CHAPTER SIX

RESURRECTING COMPLIANCE IN ASSESSING LAW'S IMPACT ON BEHAVIOR

Tonya Putnam
Department of Political Science
Columbia University

** DRAFT **
January 30, 2020
(please do not circulate or cite without permission)

Legal compliance has gotten a bad rap in international relations (IR) scholarship. Compliance – defined here neutrally as being on the "legal" side of a legal/illegal binary – was for close to two decades considered essential for demonstrating that international treaty commitments can influence the behavior of states and governments.[1] During this period, evidence indicating that there is no necessary relationship between patterns of compliance with international legal obligations, and changes in underlying behavior began to accrue. For example, it was demonstrated that a state may comply with international treaty obligations without altering its behavior if the treaty reflects its conduct prior to joining.[2] In addition, a government with a poor record of compliance on a particular issue can improve its behavior substantially, and yet still fall short of a collectively defined compliance threshold.[3] Similarly, if a state's conduct exceeds the minimum required by a treaty, that conduct can worsen without necessarily tipping the state into non-compliance.

As a result of these and similar revelations, compliance lost considerable ground as a variable of interest among empirical IR scholars. In its place, IR researchers have been exhorted to leave compliance to the lawyers, and, instead, to focus on indicators that more directly describe the substance of rule-governed behavior, and relate it to purposive goals formalized in treaties and other legal instruments.[4]

---

[1] This period dates from roughly the early 1990s to the late 2000s.
[2] Downs, Rocke, and Barsoom 1996.
[3] Raustiala 2000, Raustiala and Slaughter 2002.
[4] Martin (2014) "Against Compliance". See also Hafner Burton and Tsutsui 2007, Howse and Teitel 2010.

This chapter makes two arguments. The first is that discarding "compliance" as a variable of interest in empirical studies of international law is an over-correction that *actually impedes* the ability to theorize and observe how law affects behavior. To state matters bluntly, efforts to address the inferential challenges that flow from using the compliance binary as an exclusive outcome variable by tossing out the concept of compliance altogether make no sense. They are tantamount to excising law and legal process from efforts to study law's influence and effects. This is because law exists in no small measure to describe conduct that is permitted or required in the purposeful pursuit of policy objectives and to differentiate it from that which is prohibited. Attempting to describe law's influence (or lack thereof) on behavior without a concept of legal compliance is thus a bit like trying to describe the effects of a sudden rainstorm without the concept of 'wet' or 'dry'.

The second argument is that conventional IR approaches to legal compliance embody flawed assumptions about its character and functions. One such assumption is that compliance thresholds are set exogenously in formal lawmaking processes. Another is that compliance thresholds are easily identifiable from the texts of legal rules, and thus readily support the formation of shared conjectures about when rules apply and what they permit or require. In contrast, I argue that even clearly drafted legal texts invariably require some interpretation when applied to concrete situations, and that the indeterminacies that inhere in rules drafted for general use often allow for more than one reasonable interpretation in application. Ascertaining what constitutes the more 'reasonable' interpretation in light of legal texts and prior patterns of usage is precisely what occurs in legal processes. It follows that operative thresholds for legal compliance are thus at least partly endogenous to patterns of prior use. As a consequence, it is problematic to assume that compliance thresholds remain fixed in time, or that they do not vary across settings, or across different configurations of actors. Understanding how legal compliance thresholds emerge, how they vary, and also under what conditions this variance is likely to be tolerated or contested, and by whom, is, therefore, important – and intensely *political*.

As I underscore in earlier chapters, law is a means to enable forward-looking collective action, to formalize rights, and to regulate the exercise of political power and the exercise of social control. In these ways law invariably influences the distribution of social burdens and benefits within and across polities. Insofar as laws are generated by, and operate within, systems comprised of other rules, standards, and procedures, it is also the case that efforts to use law frequently implicate a variety of considerations that directly and indirectly condition when, how, and for what purposes

specific rules may be invoked.[5] For this reason, identifying and describing compliance thresholds can easily give rise to debate – especially in situations where law is doing some or all of the 'heavy lifting'[6] in efforts to guide, assess, coordinate, or coerce behavior.[7]

At the level of individual actors and actions, compliance with legal rules may be assessed in purely binary terms. Alternatively, it may be described and conceptualized along a continuum. Here the relevant query is not simply whether behavior is legally compliant, but also by what margin. Attention to directional trajectories of behavior toward or away from hypothesized thresholds for compliance adds another layer of information. Importantly, *none of these modes of inquiry is feasible without some conjecture about the existence of a compliance threshold*, coupled with a theory about where along the continuum of possible behavior that threshold lies. Without such a conjecture, the concept of legal rule-following itself lacks coherence.[8]

Policy makers and social scientists often prefer to engage with arguments about legal compliance at higher levels of aggregation than is the case for lawyers. The focus thus is less on individual instances of rule following (or its absence), and more on identifying central tendencies, or patterns, of behavior. However, what happens at the macro level of compliance cannot be explained without understanding the underlying processes by which the legal "sausage" is made. Here compliance – and, more specifically, efforts to construct and maintain compliance thresholds – is a key ingredient Consequently, it is worth asking IR scholars to think carefully about what the creation and maintenance of international legal compliance entails, and what attention to those thresholds can (and cannot) tell us about rule-governed behavior.

To explore these ideas further, I make use of several theoretic propositions developed in prior chapters. These include ideas about legal indeterminacy, the strategic construction of legal decision environments, practice-based time variance in legal rules, and the rule elaboration functions

[5] The effects of individual legal rules often cannot be observed or explained in any rigorous causal sense (and still less predicted) without attention to variance in how different legal systems operate, along with who, or what, has agency within them. See Chapters 1 and 2 for additional discussion of these claims.
[6] For further discussion of this idea see page 15 below, page 24 *et seq.*, and Chapter 1.
[7] Even where law and legal obligations remain in the background – in the guise of internalized norms and/or accepted institutional constraints – conjectures about legal compliance function as orientating points, or baselines. These, in turn, help actors to frame issues, to specify menus of choice, and to estimate costs and benefits to various courses of action.
[8] For compliance to be a theoretic possibility, parties to legally mediated transactions, together with relevant legal authorities, must be prepared to accept some configuration of acts, forbearances, and/or mental states as satisfying the minimum requirements of applicable legal rules. For those that have not internalized the objectives underlying legal rules, the absence of such a possibility may weaken incentives to try to adhere to rules in question. From the other direction, if all possible behavior is considered compliant, then the rule has no meaning.

of international courts and tribunals. Together they point toward a more dynamic and more politically rich understanding of legal compliance as a concept, and of its role in distributional politics.

After first offering a brief review of the theoretic argument and highlight some key implications for conceptualizing compliance, I then sketch the emergence and subsequent decay of research attention to legal compliance. Next I provide a more detailed critique of how "compliance" has been conceptualized in the IR literature with emphasis on the stakes involved in assuming that compliance thresholds are common knowledge, or readily discoverable as such. In the last section I focus on reclaiming compliance as a variable in the empirical study of international law, though transformed to emphasize what it can tell us about the applicability, scope, and content of legal rules as mediated by bargaining within and among interpretive communities.

Legal Indeterminacy as a Compliance Complicator

A central argument in this book is the idea that law is a privileged category of norm. Across a broad range of societies and political regime types, invoking law often triggers a range of psychological and social propensities for obligated actors to factor its existence into strategic behavior. This also feeds expectations that others will act similarly.[9] My argument layers this narrow but highly general form of normativity onto its capacity to reset collective baselines for deliberation, and with the boundedly rational character of individuals and organizations, which prompts them to adopt heuristics as decision tools. Deference to law – or more accurately, contextually credible claims about the applicability, scope, and meaning of legal rules – comprise one such heuristic.

The additional claim that legal rules are fundamentally indeterminate at a general level complicates this, however. Here the main idea is that the limits of linguistic precision, combined with the need to draft legal rules in general terms but apply them in specific circumstances, together make it impossible to reach definitive findings about the applicability, scope, and meaning of specific legal rules that hold across settings, and over time. How can both of sets of ideas apply simultaneously and still allow law to function? The answer is to be found – counter-intuitively – in what I call "amplifiers" of indeterminacy.[10] By systematically undercutting any potential for legal rules to operate cohesively and coherently across diverse settings, these amplifiers help to create

[9] Hart 1994, Raz 2009. See Chapter 2 for a full discussion. This normativity and its accompanying inter-subjectivity shapes behavior whether actors are seeking to follow legal rules, or to circumvent or alter them.
[10] See Chapter Two for a full discussion.

conditions for localized versions (or vernaculars) of more broadly applicable legal rules to take root and thrive.

The first amplifier is structural. Most legal rules are simultaneously in use by multitudes of actors in a variety of settings. Opportunities for consultation with ultimate authorities (assuming such authorities even exist) about the applicability, scope, and meaning of particular rules are generally rare, and may involve substantial costs. It follows that decisions involving law-governed behavior by necessity occur with incomplete information about how others have used the same (or similar) rules. This allows for multiple plausible interpretations of individual rules, or understandings of how different legal rules interact, to operate in parallel.

Two additional amplifiers are agent-centric. The first of these is bounded rationality and the accompanying reliance on decision heuristics to which it gives rise. Importantly, credible claims of lawfulness (or non-lawfulness) can themselves serve as decision heuristics by offering short cuts for formulating and evaluating choices in complex situations. At the same time, other ubiquitous heuristics, such as 'satisficing' and 'local search,' feed indeterminacy and underlying variance in applied legal rules by prompting decision-makers to prefer adequacy over optimality.[11] It follows that observed patterns of law's successful invocation, interpretation, and application in similar settings or domains to achieve objectives similar to those of the agent will often suffice as a workable hypothesis of legality, and observed patterns of local failure will have the opposite effect.

The second agent-centric amplifier of legal indeterminacy is the strategic nature of users of law. Those who seek to use law to guide their own behavior, or that of others, may be inclined to utilize whatever interpretative leeway they can find (or create) in their favor. The presence of strategic actors thus might (in theory) render the ubiquity of law's indeterminate features into a veritable playground possibility. In practice, however, this potential is often heavily disciplined both by interpretative and procedural constraints in the law itself, and by the oversight and intervention of interpretative communities of engaged observers.[12]

An important dimension of this analysis thus involves anticipating what relevant interpretive communities – conceived broadly to encompass formal and informal authorities on law, as well as users of law and those who interact with them by choice or by necessity – are willing to accept as "compliant" in the situations they encounter. This can vary according to the law at issue, the actors

[11] Lindblom 1959, Simon 1972, Forester 1984.
[12] Higgins 1994.

involved, and the circumstances in which the rule is being applied.[13] Importantly, it can also vary according to the composition of the interpretive communities – which may not be entirely predictable at the outset. In sum, even within particular geographic and issue domains, law's influence can be conditional on many other factors – which is another way of saying that "all else" often is *not* equal where legal rules are being interpreted and applied.

THE RISE AND FALL OF COMPLIANCE

The deep roots of IR fascination with international law compliance can be traced to 1980s explorations of institutions and regimes as behavior-altering features of the international system.[14] In 1992, two international lawyers, Abram Chayes and Antonia Handler Chayes, published a game-changing article, entitled "On Compliance," containing the following bold and testable-hypothesis-rich claim: "When nations enter into an international agreement […] they alter their behavior, their relationships, and their expectations of one another in accordance with its terms."[15] With this, a decade-spanning research agenda was launched.[16]

Chayes and Chayes formulated their claim to counter a (then) prevalent view among IR scholars with a structural-materialist bent that international law is "epiphenomenal" to international politics. The idea here was that legal rules are mere artifacts of politics among states; they exert no independent influence on behavior. This stance served to subordinate international law to more imperative concerns, like power maximization and "self-help," that were said to be generated by the very "structure" of the international system.[17] From this perspective it was naïve – and even dangerous – to behave as if international law might alter or constrain "rational" state behavior.[18]

In a notable departure from this mode of thinking, "On Compliance" takes Louis Henkin's oft-quoted observation in *How Nations Behave* (1979) that "almost all nations observe almost all principles of international law almost all of the time" as a point of departure[19] Far from being an exercise in arcane doctrinal argumentation, Henkin's account explores a variety of social

[13] Chayes and Chayes 1993.
[14] Raustiala and Slaughter 2002. See in particular Krasner 1982, Snidal 1985, Keohane 1984, Young 1989.
[15] Chayes and Chayes 1992:176.
[16] Note, however, in retrospect it is clear that Chayes and Chayes conceded too quickly that "no rigorous non-tautological answer is possible" to the question of whether structural realist accounts of decision-making provided a better explanation than a propensity to comply with international law. In the past 25 years numerous advances in qualitative, quantitative, and experimental methods have made the reconstruction of decision processes more tractable.
[17] Waltz 1979, Mearsheimer 1994.
[18] Mearsheimer 1994, Krasner 2002.
[19] Henkin 1979:47.

mechanisms to account for this pattern, including the importance to states of maintaining "credit" (or credibility) among international interlocutors, and to build reputations as followers of international law.[20] Having adopted the position that compliance with international obligations is the norm among states, Chayes and Chayes then took the unusual (for the time) tack of asking not why states would ever reasonably opt to keep their international commitments, but why they sometimes fail to do so? [21]

At the core of the Chayes and Chayes' argument is the claim that international law changes behavior not due to any radical shifts in the likelihood of being sanctioned for violations, but by changing in expectations about how *other* obligated entities are likely to behave. Furthermore, using examples from international trade to environmental law to arms control, Chayes and Chayes argued that states' compliance failures often are not purposeful or undertaken to maximize short-term gains. Rather, they are more often attributable to a triad of factors: ambiguity of treaty terms, states' capacity limitations, and changes in underlying social or economic conditions (i.e. unanticipated shocks). Ignoring these causes of non-compliance, Chayes and Chayes cautioned, can tip policy responses toward coercive sanctions in situations where strategies to "manage" sources of non-compliance would be more effective.[22]

But, all was not settled in the land of compliance studies. In 1996 Downs, Rocke, and Barsoom, responded with the provocation: "Is Good News About Compliance Good News About Cooperation?" Their answer was that frequently it is not. They argued that many international agreements are quite "shallow," meaning they ask little of members in terms of changed behavior. They insisted that agreements with this profile can easily generate high rates of compliance, but with few real "effects." They argued further that this apparent preference for agreement shallowness is endogenous to the paucity of reliable mechanisms for enforcing legal rules in the international system. Using a game-theoretic model, they showed that states will not rationally agree to alter their

[20] Henkin 1979:52. Although Henkin's treatment did not use the lingo of contemporary social science (some of which had not yet been invented), his treatment nevertheless reflects awareness of analytic factors such as observation bias, attribution errors, and counter-factual reasoning.

[21] It is worth underscoring is that these were no mere armchair propositions formulated inside the Ivory Tower. Henkin and both of the Chayes' all had extensive experience as government lawyers working on international affairs. In the 1950s Louis Henkin worked as a lawyer in the U.S. Department of State in the Office of European Regional Affairs (which became NATO), and later in the United Nations Office. Abram Chayes served as the State Department's Legal Advisor during the Kennedy Administration and remained heavily involved in arms control efforts until the mid 1980s. Antonia Handler Chayes was an Assistant Secretary of the Air Force, and then Under-Secretary of the Air Force during the Carter Administration.

[22] Chayes and Chayes 1995. Subsequently this logic was dubbed the "managerial approach."

own behavior in costly ways without assurance that others' decisions to defect will be punished to the point of erasing expected gains from defection.[23] Accordingly, they found that although the managerialists were not necessarily wrong about general patterns of rule-following, their optimism regarding international law's capacity to transform international relations was misplaced.

Compliance as a puzzle and a challenge

Downs, Rocke, and Barsoom's argument about the frequent shallowness of treaty commitments, and its implications for rates of compliance, struck old-school international law skeptics, together with many political economists who had not previously thought much about international law, as intuitively plausible. But it also raised difficult questions. Chief among them was why states would repeatedly pay the costs of enacting, and joining, international agreements that were presumed to have little or no effect on behavior? Efforts to address this and related puzzles spurred further productive inquiry into the mechanisms and processes by which internal law might alter the incentives and preferences of agreement members, or the expected payoffs from participation.

In this vein, one group of scholars explored whether and to what degree international agreements vary in terms of their "legalization" – meaning how much (or little) they reflect the characteristics of an ideal type of "hard" law. The approach aimed to develop, and locate international agreements within, a typology of three scaled characteristics – obligation, precision, and delegation of dispute resolution.[24] A key finding, for the purposes of this discussion, was that international agreements do indeed vary considerably in the degree to which their substantive obligations are expressed in imperative terms (from "expressly non-legal norms" to preemptory rules) and in the precision with which those obligations are formulated (from vague principles to precise, "highly elaborated" rules).[25] The implications for arguments over legal compliance were unclear, however.

From one perspective the existence of a "bindingness" continuum could be viewed as reinforcing the Downs *et al* endogeneity thesis if it turns out that higher degrees of obligation track agreement "shallowness." At the same time, the very fact that states seek to enact international

[23] Downs, Rocke, and Barsoom 1996:384-385. Subsequently this debate appears to have inspired Goldsmith and Posner to dust off the old "international law as epiphenomenon" arguments and to elevate them to grand theory.
[24] Abbott *et al* 2000:401-402. Note the definition used in this exercise is maddeningly circular in that the "hard law" ideal type is defined wholly in terms of the three characteristics.
[25] Abbott *et al* 2000:404.

agreements that utilize varying degrees of obligation suggests that international agreements do not always (or in all respects) aim to formalize behavior that would occur anyway. If that were the case, there would be no downside to opting for full bindingness all the time. A vibrant literature on the use of so-called "soft law" and international cooperation in issue areas ranging from environmental policy to international trade and investment indicated strongly that this was not the case.[26]

A different group of scholars took a more openly functionalist approach to asking whether legal compliance might be in part related to institutional design. Here the key idea was that different kinds of problem structures associated with collective action – including the gnarly Prisoners' Dilemma-type problems highlighted by Downs *et al* – can be addressed more optimally or less using different design elements. For example, if agreement drafters anticipate strong incentives to defect, they can draw from a toolbox of options for changing, or mitigating, those incentives. These might include the inclusion of monitoring provisions, options for centralizing (or decentralizing) certain decision making powers, manipulating membership criteria, adjusting the substantive scope of the agreement, or including provisions for dispute resolution. Where employed skillfully, this literature finds that international treaty design can alter the breadth, depth, and quality of cooperation.[27]

Underlying these conclusions, and the models on which they rely, however, are some unexamined assumptions. One is that international agreements, in fact, reflect shared understandings about their own applicability, scope, and meaning – or to the extent they do not, that embedded design features are capable of imposing a workable degree of uniformity. Another is that such understandings formalized in written rules remain static over time. Relaxing these assumptions in line with the ideas about law and legal indeterminacy elaborated here, however, suggests that agreement design may have a more modest and time-limited role than this literature claims.

Still within the rationalist-materialist idiom, but more in line with the theses forwarded by Chayes and Chayes, several IR scholars took up the question of how patterns of compliance (and non-compliance) with international agreements convey information about a state's "type."[28] Here the driving intuition is that a reputation for keeping international commitments is valuable beyond whatever direct gains accrue from cooperation, and is thus worth building and safeguarding.[29]

---

[26] Dupuy 1990, Boyle 1999, Abbott and Snidal 2000, Shaffer and Pollack 2009, Guzman and Meyer 2010, Brummer 2010. Cf. Klabbers 1996 and 1998.
[27] Koremenos, Lipson, and Snidal 2001, Fearon 1998.
[28] Martin 2005, Morrow 2007.
[29] Keohane 1996, Simmons 2000, Lutz and Sikkink 2000, Simmons and Hopkins 2005, Kelley 2007, Guzman 2002, 2005, 2008, Kydd 2009.

Specifically, it was hypothesized that states' observable records of compliance feed others' expectations for future trustworthiness, which, in turn, influences whether opportunities for international cooperation arise, and on what terms. However, the ability of states to form multiple, non-fungible reputations across different clusters of states and across different domains, along with questions about to what or whom statecraft-related reputations accrue (states, governments, individuals), all complicated efforts to theorize reputation as a general mechanism underlying compliance and non-compliance.[30]

Other research explored whether making formal treaty commitments, or the anticipation of doing so, is itself associated with changes in state behavior.[31] A key finding in this work was that formal commitment making serves both to "screen" out insincere joiners, while also incentivizing compliance *ex post*.[32] These effects were quickly understood as conditional—first, on the degree to which the agreement in question has barriers to joining (in the form of either *ex ante* preconditions, or reasonable fears of *ex post* sanctioning of violations), and second, on the ability of treaty membership to generate enough shared gains to trigger self-enforcing dynamics. Still, this line of argument went some way toward explaining high degrees of compliance within multilateral economic agreements.[33] But what of other issue areas where agreements tend to have fewer preconditions for membership, and the likelihood of being punished for violations is lower?

In 2002 legal scholar Oona Hathaway turned international human rights scholarship on its head with an article that presented findings from a series of simple models regressing ratifications of human rights treaties against topic-specific outcome data drawn from a variety of sources, along with state-level controls borrowed from earlier cross-national studies of political violence.[34] She found (contra her own expectations) that patterns of state compliance were decidedly mixed – even among established democracies. This prompted a flurry of research activity probing both her models

[30] Keohane 1996, Downs and Jones 2002, Brewster 2009. But see Kelley 2017.
[31] Depending on the setting, anticipation of legal commitment can prompt better behavior – or worse. "Better" behavior might occur in anticipation of joining institutions with membership pre-requisites (e.g. the WTO, the European Union, some military alliances). By contrast, prospective signatories to armistices often behave worse in anticipation of agreements in order e.g. to grab as much territory as possible, or to weaken known pockets of resistance in areas of contested control (Fortna 2003). To avoid incentivizing this type of harmful behavior, environmental agreements often set individual country benchmarks at points well prior to the point of obligation [need citation]. Human rights commitments do not appear to encompass either type of anticipatory effects, except where other benefits are made conditional on human rights practices (Hafner Burton 2011).
[32] Von Stein 2005, Simmons and Hopkins 2005.
[33] Neumeyer and Spess 2005, Elkins, Guzman, and Simmons 2006, Tobin and Rose-Ackermann 2011; Hafner Burton 2005, Buthe and Milner 2008, Yackee 2008, Tobin and Busch 2010, Baccini and Urpelainen 2014.
[34] Hathaway 2002, Mitchell and McCormick 1988, Poe and Tate 1994.

and the underlying data.[35] This work produced important qualifications to, and extensions of, these initial findings. These include, notably, differences among regime types and regime trajectories, and how such characteristics influence not just *ex post* incentives for commitment keeping, but also selection into treaty commitments.[36] It also further solidified the convention of using cross-national patterns of ratification as a baseline for addressing questions about international human rights commitments and subsequent behavior.[37]

As this was underway, attention to two additional factors basic to international politics also began to figure more prominently in inquiries about compliance with international law: power and agency. Institutionalist treatments of international cooperation have been often faulted by realists for giving insufficient attention to power differentials among states.[38] In parallel to many of the developments just described, IR constructivists opted for a fundamentally different approach to theorizing the origins and socio-behavioral underpinnings of norm and rule following that nevertheless kept power and agency in focus.[39]

Instead of reducing social interactions to information and material incentives, constructivist scholars looked to the ability of norms, ideologies, identities, and peer associations to shape behavior and strategic choice. Much of this work also sought to explain stability and change in the attitudes, beliefs, and preferences of instrumental actors that many rationalist treatments specify by assumption (and often treat as fixed by methodological convention). Phenomena of specific interest in this scholarship include cycles of norm emergence, socialization, and (sometimes) internalization, along with patterns of norm diffusion -- all of which can be shaped by various facets of international power and influence.[40]

On the agency front, some materialist-oriented IR scholars also began to question the expectation that treaty commitments alone are sufficient to create or sustain legally compliant behavior.[41] This prompted a new wave of inquiry into patterns of domestic level implementation and

[35] Goodman and Jinks 2003, Simmons 2009, Simmons 2010:289.
[36] Moravscik 1998, Hafner Burton 2005, Hafner Burton and Tsutsui 2007, Simmons 2009, Hill 2010.
[37] Hafner Burton and Ron 2009, Shaffer and Ginsburg 2012.
[38] Wendt 1992, Jervis 1999.
[39] Occasionally constructivists crossed over to spar with rationalist scholars directly. See e.g. Finnemore and Toope 2001.
[40] Finnemore and Sikkink 1998, Checkel 1999, Payne 2001, Goodman and Jinks 2003 and 2013, Barnett and Duvall 2004, Cole 2009, Risse, Ropp, and Sikkink 2013.
[41] Perhaps ironically this stance among empirical scholars, who often have (at best) modest expectations for the transformative power of international law, reflects both a hefty dose international legal formalism, along with blind faith in *domestic* law and legal institutions.

uptake of international legal commitments.[42] Whereas some studies focused on variance in domestic implementing legislation or constitutional adjustments,[43] others examined the relationship between states' incorporation of international legal commitments into domestic law and subsequent patterns of political and legal mobilization.[44] In a number of settings and issue domains, scholars found associations between international legal commitments the emergence of domestic-level constituencies organizing to demand that their own governments, and occasionally those of other states, comply with international legal obligations.[45]

Compliance as a distraction and a limiting idea

Alongside, and in dialogue with, the many studies aiming to test for compliance with international legal rules, a more research design-oriented set of criticisms concerning the limitations of compliance as a concept and a metric was taking shape. Like much of the literature just described, this line of argument was implicitly bullish on the idea that, under at least some conditions, international law and legal commitments can change states' preferences and behaviors. At the heart of this critique is whether "compliance" is adequate to capture and describe those changes, or to observe their absence.[46]

Victor, Raustiala, and Skolnikoff (1998), Victor (1999), and Raustiala (2000) pointed out that law's "effects" and "effectiveness" are what policy makers and social scientists alike often care about, but that "compliance" as a binary fails as a proxy for these concepts.[47] They insisted not only that legal compliance is possible without observable effects on behavior, as Downs *et al* had argued, but law also can be "effective" – meaning it can prompt states to adjust their practices in line with the law – without the states concerned ever achieving compliance.[48] It follows that focusing on

[42] [Add citations!]
[43] Simmons and Elkins 2008, Elkins, Guzman, and Simmons 2008, Ginsburg, Chernykh and Elkins 2008, Putnam 2009 and 2016.
[44]. Neumayer 2005, Hathaway 2007, Simmons 2009, Lupu 2013b and 2015. See also Merry 2006.
[45] Kahler 2000, Dai 2005, Simmons 2009; Keck and Sikkink 1998, Hertel 2006.
[46] Raustiala 2000:388. See also Raustiala and Slaughter 2002, Howse and Teitel 2010, Martin 2013.
[47]. Note, in much of this volume I refer to "legal rule-following" or law-mediated behavior when seeking to describe or theorize how legal rules may influence how things are done and the choices actors make. This is in order to retain the potential for differences to exist between efforts to follow legal rules and legal compliance.
[48] Raustiala and Slaughter 2002, citing Keohane *et al* 1993, Young 1994. Note that the concept of "effects" is potentially much broader insofar as it can encompass not only consequences *intended by* users of law, but also other consequences (whether helpful or perverse) that were not intended.

compliance can lead researchers to systematically underestimate the impact of law and legal rules on behavior – or even to miss it altogether.[49]

In a still more unwaveringly critical account of compliance as a variable in IR research on international law and institutions entitled "Against Compliance," Lisa Martin echoes earlier points about legal effects and effectiveness and also insists that in the policy world compliance with legal rules is rarely an end in itself.[50] To the contrary, she argues that policy makers are far more concerned with progress toward achieving the substantive goals that international agreements are enacted to address – such as lowering tariffs, securing better treatment of women, reducing incidences of torture, and other pressing needs. Martin contends further that many IR scholars who use the language of "compliance" in their theoretic discussions have actually focused on measuring law's effects and effectiveness in their empirical work—and appropriately so.

Martin dismisses compliance as a concern of lawyers and legal scholars that she describes as "unusually ill-suited to […] the identification and measurement of causal relationships.[51] Furthermore, she states that the problems associated with using compliance as an indicator "cannot be addressed by more careful treatment of compliance; [they] can only be addressed by *dropping compliance as a central concept in the study of institutional effects*" [emphasis added].[52] In a sweeping jab, Martin calls the uptake of compliance in IR research a "cautionary tale" about the dangers of insufficiently considered inter-disciplinarity.[53]

RETHINKING COMPLIANCE

As discussed above, "compliance" can be approached as either a binary indicator *or* as a continuous variable. Whenever a continuous variable is operationalized as a binary indicator, information about the existence, amount, and directionality of change may be lost.[54] However, whether this discarded information is essential to a given explanatory task depends on the question being asked, and on the accompanying research design. As argued above, for legal compliance to be

[49] Raustiala 1998. Martin (2014) adds that a focus on compliance, as opposed to effects, also can filter out change among non-committed states that might be plausibly attributable to treaty regimes operating in a given region or issue domain.
[50] Martin 2014:595-596. After providing several confirming examples, she offers one counter-example of Japanese bureaucrats appearing to value compliance for its own sake.
[51] Martin 2014:591.
[52] Martin 2014:592.
[53] Martin 2014:606.
[54] Note, "effectiveness" can also be made into a binary outcome, though this is not often done in current social science due to the difficulty of defining what "effective" means.

a coherent concept it must incorporate a threshold at which regulated behavior is adequate in light of the law. But it need not be limited to a binary. Nor—and this is important element of the argument—can it be assumed that all relevant actors will agree on where the threshold lies, *particularly in applied cases* and of what its contents consist.

In the litany of identified problems associated with using compliance as a metric, the difficulty of locating compliance thresholds has been the focus of relatively little consideration.[55] This can be attributed to the convention of modeling compliance thresholds either as common knowledge,[56] or as reliably determined by some specific (and often data-friendly) actor or institution.[57] This research practice filters out important interactions both within and across states in which the applicability, scope, and meaning of legal rules are debated. Although this convention is doubtlessly a boon for efforts to model compliance formally and statistically, it does so *at the cost of excising distributive politics as a theme of law-based governance*.

In summary, the ability to locate this threshold is quite clearly *not* (despite Martin's claims to the contrary) irrelevant to the task of assessing the influence of law and legal obligations. Nor, as I argue below, are such efforts on the sidelines of politics. One can, thus, concur with the expressed need for caution when using compliance as an outcome in studies that aim to draw inferences about legal effects, but still reject the scorched earth assessment that legal compliance has no analytic or explanatory use for social science.

So, You Want to Measure Compliance?

To understand why it is necessary to include compliance into explanatory models of legal effects, scholars need to think hard about two things: "law" and "compliance." The first question to ask is: what work (if any) is law, or legal obligation, doing in the theory?

In some situations, the existence, applicability, scope, or meaning of law and norms of legal practice may be in the foreground of an interaction, and thus constitute a contested outcome of interest. At other times, the legality or illegality of conduct may not be in question, but there may still be a need to assess the severity of violations, or the margin by which a state is exceeding its legal

[55] Kingsbury 1998:348. See, however, the work of Tallberg 2002, Elgstrom and Jonsson 2004, Tallberg and Jonsson 2005, Smith 2006, Tallberg and Smith 2014. This excellent body of work, which mainly concerns international trade and aspects of European integration, frames itself as "compliance bargaining." But it uses a relatively narrow definition that focuses mainly on compliance with the judgments and directives of international tribunals. My aim here is far broader.
[56] Morrow 2007, Howse and Teitel 2010.
[57] Simmons 2000, Hafner Burton and Ron 2009, Hill and Moore 2013.

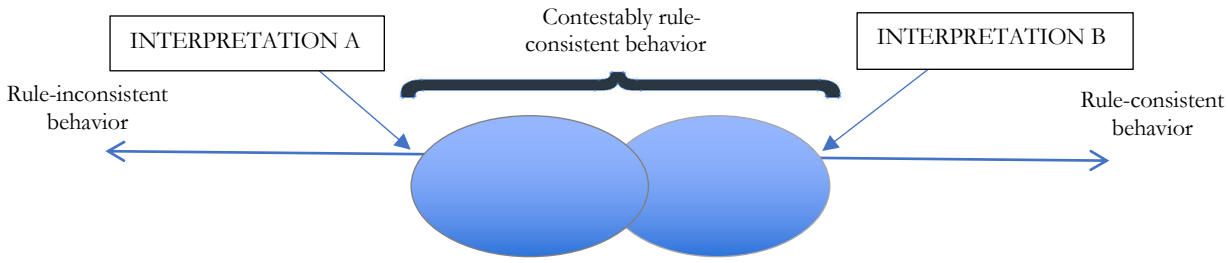
obligations. In each of these variants, the normative expectation that 'legal rules are made to be followed' is in play. Likewise operating in the background is the idea of compliance with the law as a theoretical possibility and as an operative baseline.

Next we need to ask what is meant by "compliance." The simple definition that begins this chapter should now strike the reader as merely trivially accurate. Why? Because laws are drafted using general language, which requires interpretation in light of the messy details that almost always accompany efforts to use law to guide or assess actual behavior. Consequently, different interpreters of law can reach different results about which rules apply and what conduct is allowed or required, even without bad faith or overt biases as a factor. Where legally mediated decisions, or assessments, are not challenged – whether due to lack of awareness or lack of opportunity – they become part of the corpus of (tacitly) accepted practice. The point is not that generalization and prediction are impossible, or that it makes sense to approach every situation is *sui generis*. Rather, it is that when purposeful actors use law as a means to guide or assess behavior, much turns on how interpreters of law configure and weigh the details on which law is applied. Furthermore, this also extends to the identification of operative compliance thresholds.

Chayes and Chayes highlight many of these issues. After overlaying what they call the "ambiguity of treaty language" with variation in states' capacity to do the bidding of legal rules, they observe that "more often than not there will be a considerable range within which parties may reasonably adopt differing positions as to the meaning of [an] obligation."[58] They argue further that compliance is thus most usefully thought of as a "zone" (as distinct from a line) "that is subject to broad variance across regimes, times, and occasions."[59] Acknowledging – and endogenizing – this aspect of rule-based governance is critically important for theory development, since different answers may require different ideas about how international law and legal obligation work (or fail to work) in various settings and domains.[60]

[58] Chayes and Chayes 1993:189. Subsequently Chayes and Chayes (1995) and others in the "managerial school" were drawn more to addressing various problems of compliance incapacity and capacity-building. It is possible the point about the shifting nature of compliance thresholds seemed so self-evident as to not require extensive elaboration.
[59] Chayes and Chayes 1993:189 and 201. For example, within the sphere of arms control, Chayes and Chayes note that the United States appeared to demand strict compliance with the Non-Proliferation Treaty, but took a more relaxed approach toward deviations from the letter of the Anti-Ballistic Missile Treaty. This is bad news for data coding protocols, but good news for fans of bargaining. See Howse and Teitel 2012, Cronin Furman 2020.
[60] Cronin Furman 2020.

Figure 6.1: Legal compliance as a zone of acceptability with variance in rule interpretation



Still, an important lacuna of the Chayes and Chayes argument is its implicit assumption that thresholds for compliance are known to obligated subjects—that a shared (or common) conjecture about what the law requires is in operation. Indeed, much of the "play" in their informal model of compliance involves managing pragmatic risks that violations will be observed and punished. As Figure 6.1 illustrates, however, allowing for more than one plausible conjecture about where operative thresholds for legal compliance are located may be can drastically reduce the range of legally permissible options for coordinating behavior *ex ante*, or assessing it *ex post*.

Further complicating matters is the fact that some legal rules lend themselves to easier assessment of compliance and non-compliance than others, either because they are relatively narrow in scope, or because they contain mainly provisions of a procedural nature.[61] Interpreting international agreements that have a broad substantive scope and few internal (textual or practice-based) benchmarks or guidelines can be considerably more difficult as a technical matter that those that are narrowly tailored and contain specific metrics.[62] Much also depends on the nature and extent of applied practice authoritatively associated with particular legal rules. The higher the variance, the greater the leeway for self-serving forms of discretion (a.k.a. politics) to factor into when and how they are used. I expect this leeway to increase with the level of aggregation at which compliance is being considered.

[61] By analogy, it is easy to identify statutory thresholds for compliance with posted road speed limits. It can far more difficult to operate at, or deviate i measured ways from, the compliance threshold of a legal rule requiring "safe driving."
[62] Abbott and Snidal 2000.

Substance versus procedure

Complex goals – for example, reducing barriers to international trade, stopping the global spread of communicable diseases, or restricting the use of military force in international affairs – cannot be simply commanded into existence. Breaking down complex objective (such as building a suspension bridge across the Hudson River, or reducing global outputs of greenhouse gases) into smaller, workable pieces is key to competent engineering. It is likewise a feature of effective policy design and oversight, insofar as achieving, or progressing toward, collectively defined goals often requires the participation of many actors engaged in a range of inter-connected tasks.

International agreements have a wide variety profiles and features. Agreements that are heavy on substantive goals but light on procedure may reflect a sense that there are many possible pathways toward a shared objective, and many possible sets of starting conditions. [63] Under these circumstances, a loose approach to procedure may be expected to yield betters results than a one-size-fits-all approach, particularly if the latter approach prompts fewer states to join.[64] However, lack of procedural detail also might reflect the inability of enacting states to agree on key aspects of substance.[65] Other agreements may be less clear about their substantive objectives, but contain a high proportion of procedural content.[66] Emphasis on procedures and benchmarks can be useful where the scope or content of a law's objective is difficult to specify in detail, or where there is recognition that progress toward shared substantive goals requires coordinating or sequencing how tasks are carried out. Still others contain a mix of provisions on procedure and substance.

Keeping tabs on compliance with procedures and intermediate benchmarks might appear boring and technical—and also far from larger policy goals. Indeed, critics of compliance as an indicator of legal effects are quick to point out that where laws contain a high proportion of procedural provisions, satisfying these procedures can easily become a goal in its own right, causing those concerned to lose sight of larger objectives. However, in complex undertakings structured by legal rules, the ultimate policy objective(s) may not be known to, or shared by, those actually carrying out the tasks required. Under these conditions the thin normativity of legal obligation, also possibly backed by the threat of sanctions for non-performance, can be a cheap substitute for scads of

[63] Many multilateral human rights agreements have this characteristic.
[64] The question, then, would be whether the decision not to join was due to a substantive commitment to a different pathway to the goal, or to a concern about potentially being assessed as non-compliant.
[65] Drezner 2008:72-85, Steinberg 2004.
[66] Abbot and Snidal 2000. Examples here include extradition treaties, and framework conventions, such as the U.N. Framework Convention on Climate Change.

onerous persuasion. In addition, even where participants are on board with a policy objective, the specific tasks for which individuals (or bureaucratic units) are responsible many not yield, or even visibly contribute to that objective. In such instances, the existence of intermediate compliance benchmarks is essential.

Where reasonable efforts to following legal rules fail to bring about the results envisioned by the law's creators or stewards, a gap may emerge between individual states' records of intermediate (i.e. technical) compliance, and overall assessments of the law's effects and effectiveness.[67] In such situations, the danger misapprehending aggregate results, and drawing flawed inferences about actors' willingness and capacity to fulfill legal obligations may be substantial.

The politics of aggregation and assessment

For an entity to be considered generally "compliant" with international law, or with particular domains of law, does not require strict adherence to every rule all the time.[68] For this reason, Chayes and Chayes insisted that the pragmatic aim of legal rules and processes is to contain "deviance" within reasonable bounds.[69] As a result, they argue, compliance is more helpfully imagined as a "zone" than as a clean threshold.[70] Importantly, where zones of acceptability begin and end also may be influenced by what others typically do (or don't do) in relation to the law—that is, according to associated norms and practices in the setting or domain in question.[71]

Assessing legal compliance at higher levels of aggregation is a less technical exercise, although individual data points may still the be a product of technical operations and assessments. Consequently, a general assessment that a state (or a corporation, or an individual) is on the whole "law-abiding" may say little about its behavior with regard to any specific law or cluster of laws. Even within a given domain, an assessment that, overall, a state observes its international human

---

[67] Buzas 2018.
[68] Downs and Jones 2004.
[69] Chayes and Chayes 1993:197-198. Indeed, not all violations of legal rules and procedures are equal. Some forms or degrees of non-compliance are more acceptable than others, and, likewise, some reasons for failing to adhere to legal rules are more likely to be excused. More substantively, types and magnitudes of harm vary, as do types of victims from failures to follow legal rules.
[70] Risk management, for example, focuses on the relationship between the probability something will occur and the risks it poses if it does occur. Deviance in low probability/low risk situations can be widely tolerated without damaging a regime, as can some kinds of high probability/low risk behavior. However, low probability/high risk actions (like using nuclear weapons) and high probability/high risk acts (like transporting hazardous chemicals) are approached with much different margins of acceptable deviance in legal-regulatory frameworks (Putnam 2006). Indeed legal and policy tools can be used to shape both probability and risk.
[71] According to Chayes and Chayes (1993:198) determinations of compliance and non-compliance "are often contestable and call for complex, subtle, and frequently subjective evaluation."

rights obligations, or its BITS-embedded commitments regarding foreign investment, may say little about the government's efforts or achievements toward ensuring respect for e.g. the right to education, or its willingness to address every complaint regarding allegedly discriminatory taxation.

For example, according to the International Civil Aviation Organization (ICAO) there are (or have been) 57 different multilateral treaties, or formal (requiring separate ratification) treaty modifications in force since 1929 with the broad purpose of ensuring the "sustainability" and safety of international air travel and transport. Many, many factors are relevant to achieving this objective. There are, for example, agreements for ensuring the airworthiness and maintenance of aircraft; sharing international airspace; procedures for making "unscheduled" (emergency) use of airport facilities; guidelines and restrictions for the transport of international cargo; and also legal instruments aimed at preventing the hijacking and weaponization of aircraft.[72] Some of these rules and standards are embodied in stand-alone conventions, whereas others are part of larger covering agreements with detailed annexes.

Despite (or perhaps because of) the sizeable number of agreements comprising the international civil aviation regime, there is no single formula for determining whether individual countries are "in compliance" with international rules. In line with Henkin's edict, one might presume that states are in compliance with most aviation safety rules most of the time – otherwise they would be regulated out of international markets.[73] But what exactly goes into that evaluation, and where operative thresholds of non-compliance are located, can be far from transparent.

Of course, some instances of non-compliance are costlier than others. A single shoddy mandatory pre-flight safety check that leads to an aircraft running out of fuel in the middle of a scheduled route is far more serious than multiple instances in which flights that will not cross any body of water are allowed to depart without adequate numbers of life jackets on board for all passengers and crew. In the first instance non-compliance is both highly consequential, and highly visible; in the second it is not.[74] Airline passengers' assessments of an airline's safety may not give much weight to non-visible instances of technical non-compliance even if they learn of them. But if

[72] 48 are currently in force. https://www.icao.int/secretariat/legal/LEB%20Treaty%20Collection%20Documents/composite_table.pdf.

[73] But see the Good Jobs First Violation Tracker for aviation safety violations cited by the Federal Aviation Association and other U.S. regulators. Between 2000 and 2017, more than 2,500 separate primary violations were recorded; fewer than 100 were attributed to non-U.S. air carriers. https://violationtracker.goodjobsfirst.org/prog.php?major_industry_sum=airlines&order=primary_offense&sort=asc.

[74] http://www.bbc.com/news/world-latin-america-38165757

large numbers of such violations begin to raise alarm among industry experts or consumer watchdogs (perhaps because of what they suggest about the home government's willingness or capacity to exercise oversight), or other governments' regulators issue high fines or threaten licensing suspensions, the violations will no longer appear merely "technical."

With aggregation comes discretion (or room for judgment) with respect to how to organize information, how to weigh, or discount, certain features or gaps, and how to identity relevant patterns or trends.[75] (Of course, there are often "experts" that can be consulted on matters ranging from gritty details to broad trends. But, unless those experts are unified in their opinions, there still may be wiggle room to choose to value some over others if the goal is to tip the evaluation a particular way.[76]) The higher the degree of aggregation, and the broader the legal frame considered, the greater the potential for details to get washed out, and for "good" behavior in one matter to be allowed to compensate for "bad" behavior in others—or vice versa, depending on the more general tenor of the relationship between the assessor and the assessed.[77] In other words, aggregation makes space for politics and political factors to influence assessments. For this reason, whether a government's performance in a particular law-governed domain is evaluated as "acceptably" compliant, and how specific instances of non-compliance are characterized, may reveal as much about an entity's status within relevant political and interpretive communities as about any objective qualities of its practices.[78]

RECLAIMING COMPLIANCE

Where does this leave us with respect to compliance as concept? As a practical matter, governance with legal rules presumes the existence of some set (or sets) of behaviors sufficient to meet requirements established in the law.[79] To ignore compliance as an orienting concept would be to make rule-governed behavior incoherent in settings where coordination is not self-enforcing. It also excises one of the law's more powerful tools for shaping behavior: its ability to force

[75] Legal frameworks incorporate this by specifying ranges of offense (e.g. misdemeanors versus felonies) and ranges of penalty (e.g. due to mitigating or aggravating circumstances).
[76] Jasanoff 1995.
[77] Note also that where a given state is considered to be in compliance with its international human rights obligations, it does not necessarily mean that no violations of the rules ever occur in settings under its jurisdiction. But it does generally mean that rate of violations is relatively low, and that when violations occur they are taken seriously and addressed in a competent manner.
[78] For this reason, scholars should be cautious of measures of reputation crafted from highly aggregated assessments of compliance with legal rules.
[79] Simmons 2009 and 2010. Raustiala and Slaughter 2002.

deliberation in close (or hard) cases over whether legal rules and standards are being met. Finally, even legal "effectiveness" relies on a notion of compliance as an orienting point, and for calibrating and assessing behavioral changes. It follows that some concept of compliance—or more accurately its *potential*—is necessary for theorizing law-mediated behavior.

The critics described above are therefore correct about what compliance is *not* (a reliable indicator of the existence or character of legal effects). But they have failed to appreciate what compliance is: an essential element of rule-following logics and law-based governance. Still, how we think about compliance as a potentiality can (and should) differ in important regards depending on how much and what type of work the threshold is doing in legal and political interactions around rules and rule following. Although this too is obviously a spectrum, thinking about ideal types at the extremes will help to make a few key points.

Easy cases

The more starkly a law-governed act, or set of actions, appears to depart from the law's minimum requirements, the easier it is to assign it a "non-compliance" label. In situations where locating an actual or anticipated act, or pattern of actions, on one side of the compliance binary is simple (meaning it falls near one of the extremes of the spectrum in Figure 1), it is a fair bet that legal obligation is, a best, a marginal consideration to the *author* of the actions. Still, even "easy" classifications of non-compliant behavior can require a reasonable estimate of operative compliance thresholds—if, for example, there is a desire to describe not just the fact of non-compliance, but also its magnitude. Such determinations are quite important to the assignment and authorization of remedies.

Less obviously, there are also implications for governance politics that follow from being unambiguously in compliance with legal obligations under a given rule, statute, treaty, or regime. For some states in some situations, such awareness might generate internal pressure for reductions in socially desirable behavior—for example, where the issue involves making greater-than-required investments in the provision of public goods, or exhibiting greater-than-required forbearance in the use of a common pool resource. For others, awareness of that their own practices substantially exceed those of the legal minimum might prompt efforts to make the collectively recognized threshold for compliance more demanding.[80] Still, only in the first case might it be necessary to

[80] Chayes and Chayes 1993:204

identify a threshold of compliance with any precision (and then only for the purpose of planning and implementing shared downward adjustments).

On the opposite side of the spectrum, it is also possible to easily spot behavior that is obviously not in compliance with the law. For example, in the book's introduction, I describe two incidents where internationally banned chemical weapons were used to indiscriminately massacre unarmed civilians during an armed conflict. The fact that *some* serious breach of compliance with international law occurred was not seriously in doubt. Attributing responsibility and gathering options for possible responses, nevertheless, requires connecting identified wrongs to specific obligations (via requirements for legal necessity and proportionality). Such exercises are not merely academic. How acts are labeled – e.g. as war crimes, weapons treaty violations, crimes against humanity, genocidal acts, or some combination of these – can determine which states and organizations are brought into play, what types of legal permissions for carrying out reprisals are unlocked, and how all of it is processed politically.

A related use of compliance thresholds involves efforts to estimate the severity of non-compliance. Here the existence of compliance binaries serves as a trigger, and as a gauge.[81] All else equal, higher degrees of suspected non-compliance with (or violations of) international legal obligations, might be expected to correlate with greater pressure inside or outside the state for some type of sanction (or nullity in the case of power-conferring rules).[82]

The docket of the International Criminal Court (ICC) offers a case on point. The ICC's limited investigatory and prosecutorial capacity means that in practice only truly egregious offenses are likely to become the focus of proceedings.[83] The process of authoritatively labeling alleged offenses, matching those offenses to specific actors and evidence, and formally weighing defendants' efforts at exculpatory arguments all require that a threshold of appropriateness exists somewhere out there. Although little may be achieved in terms of rule elaboration with regard to the primary charged offenses, such proceedings still can be valuable as an exercise in publicly reinforcing the existence and applicability of relevant laws and norms of practice. Moreover, even in proceedings

[81] Note, however, that once permission has been triggered, switching to a non-binary sense of the magnitude of, and reasons for, the violation in question makes for better responses.
[82] Hart 1994.
[83] Cronin Furman 2013.

before the ICC and other tribunals brought on the basis of fairly clear violations, other, harder issues also can arise.

On the substantive side there can be opportunities for litigants (or their lawyers) to explore the limits of established legal categories of offense by bringing previously unrecognized categories of conduct into consideration. Thus, for example, the International Criminal Tribunals for Yugoslavia and Rwanda authoritatively acknowledged sexual violence as falling within international legal definitions of "war crimes."[84] Also, in the 1940s and beginning again the 1990s international tribunals have wrestled with the boundaries of command responsibility as a doctrine of international criminal law.[85] Efforts to deal with nominally "easy" cases of non-compliance have also forced deliberation and over many legally (and politically) charged questions of international procedural due process – for example, rules of evidence and testimony, along with issues concerning the arrest and detention of defendants, and anonymity rights for witnesses.[86]

"Close" cases and "hard" cases

In situations where law is doing at least some heavy lifting to guide behavior, determinations of compliance and non-compliance can be far more messy and contestable. But what makes a decision "hard" or "close"? And why might the distinction matter for efforts to sort out the legal versus political character of determinations of "acceptable" compliance, or its absence? To call a matter of legal interpretation a "close" decision with respect to compliance implies that the outcome – whether "thumbs up" or "thumbs down" – turns on relatively minor points of fact or circumstance, or on interpretative nuance. The difference, in short, is *technical.*

By contrast, to say a situation presents a "hard" case means that it requires sorting out conflicting legal rules, or making a legally principled decision in light of two or more legally (and perhaps also morally or culturally) incommensurable choices.[87] Decisions that require tradeoffs between laws to protect personal data privacy and those that regulate public security are one example. Efforts to balance economic development rights with environmental rights, or to balance claimed rights in intellectual property with public heath laws provide many others.[88] Making

[84] See the *Tadic, Celebici and Furundzija* decisions involving conduct in former Yugoslavia, and the *Akayesu* decision involving conduct in Rwanda.
[85] Green 1995:178 *et seq.*, Cronin Furman 2013.
[86] DeFrancia 2001, Ambos 2003.
[87] Dworkin 1974.
[88] Newman 2008, Adams 2003, Stiglitz 1999.

tradeoffs of these types can be especially difficult *as a matter of law* (and politics as well) where the rules themselves contain nebulous concepts or difficult to define standards.

Recall that law can shape behavior wherever social beings have choices to make. In decisions that range from how fast to drive and where to cross the street, to whether to order a missile attack on a foreign country's airbase, law can serve as an important (albeit imperfectly dictatorial) decision heuristic. It may matter most in situations where decision makers are cognitively or organizationally overwhelmed by the range of substantive or normative options before them, or simply opt to focus on other things. When backed by a threat of sanctions for non-compliance, assessments when the law applies and what it requires can also help to tip individual-level incentives toward more collectively desirable forms of behavior than would obtain in the law's absence.

In "close" cases and "hard" cases, contestation over where to find the break point between acceptable compliance and non-compliance, can easily begin to feel like bargaining. But, law and legal processes (unlike political ones) require a type of non-arbitrariness by which law-based outcomes must be traceable back to a particular legal rule (or source) and to established processes in a legally discernable (and defensible) way.[89] Still, the general indeterminacy of particular rules, as theorized in Chapter Two, means that what is legally defensible can vary for the same legal rule by setting or by domain if its associated norms of practice differ. At times, therefore, the "closeness" or "hardness" of a determination of legality, or an assessment of compliance, involves deciding among several potentially plausible interpretations of what is allowed, or required that may have roots in different settings or problem structures. In these types of cases, deciding which interpretation of the law is controlling turns importantly on successfully claiming jurisdiction over the conduct or transaction at issue.[90]

Questions of compliance in close or hard cases can be further complicated by "moving goal posts"—the potential for thresholds of compliance acceptability to shift over time as modal interpretations of legal rules and norms of practice change. Such shifts can be upward, as when the practices of leading states consistently exceed what are understood as minimum obligatory levels of compliance). Where this occurs, the behavior of states once comfortably within the zone can come

[89] Fuller 1969, Hadfield and Weingast 2012. Note, however, that legal reasoning, like political reasoning, is not blind to endogeneity. Thus, an entity that tries to cajole an acceptance of compliance sufficiency on the basis of its own relative incapacity is likely to have success only to the extent that the lack of capacity at issue was not of its own making in an effort to extract resources from others, or to cultivate lenience.
[90] Putnam 2009 and 2016.

under critical scrutiny.[91] Collective understandings concerning compliance thresholds also can shift downward where patterns of state practice prompt applied rules and standards to become more permissive, or riddled with allowed exceptions. The tactic of (re)interpreting law to fit behavior – whether one's own or that of others – is tempting where the immediate costs of changing course, or taking steps to ensure that others do so, are politically prohibitive.[92]

Similarly, legal-political concepts like a "clean environment," may have different applied definitions depending on whether it used in a country in the top tier of economic development, or one closer to the bottom. The short-term interests and preferences of the user, which are likely to differ depending upon whether she is a chemical manufacturer, or she heads an environmental CSO, might also be a factor. This type of contextual variance affects vast ranges of concepts mediated by international law. To give just a few additional examples, consider "capital adequacy," "trade barrier," "threatened species," "torture," and "self-defense."

This suggests that "hard" cases and "close" cases may present greater opportunities for various interpretive communities to intervene as effective arbiters of compliance. As explained in Chapter Two, formal authorities are not necessarily coterminous with interpretative communities. Chayes and Chayes observe that "if there is no authoritative arbiter (and even sometimes when there is), discourse among parties, often in the hearing of a wider public audience, is an important way of clarifying the meaning of rules."[93] Indeed, a growing literature on international courts and tribunals suggests that international judicial authorities are often quite aware of this potential disjuncture, and may take measures to ensure that written opinions situate and justify outcomes in terms of broad legal or institutional goals and prior practice.[94]

Furthermore, although estimating at what point *precisely* an entity's observed failure to fulfill some proportion of a law's provisions likely will tip it into non-compliance may be a difficult task, it is a task that lawyers are asked to perform all the time. Indeed, from the perspective of legal practice, every interpretation and application of a legal rule is implicitly a test of a legality hypothesis, the

[91] In Chapter Four I show, using a longitudinal examination of the 1948 Convention for the Prevention and Punishment of Genocide that one cannot simply assume the legal obligations embedded in international treaties remain fixed in time.
[92] See Chapter 4 for a demonstration that even absent formal amendment, the applicability, scope, and meaning of legal rules can shift in light of practice and how relevant interpretative communities interpret and apply them.
[93] Later they add, "What is 'acceptable' in terms of compliance thus may reflect the perspectives and interests of engaged participants in the ongoing political process, rather than some external scientific or market-validated standard." Chayes and Chayes 1993:190 and 202.
[94] Alter 2014, Creamer 2016, Bausch and Pelc 2019.

outcomes of which yield confirmatory or contrary evidence about when and to whom law applies, and what it allows or requires.[95] And, as noted in Chapter Two, the overwhelming bulk of this testing-by-use takes place far removed from the context of any courtroom, and much of it is not especially contentious.

Asking New (and Better) Empirical Questions With, and About, Compliance

At the beginning of this chapter I suggested that rather than thinking about compliance primarily as a metric for rule following, it might be more productively examined as a way to describe structures of expectation around law and legal rule following. (Indeed, one way to think about legal "violations" is rule-following that was expected, but failed to happen.) I will now elaborate this idea a bit more as a path to articulating questions for the future research on legal compliance and its associated politics that have been enabled by the analysis in this and prior chapters.

One important takeaway from the analysis is that "compliance" operates as a narrowly legal concept only when it is at its most micro and technical. If we imagine (not implausibly) that all legal rules have a bit of haziness at the margins concerning when or to whom they apply, with respect to what scope of activity, and with what requirements or permissions, fissures of indeterminacy appear. Some of all of these fissures may then become magnified through the interpretation and application of the rule in concrete settings. And, through these cracks all manner of distributional concerns, ideological preferences, and other sources of politically laced variance can begin to seep in.[96] If, in addition, a need arises to somehow aggregate evidence on rule following to form general assessments, or to balance the imperatives or allowances of one rule against those of other rules – suddenly we may find ourselves off to the races with politics shifting and buffeting what had previously been assumed to be a fairly neutral exercise in evaluation.

Efforts to achieve compliance with individual laws are measurable only if thresholds of acceptability are identifiable (at least in theory) at whatever level of aggregation the assessment game is being played. The potential for legal and political understandings of compliance to differ – or, more accurately, for some assessments of compliance to contain a higher proportion of political versus legal-technical inputs – means that there can be space for bargaining. All else equal, we might

[95] Putnam 2016.
[96] Shapiro and Stone Sweet 2002, Friedman 2006.

expect this space to be wider in situations in which the outcome turns on standards, as opposed to rules, and where it turns on substance versus procedure.

Still, one might protest that notwithstanding the often quirky pathways that lead litigants to international courts, or to high-level domestic courts, the fact that authoritative statements of international legal rules are surely important narrowing the range of discretion for how compliance will be assessed going forward. Maybe. But maybe not. For example, state decisions to adopt and implement the rulings of the WTO's dispute settlement mechanism, and in particular its Appellate Body, are often presumed to carry substantial weight. But, the WTO is very high capacity tribunal. Not only does it have compulsory jurisdiction over all treaty member states, the international trade regime itself is of sufficiently great value to its leading members to discourage credible threats of exit. Even here, however, it is far from guaranteed that Appellate Body decisions will be taken up and implemented in ways that escape contestation among interpretative communication within or outside the international trade domain.[97]

When attempting to ask (and answer) research questions about compliance, it helps to be clear about whether the primary objective is to study rule following and behavioral expectations around legal obligation, or to track progress toward some substantive goal that the existence of legal rules is expected to facilitate (or complicate). Each requires a different approach.

If a researcher's main area of concern is general "cooperativeness" within international legal frameworks, irrespective of effects, then technical compliance is a perfectly adequate indicator. If, however, the main focus of interest is whether and how governments or other subjects of international law change their attitudes or behavior in accordance with what legal rules require, then some concept of compliance as an orienting point is still necessary as means to tame the wide-ranging potentialities of law's indeterminacy. Without it even the standard fallback – legal "effectiveness" – which implies measurable movement toward a goal, loses some of its focus and punch. Of, course one might choose to try to anchor a metric for evaluation in the substantive policy aims of the law (to the extant such metrics can be specified). But, doing so puts a fairly maximalist burden on treaty members and other collaborators. For example, if the point of a treaty is to mitigate global warming, a failure to actually achieve the policy goal is a high bar for legal

[97] Moreover, as I argue in Chapter Five, among longtime members of the organization, the overwhelming bulk of claims initiated primarily concern fairly "hard" or "close" questions—unlike claims that are primarily about enforcement, which are generally "easy." Hard and close cases generally generate contestable outcomes.

compliance, as compared to reaching, or adhering to, more graduated steps toward that goal. Having the possibility of being in compliance as an attainable objective allows states to cooperate to do *something*, without the pressure or burden to do *everything*.[98]

So, what might embracing this fuzzier, more malleable, and more variable understanding of compliance require in terms of research design and data? In the spirit of King, Keohane, and Verba's injunction to first think big and then accommodate creatively, I offer a few preliminary thoughts. One essential thing to keep in mind is that instead of just two moving parts of interest (law-governed behavior, observable and unobservable outcomes), there is now a third: *a potentially shifting zone of legal and political acceptability*.

In order to be able to ask better observational questions about compliance and its associated politics, attention to temporal baselines is critical across the board. Obviously, reliable longitudinal data are needed for the law-governed practices of interest themselves, or reasonable proxies. Also needed is appropriately detailed information about the political and economic settings where law and legally constituted agents operate. (The broad country-level measures that many large-N studies of selection into treaties rely upon to control for things like economic factors, regime type, along with binary indicators for armed conflict or legal traditions, simply may be too coarse to capture the effects of rule following.)

A third desideratum is data about patterns of assessed compliance by country (or specialized sub-state institution) and assessor(s). In the human rights sphere, for example, annual U.S. State Department Reports, often cross-checked with reports from organizations like Amnesty International and Human Rights Watch, have long served this objective. The relatively recent rise of international ratings and rankings across a wide range of topics and settings may help to provide additional data for scholars as additional sources and layers of assessment. That said, the principle of digging into the sources and methods of all such measures to estimate the reliability, and to understand any quirks and blind spots is critically important to the integrity of such exercises.[99]

For qualitative and large-N statistical work alike, data are needed in each of these areas to ensure that change on one or more dimensions is not masking change or stasis in others. For example, evidence that a country's observed practices on the protection of endangered elephant

[98] Both of these questions differ from narrower policy questions about whether a given law leads to specific (or better) results than a different law, or no law at all.
[99] Cooley and Snyder 2015, Clark and Sikkink 2014, Andreas and Greenhill 2011.

habitats are moving from a more collectively satisfactory position within the zone of contestable compliance to a less satisfactory position, or, alternatively that its practices have crossed the lower bound into non-compliance, is not necessarily evidence that legal rules are not influencing behavior in a positive way. Why might this be so?

One possibility is that circumstances or projections about future conditions that prevailed when a state made its international legal commitment may have changed in unanticipated ways.[100] Where those changes are negative – a massive drought may have hit, or a war may have started – the state's behavior may worsen by the substantive measures. In some instances this may tip a state previously assessed as compliant into non-compliance. Alternatively, it may narrow a state's degree of over-compliance, or arrest (or reverse) a non-compliant state's prior progress toward compliance. But, it still may be the case that subsequent damage would have been worse in the absence of the law, and the domestic legal or institutional follow-through it prompted, or reinforced, prior to the change in circumstances. As such, some ability to describe or estimate this type of counter-factual is a necessary part of an explanatory account of legal effects.

Consider also that an absence of changes in behavior is not necessarily evidence that legal rules are having an effect if shared understandings of when the legal rules apply, or what they entail, are changing through subsequent patterns of invocation and use. Where legal expectations are becoming more demanding (or shifting the zone of contestation to the right in Figure 6.1), a state's failure to change its practices in step with this shift would detract from its degree of compliance, all else equal. Thus approaches that assume fixed understandings of what legal rules entail may systematically overlook the potential for international legal commitments that may seem shallow and superfluous at one point in time to grow teeth.

Finally, even if underlying circumstances stay the same, and the rules stay the same, the width or narrowness of the operative zone of compliance may vary. On the upper end of the spectrum, where a state is systematically "over-complying" – doing far more than the minimum needed to comply with the law – this may prompt directed efforts to shift the ultimate objective of the law upward to reflect the state's actually level of performance. Or it may prompt efforts to shift or narrow the zone of acceptance within extant understandings of the rule or standard – i.e. the "moving goal post" issue discussed above.[101] Alternatively, in circumstances where legal rules appear

[100] This is known more generically as a "negative time inconsistency" problem.
[101] Chayes and Chayes 1993.

too stringent, in order to avoid the label of widespread non-compliance, which might threaten the viability of a treaty or regime, exceptions of a standing or temporary nature may be allowed.[102]

*

The ideas and criticisms laid out in this chapter suggest several potentially fruitful questions by which to explore compliance as a feature of law. First, how are compliance thresholds defined in practice? How and why do they change, and when are they sticky? Not only are aggregate compliance thresholds difficult to estimate with any precision, they also cannot be counted on to remain fixed. This is so even if (for the sake of argument) underlying behavior were to stay roughly constant. Available information about that behavior may change, the environment in which the behavior occurs may change, the composition of interpretative communities and associated standards may change, and actors affected by evaluations of compliance and non-compliance may attempt to alter any or all of these factors in a bid to alter those evaluations.[103]

Second, how do interpretive communities form around particular laws and legal issues? Under what kinds of conditions do they fluctuate in terms of size, diversity, or intensity of activity? Work on compliance constituencies, epistemic communities, and advocacy strategies and networks is all relevant here. Law-governed activities that are international, or transnational, whether by nature or by choice, can build up broad and potentially diverse interpretive communities. Also relevant is work that aims to understand for what kinds of issues, and among what kinds of actors assessments of compliance and non-compliance function as status, or market swaying indicators – either in isolation, or as part of some more complicated ranking or valuation exercise.[104]

Third, under what conditions might the outcomes of bargaining over (or around) compliance affect shared understandings of a law's applicability or meaning *in practice* even if the outcome turns on something other than legal reasoning? In some instances tracking the substantive and procedural breaking points of "compliance" may be a useful (if approximate) barometer of politically "safe" practices, or, alternatively as a gauge of varying political pressure and shifting standards of legal acceptability.

[102] Hafner Burton *et al* 2011, Pelc 2016, Carnegie and Carson 2018..
[103] Risse, Ropp, and Sikkink 2013, Cooley and Snyder 2015, Kelley and Simmons 2015.
[104] Ruggie 2004, Vogel 2005, Cooley and Snyder 2015.

Fourth, under what conditions do gaps emerge between the letter and the spirit of international legal requirements, and when and how are those gaps reinforced or undermined by individual (i.e. 'state to state,' or 'non-state actor to state') or collective assertions of non-compliance – or their absence? Here, and with respect to the prior question, there is already some excellent work in the domain of transitional justice and bargaining among international *demandeurs* and reticent domestic governments for legal accountability in the aftermath of mass post-atrocities.[105]

CONCLUSION

Compliance – particularly in its purely binary form – may indeed leave much to be desired as a stand-alone indicator of legal effects and effectiveness in causal analysis. Even so, this does not mean that compliance has no meaningful place in efforts to theorize, and to account for, patterns of legally mediated behavior. Insisting on such a position requires overlooking critically important, and politically consequential features of rule-governed behavior.

This chapter argues that it is not possible to use law to guide behavior, or as a metric to assess it, without at least some working hypothesis for what it would take to comply with the legal rules in play. Generating such working hypotheses is, in effect, an exercise in formulating expectations about actors' applied interpretations of legal rules. Of course having firm ideas about what the law requires, or anticipating what others think the law requires, is often only one factor in more complex explanatory equations. Even so, where researchers want to test for the causal influence of law and legal obligations, getting this right is important.

Another key takeaway is that conventional approaches to theorizing compliance with international legal rules systematically excise this variance by means of convenient but deeply inaccurate assumptions that treat legal rules as exogenously given, linguistically determinate, and with fixed meanings.

The implications for how researchers think about compliance are obvious and important: If patterns of legal practice vary in terms of when and how they understand specific laws to apply, or what they permit or require, then assuming (or imposing) a single 'bright line' threshold for what counts as compliance fails to reasonably describe the object of study. Moreover, if questions about legal compliance in the narrow (lawyers') sense appear hopelessly remote from the interests of

[105] Cronin Furman 2020.

causally oriented social science, then this is because causally oriented social scientists have not bothered to look closely enough at the processes by which legal compliance thresholds take shape and change to appreciate their thoroughly political character.

Attention to variance claims about legal compliance, and what they reveal about what is, and is not, socially or politically tolerable to different actors and interpretative communities, can be informative on a wide range of questions. In short, scholars should take care not to confuse weak efforts to describe and operationalize compliance with its value as a concept.

Ultimately, when states and other subjects of international law act in good faith in efforts to carry out their legal obligations, it can nevertheless occur that the substantive policy objectives behind those obligations fail to bear fruit. This may be even more likely when good faith cannot be assumed. In either situation there may be gaps between what I have called aggregate levels of compliance, and more technical levels of compliance. It follows that scholars should aim to be clear about whether only one type compliance, or both, is the focus of empirical study.[106]

REFERENCED WORK

Abbott, Kenneth W, Robert O Keohane, Andrew Moravcsik, Anne-Marie Slaughter, and Duncan Snidal. 2000. "The Concept of Legalization." *Power and Governance in a Partially Globalized World* 132-51.

Abbott, Kenneth W., and Duncan Snidal. 2000. "Hard and soft law in international governance." *International organization* 54.3: 421-456.

Alter, Karen J. 2014. *The New Terrain of International Law: Courts, Politics, Rights* Princeton University Press.

Ambos, Kai. 2003. "International criminal procedure:" adversarial"," inquisitorial" or mixed?." *International Criminal Law Review* 3.1 1-37.

Andreas, Peter, and Kelly M. Greenhill, eds. 2011. *Sex, drugs, and body counts: The politics of numbers in global crime and conflict*. Cornell University Press.

Baccini, Leonardo, and Johannes Urpelainen. 2014. "International institutions and domestic politics: can preferential trading agreements help leaders promote economic reform?." *The Journal of Politics* 76.1: 195-214.

Barnett, Michael, and Raymond Duvall, eds. 2004. *Power in Global Governance*. Cambridge University Press.

Boyle, Alan E. 1999. "Some Reflections on the Relationship of Treaties and Soft Law." *International & Comparative Law Quarterly* 48.4: 901-913.

Brewster, Rachel. 2009. "The limits of Reputation on Compliance." *International Theory* 1.2: 323-333.

Brummer, Chris. 2010. "Why Soft Law Dominates International Finance—and Not Trade." *Journal of International Economic Law* 13.3: 623-643.

Busch, Marc L., and Krzysztof J. Pelc. 2019. "Words Matter: How WTO Rulings Handle Controversy." *International Studies Quarterly* 63.3: 464-476.

e

Büthe, Tim, and Helen V. Milner. 2008. "The politics of foreign direct investment into developing countries: increasing FDI through international trade agreements?." *American journal of political science* 52.4: 741-762.

Búzás, Zoltán I. 2018. "Is the Good News About Law Compliance Good News About Norm Compliance? The Case of Racial Equality." *International Organization* 72.2: 351-385.

Carnegie, Allison, and Austin Carson. 2018. "The Spotlight's Harsh Glare: Rethinking Publicity and International Order." *International Organization* 72.3: 627-657.

Chayes, Abram, and Antonia Handler Chayes. 1995. *The New Sovereignty*. Harvard University Press.

Chayes, Abram, and Antonia Handler Chayes. 1993. "On Compliance." *International organization* 47:2, 175-205.

Clark, Ann Marie, and Kathryn Sikkink. 2013. "Information effects and human rights data: Is the good news about increased human rights information bad news for human rights measures?." *Human Rights Quarterly* 35:3, 539-568.

Cole, Wade M. 2005. "Sovereignty relinquished? Explaining commitment to the international human rights covenants, 1966-1999." *American Sociological Review* 70.3: 472-495.

Cooley, Alexander, and Jack Snyder, eds. 2015. *Ranking the World*. Cambridge University Press.

Cronin-Furman, Kate. 2020. "Human Rights Half Measures: Avoiding Accountability in Postwar Sri Lanka." *World Politics* 72.1: 121-163.

Cronin-Furman, Kate. 2013. "Managing expectations: international criminal trials and the prospects for deterrence of mass atrocity." *International Journal of Transitional Justice* 7:3, 434-454.

Dai, Xinyuan. 2005. "Why comply? The domestic constituency mechanism." *International Organization* 59.2: 363-398.

DeFrancia, Cristian. "Due process in international criminal courts: why procedure matters." *Virginia Law Review* (2001): 1381-1439.

Downs, George W., and Michael A. Jones. 2002. "Reputation, compliance, and international law." *The Journal of Legal Studies*31.S1: S95-S114.

Downs, George W., David M. Rocke, and Peter N. Barsoom. 1996. "Is the good news about compliance good news about cooperation?." *International Organization* 50.3: 379-406

Drezner, Daniel W. 2008. *All politics is global: Explaining international regulatory regimes*. Princeton University Press.

Dupuy, Pierre-Marie.1990. "Soft law and the international law of the environment." *Mich. J. Int'l L.* 12: 420.

Dworkin, Ronald. 1974. "Hard Cases." *Harv. L. Rev.* 88: 1057.

Elkins, Zachary, Andrew T. Guzman, and Beth A. Simmons. 2006. "Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960–2000." *International Organization* 60.4: 811-846.

Fearon, James D. 1998. "Bargaining, enforcement, and international cooperation." *International organization* 52.2: 269-305.

Finnemore, Martha, and Kathryn Sikkink. 1998. "International norm dynamics and political change." *International organization* 52.4: 887-917.

Finnemore, Martha, and Stephen J. Toope. 2001. "Alternatives to "Legalization": Richer Views of Law and Politics." *International Organization* 55.3: 743-758.

Forester, John. 1984. "Bounded Rationality and the Politics of Muddling Through." *Public Administration Review*: 23-31.

Fortna, Virginia Page. 2003. "Scraps of paper? Agreements and the durability of peace." *International Organization* 57.2: 337-372.

Friedman, Barry. 2006. "Taking Law Seriously." *Perspectives on Politics* 4.2, 261-276.

Fuller, Lon L. *The Morality of Law*. Vol. 152. Yale University Press, 1969.

Ginsburg, Tom, Svitlana Chernykh, and Zachary Elkins. 2008. "Commitment and diffusion: How and why national constitutions incorporate international law." *U. Ill. L. Rev.*: 201.

Goodman, Ryan, and Derek Jinks. 2003. "Measuring the effects of human rights treaties." *European Journal of International Law* 14.1: 171-183

Green, Leslie C. 1995."Command Responsibility in International Humanitarian Law." *Transnat'l L. & Contemp. Probs.* 5: 319.

Guzman, Andrew T. 2008. *How International Law Works: A Rational Choice Theory*. Oxford University Press.

Guzman, Andrew T. 2005. "Reputation and International Law." *Ga. J. Int'l & Comp. L.* 34: 379.
Guzman, Andrew T. 2002. "A Compliance-Based Theory of International Law." *Calif. L. Rev.* 90: 1823.
Guzman, Andrew T., and Timothy L. Meyer. 2010. "International Soft Law." *Journal of Legal Analysis* 2.1: 171-225.
Hadfield, Gillian, and Barry R. Weingast. 2013. "Microfoundations of the Rule of Law." *SSRN Electronic Journal SSRN Journal.*
Hafner-Burton, Emilie M. 2011. *Forced to be good: Why trade agreements boost human rights*. Cornell University Press.
Hafner-Burton, Emilie M., Laurence R. Helfer, and Christopher J. Fariss. 2011. "Emergency and escape: explaining derogations from human rights treaties." *International Organization* 65.4: 673-707.
Hafner-Burton, Emilie M., and James Ron. 2009. "Seeing double: Human rights impact through qualitative and quantitative eyes." *World Politics* 61.2: 360-401.
Hafner-Burton, Emilie M., and Kiyoteru Tsutsui. 2007. "Justice lost! The failure of international human rights law to matter where needed most." *Journal of Peace Research* 44.4: 407-425.
Hafner-Burton, Emilie M., and Kiyoteru Tsutsui. 2005. "Human rights in a globalizing world: The paradox of empty promises." *American journal of sociology* 110.5: 1373-1411.
Hart, Herbert Lionel Adolphus, and Leslie Green. 2012. *The concept of law*. Oxford University Press.Howse, Robert, and Ruti Teitel. 2010. "Beyond compliance: rethinking Why International law really matters." *Global Policy* 1.2: 127-136.
Hathaway, Oona A. 2002. "Do human rights treaties make a difference?." *The Yale Law Journal* 111.8: 1935-2042
Henkin, Louis. 1979. *How Nations Behave: Law and Foreign Policy*. Columbia University Press.
Hertel, Shareen. 2006. *Unexpected Power: Conflict and Change Among Transnational Activists*. Cornell University Press.
Higgins, Rosalyn. 1994. *Problems and Process: International Law and How We Use It*. Oxford University Press.
Hill, Daniel W. 2010. "Estimating the effects of human rights treaties on state behavior." *The Journal of Politics* 72:04, 1161-1174.
Jasanoff, Sheila. 1995. "Procedural Choices in Regulatory Science." *Technology in Society* 17.3: 279-293.
Jervis, Robert. 1999. "Realism, Neoliberalism, and Cooperation: Understanding the Debate." *International Security* 24.1: 42-63.
Kelley, Judith G. 2017. *Scorecard diplomacy: Grading states to influence their reputation and behavior*. Cambridge University Press.
Kelley, Judith. 2007. "Who Keeps International Commitments and Why? The International Criminal Court and Bilateral Nonsurrender Agreements." *American Political Science Review* 101.3: 573-589.
Keohane, Robert O. 1996. "International Relations and International Law: Two Optics." *Harv. Int'l. LJ* 38: 487.
Keohane, Robert O. 1984. *After Hegemony Cooperation and Discord in the World Political Economy,* Princeton.
Kingsbury, Benedict. 1997. "The concept of compliance as a function of competing conceptions of international law." *Mich. J. Int'l L.* 19: 345.
Klabbers, Jan. 1998. "The Undesirability of Soft Law." *Nordic Journal of International Law* 67: 381-391.
Klabbers, Jan. 1996. *The Concept of Treaty in International Law*. Vol. 22. Martinus Nijhoff Publishers.
Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "The rational design of international institutions." *International organization* 55.4: 761-799.
Krasner, Stephen D. 2002. "Realist views of international law." *Proceedings of the ASIL Annual Meeting*. Vol. 96. Cambridge University Press, 265-268.
Krasner, Stephen D. 1982. *International Regimes*. Cornell University Press, 1983.
Kydd, Andrew. 2009. "Reputation and cooperation: Guzman on international law." *International Theory* 1.2: 295-305.
Lindblom, Charles E. 1959. "The science of" muddling through"." *Public Administration Review*: 79-88.
Lupu, Yonatan. 2015. "Legislative veto players and the effects of international human rights agreements." *American Journal of Political Science* 59.3: 578-594.

Lupu, Yonatan. 2013a. "Best evidence: The role of information in domestic judicial enforcement of international human rights agreements." *International Organization* 67.3: 469-503

Lutz, Ellen L., and Kathryn Sikkink. 2000. "International human rights law and practice in Latin America." *International Organization* 54.3: 633-659.

Martin, Lisa. 2013. "Against Compliance" in *Interdisciplinary perspectives on international law and international relations: the state of the art*. (Dunoff, Jeffrey L., and Mark A. Pollack, eds.) Cambridge University Press, 591-610

Martin, Lisa L. 2005. "The president and international commitments: Treaties as signaling devices." *Presidential Studies Quarterly* 35.3: 440-465.

Mearsheimer, John J. 1994. "The False Promise of International Institutions." *International Security* 19.3: 5-49.

Merry, Sally Engle. 2009. *Human rights and gender violence: Translating international law into local justice*. University of Chicago Press.

Mitchell, Neil J., and James M. McCormick. 1988. "Economic and political explanations of human rights violations." *World politics* 40.4: 476-498.

Morrow, James D. 2007. "When do states follow the laws of war?." *American Political Science Review* 101.3: 559-572.

Neumayer, Eric, and Laura Spess. 2005. "Do bilateral investment treaties increase foreign direct investment to developing countries?." *World development* 33.10: 1567-1585.

Newman, Abraham. 2008. *Protectors of privacy: Regulating personal data in the global economy*. Cornell University Press.

Pelc, Krzysztof J. 2016. *Making and bending international rules: The design of exceptions and escape clauses in trade law*. Cambridge University Press.

Poe, Steven C., and C. Neal Tate. 1994. "Repression of human rights to personal integrity in the 1980s: A global analysis." *American Political Science Review* 88.4: 853-872.

Putnam, Tonya L. 2016. *Courts without borders: Law, politics, and US extraterritoriality*. Cambridge University Press.

Putnam, Tonya L. 2009. "Courts without borders: domestic sources of US extraterritoriality in the regulatory sphere." *International Organization* 63.3: 459-490.

Raustiala, Kal. 2000. "Compliance & (and) Effectiveness in International Regulatory Cooperation." *Case W. Res. J. Int'l L.* 32: 387.

Raustiala, Kal, and Anne-Marie Slaughter. 2002. "International Law, International Relations and Compliance" in *Handbook of International Relations*, Walter Carlsnaes, Thomas Risse, and Beth A. Simmons, eds, Sage Publishers.

Raz, Joseph. 2009. *Between authority and interpretation: On the theory of law and practical reason*. OUP.

Risse, Thomas, Stephen C. Ropp and Kathryn Sikkink, eds. 2013. *The persistent power of human rights: from commitment to compliance*. Vol. 126. Cambridge University Press.

Shaffer, Gregory, and Tom Ginsburg. 2012. "The empirical turn in international legal scholarship." *American Journal of International Law* 106.1: 1-46.

Shaffer, Gregory C., and Mark A. Pollack. 2009. "Hard vs. Soft Law: Alternatives, Complements, and Antagonists in International Governance." *Minn. L. Rev.* 94: 706.

Shapiro, Martin, and Alec Stone Sweet. 2002. *On Law, Politics, and Judicialization*. Oxford University Press.

Simmons, Beth. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13: 273-296.

Simmons, Beth A. 2009. *Mobilizing for human rights: international law in domestic politics*. Cambridge University Press.

Simmons, Beth A. 2000. "International law and state behavior: Commitment and compliance in international monetary affairs." *American Political Science Review* 94.4: 819-835.

Simmons, Beth A. 1998. "Compliance with international agreements." *Annual review of political science* 1.1: 75-93.

Simmons, Beth A., and Daniel J. Hopkins. 2005. "The constraining power of international treaties: Theory and methods." *American Political Science Review* 99:4, 623-631.

Simon, Herbert A. 1972. "Theories of Bounded Rationality." *Decision and Organization* 1.1: 161-176.

Snidal, Duncan. 1985. "Coordination versus prisoners' dilemma: Implications for international cooperation and regimes." *American Political Science Review* 79.4: 923-942.

Steinberg, Richard H. 2004. "Judicial lawmaking at the WTO: Discursive, constitutional, and political

constraints." *American Journal of International Law* 98.2: 247-275.
Tallberg, Jonas. 2002. "Paths to compliance: Enforcement, management, and the European Union." *International organization* 56.3: 609-643.
Tallberg, Jonas, and James McCall Smith. 2014. "Dispute settlement in world politics: States, supranational prosecutors, and compliance." *European Journal of International Relations* 20.1: 118-144.
Tallberg, Jonas, and Christer Jönsson. 2005. "Compliance Bargaining." *European Union Negotiations: Processes, Networks, and Institutions*.
Tobin, Jennifer L., and Marc L. Busch. 2010. "A BIT is Better Than A Lot: Bilateral Investment Treaties and Preferential Trade Agreements." *World Politics* 62.1: 1-42.
Vogel, David. 2007. *The market for virtue: The potential and limits of corporate social responsibility*. Brookings Institution Press.
Von Stein, Jana. 2005. "Do treaties constrain or screen? Selection bias and treaty compliance." *American Political Science Review* 99.4: 611-622.
Waltz, Kenneth N. 1979. "Theory of International Politics". Long Grove.
Wendt, Alexander. 1992. "Anarchy is What States Make of It: The Social Construction of Power Politics." *International Organization* 46.2: 391-425.
Yackee, Jason Webb. 2008. "Bilateral Investment Treaties, Credible Commitment, and The Rule of (International) Law: Do BITs Promote Foreign Direct Investment?." *Law & Society Review* 42.4: 805-832.
Young, Oran R. 1989. *International cooperation: Building Regimes for NaturalRresources and the Environment*. Cornell University Press.