselves vulnerable to renegotiation if and when the ascendant state becomes more powerful, or give some larger share of the gains from cooperation to the ascendant state now. Nor does it matter whether a state's expectations ultimately turn out to be correct. At the time of negotiation, each state has expectations about its own future bargaining power and that of its negotiating partners. States bargain on the basis of those expectations, regardless of whether those expectations ultimately turn out to have been wise judgments.¹²³

A brief numerical example illustrates the point (a somewhat more formal argument appears in footnote 126). An ascendant state, *A*, has an outside option worth 40 during the period in which it is negotiating an agreement with a declining state, *D*, to divide the surplus from cooperation, which is equal to 100 per period. Suppose, for simplicity, that *A* and *D* know for certain that the value of *A*'s outside option in the second period will be 65, and *A*'s discount rate is .9. A fifty-fifty division of the surplus is therefore vulnerable to renegotiation in the second period unless the agreement has exit costs of at least 15 ($65 - 50 = 15$).¹²⁴ If exit costs are at least 15, *A* can do no better than 50 in the second period by exiting and therefore does not

122. See Douglass C. North & Barry R. Weingast, *Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England*, 49 J. ECON. HIST. 803 (1989) ("For economic growth to occur the sovereign or government must not merely establish the relevant set of rights, but must also make a credible commitment to them.")

123. States' expectations about their own options are, perhaps intuitively, likely to prove more accurate than their forecasts about the fortunes of others. In developing expectations about others, states will have to rely on whatever information they can obtain publicly or through other means. In part due to concerns about the information on which they base their beliefs about the future, states may discount their expectations according to how reliable they think their forecasts are. In particular, states may discount expectations that reach further into the future on the grounds that such expectations are less likely to be accurate than short-term expectations.

124. I choose a fifty-fifty division in this example only because it corresponds to the Nash bargaining solution.

have a credible threat to exit.¹²⁵ Exit costs of 15 or more, then, prevent renegotiation. But will *A* agree to exit costs of 15 and a fifty-fifty division?

A simple comparison of utilities indicates that the answer is no. If *A* agrees to the proposal, then its utility is:

$$U_A = 50 + 50(.9) = 95.$$

If *A* rejects the proposal, preferring its outside option from the beginning, then:

$$U_A = 40 + 65(.9) = 98.5.$$

Clearly, in such a situation *D* will have to compromise either on its preferred level of exit costs or on substance in order to get an agreement. Specifically, *D* will have to ensure that *A*'s present value from the deal is equal to 98.5, so that *A* receives at least as much by agreeing as it does by exercising its outside option from the beginning. This can be done either by increasing *A*'s share of the surplus (here to $98.5/1.9$, or roughly 51.8), or reducing the level of exit costs to allow renegotiation. In the scenario in which exit costs are low enough to allow renegotiation, assuming the agreement calls for a fifty-fifty division in the first round, exit costs can be no higher than roughly 11.1 ($98.5 = 50 + (.9)(65 - c)$, where c is equal to exit costs). Note that exit costs still matter even when renegotiation occurs—they reduce the value of *A*'s outside option and therefore reduce what *A* gets in a renegotiation.

D's preferred concession is that which maximizes its utility. *D*'s utility from a deal that divides the benefits from cooperation fifty-fifty but has exit costs that permit renegotiation in the second round is:

$$U_D = 50 + \delta_D(100 - (65 - c)),$$

where c cannot be higher than 11.1 and therefore *D*'s utility in the second period cannot be higher than 46.1 (δ_D is *D*'s discount rate). *D*'s utility from conceding on substance but retaining exit costs high enough to prevent renegotiation is roughly:

$$U_D = 48.2 + \delta_D(48.2).$$

Clearly, *D* prefers conceding on the substantive terms of an agreement for a wide range of values of δ_D . For sufficiently low values of δ_D (or for different distribution of the benefits under the initial agreement), however, *D* will prefer to concede on exit costs and capitalize on its ability to lock in a larger benefit in the first round.¹²⁶

125. This analysis assumes that *D* has all of the structural bargaining power in a renegotiation, such that *A* receives only its outside option. In most bargaining protocols, however, *A* would do better than its outside option, thereby even further improving the position of the ascendant state.

126. This example can be expressed more formally as a game with two phases in which states bargain over and implement an agreement characterized by a distribution of benefits, x , and a level of exit costs, c . Because a formal model does not improve significantly on the insights discussed above, I only sketch what such a model would look like. In each phase, states have the opportunity to negotiate or renegotiate

This numerical example highlights a significant point about the stability of international law, one that is perhaps obvious but sometimes overlooked. In the example, exit costs have distributional ramifications. Even if they do not prevent renegotiation, exit costs increase *D*'s payoff in round two while reducing *A*'s payoff by a similar amount. Exit costs accomplish this redistribution from the ascendant to the declining state by preventing renegotiation or, where renegotiation does occur, reducing what an ascendant state can get by exiting an agreement. Put differently, increasing the stability of international law through higher exit costs has distributional ramifications. More generally, rigid formal mechanisms—whatever benefits they may have in terms of expanding the possibilities and benefits of cooperation, as when they permit more accurate reliance—*have distributional consequences that undermine the ability of states to create stable international legal rules.*

The broader insight—that expected future shifts in power can affect bargaining over the tradeoff between form and substance—is crucial to understanding why international agreements often lack terms that seem to make

an agreement, which is then immediately implemented. If bargaining concludes without an agreement in place, each state receives in each phase the value of its outside option (a_i, d_i). At the beginning of Phase Two, *A*'s outside option becomes higher by some amount δa , such that $a_2 = a_1 + \delta a$. *A*'s reservation price therefore is:

$$U_A = a_1 + \delta_A(a_1 + \Delta a). \tag{1}$$

A's utility from entering into an agreement is:

$$U_A = x + \delta_A(\max\{x, a_1 + \Delta a - c\}). \tag{2}$$

In other words, *A* gets x in Phase One if it agrees and then gets the larger of the two bracketed expressions in Phase Two. The second expression in brackets is the value of *A*'s outside option less the cost of exiting the agreement. For simplicity, I assume that in a renegotiation *A* gets only the value of its outside option less the cost of exiting. This amounts to assuming that *D* has all the structural bargaining power in a renegotiation. In order for *A* to participate in the agreement, equation (2) must be larger than equation (1). Simplifying that inequality yields *A*'s participation constraint, or the formula that must be satisfied if *A* is to enter the agreement:

$$(x - a_1) + \delta_A(\max\{x, a_1 + \Delta a - c\} - (a_1 + \Delta a)) \geq 0 \tag{3}$$

The left-hand side of equation (3) is *A*'s net utility from participating in the agreement. Inspecting equation (3) indicates that an increase in x , not surprisingly, increases *A*'s utility. An increase in c , on the other hand, reduces *A*'s welfare from the agreement. It does this in two ways. Most obviously, in the event that renegotiation occurs (that is, when $x < a_1 + \delta a - c$) *A*'s payoff in Phase Two is reduced from its outside option by the level of exit costs. Second, if c prevents renegotiation (if $x < a_1 + \delta a$, but $x \geq a_1 + \delta a - c$) c reduces *A*'s payoff in Phase Two by the difference between x and some larger renegotiated amount (e.g., $a_1 + \delta a$).

Assuming renegotiation is always successful, *D*'s per-period payoff from an agreement is 1 minus *A*'s per period payoff. *D*'s reservation price is defined analogously to *A*'s in equation (1). *D*'s participation constraint is therefore:

$$(1 - x - d_1) + \delta_D(\min\{(1 - x), (1 - a_1 - \delta a + c)\} - d_2) \geq 0 \tag{4}$$

Again, the left-hand side of the equation is *D*'s net utility from the agreement. Comparing the sign on c in equations (3) and (4) indicates that c is distributional. That is, exit costs reallocate the benefits of cooperation from ascendant states to declining states. Comparing the inequalities above that control, Phase Two payoffs also indicate that higher exit costs correspond to higher levels of stability in legal rules. This leads to the conclusion that stability in legal rules has distributional implications—stability favors declining states.

sense when judged by present conditions. Potentially ascendant states will either demand substantive rules that are not justified by their present alternatives to cooperation, or they will demand that they be able to easily exit the agreement in the future. How, then, will states decide whether it is better to concede on form or substance? In the remainder of this section, I discuss some of the variables that affect how states trade off form and substance in the presence of shifting power. I continue this discussion in Part IV, where I consider broader implications of the theory for debates in the international legal literature.

C. The Tradeoff

1. Discount Rates

In order to understand how states will make the tradeoff, it is useful to keep in mind that the substance of an agreement and the agreement's exit costs operate at different points in time. The substance of an agreement comes into play as soon as states implement their cooperative policies. Exit costs do not factor in unless and until a state's outside option improves, something that is unlikely to happen until some time has passed. The result is that substantive provisions have immediate value to the parties, but provisions governing exit costs only affect a state's share of the surplus in the future. Thus, if a declining state yields on the substantive provisions of an agreement, it is giving up a present benefit—a larger share of the gains from cooperation immediately upon the agreement taking effect—in exchange for a long-term benefit—a reduction in the likelihood of renegotiation. This tradeoff will not be worthwhile if a declining state heavily discounts the future. A state with a short time horizon would prefer to lock in what benefits it can today at the expense of being able to cooperate on better terms in the future.

In other words, where states have different discount rates, states that heavily discount the future will prefer to concede on form in order to capture benefits on substance, while states that are patient will prefer to prevail on form. This suggests an empirical prediction: perhaps counterintuitively, the presence of different discount rates suggests that we should see higher exit costs—or more constraining agreements—where rising powers discount the future at a higher rate than their declining partners, but lower exit costs where rising powers are relatively more patient.

To illustrate how discount rates can affect the tradeoff between exit costs and substance, recall the numerical example given above and imagine that *D* does not discount the future at all (that is, $\delta_D = 1$; recall that *A*'s discount rate was .9 above). In such a situation, *D* prefers to concede on substance and obtain exit costs that prevent renegotiation. To see this, recall that *D* must give *A* roughly 51.8 out of 100 in order to obtain exit costs sufficient to prevent renegotiation. This leaves *D* with 48.2 in each period:

$$U_D = 48.2 + 48.2(1) = 96.4.$$

If instead *D* elects to concede on exit costs, effectively allowing a renegotiation in the second period, *D* does worse even if we assume, as above, that *D* is able to obtain the highest exit costs to which *A* would possibly agree while still agreeing to a fifty-fifty division:

$$U_D = 50 + 46.1(1) = 96.1.$$

D's utility in such a situation is still lower than its utility in prevailing on exit costs. *D* thus prefers to prevail on exit costs and concede on substance when it is the more patient party. However, as *D* becomes less patient than *A*, the opposite result occurs. Using a discount rate of .8 for *D*, *D*'s utility to conceding on substance to obtain a renegotiation-proof agreement is:

$$U_D = 48.2 + 48.2(.8) = 86.76.$$

D's utility to conceding on exit costs is:

$$U_D = 50 + 46.1(.8) = 86.88.$$

Therefore, as *D* becomes less patient than *A*, *D* prefers to concede on exit costs and allow *A* to renegotiate.¹²⁷ This insight has potentially dramatic implications. One might think that in many situations declining states are likely to discount the future more heavily than their ascendant peers precisely because—or for the same reasons that—their situation is worsening. Similarly, being an ascendant state may often be correlated with placing a high value on the future—the same factors that cause a state to invest in developing its outside options for the future may make the state more patient overall.

This suggests that in many cases we should expect low exit costs to be the norm. Those states most likely to favor high exit costs (declining states) will not push for high exit costs because they do not value the future as much as ascendant states. Put differently, where relative changes in power are correlated with relative valuations of the future, we should expect to see low exit costs. In such situations, high exit costs are not terribly important to *either ascendant or declining states*.

2. *Distributional Tension and the Benefits of Reliance*

The distributional tension on a given issue will also play a significant role in how ascendant and declining states bargain over form and substance. Sophisticated bargaining models in political science generally assume that a

127. Again, one can see this intuition worked out somewhat more formally by inspecting equations (3) and (4). See *supra* note 126. In order for an agreement to be possible, both equations (3) and (4) must be satisfied; that is, both states must have an incentive to participate in the agreement. Recall that in equation (3), increasing exit costs reduces *A*'s utility, while in equation (4) *c* is a benefit to *D*. Because *c* only matters in the second period, it is always modified by a state's discount rate. Thus, as *D* becomes more impatient and *A* becomes more patient, higher exit costs become less valuable for both parties.

state's offers can be continuous.¹²⁸ Under such an assumption, there are an infinite number of distributions (substantive rules) that can be supported by an agreement. In the real world, though, offers are not continuous. Instead, there are usually a number of discrete alternatives that the parties are considering. How many alternatives there are depends on a number of factors, including the subject matter and political sensitivities. Bilateral arms control agreements that limit the number and use of certain categories of weapons, such as the Treaty on the Limitation of Anti-Ballistic Missile Systems¹²⁹ or Treaty on the Reduction and Limitation of Strategic Offensive Arms¹³⁰ approximate the assumption of continuity much more closely than multilateral arms control agreements, such as the NPT, that attempt to outright ban a certain category of weapons. The purpose of the latter tends to be to ban certain types of conduct for either all states or groups of states, and is therefore not as susceptible to small changes in legal rights.

Where the distributional tension on an issue is high—where the available options allocate the benefits from cooperation in dramatically different ways—states may be more reluctant to yield on substance in exchange for form. In effect, yielding on substance is too costly, so states will prefer to settle for their less preferred level of exit costs, rather than compromise their preferred substantive rules. This will be particularly true when a marginal increase in exit costs does not result in a large decrease in the probability of renegotiation. At the opposite extreme, where provisions imposing exit costs are effective at deterring renegotiation, states will be more likely to insist on their preferred form. In such situations, yielding on matters of form is too costly unless justified by significant concessions on substance.

A second issue is that in many situations high exit costs may actually increase the overall benefits from cooperation. I have so far assumed that high exit costs do not affect the overall size of the benefits from cooperation. However, stability of legal rules is often thought to improve the ability of all parties to more accurately rely on their partners' behavior.¹³¹ More accurate reliance may permit states to allocate resources in a way that is more efficient over the long term, thus increasing overall welfare as against the situation in which exit costs are lower and states cannot efficiently invest in reliance. As discussed above (and in footnote), raising exit costs alone may not be Pareto-improving because increased exit costs reduce the welfare of ascendant states. Whether increased exit costs are Pareto-improving would

128. See, e.g., Powell, *supra* note 3, at 86–88.

129. Treaty on the Limitation of Anti-Ballistic Missile Systems, U.S.-U.S.S.R., May 26, 1972, 23 U.S.T. 3435, T.I.A.S. No. 7503.

130. Treaty on the Reduction and Limitation of Strategic Offensive Arms, July 31, 1991, 16 U.N. Disarmament Y.B., App. II at 450 (1991).

131. See Helfer, *Exiting Treaties*, *supra* note 34, at 1600; Swaine, *supra* note 104, at 2074 (“Where parties are free to exit a relationship at any point and for any reason, they will under-invest in reliance—that is, fail to depend upon the relationship’s perpetuation in ways that might be efficient.”).

instead turn on whether an ascendant state's share of the benefits from cooperation was larger than its costs of being constrained. If transaction costs are low, however, transfers should exist that allow states to bargain around the costs to an ascendant state of making exit difficult. There is thus reason to think that all else equal, where stability increases overall welfare, states will opt for high exit costs. Of course, transaction costs are often not low. In such situations, bargaining around the costs imposed on ascendant states by high exit costs may not be possible. Where ascendant states cannot be compensated for their loss from high exit costs, states may have to settle for an agreement with low exit costs that is not on the Pareto frontier.

3. *Asymmetric Information and Expectations*

Asymmetric information and expectations about future changes in bargaining power can be another reason that states fail to reach agreements with high exit costs in the face of shifting power.¹³² As in litigation in which both parties are optimistic about their chances for success, when both states are optimistic about their future prospects the space for an agreement shrinks.¹³³ Consider a case where State A believes that it is going to become more powerful over time, but State B believes the current configuration of power is stable. State A will be unwilling to agree to high exit costs and substantive terms that reflect current bargaining power because it views such an agreement as giving away its future leverage. For its part, State B will be unwilling to offer significantly better substantive terms than are justified by State A's current bargaining power.

One possibility in such a situation is that the states fail to reach an agreement or that an agreement is delayed. The benefit of delay is that over time information may be revealed or expectations aligned. By waiting, a potentially ascendant state will have its outside option revealed and, if power has become more stable, states can then bargain from the same informational position. But delay is also costly, as it means putting off cooperation until such time as power configurations (or beliefs about power configurations) permit an agreement.

Another possibility is that the states will reach an agreement that contains low exit costs and substantive terms that reflect today's bargaining power. Such an agreement is value-creating from each state's perspective. State A's future stream of benefits from its ability to renegotiate has not been significantly reduced by the current agreement. It therefore does not

132. Cf. John P. Gould, *The Economics of Legal Conflicts*, 2 J. LEGAL STUD. 279 (1973) (developing the theory that trials result because of mutual optimism by the parties); Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399 (1974) (same).

133. *Id.* Interestingly, having two optimistic parties may cause the space for an agreement to shrink on the substantive terms of an agreement but may also remove exit costs as an issue in the negotiation. If both parties believe they are becoming more powerful, both sides will prefer low exit costs. Similarly, if both sides believe they are declining, both sides will prefer high exit costs, again removing exit costs as an issue.

view its concession on the substantive rules as particularly costly. State B has conceded on the issue of exit costs, but does not view this concession as costly because it does not think State A is likely to get stronger and develop a credible threat of exit.

Finally, it is worth mentioning that power may be shifting so rapidly as to preclude agreement altogether (or at least an agreement that has nonzero exit costs). For example, a declining state may not be able to offer a patient ascendant state that expects to see the value of its outside option double in the next period enough to make an agreement worth its while. The value of the ascendant state's outside option today plus the present value of its outside option tomorrow may exceed the gains from cooperation that are left after the declining state receives the minimum necessary for it to cooperate.

This section has argued that in the presence of shifting power states will often be forced to concede on either exit costs or the substantive provisions of an agreement. How states will make that tradeoff is a complicated question. I have argued that variables such as the difference in states' discount rates, exogenously determined distributional conflicts, and uncertainty and asymmetric information determine whether an agreement is reached quickly and what shape it takes.

IV. IMPLICATIONS

The theory of bargaining presented in this Article has major implications for the ways in which international agreements are negotiated. In this Part I discuss how my theory helps explain a number of puzzles in international law.

A. *Rogue States and Legalization*

One of the central puzzles of international law is why so-called "rogue" states are willing to sign on to legally binding commitments that they have little intention of honoring.¹³⁴ Both statistical and anecdotal studies have documented this oddity. For example, countries with poor human rights records ratify human rights treaties at nearly the same rate as those states with good human rights records,¹³⁵ and rogue states sign binding arms control agreements that they almost immediately violate.¹³⁶ The fact that rogue states with little intention of honoring their commitments at times join international agreements at the same rate as well-behaved states implicates

134. See Oona A. Hathaway, *The Cost of Commitment*, 55 STAN. L. REV. 1821, 1844–45 (2003) (discussing the small difference between the rates at which states with good human rights practices and states with bad human rights practices ratify human rights treaties) [hereinafter Hathaway, *Commitment*]; Oona A. Hathaway, *Do Human Rights Treaties Make a Difference?*, 111 YALE L. J. 1935, 1940 (2002).

135. Hathaway, *Commitment*, *supra* note 134, at 1844–45.

136. See, e.g., Arsalan M. Suleman, *Bargaining in the Shadow of Violence: The NPT, IAEA, and Nuclear Non-Proliferation Negotiations*, 26 BERKELEY J. INT'L L. 206, 231 (2008).

one of the central debates in international legal circles: whether international law does indeed affect the behavior of states, and by extension, the value of international law to regulating world order.

Proponents of the view that international law does help order the behavior of states have advanced various theories to explain why states with little intention of honoring their international commitments would nevertheless sign up to an agreement that is binding and frequently very difficult to exit. Where human rights treaties are involved, for example, Hathaway has argued that states will join international agreements that are binding because the expected domestic costs of compliance and enforcement—the “costs of commitment”—are low, while the collateral benefits of joining, such as insulation from criticism or eligibility for foreign assistance, are high.¹³⁷ A reputational theory of international law suggests a different reason for these seeming anomalies. A state that already has a poor reputation for compliance with international law loses little by joining a legally binding agreement with which it does not intend to comply.¹³⁸ Such a state’s reputation for compliance cannot get significantly worse, and so the marginally higher cost associated with violating or illegally exiting the agreement is unlikely to deter a state from signing a legally binding commitment.¹³⁹

Both of these theories leave significant room for further explanation. The “cost of commitment” theory, for its part, is explicit in identifying only one set of costs associated with joining an international agreement—the domestic costs of enforcement.¹⁴⁰ This theory thus does not delve deeply into other associated costs, such as reputational costs for noncompliance. And while the reputational theory does address some of these costs (or the lack thereof) from the standpoint of the rogue state, it does not explain why states partnering with the rogue state prefer a legally binding agreement. If a state with a good reputation for compliance and a state with a bad reputation for compliance enter into an agreement in which both understand that the rogue state’s reputation for compliance is already so poor that reputational sanctions are unlikely to deter illegal exit, one might expect the result to be a nonbinding agreement. The logic behind this prediction is that in such a situation the rogue state’s exit may be undeterrable. Raising exit costs is unlikely to affect the rogue state’s behavior meaningfully. High exit costs are therefore a net cost to the state with a good reputation; they constrain the well-behaved state while leaving the rogue state relatively unconstrained in seeking a renegotiation.

The theory presented in this Article helps further unlock the puzzle of why rogue states sign binding agreements with which they have little inten-

137. Oona A. Hathaway, *Between Power and Principle: An Integrated Theory of International Law*, 72 U. CHI. L. REV. 469, 514 (2005).

138. GUZMAN, *supra* note 54, at 86.

139. *Id.*

140. Hathaway, *Commitment*, *supra* note 134, at 1824.

tion of complying. The key insight is that states are negotiating two issues—exit costs and substance—and differing discount rates mean that they value prevailing on the two issues differently. Exit costs are essentially a forward-looking provision, valuable to a patient state, while substance is valuable to both states. In short, differential discount rates among ascendant and declining states can explain high exit costs in situations in which violations may seem largely undeterrable from an *ex ante* perspective—a situation left unexplained by current theories.

To illustrate, suppose that an ascendant rogue state retains some minimal reputational capital, such that signing a binding agreement imposes a cost on the rogue state should it choose to illegally exit the agreement. An ascendant rogue state, because it substantially discounts the future, is willing to hobble its ability to renegotiate tomorrow by agreeing to high exit costs today. Giving away the ability to renegotiate might very well lock in its ability to obtain a better substantive deal today from the well-behaved nation, something to which it attaches considerably more importance. On the other hand, the concession on high exit costs still has value to the well-behaved state, which has a longer time horizon, even if it does not to the rogue state. The reason is that the constraint on the rogue state is real. In five years, the rogue state will have to take into account the reputational sanctions associated with the higher level of exit costs when deciding whether to threaten to illegally exit the agreement.

Consider in this vein negotiations between North Korea and the United States over nuclear disarmament. In such negotiations, one might think of North Korea as the ascendant state. As its nuclear capability becomes more developed, its bargaining power vis-à-vis the United States and its own neighbors increases. North Korea, though, arguably has a short time horizon created by its precarious economic situation.¹⁴¹ North Korea has entered into agreements with the United States or its surrogates several times in an effort to exchange its nuclear capability for various forms of basic economic aid. For example, following the 1994 Agreed Framework to Negotiate Resolution of the Nuclear Issue on the Korean Peninsula with the United States,¹⁴² a non-binding document, North Korea signed the so-called “Supply Agreement”¹⁴³ with the Korean Peninsula Energy Development Organization (“KEDO”), an international organization consisting of the United

141. This example is not meant to suggest that rogue states will necessarily or even usually be ascendant. There is good reason to think that states that highly discount the future are more likely to be declining states. Because they are not concerned with their future, states that discount the future are unlikely to invest in developing outside options that will allow them to renegotiate at a later time.

142. Agreed Framework to Negotiate Resolution of the Nuclear Issue on the Korean Peninsula, U.S.-N. Korea, Oct. 21, 1994, 34 I.L.M. 603 [hereinafter *Agreed Framework*].

143. Agreement on Supply of a Light-Water Reactor Project to the Democratic People's Republic of Korea Between the Korean Peninsula Energy Development Organization and the Government of the Democratic People's Republic of Korea, KEDO-N. Korea, Dec. 15, 1995, available at <http://www.kedo.org/pdfs/SupplyAgreement.pdf> [hereinafter *Supply Agreement*].

States, South Korea, and Japan.¹⁴⁴ In 2002, all three states participating in KEDO denounced North Korea for violating its arms control agreements, and KEDO suspended heavy oil shipments to North Korea.¹⁴⁵

A reputational theory of international law suggests that North Korea is willing to enter into binding agreements despite an intention not to fully comply because it has a poor reputation which it does not value.¹⁴⁶ Having violated so many of its commitments in the past, North Korea lacks the ability to signal to cooperative partners through the signing of a binding agreement its intention to comply. This theory does not, however, explain why the KEDO nations would prefer a binding agreement with North Korea. Not expecting the binding provision to provide a significant deterrent to violation and illegal exit in either event, it is not clear why the KEDO nations would not prefer a nonbinding agreement as well. A nonbinding agreement would allow the KEDO nations to withdraw from the agreement more easily and save any reputational costs associated with their own possible violations.¹⁴⁷ Moreover, even if North Korea does retain some reputational capital, the high likelihood of violation may mean that a legally binding agreement imposes a net loss on the parties—little of North Korea's noncompliant behavior is deterred, and so there is little benefit to the KEDO nations, but North Korea still suffers a reputational sanction.

The theory of exit costs offered here explains the rationality of a binding agreement from both the KEDO nations' and North Korea's perspectives. As a state that both has little reputational capital and places little value on the future, North Korea is perfectly willing to enter a binding agreement in exchange for immediate economic relief. The KEDO nations, which take a longer view of events, want a binding agreement because they are looking ahead to a time when those exit costs may influence North Korean behavior.

B. Weak Agreements as a Prelude to Rigid Agreements

A frequent refrain in the debate about form and substance is that relatively less constraining formal provisions—provisions that are more “flexible”—are used as a precursor to agreements that contain more stringent formal provisions.¹⁴⁸ Scholars have argued that “flexible” agreements are

144. Masahiko Asado, *Arms Control Law in Crisis? A Study of the North Korean Nuclear Issue*, 9 J. CONFLICT & SEC. L. 331, 338 (2004).

145. *Id.* at 340.

146. GUZMAN, *supra* note 54, at 86.

147. North Korea has accused the United States of violations of the Agreed Framework and the Supply Agreement by, *inter alia*, suspending oil shipments under false pretenses. Charles J. Moxley, Jr., *The Sword in the Mirror—The Lawfulness of North Korea's Use and Threat of Use of Nuclear Weapons Based on the United States' Legitimization of Nuclear Weapons*, 27 FORDHAM INT'L L.J. 1379, 1385–86 (2004).

148. See, e.g., Abbott & Snidal, *supra* note 4, at 446–47; David J. Bederman, *Review Essay: Constructivism, Positivism, and Empiricism in International Law*, 89 GEO. L.J. 469, 484 (2001). These studies often use slightly different terminology than that used in this paper, referring to agreements that have weak formal provisions as “soft law” and those with strong formal mechanisms that constrain behavior as “hard law.” For reasons discussed elsewhere, I consider the distinction between hard and soft law to refer only to the

used, for example, to allow states to experiment with particular substantive rules and learn how they operate in practice.¹⁴⁹ Similarly, soft law is sometimes seen as a second-best option to hard law, but one that can usher reluctant states from non-cooperation to an eventual hard law agreement.¹⁵⁰ Other scholars have countered by pointing out that there are a number of reasons why relatively more flexible agreements may be first-best, or at least value-creating.¹⁵¹ Flexibility provided by formal provisions can be crucial to making an agreement possible because the flexibility reallocates risk. For example, escape clauses, which allow a state to avoid its substantive obligations under certain circumstances, may promote agreement by providing insurance to domestic politicians, allowing them to make politically necessary decisions without the fear of violating an international obligation.¹⁵²

By explicitly considering renegotiation in light of shifting power and its effect on the initial bargaining over an agreement, the theory presented in this Article bridges the divide between these two viewpoints. I have argued that where one state is increasing in power, a declining state will likely have to make one of two concessions in order to obtain an agreement. On the one hand, the declining state can make the terms of a deal substantively more attractive. On the other hand, the declining state can concede lower exit costs, leaving itself vulnerable to renegotiation but locking in preferred substantive rules for the immediate future. Where an agreement reflects the latter option, states enter into an agreement with the expectation that it will be renegotiated over time to reflect changes in power.¹⁵³

At the same time, if power relationships become more stable over time, it may be possible during a future renegotiation to modify the formal provisions of an agreement. Weak formal mechanisms can be strengthened as states come to expect that the present configuration of power is stable.¹⁵⁴ Consider the following example. As a result of the fact that one state is becoming stronger, two states include provisions reducing exit costs in an agreement they are negotiating. As the ascendant state becomes more powerful, it capitalizes on its ability to threaten exit by demanding that the substantive rules be renegotiated in its favor. At some point, however, the

decision whether to make an agreement legally binding. See Meyer, *Delegation*, *supra* note 15, at 890–91; see also Raustiala, *supra* note 5, at 582.

149. See Koremenos, *supra* note 4, at 292; Setear, *supra* note 25, at 214.

150. Abbott & Snidal, *supra* note 4, at 446–47; Martha Finnemore & Stephen J. Toope, *Alternative Views of 'Legalization': Richer Views of Law and Politics*, 55 INT'L ORG. 743, 750–51 (2001).

151. See, e.g., Helfer, *Exiting Treaties*, *supra* note 34, at 1599–1601; Koremenos, *supra* note 4, at 291; Raustiala, *supra* note 5, at 583; Edward T. Swaine, *Reserving*, 31 YALE J. INT'L L. 307, 311 (2006); Sykes, *supra* note 34, at 289.

152. See Sykes, *supra* note 34, at 279.

153. Alternatively, where the declining state believes that the ascendant state overestimates the probability that it will develop a credible threat to exit in the future, compromising on exit costs may make sense because neither side views itself as making a significant concession. The ascendant state has preserved its future bargaining power and the declining state has obtained its preferred substantive rules at what it considers to be minimal risk.

154. Cf. Powell, *supra* note 3, at 127.

ascendant state may cease getting more powerful (or the declining state may cease getting weaker). At this point in time, the two states may decide to amend the formal provisions of the agreement to raise exit costs.

This evolutionary dynamic provides a theoretical justification for believing that weak formal mechanisms will often be a precursor to more stringent ones. Shifting power causes states to negotiate and renegotiate legal rules over a number of years. At any one point in time, an agreement with weak formal provisions may be first-best given the parties' expectations about the stability of power relationships. But when the parties believe the underlying distribution of power is indicative of the distribution of power for the foreseeable future, they can settle on legal rules that can be incorporated into a more formally stringent agreement.

One might wonder why states would feel the need to raise exit costs if they expect power to be stable. On this view, once the substantive terms of an agreement reflect the power balance between the two states, neither state will be able to do better by exiting in the future. Thus, exit costs are not necessary to dampen the effects of shifting power. While correct as far as it goes, there still could be several reasons why states would want to raise exit costs. First, a doctrinal provision increasing exit costs might also increase the costs of opportunistic violation. Moving from soft law to hard law has this dual effect; the increase in the reputational sanction for violative conduct deters both illegal exit as well as illegal violations. Second, states may be uncertain or have asymmetric beliefs about the likelihood that power is stable. For example, an ascendant state may have private information that it is no longer becoming more powerful because, say, its domestic energy production has dropped or leveled off. It thus no longer attaches importance to low exit costs. The (erstwhile) ascendant state may thus be able to extract a last round of concessions on the substantive rules by agreeing to higher exit costs.

The example of the NPT is illustrative. The central substantive provision of the NPT, which came into force in 1970, is that only five states may legally possess nuclear weapons—the United States, the Soviet Union (now Russia), the United Kingdom, France, and China.¹⁵⁵ As the rule was crafted, it was an effort to freeze membership in the nuclear club at those states that already possessed a nuclear capability at the time the agreement came into effect. Formally, the NPT is a legally binding agreement that contains a sunset provision. After 25 years, the parties to the agreement had to decide by a majority vote whether to renew the treaty.¹⁵⁶ As others have noted, this sunset provision reflected the parties' uncertainty as to what the future held, especially with respect to how the blanket rules against the expansion of the

155. NPT, *supra* note 21, art. IX(3). The provision actually provides that a "nuclear-weapon State is one which has manufactured and exploded a nuclear weapon or other nuclear explosive device prior to January 1, 1967." That definition only captures the five mentioned states.

156. *Id.* art. X(2).

nuclear franchise would work over time.¹⁵⁷ Moreover, there was a fear in the 1960s that many more states would develop a nuclear capability.¹⁵⁸ If that fear had been realized in spite of the NPT, many non-nuclear weapons states within the NPT may have sought to amend the NPT or exit completely. Likewise, states that had developed a nuclear capability between 1970 and 1995 might have sought an amendment allowing them to legally possess nuclear weapons.

Instead, by 1995 the shifts in power as a result of the expansion of the nuclear franchise were considerably less drastic than had been feared. Although India, Pakistan, and North Korea all developed a nuclear capability in the years shortly after the 1995 renewal conference (and Iran may have as well), by 1995 it appeared relatively certain that nuclear weapons were not going to become prevalent among states. Indeed, in conjunction with the NPT's unilateral right of exit,¹⁵⁹ the technological changes since 1995 that allowed for a "latent" nuclear capability—the industrial and technological ability to produce a nuclear weapon in a relatively short period of time—have caused a leveling up in power relationships. States were comfortable both that they could maintain their latent capabilities under the NPT's existing rules, and that there was no need to renegotiate the treaty to broaden membership in the nuclear club. Thus, in 1995 the parties to the NPT renewed the agreement indefinitely, removing that source of leverage.¹⁶⁰

C. *Multilateral versus Bilateral*

Finally, the tradeoff between exit costs and renegotiation has implications for the negotiation of bilateral versus multilateral agreements. Bilateral agreements are the easiest place to see the effect of shifts in power on the initial design of an agreement. As one state becomes more powerful, the other state must react. Behaving rationally, both states take this into account at the time of renegotiation and the descending state internalizes the full costs and benefits of the concessions it makes to the ascendant state.

Of course, even without shifting power, generating agreement in the multilateral context is already considerably harder than in the bilateral context. For example, Gilligan has demonstrated that where all states must adopt the same standard in an agreement, there will be a tradeoff between the breadth of membership and the depth of cooperation.¹⁶¹ The fact that broad multilaterals often are accompanied by shallower substantive terms can cause states desiring deeper cooperation to prefer agreements with

157. See Koremenos, *supra* note 4, at 304–12.

158. *Id.* at 305.

159. NPT, *supra* note 21, art. X(1).

160. Koremenos, *supra* note 4, at 312. The right to unilaterally exit the agreement remains and was exercised by North Korea in 2003.

161. Michael Gilligan, *Is There a Broader-Deeper Tradeoff in International Multilateral Agreements?*, 58 INT'L ORG. 459, 461 (2004).

smaller membership, such as bilateral or regional agreements.¹⁶² Generating a multilateral agreement where the benefit from cooperation is a public good leads to a familiar free-riding problem, in which each state would prefer to remain outside of the agreement so that it can reap the benefits of cooperation without paying the associated costs. Finally, in truly multilateral agreements the reciprocal withdrawal of benefits is not an effective sanction.¹⁶³ Fears about enforceability and compliance can lead otherwise cooperation-minded states to avoid investing in cooperation.

Shifting power exacerbates these problems. Where renegotiation is purely distributional, as it is likely to be in the presence of shifting power, retaining the cooperation of an ascendant state in a multilateral agreement is itself a public goods problem. The ascendant state's continued membership in the agreement is a public good because individual states cannot be excluded from receiving the benefits from the ascendant state's cooperation. Concessions or side payments designed to satisfy an ascendant state thus suffer from the very same problem that enforcement actions do: the state making the concession or side payment cannot capture the full benefit of its actions. The result is that even where concessions or side payments might be optimal from the standpoint of global welfare, such concessions or side payments will be undersupplied.

These difficulties with renegotiation will have ramifications for the initial design of institutions. As I have argued in this Article, potentially ascendant states will demand one of two things in exchange for cooperation. They will either expect more favorable substantive rules than their current bargaining power would seem to justify, or they will expect low exit costs. As noted above, however, where an agreement produces a public good, the need for broad membership will tend to drive the substantive terms toward shallower cooperation.¹⁶⁴ Moreover, certain states may wield effective veto power over the terms of a multilateral agreement.¹⁶⁵ The result is that it may be impossible to provide ascendant states with their preferred substantive rules. Where the substantive rules necessary to attract an ascendant state would be blocked by a veto-wielding state, states may have to look to other avenues to attract the ascendant state.

This yields an empirically testable proposition: in the presence of shifting power we would expect to see multilateral agreements with lower levels of exit costs relative to bilateral agreements entered into by some subset of the parties to the multilateral agreement. The logic of this proposition is

162. See Laurence R. Helfer, *Towards a Human Rights Framework for Intellectual Property*, 40 U.C. DAVIS L. REV. 971, 975 (2007) (noting that in the area of intellectual property developed nations are moving increasingly to regional or bilateral agreements that provide greater protections than even the most protective broad multilateral agreements); Chris Brummer, *The Ties That Bind? Regionalism, Commercial Treaties, and the Future of Global Economic Integration*, 60 VAND. L. REV. 1349, 1350 (2007).

163. Guzman & Meyer, *International Soft Law*, *supra* note 19, at *22.

164. See Gilligan, *supra* note 161, at 461.

165. Raustiala, *supra* note 5, at 598–99.

straightforward. The need for broad membership—sometimes termed the “participation constraint”—in an agreement constrains states’ ability to vary the substantive terms.¹⁶⁶ A substantive rule more favorable to an ascendant state may redistribute the gains from cooperation away from one or more other states. These states, in turn, might be unwilling to join the agreement as a result of the redistribution. Where these states are necessary to the success of the agreement—as in the public goods context—the participation constraint will rule out making substantive rules more favorable to the ascendant state. Thus, attracting the ascendant states to the agreement will mean agreeing to reduce the level of exit costs.

Moreover, this same logic also suggests that ascendant states should actually exit multilateral agreements more frequently than bilateral agreements. The reason for this further prediction is two-fold. First, and unremarkably, we would expect exit rates to be higher in agreements with lower exit costs, all else equal. Thus, a prediction that exit costs will be lower leads directly to a prediction that observed exits should be higher. Second, in this situation, renegotiation of an agreement’s substantive terms is likely to fail for the same reasons that the initial negotiation could not result in substantive rules favorable to the ascendant state. The participation constraint may dictate allowing a single ascendant state to exit, rather than change the substantive rules that ensured broad membership in the first place.¹⁶⁷

The example of climate change negotiations is instructive. Climate change negotiations have just recently completed their third major iteration. The first was the 1992 UNFCCC; the second was the Kyoto Protocol to the UNFCCC; and the third was the Copenhagen negotiations in December 2009. One of the central issues in these negotiations has been the extent to which groups of states would accept legally binding emissions targets. Making emissions targets nonbinding, of course, is one way to reduce the sanctions both for not meeting the targets and for *de facto* exiting from the targets.

One way to operationalize bargaining power in the climate change area is through relative emissions levels. Because climate change is a public bad, states or groups of states that produce a significant percentage of the world’s emissions may be necessary members of a regime in order for the regime to be effective. It follows that as a state or group of states produces more emissions, it will be able to demand more concessions in exchange for its participation in a climate change regime. Moreover, developing nations in particular have much to lose, in terms of the ability to industrialize their economies, by agreeing to binding emissions caps. Developing states there-

166. See BARRETT, *supra* note 47, at 201–05.

167. A third possibility, of course, is that the presence of shifting power will make it impossible to satisfy the participation constraint at the outset. In such situations, we would not necessarily observe either lower exit costs in multilateral agreements or higher rates of exit. Instead, states would simply not enter into agreements in situations in which exit costs would otherwise have to be low.

fore have a two-fold interest in resisting climate change negotiations. The first, which has been recognized throughout the climate change negotiations, is that emissions caps could harm their economic prospects.¹⁶⁸ But the second is that developing states might reasonably expect to be able to negotiate a better climate change agreement at a time when they are having a more negative impact on climate change.

The table below illustrates changes in relative emissions levels starting in 1990.¹⁶⁹ The trend is clear, with developing nations India and China soon expected to produce roughly one-third of the world's carbon dioxide emissions,¹⁷⁰ while developed nations such as the United States and European countries expect to see their share of emissions fall from nearly forty-five percent to less than thirty percent by 2030. One can also see that from 1990 onward the United States is becoming more "powerful" relative to European nations, as its share of global emissions falls slower than Europe's does.

TABLE 1: RELATIVE CONTRIBUTIONS OF ANNUAL CARBON DIOXIDE EMISSIONS BY COUNTRY/REGION (APPROXIMATE PERCENT OF WORLDWIDE EMISSIONS)¹⁷¹

	1990	2003	2004	2010	2015	2020	2025	2030
United States	23.5%	22.7%	22.0%	20.1%	19.4%	18.8%	18.7%	18.5%
OECD Europe	19.3%	16.9%	16.3%	14.6%	13.4%	12.4%	11.6%	10.9%
China	10.5%	15.3%	17.5%	21.1%	22.4%	23.9%	25.0%	26.2%
India	2.7%	4.1%	4.1%	4.2%	4.4%	4.7%	4.9%	5.0%
Japan	4.8%	4.9%	4.7%	4.1%	3.8%	3.5%	3.3%	3.0%
Africa	3.1%	3.5%	3.4%	3.7%	3.8%	3.9%	3.9%	3.9%

The bargaining over the form and substance of emissions targets illustrates the prediction that in the presence of shifting power, multilateral agreements are likely to have low exit costs. The UNFCCC, for its part, contained only a nonbinding target that states stabilize their emissions at 1990 levels by the year 2000.¹⁷² As the theory put forth in this Article suggests, the United States made substantial concessions on substantive points to ensure that the targets would be non-binding.¹⁷³ In the Kyoto

168. Charlotte Booncharoen & John Gase, *International Commitment Toward Curbing Global Warming*, 4 ENVTL. L. 917, 932 (1998).

169. Eric A. Posner & Cass R. Sunstein, *Should Greenhouse Gas Permits Be Allocated on a Per Capita Basis?*, 97 CAL. L. REV. 51, 61 (2009) (quoting International Energy Agency statistics from 2006).

170. *Id.*

171. *Id.*

172. United Nations Framework Convention on Climate Change art. 4.2, May 9, 1992, 1771 U.N.T.S. 107 (hereinafter UNFCCC). Although the UNFCCC is itself a binding agreement, the emissions reductions obligations are specifically couched in nonbinding terminology.

173. First, the United States did agree to a specific reduction target and timetable: stabilizing output at 1990 levels by the year 2000. The UNFCCC is thus not a case in which the obligation reached was

Protocol, the tradeoff was made in the opposite direction. The United States agreed to binding targets, but extracted significant concessions on the substance of the emissions control regime established by the Kyoto Protocol.¹⁷⁴ But the United States failed to ratify the Kyoto Protocol and thus ultimately failed to agree to binding targets. Moreover, in neither agreement did developing nations—the fastest-growing producers of greenhouse gas emissions and thus ascendant states in this Article’s terminology—accept any binding obligations to reduce emissions, despite the United States’ insistence that they do so.¹⁷⁵

Fortunately, if the chart above reflects an accurate estimate of how states’ greenhouse gas emissions will be allocated in the future, the prospects for an agreement in the next several years may be brighter. The chart indicates that while major changes in the allocation of emissions occurred or are expected

couched purely in terms of principles, although the UNFCCC does contain quite a bit of language calling on signatories to abide by various principles. Second, the United States dropped its resistance to providing financial support to developing countries. The text of the UNFCCC creates a legal obligation for developed states to provide financial aid to developing nations, and outside of the UNFCCC, the United States committed to a seventy-five-million-dollar figure. Finally, developing nations made no specific commitments—binding or otherwise—with respect to emissions targets. The United States had strongly favored some commitment from developing states. Shardul Agrawala & Steinar Andresen, *Indispensability and Indefensibility? The United States in the Climate Treaty Negotiations*, 5 GLOBAL GOVERNANCE 457, 461–62, 465 (1999).

174. The Kyoto Protocol included specific targets for Annex I countries, including, for the United States, a seven percent reduction of greenhouse gas emissions from 1990 levels by 2008–2012. These targets reflected a twin American victory. First, the time period agreed upon reflected exactly the American proposal, designed to ameliorate the effects of short-term fluctuations in economic or weather conditions that might affect a state’s ability to meet a specific deadline. Second, the target itself reflects agreement on a standard very close to the American standard. Agrawala & Andresen, *supra* note 173, at 465. Perhaps the greatest success for the United States, however, came in the form of the Kyoto Protocol’s emissions trading rules. A major emissions trading program was the chief concession that the United States demanded in exchange for accepting binding emissions targets. David M. Driesen, *Sustainable Development and Market Liberalism’s Shotgun Wedding: Emissions Trading under the Kyoto Protocol*, 83 IND. L.J. 21, 34–35 (2008); James H. Searles, *Analysis of the Kyoto Protocol to the U.N. Framework Convention on Climate Change*, 21 INT’L ENV’T REP. 131, 133 (Feb. 4, 1998). The Kyoto Protocol involves three distinct emissions trading programs. Under Article 16, developed countries may trade their national allowances at the state-to-state level. Moreover, Article 6 allows the parties, or private actors within the state-parties, to buy and sell emissions reductions units created by projects within a country (this method of emissions trading is known as “joint implementation”). For example, a private firm that creates a carbon sink may essentially sell the resulting emissions reductions to a firm in a different country, with the latter country receiving credit for the reduction. A similar mechanism, known as the Clean Development Mechanism (“CDM”), is established by Article 12. Under CDM, developed countries or private firms within developed countries can purchase certified emissions reductions from firms in developing countries, despite the fact that developing countries have no emissions reductions obligations under the Kyoto Protocol. Thus, under the agreement, states can trade their legally allotted allowances with one another, and they can also invest in reductions in other countries, both developed and undeveloped. Collectively, these measures are known as the Kyoto Protocol’s “flexibility mechanisms.” Driesen, *supra*, at 34–35.

175. A continuous sticking point for the United States during climate change negotiations, that in part led to its failure to ratify the Kyoto Protocol, has been the inclusion of binding targets for developing as well as developed nations. In 1997, by a vote of ninety-five to zero, the Senate passed the Byrd-Hagel Resolution, which provided in relevant part that the United States should not agree to any binding emissions targets “unless the protocol or other agreement also mandates new specific scheduled commitments to limit or reduce greenhouse gas emissions for Developing Country Parties.” S. Res. 98, 105th Cong. (1997) (enacted).

to occur between 1990 and 2010 or 2015, after 2015 the rate of change is expected to slow considerably. If states expect emissions levels to stabilize in coming years, then the prospect of future renegotiations may be sufficiently reduced such that states are able to get an agreement that places meaningful emissions limits—with high exit costs—on all relevant states.

V. CONCLUSION

Existing theories of the design of international agreements do not account for the fact that states' bargaining power changes over time. However, power is constantly shifting between states, often in dramatic ways. These shifts in power affect the design of international agreements. States take into account their expectations about future power, and if a state expects to be stronger tomorrow, then agreeing to constrain its future behavior represents an opportunity cost for which it should expect to be compensated. For a variety of reasons, including uncertainty, the extent to which states discount the future, and commitment problems, states may not wish or be able to reach an agreement that is renegotiation-proof. States thus must decide how to trade off the substance of an agreement (the distribution of benefits) against the level of exit costs. They must determine, in effect, how much to take their future interactions out of the realm of politics and put them into the realm of law. Understanding how states make this tradeoff—how they cope with their shifting prospects for the future—can help us unlock ways to generate value-creating agreements in crucial areas of international cooperation.

