Why Do Some International Institutions Contain Strong Dispute Settlement Provisions? Evidence from Preferential Trade Agreements

Todd Allee
Department of Government and Politics
University of Maryland
tallee@umd.edu

Manfred Elsig
World Trade Institute
University of Bern
manfred.elsig@wti.org

Paper presented at the
7th Annual Conference on the Political Economy of International Organizations
Princeton University
January 16-18, 2014

Introduction and Overview

Dispute settlement bodies (DSBs) play a crucial role across contemporary international organizations (IOs) ranging from the United Nations system to the World Trade Organization to regional organizations. These institutions – which take the form of courts, tribunals, and arbitration venues, among others – can clarify standards and obligations, resolve disputes between states as well as other actors, punish violations of international law, and enforce state obligations. In doing so, they have the ability to fundamentally recast interactions among even powerful states and alter the behavior of major world leaders.

Examples of the importance of these bodies abound. The European Court of Justice (ECJ) has been portrayed by observers as a major driving force behind European integration (*e.g.*, Stone Sweet 2004; Stone Sweet and Brunell 1999). In fact, Karen Alter has called the ECJ "…perhaps the most active and influential international legal body in existence" (2000: 491). Likewise, the dispute settlement mechanism (DSM) of the World Trade Organization (WTO) was recently called the "pride of the WTO" by one notable observer, who echoed a common sentiment that legal dispute settlement in IOs like the WTO "gives every member, big or small, a platform and a voice."[1] International dispute bodies can also have important distributional consequences which can lead to attempts to change or exit the organization. For instance, the International Centre for the Settlement of Investment Dispute (ICSID) has faced recent charges of being predisposed in favor of northern investors, leading countries such as Venezuela, Bolivia, and Ecuador to withdraw from ICSID's jurisdiction. Finally, for individual leaders, the emergence of the International Criminal Court (ICC), with its power to indict past and present state leaders for war crimes, has led sitting presidents such as Sudan's Omar al-Bashir

[1] Jagdish Bhagwati, "The Broken Legs of Global Trade." *The American Interest*. June 1, 2012. Available at: blos.the-american-interest.com/bhagwait/2012/06/01/the-broken-legs-of-global-trade/ (accessed August 26, 2013).

dramatically alter their behavior, such as traveling only to countries that will not turn them over for prosecution.

Yet not all international institutions – be they treaties, conventions, or formal organizations – are accompanied by such noteworthy legal dispute settlement bodies. The formal dispute settlement mechanisms of the International Labor Organization, for instance, go virtually unnoticed. Other international organizations rely primarily on informal, diplomatic mechanisms of dispute settlement calling for conciliation, mediation and good offices to address disputes, with the Association of Southeast Asian Nations (ASEAN) being a prominent example. Finally, many international organizations contain no explicit dispute settlement mechanism at all.

In sum, contemporary international organizations exhibit considerable variation in terms of what they specify can be done in the event of a dispute among members. The story, in fact, is much richer than has been appreciated to date. The question is not just whether an international organization has a dispute settlement body (Koremenos 2007). Rather, one should consider important questions such as: how much authority is delegated to judges and arbiters, do aggrieved parties have the full ability to pursue their claims, and can sanctions be used to implement awards? These questions can be translated into important new variables that more fully capture characteristics of dispute settlement across international institutions. They also can serve as subcomponents of an overall index of how "strong" dispute settlement is within any given IO. Our primary goal in this paper, therefore, is to explain why some international institutions have stronger dispute settlement than others. Among the contributions we make while pursuing this goal are: 1) to produce a richer typology of features of dispute settlement mechanisms, and 2) to present original data on these features across nearly 600 international institutions whose DSMs vary in multiple ways.

Several theoretical literatures, which for the most part have not been translated into quantitative tests, provide a valuable departure point for understanding this underappreciated variation (in dispute settlement mechanism). Most notable are more formalized arguments about the need for "deep" cooperative agreements to be backed by strong, legal enforcement (*e.g.*, Downs, Rocke, and Barsoom 1996; Yarbrough and Yarbrough 1997). According to this line of argument, powerful dispute settlement bodies, when they exist, will reflect features of the founding treaties or agreements, including most notably their depth but also the number of parties to the agreement (e.g., Koremenos et. al. 2001). Although persuasive, these ideas are mostly backed by selected case studies or very simple quantitative tests, if at all (Duffield 2003). Several literatures from international law also are relevant. One is the debate over the relative influence of state power vis-à-vis prominent international legal institutions (*e.g.*, Brewster 2006; Helfer and Slaughter 1997; Goldsmith and Posner 2005). From this perspective, one might conclude that dispute settlement design will reflect the preferences of (powerful) states that are part of an international institution, and the relative bargaining power among them. Also relevant are long-standing ideas about different regional attitudes toward "hard" law, which we ultimately find to be more nuanced and surprising than prevailing stereotypes might suggest.

To gain insights into how states design international dispute settlement mechanisms, and why some are more powerful than others, we examine variation across preferential trade agreements (PTAs) in terms of their dispute settlement rules and the overall strength of dispute settlement provisions. PTAs possess several desirable features: they are numerous, politically relevant, and vary widely across values of possible outcomes (DS provisions) and explanatory variables (agreement depth, number of signatories, countries, and regions). Most important for us is the fact that they contain multi-faceted dispute settlement procedures, which allows us to

compile rich data on the various dispute settlement features that each PTA contains. The DSMs within many agreements have been utilized far more frequently than is commonly understood, too.[2] Thus we believe it is important to study dispute settlement design and that the lessons drawn from our investigation of PTAs will be directly applicable to formal IOs as well as less formal international institutions.

Empirical tests using our original data produce several novel findings. Most notably, we find that agreement depth is a robust, positive predictor of strong rules for dispute settlement. Agreements that ask states to change policies the most – in our case to liberalize trade – provide substantial market access and harmonize trade rules, also give them powerful tools to enforce those obligations. Other agreement features that might increase the need for stronger enforcement, such as larger memberships and plurilateral agreements, also are positive predictors of stronger dispute settlement rules. Yet state and regional preferences also play a large role. PTAs signed in the Americas, particularly those to which the US is a party, contain stronger dispute settlement rules. But this is not true for European and EU agreements, which until recently had relatively weak dispute settlement. Surprisingly, agreements among Asian partners appear if anything to be more, and not less, likely to contain stronger dispute settlement rules, contradicting widespread perceptions.

The remainder of the paper proceeds as follows. We begin by surveying literatures on international dispute settlement bodies, with an eye toward approaches and findings that might help us understand dispute settlement variation in PTAs. We then discuss PTAs as worthwhile venues to understand variation in dispute settlement rules, and critique and build off past work

[2] Throughout Latin America, for instance, dozens of disputes among members of the Andean Community, MERCOSUR, CARICOM, NAFTA, and the Central American Common Market have been taken before the organizations' legal dispute settlement procedures.

on PTA dispute settlement. Next we engage in a thorough discussion of possible explanations for strong dispute settlement in international institutions and provide details on variable measurement and empirical tests. After presenting and discussing our empirical findings, we conclude by outlining future avenues for research.

Dispute Settlement in International Institutions

Dispute settlement, broadly conceived, has been a major focus of contemporary scholarship on international law and international organizations (Dunoff and Pollack 2012). Several of the most prominent international courts, tribunals, and legal dispute settlement bodies have attracted significant attention, among them the ECJ and the WTO DSM. In the short section that follows, we briefly survey various literatures on international dispute settlement, with an eye toward insights and approaches that assist us in explaining variation in the strength of dispute settlement in PTAs.

Generalized international courts, at both the regional and global level, have attracted considerable scholarly interest, for obvious reasons. At the global level, the International Court of Justice (ICJ) occupies a prominent position. With comparatively fewer cases but detailed awards, most research on the ICJ has come from legal scholars. Nevertheless, large-n empirical studies of ICJ outcomes are becoming more common. As is the case with the ECJ, some now question the assumption of judicial independence and find that ICJ judges are subject to national biases and other predispositions (Posner and de Figueiredo 2005). In studies of outcomes of multiple international courts, then, a common back-and-forth is whether international dispute settlement bodies are truly independent from the states that created them.[3] Yet we caution

[3] The idea has been explored more generally in studies that apply principal-agent analysis to IOs (see Elsig and Pollack 2014).

against focusing too much on the outputs (rulings) from international courts and legal bodies, since these represent the "tip of the iceberg" and could reflect exceptional instances brought before the most exceptional courts. Indeed, one of the most interesting ICJ studies for our purposes is Fisher's (1982) study of why four pairs of states brought disputes before the ICJ. She finds motivations ranging from avoiding the appearance of asymmetric bullying to alliance ties explain why states would resolve disputes before the ICJ. This is suggestive for studying the design of DSMs: states anticipate the reasons why they might need to resolve disputes and design dispute settlement mechanisms accordingly.

At the regional level, the literature on the ECJ is particularly vibrant, with many arguing that through important rulings the ECJ has clarified member state's obligations under the treaties and helped to propel integration forward (*e.g.*, Alter 1998, Stone Sweet 2004, Mattli and Burley 1993). Varying arguments have been made about an independent/supranational ECJ playing both a "lawmaking" and/or an "enforcement" role. Intergovernmentalists and ECJ skeptics have countered by arguing that the ECJ is not fully independent from its member states (Garrett et. al 1998) and that ECJ judges are constrained by the threat of veto from member states and thus neither as independent nor powerful as commonly thought (Carrubba et al. 2008). Notwithstanding the above debate, one point to emphasize is that DSBs and regional courts like the ECJ typically emerge out of international agreements and are intended to promote compliance with those agreements. Courts and DSBs of other regional organizations have attracted comparatively less attention, although this is changing as many of them (Andean Pact, Mercosur, CARICOM) become more active and prominent. One lesson to be drawn from investigations of across regional bodies is the potential for institutional diffusion effects within

and across regions, as well as for some bodies (ECJ) to be a model for other regions (Alter and Helfer 2010).

Human rights courts and tribunals represent another frontier from which lessons may be drawn. The European Court of Human Rights (ECHR) is particularly powerful (having mandatory jurisdiction for signatories to the European Convention on Human Rights) as well as active, with thousands of judgments issued annually and tens of thousands of pending applications (see Voeten 2008, who also focuses on the question of whether ECHR judges are impartial). Next on the list is the Inter-American Court of Human Rights (IAHCR). In both cases, the courts have been tasked with enforcing the noteworthy regional human rights convention (or charter, in the case of the IAHCR), highlighting/suggesting a link between the depth of agreements and the strength of enforcement. Enforcement is central to the human rights issue area; in fact, 44 of the 66 articles in the European Convention on Human Rights deal with enforcement (Karns and Mingst 2010, 482), which illustrates the general importance of dispute settlement clauses in international agreements. The European and Inter-American tribunals, when placed beside regions with weaker rules on human rights, also highlight the importance of potentially differing regional approaches to dispute settlement and enforcement. Recently the world has witnessed the development of ad hoc tribunals to punish war crimes in Bosnia, Rwanda, Sierra Leone, and Cambodia, which culminated in the creation of a more formalized International Criminal Court. Once again, the ad hoc tribunals were created to enforce a set of human rights and laws of war that had been enshrined in a variety of international agreements – once again raising the issue of legal dispute bodies as "enforcement" bodies. Moreover, the fact that the ICC was created after the development of the ad hoc tribunals demonstrates that states

can learn from past experiences and that positive experience with one dispute settlement venue can lead to the creation of similar ones.

Dispute settlement bodies across other, diverse issue areas generate additional insights and provide building blocks for our investigation. Many international environmental agreements have innovative dispute settlement clauses, the most notable of which might be the International Tribunal for the Law of the Sea, which emerged from the UN Convention on the Law of the Sea. "Compliance" with environment agreements like the UN Convention on the Law of the Sea is the central issue in this issue area (see Young 1994), and dispute settlement mechanisms also can be thought of as devices for achieving compliance (see von Stein 2012). Also, several territorial disputes have been successfully resolved through international arbitration and adjudication. One finding is that states with democratic and legalized domestic political systems are more likely to pursue these types of international dispute settlement (Allee and Huth 2006), and these same states should be most likely to include strong dispute settlement in their international agreements. Finally, studies of investment arbitration find that powerful states are actually more likely to include strong arbitration provision in their bilateral investment treaties (Allee and Peinhardt 2010, 2014). This suggests that asymmetries among treaty signatories could lead to stronger, and not weaker, dispute settlement provisions.

Last but not least, the sizeable literature on WTO dispute settlement encapsulates a number of debates that are directly relevant to our efforts.[4] The WTO's DSM is perhaps the most comprehensive among all IOs and contains innovations such an appeals court (i.e., the Appellate Body) and rules on implementation (i.e., retaliation). As such, it might serve as a model for dispute settlement clauses in PTAs and other agreements – particularly in the post-Uruguay Round period and for member states most comfortable with the WTO DSM.

[4] For an extensive discussion of research on WTO dispute settlement, see Bernauer, Elsig, and Pauwelyn 2011.

Conceptually, the WTO DSM has been portrayed as both a clarification body as well as an enforcement body (Sattler and Bernauer 2011), as well as one that resolves "incomplete contracts" (*e.g.*, Horn, Maggi, and Staiger 2010; Maggi and Staiger 2008). Empirical studies look at various "stages" in the process, including dispute initiation, dispute escalation, and dispute outcomes. Power asymmetries are a common theme in these studies (*e.g.*, Elsig and Stucki 2011; Sattler and Bernauer 2011). It was believed that the DSM would be attractive to smaller states by providing them with a legal forum to challenge large-state trade policies and counteract power-based bargaining, but this has not necessarily been borne out. Moreover, it is surprising that the United States was so supportive of such a strong dispute settlement procedure (e.g., Croome 1995; Brewster 2006, Elsig 2013). Consideration of the DSM, then, leads us to acknowledge some of the unique features that could be included in dispute settlement provisions (sanctions, etc.) and to appreciate how WTO members negotiating other agreements (i.e., PTAs) might have differing preferences for a strong dispute settlement body on the DSM model.

Despite the insights and quality of past research, we identify several limitations with the way in which international dispute settlement bodies are often studied. First, scholars typically explore only the most prominent and most "successful" institutions. This is justifiable in many respects, but it also unintentionally obscures broader patterns, ignores the remaining universe of less prominent international institutions (the less prominent DS bodies and IOs with no DS mechanism), and raises concerns with selection bias and generalizability. A second feature is that scholarship focuses disproportionately on debates where two worldviews or opposing perspectives collide. Questions thus produce one toward either side of a bifurcated answer, such as whether a court is independent or whether a dispute settlement body benefits some states more than others. These are interesting debates, but they can prevent us from exploring richer and

more nuanced questions about dispute settlement bodies – including their origins and how they vary. Finally, and perhaps most importantly, scholarship tends to be quite atomized, with literatures on one court or tribunal not engaging or speaking to literatures on others. Along with an increasing chorus of others, we argue that a more comparative perspective is needed (Haftel 2007).

Thankfully, several positive lessons are drawn from previous studies to assist us in moving forward. The first is that for most international institutional settings "dispute settlement" is perhaps better thought of as "enforcement" (see Alter 2013). It is clear that most dispute settlement bodies that are created arise out of treaties and other international agreements. Their purpose, then, is to ensure that state signatories uphold the terms/obligations enshrined whatever agreement they sign or organization they created. Any resulting "disputes" will revolve around claims by one signatory or member (a complainant) that another is not meeting its obligations. A second lesson is that to understand the design of these enforcement bodies, one should look in multiple directions and at multiple "levels." Features of the agreement, such as its depth and number of signatories, are a logical starting point. The region in which an international institution is located is another, as is the identity and preferences of powerful states. Internal political characteristics of member might also explain their design preferences. Finally, asymmetries among members might also lead to different enforcement provisions, and studies of other dispute settlement bodies suggest that these relationships are not always as expected.

Dispute Settlement in PTAs

Our primary mission is to identify and explain variation across dispute settlement provisions, namely the strength of rules regarding enforcement of obligations. We seek out a

population of international institutions that contain dispute settlement provisions of varying types (DV), and whose values for possible explanatory variables (agreement depth, number of signatories, different regions, different states) also vary widely. Yet we also want cases that hold constant confounding variables like salience and issue area, since too much heterogeneity across cases can lead to simplistic explanations or garbage-can-type regression models. PTAs provide a useful match for these criteria, and have the added benefit of being increasingly popular and highly policy relevant. Their dispute settlement mechanisms have been used far more than is commonly realized – producing both notable awards from dispute settlement bodies but also many negotiated or "out of court" settlements.[5]

We define PTAs as agreements between states or regional organizations that provide reciprocal preferential market access for members' goods and services (see Dür et al. 2014). Within this broad definition, they can range from trade agreements between two parties to agreements consisting of multiple members. Some are regionally concentrated (e.g. NAFTA), other are characterized by membership that is geographically dispersed (e.g. Chile-Singapore PTA). Some are partial trade agreements that liberalize selective industry sectors (e.g. the automobile sector), others are attempts to provide duty-free access across the board (free trade agreements). Finally, some trade agreements develop within a broader economic or political integration process, such as in the case of Customs Unions (e.g. South-African Customs Union) or as important elements of Economic and Monetary Unions (e.g. the European Union).

PTAs have become a worldwide phenomenon. Since 1947, there have been around 790 such agreements concluded (Dür et al. 2014). The rise since the early 1990s has been impressive

[5] Disputes adjudicated by NAFTA panels and courts associated with the Andean Community and European Union have been studied heavily by legal scholars, for instance. However, the EU has addressed several disputes with its trading partners through formal consultations, the first step of formal dispute settlement procedures, and multiple EFTA disputes have been resolved at the Committee level, which is part of the agreement's dispute settlement procedures.

and the once-called "new regionalism" (Mansfield and Milner 1999) is no longer restricted to regions or political or economic systems. In fact, at present only one WTO member has not yet officially concluded a PTA (Mongolia). Most countries participate in multiple agreements and it seems that this evolution is continuing while mega-PTAs are showing up the horizon, such as the US-EU trade and investment agreement (TTIP).

Large-n empirical scholarship typically has treated PTAs as uniform, employing a dummy variable to indicate whether or not a pair of states has signed a PTA. In recent years there has been considerable progress toward understanding important differences across PTAs (Dür, et al. 2014; Kucik 2012; Mansfield and Milner 2012). The creation of new, richer data sets is a crucial step in this regard (Dür et. al. 2014). Some work has focused on the extent to which PTAs contain flexibility provisions (Baccini et. al. 2013; Kucik 2012). Similarly, scholars now focus on particular subsets of the agreements, such as provisions on investment (Dür and Baccini 2012), services (Kim and Manger 2013) or non-trade concerns such as human rights clauses (Hafner-Burton 2008, Spilker and Böhmelt 2013). For our purposes, one of the most important developments is attempts to ascertain the "depth" of different PTAs (Dür et al. 2014; also Baier and Bergstrand 2007).

When we probe the dispute settlement features of PTAs, what is common in most agreements is that they foresee a range of instruments to settle disputes, which we describe in detail in other work (see Allee and Elsig forthcoming). Most stipulate a type of consultation procedure in case of disagreement among partners. If one side is of the view that the other party has not fulfilled its obligations under the treaty, parties are asked to seek amicable solutions to settle disputes. This is the core basis for cooperating states, but is fairly standard across all treaties. Significant differences exist, however, when the emphasis shifts toward forms of legal

dispute settlement: the degree of involving third parties in dispute resolution, the creation of ad hoc vs. permanent bodies, the emphasis of conciliation over litigation, the relationship to outside fora to settle disputes (e.g. WTO), and the tools to induce losing parties into compliance, just to mention a few (see Allee and Elsig, forthcoming).

Dispute settlement was actually one of, if not the first, aspects of PTAs in which variation was explored and subjected to empirical testing. Smith's (2000) study represents a valuable early contribution. He examined an important subset of PTAs (62 regional organizations) to see how legalistic their DSMs are – creating a simple scale ranging from "diplomacy" to "legalism". His focus on regional organizations and trade institutions, while reflecting the wave of regionalism at the time, has become somewhat less applicable given the incredible rise in bilateral PTAs. His emphasis on legalism and legalization similarly reflects the conceptual interests of IR scholars at the time (Abbott et. al 2000), but it also obscures alternate dimensions of DSMs. Smith coded variables that are undoubtedly important, such as whether DSMs allow for third-party adjudication and whether adjudication occurs through standing bodies. Yet these variables also could be thought of as reflecting other concepts, such as "delegation." Some of Smith's other "legalism" variables, such as private standing and whether a dispute is explicitly binding, seem a bit unnecessary given that in our view they do not actually vary across agreements.[6] Recent work by Jo and Namgung (2012) updates Smith's data collection and builds off his framework, although they actually narrow his variable measurement and legalism continuum. They add more explanations to those put forward by Smith and overall their piece is

[6] For instance, in customary international law any judgment from an international body would be "binding," and most PTAs include either explicit language that awards are final and binding or less directly language about how the parties should resolve the dispute in accordance with the recommendations of the panel. Explicit mentions that awards are not binding are extremely rare; we find only two PTAs in which the parties explicitly say that awards are not legally binding.

also a valuable contribution, yet it does not address our central concerns about capturing greater variation in DS among a wider range of PTAs.

Various descriptive studies in recent years have made major contributions to improving our understanding of variation across DSMs in PTAs. Several case studies have examined DSMs in comparatively less prominent PTAs (Biukovic 2008; Gantz 2007). Others have looked more comparatively at DSMs in PTAs from a primarily legal viewpoint (Donaldson and Lester 2009; Morgan 2008; Porges 2011). Porges (2001) skillfully discusses a broad set of DSM provisions in PTAs and continues to label DSMs on a simple scale from diplomatic to legal, with a distinction among the later made between ad hoc and standing arbitration. The most noteworthy recent study is a WTO working paper from Chase et al. (2013). The majority of their paper is a "nuts and bolts" probe of all of the major dispute settlement provisions in PTAs. Theirs is a rich and stimulating survey; they provide descriptive tables and case examples for all of these provisions, but do not attempt any explanation or prediction. We see two limitations with their paper. First, they consider an incomplete 226 PTAs – only those notified to the WTO – compared to the nearly 600 for which we compile and analyze data. Second, at a broader analytical level, they default to thinking about PTAs on a simple three-part continuum from "political" to "judicial," in which nearly all recent PTAs are classified in the middle category of quasi-judicial. Nevertheless, all of the aforementioned valuable studies push scholarship on PTAs, and institutional design, in the right direction.

Our analytical efforts are driven by three goals. First, we attempt to identify the texts of as many PTAs as possible, including bilateral, regional, and other plurilateral agreements. Second, we strive to code as many variables relevant to dispute settlement provisions as possible.

We end up coding 32 variables related to dispute settlement for 587 agreements signed between 1945 and 2009.[7] Third and most importantly, we synthesize in theoretically-driven ways the variables we have compiled, eschewing reliance on a single idea of legalization and instead emphasizing multiple subcomponents that reflect "strong" enforcement provisions within PTAs. This idea of PTA enforceability matches depictions of other dispute settlement bodies (ECJ, WTO, ICC) and links nicely to discussions of credible commitments and treaty compliance.

Six components go into our overall index of the strength of dispute settlement provisions within a given PTA – all of which should aid in enforceability of a treaty. A first component follows directly from existing work on DSMs in PTAs to capture the extent to which dispute settlement authority is delegated to a third-party, legal body. The lowest category in the three-point scale (coded with a 0) is when there is no legal dispute settlement at all. This category includes the scenario in which there are no provisions for dispute settlement as well as those in which dispute settlement provisions exist but specify only consultations and/or mediation. This category comprises a slight majority (337 of 589, or 57% overall) of all PTAs, as depicted in Table 1a. The remaining two categories sub-divide the 247 instances in which some type of legal dispute settlement (arbitration or adjudication) is specified in the PTA. The key distinction is between legal settlement through a standing body (the highest category) and settlement through an ad hoc arbitral panel (Porges 2011). By creating a standing body, PTA members delegate substantial authority to a more autonomous, established actor, which leads to more coherent interpretation of trade law and the development of case law around which expectations will converge. Standing bodies also typically possess greater resources, which helps with the carrying out of proceedings and the enforcement of awards. Therefore, this ***delegation to a legal third party*** variable is coded 2 if it creates a standing body for dispute settlement and 1 if the

[7] See Dür, et.al. 2014 for a discussion of which treaty texts were available for coding.

PTA specifies only ad hoc arbitration.[8] Table 1a contains the complete distribution for this delegation variable. Among the 43% of PTAs that do allow for legal dispute settlement, most (38% overall, 224 PTAs) specify only ad hoc arbitration, whereas just over 5% (31 PTAs) delegate dispute settlement authority to a standing body (see Table 1a).

[Table 1 here]

The second component emphasizes the ability of a complainant state to choose the dispute settlement venue. We draw upon work in two literatures: 1) on trade disputes in WTO context, in which a complainant brings cases when the other party has either violated commitments or its behavior – while maybe not an obvious breach – has negatively affected the legitimate expectations of the complainant, and 2) discussions of "forum shopping" in international legal affairs (*e.g.*, Busch 2007, Drezner 2007). The logic here is that dispute settlement provisions will be "stronger" and lead to more effective enforcement of PTA obligations, when an aggrieved party can select where a PTA-related dispute will be heard. This ***complainant forum choice*** variable is also on a three-point scale. PTAs are given a 0 for this component when they fail to specify anything about multiple fora and forum choice. Approximately 80% of PTAs fall into this first category (see Table 1b). The next category (coded with a 1) indicates scenarios in which the complainant is allowed to choose the forum, yet they can only pursue settlement in one forum, thereby excluding *ex post* the use of an alternative forum. Most of the remaining cases (104 PTA, or 18% of the total) follow this "fork in the road" logic. Finally, the highest value (coded with a 2), which is quite rare, is when the complainant

[8] Some (113) PTAs also contain some reference to other international institutions and external dispute settlement bodies (primarily the GATT/WTO and the respective DSMs, as well as the International Court of Justice, in older agreements). Almost all of these also specify ad hoc arbitration (108 of 113) and a handful (2 of 113) create a standing body and thus are coded accordingly; the other three such PTAs are coded as 0 for the delegation variable due to the vague nature of the external reference.

chooses the venue and there are no restrictions on the use of multiple fora. This allows the use of multiple enforcement mechanisms to induce the other party into compliance.

A third element is the manner in which the chairman of a dispute settlement panel is chosen. This issue of ***chairman selection*** is important because arbitral panel take decisions by majority and the chair often plays a pivotal role. Effective dispute settlement is more likely to occur when there is an unbiased chair that moves the process along and is not beholden to the interests of the state parties. Furthermore, the actual process of choosing the chairman can be contentious and can slow down proceedings, thus making treaty enforcement less effective. Each state party typically appoints one panelist to a three-person panel, and the panel composition process concludes with a decision about how the third person (typically the chair) is selected. Common options for doing so are: the two parties consult and decide, the party-appointed arbitrators choose, an outside actor (an international organization/secretary-general) selects, or the chairman is chosen "by lot." The latter two options enhance enforcement the most, since the selection process is faster and less subject to pressure by the respondent state. By contrast, consultations among the two parties is the least effective and slowest selection method, since the respondent maintains much control. Multiple options are often specified in a sequential pattern, resulting in sixteen unique combinations of chairman selection across our data set (see Allee and Elsig forthcoming). For coding purposes, we are interested in whether either of the more effective options (third party selects or by lot) is specified at all, since this provides a route for swifter appointment of an effective chairman. One of these two options is included in 145 PTAs, as depicted by the values for "2" in Table 1c. The party-appointed arbitrators are allowed to select the chairman in 35 PTAs, but in only 15 of these is selection by lot or a third party also not

specified (coded as 1). All remaining PTAs, coded as 0, specify only bilateral consultations as the method of selection or have no legal dispute settlement.

A fourth component captures whether the DSP in a given treaty provides ***time limits*** for the dispute settlement process, whether overall and/or for particular stages (*i.e.*, pre- and post-award). Some of the logic is similar to that for chairman selection. The specification of time frames encourages a (comparatively) faster dispute settlement process and thus should aid with enforceability of obligations. This is important because in the absence of time limits the respondent can engage in strategic delay at various stages of the process. A total of 221 agreements specify time limits (see Table 1d).

The fifth component of overall enforceability captures the extent to which ***post-award sanctions*** can be used to effectively implement awards. Post-award enforcement mechanisms such as retaliatory measures have attracted considerable attention in the WTO dispute settlement context (Bown and Pauwelyn 2010, Zangl 2008). Even after an arbitration or adjudication ruling, a state found to have violated its obligations might be able to delay or even avoid changing its non-compliant behavior. By allowing an aggrieved complainant to punish the non-violation, institution-sponsored sanctions serve as negative inducements that make compliance more likely. Our empirical indicator for this post-award sanctions variable is a four-point, additive scale that combines values for four 0 vs. 1 indicators. The first indicator captures whether a PTA contains a sanctions provision, which is the case in 162 agreements, which comprises just under one-third of all PTAs and nearly two-thirds of those that allow for some type of legal dispute settlement (162 of 247).[9] The second indicator measures whether the complainant can choose the level of retaliation, which is allowed in 151 PTAs – a large majority

[9] We do not count cases where no arbitration or adjudication is specified, but the dispute settlement language allows for trade remedies.

of those that allow for some type of sanctions. Finally, the third and fourth indicators capture whether same-sector or cross-retaliation is allowed as well as whether monetary compensation is envisaged. The former is present in 88 PTAs, while the latter occurs in 20 agreements.[10] All four elements of the retaliation provisions, captured by respective four dummy variables, enhances the enforceability of an award against a defendant and thus strengthens dispute settlement overall. The distribution of the resulting, four-point, additive indicator of sanctioning power is shown in Table 1e.

The final component explores the ***comprehensiveness*** of dispute settlement provisions; that is, does the DSP apply broadly to all areas covered by the agreement? Across PTAs we observe various exceptions in which some areas are explicitly listed as not being subject to dispute settlement rules. Areas most commonly excluded from dispute settlement are trade remedies, safeguards, some forms of services, temporal entry of workers, SPS and TBT, competition policy, and investment. These negative exceptions weaken the enforceability of treaty commitments, particularly since they are likely inserted by a party that may be hesitant or unwilling to carry out particular obligations. The most comprehensive DSPs, then, are those that lack any such exemptions. This occurs in 169 PTAs, as depicted by "1s" in Table 1f.

We combine these six components to create a simple index of the ***strength of enforcement*** for all 589 PTAs in our data set. For our primary measure of overall enforceability, we utilize a straightforward, additive index that adds the six components discussed in the previous paragraphs. In principle this index can range from 0 to 12, but in reality the largest value maxes out at 9. This index serves as the primary dependent variable in our empirical tests. Table 1g shows the distribution of this variable across the universe of post-war PTAs.

[10] See Allee and Elsig forthcoming for a more comprehensive analysis of sanctions provisions in PTAs, including a detailed presentation of data and a discussion of patterns.

As an alternate outcome variable, we also create a more pared-down and standardized six-category index (see Table 1h). Here we take the six components depicted in Table 1 (a-f), but instead of adding the raw values for each we standardize them all on a 0-1 scale. Two of the six components (time limits, comprehensiveness) are already binary (0-1). For three of the others (forum choice, chairman selection, sanctions) we divide by the maximum value to array on a 0-1 scale. For the fourth (delegation), we (re-)code both ad hoc arbitration and creation of a standing body as 1. As a final method of measuring the overall strength of enforcement, we also combine all of the relevant variables through various types of factor analysis, although we do not report those findings due to their considerable overlap with the main findings.

Explanations for Variation in PTA Dispute Settlement Design

Our attention now turns to explaining why some PTAs, and some international institutions more generally, would have strong dispute settlement mechanisms than others. Building off the earlier discussions, we identify a range of possible explanations for this variation, focusing first on features of the agreements before considering regional and country explanations and then internal and other characteristics of the PTA members. Within each category we identify the relevant explanatory variables, discuss the logic for why each might produce strong dispute settlement, and provide details on variable measurement and operationalization.

Features of the Agreement

One way to think about variation in dispute settlement is to think about design features as being "rational" responses to the characteristic of the PTA (*e.g.*, Koremenos et. al. 2001). In

other words, the basic features of the agreement will dictate the type and strength of dispute settlement provisions that are required. The identities of the countries that signed the agreement, and their individual preferences, characteristics, and bargaining power, are largely inconsequential; all states will behave similarly by designing the type of dispute settlement that is dictated by the features of the agreement.

The first and most obvious such explanation for stronger dispute settlement is agreement ***depth***. Put simply, "deeper" agreements promise greater gains from cooperation but also a greater potential for defection. Therefore, those agreements that include the most meaningful commitments – and require the greatest and costliest policy changes – are the one that require some type of strong enforcement mechanism (Downs et al. 1996). Strong enforcement bodies, along the lines of what is described above, can compel state to uphold their commitments for fear of adverse rulings by neutral third parties, which can entail reputational costs, as well as punishment through various sanctions.

Downs et al (1996) define depth as "the extent to which (an agreement) requires states to depart from what they would have done in its absence" (383). In the PTA context, "depth" should capture the degree to which commitments are made that can potentially lead to improved market access for foreign products. Insights from the economics literature show that different design elements of PTAs can contribute to market opening and to an increase in exchange of goods, services and investments, and thus these design elements inform our operationalization. These may include direct effects (in areas such as tariff liberalization for goods, trade in services, rules that encourage mergers and acquisitions by foreign entities, and allowing foreign firms to bid for public procurement tenders) as well as indirect effects (from promoting regulatory

convergence, harmonizing standards or the protection in intellectual property rights). Our operationalization of depth combines these different elements.

Given the centrality of this concept and different possible approaches to measurement, we devise multiple indicators to capture PTA depth, all of which are drawn from Dür et al. (2014). Our primary measure is a simple additive index that potentially ranges from 0 to 7. It captures (1) whether the PTA reduces substantially all tariffs to zero and (2-7) whether it has substantive provisions on each of the following areas: competition, intellectual property rights, investment, public procurement, services, and technical barriers to trade (TBT) and/or sanitary and phytosanitary (SPS) measures. A second indicator is a fuller version of the above, and captures whether the agreement contains certain sub-provisions within each of the above areas. It ranges from 0 to 48, and more details can be found in Dür et al. (2014). A third indicator uses exploratory factor analysis to combine information across the widest possible range of indicators (see Dür et al. 2014). Finally, we also consider a measure of depth that prioritizes agreement features pertaining to domestic regulation. The logic is that commitments pertaining to member state's regulatory rules are particularly difficult to observe, monitor, and enforce – and thus are particularly likely to require stronger dispute settlement. This final operational indicator is an additive index that examines 18 PTA commitments pertaining to domestic regulation, the components of which are taken from across the services, investment, intellectual property, and SPS/TBT, competitive, and procurement provisions.

A second feature of PTAs that might determine dispute settlement design is the size of the agreement. A frequent claim is that international agreements with more members are more likely to be designed differently, namely with greater centralization (Koremenos et. al. 2001), which could be equated to strong, third-party enforcement of agreements (Allee and Peinhardt

2014). Put simply, agreements with more members might produce stronger enforcement provisions, and this can be observed in two ways, with overlapping and unique arguments for each.

An initial distinction is between bilateral agreements, with two members, and ***plurilateral*** agreements, with three or more members. For the most part, bilateral agreements do not have broader ambitions and are less likely to generate strong dispute settlement. Bilateral agreements entail the granting of specific trade preferences to one another, but they generally do not envisage future political integration or the creation of new, formal institutions. Moreover, the demand for enforcement, *ceteris paribus*, is lower in bilateral agreement. Each party is only responsible for monitoring one another, and disputes should be less complex and easier to deal with directly and informally. In contrast, we argue that plurilateral agreements possess several features that are more likely to generate stronger dispute settlement mechanisms. For one, they produce greater problems with monitoring, since each party is responsible for observing the (non-) compliance of multiple members. They also are more likely to have an eye toward future widening (bringing in new members) or deepening (further economic harmonization), which makes the creation of new (dispute settlement) institutions more likely.

It also becomes less costly to create these types of dispute settlement institutions as the number of members increases. This suggests an additional distinction between a PTA with, say, three members and one with a dozen or more members. For these reasons, we also consider the raw ***number of members*** of the PTAs. We include in our tests both relevant variables, an indicator for plurilateral agreements and a count of the number of members, since they capture slightly different logics and are positive but not perfectly correlated (.52).

Regional and Country Preferences

We also generate a series of variables that reflect whether a PTA is comprised of member from different regions or centered within a single region. There are various *ex ante* reasons to expect members within some regions to be more predisposed toward stronger dispute settlement, so we also consider these possible regional differences.

One candidate for the most exceptional region is ***Asia***. A prominent argument is that when it comes to dispute settlement, in particular, Asian countries prefer informal, non-confrontational methods of dispute settlement (Kahler 2000). Thus they should be less likely than groups of states from other regions to include strong dispute settlement in any PTAs they sign. We therefore include a dummy variable equal to 1 to indicate PTAs in which all members are from within Asia.

Predictions for other regions are perhaps less definitive; nevertheless, we have several expectations about other regions' propensities for stronger versus weaker dispute settlement in PTAs. For several reasons we expect PTAs within the ***Americas*** to be more likely to contain strong dispute settlement provisions. One reason is because of the generally higher level of cooperation within the broadly-defined region, which contains a general political organization (the Organization of American States) as well as a human rights convention, among other institutions. Another reason is because of the dominant presence of the United States, which typically has been supportive of strong trade dispute settlement rules (Brewster 2006).

Europe possesses similar dynamics in that a large actor, the European Union, dominates the trade landscape. But the expectation for "European" PTAs are perhaps less clear because of the peculiarities of EU trade policy, including a number of hub-and-spoke association agreements as well as overlapping institutions. We expect European agreements to be positively

associated with strong dispute settlement because of the overall level of legalization in Europe (i.e., ECJ and ECHR), but we expect a fairly modest result. We include an indicator variable for ***European*** PTAs; that is, PTAs signed by actors contained within Europe, both including and excluding EU agreements.

Finally, we also would expect ***intercontinental*** agreements, those signed by countries from multiple regions, to include stronger dispute settlement. The general idea is that in such situations states cannot apply regional norms or draw upon other venues for dispute settlement. In such instances states will have less familiarity with one another and a more difficult time monitoring one another and reaching informal settlements to any disputes that might arise. Thus they may want to have stronger, delegated dispute settlement available as an option.

Power and Asymmetry

The preceding discussion also suggests that power (and asymmetries) might also play a role in dispute settlement design. Thus we dig a bit deeper to see whether agreements involving powerful actors like the ***United States*** and ***European Union*** are likely to have different dispute settlement provisions than PTAs involved other states. We see two competing pressures for both major actors. On one hand, we expect both to be quite comfortable with powerful trade dispute settlement, given the strong role of courts and the rule of law with both. The US, in particular, should look favorably upon such strong mechanisms, as we discussed earlier. On the other hand, both might be leery of including powerful, legal dispute settlement in PTAs for fear that such a tool would be used against them to challenge. A partner country could effectively challenge the trade policies of both powerful actors before an enforcement body, something that would be

impossible in an anarchic environment absent such rules. On balance, however, we expect EU and US agreements, in particular, to have stronger enforcement on average.

The preceding discussions of the US and EU raise the issues of how broader asymmetries might affect PTA design outcomes. Agreements can involve very similar groups of countries, or countries that are very different in terms of levels of development and economic size. For us the most interesting combinations are those agreements that a country from the global "north" and one from the global "south." It is plausible that both types of states in this scenario would prefer stronger dispute settlement. For the wealthier, capital-exporting state, it may worry about a developing countries meeting its obligations, and in reverse, the developing country might welcome the ability to contest policies from the wealthier state that it finds objectionable or to have a general check on the abuse of power (Grant and Keohane 2005). Thus we predict that these ***north-south agreements*** are more likely to have stronger dispute settlement than the alternative pairings. PTAs only among northern states should have less demand for strong enforcement, and southern states should be less likely to surrender sovereignty to a powerful DSM in their agreements with one another.

Other Features of the Signatories

Various features of the PTA signatories also might affect PTA dispute settlement outcomes. One factor is whether the signatories are member of the WTO. ***WTO members*** are likely to have greater familiarity and comfort with powerful legal dispute settlement bodies, via their experiences with the WTO DSM, and thus will be more willing to include strong dispute settlement in their PTAs, too. To test this argument we include a dummy variable that captures whether all PTA signatories also are members of the WTO.

A couple of internal features of PTAs members also might be important. States that are characterized by a strong rule of law should be more willing to embrace stronger dispute settlement. Since we do have not yet finished compiling our preferred data on domestic rule of law, we substitute the Polity measure of democracy as a loose proxy for ***strong rule of law states***, and take the mean level across all members of the PTA. Furthermore, states that recently have become democratic should be more likely to agree to stronger dispute settlement. One explanation is that ***new democracies*** are eager to signal to a wider audience that they will uphold commitments they have made on issues like trade liberalization and harmonization (Milewicz and Elsig 2014). Those who sign PTA with states trying to consolidate transitions to democracy also are happy to "lock in" the pledges made by their partners. We include a variable that captures whether any member of the PTA has become a democracy within the past ten years.

Controls

We also include a set of control variables. One is a variable that captures the ***year*** in which the PTA was signed, employed as a control for time. This is a simple way to capture any over-time variation or trends in PTAs. We also substitute and consider period and decade dummy variables, and a dummy for the pre- and post-WTO era. We also control for ***PTA trade volume***, since PTAs involving larger amounts of trade might require stronger enforcement. Our empirical measure captures the mean value of imports and exports across all members of the PTA.

Empirical Tests

We estimate a series of multivariate regression models to evaluate the above variables as predictors of strong dispute settlement in PTAs. Given the nature of the primary and alternate

dependent variables, ordered probit estimation is used for all models, unless otherwise noted. Different methods of controlling for time are employed, although results are robust to any of these changes. We first present finding from our primary model in Table 2, followed by findings for different measures of agreement depth (Table 3).

[Table 2 here]

Several findings confirm expectations, with those variables regarding agreement features, asymmetries, and regional and country differences generating strongest and robust findings. Agreement depth is the best and most consistent predictor of strong enforcement provisions across all of the models we estimate. Similarly, plurilateral agreements consistently are associated with stronger DSPs, and north-south agreements are another strong predictor. At the regional level, the same is true of agreements in the Americas, and agreements involving the U.S., albeit with a few caveats. In contrast, European agreements do not have stronger dispute settlement, and instead have weaker rules. Finally, there is no evidence that Asian PTAs are designed with weaker dispute settlement mechanisms, and if anything, the opposite seems to be true.

Across all of the models we estimate, agreement depth is the most consistent predictor. Overall the agreement depth variable is always positively associated with strong dispute settlement, at the 99% level of confidence, across the approximately two dozens models we estimate. This strong finding holds regardless of how we conceptualize and measure agreement depth, as Table 3 shows. Similar findings are returned when we substitute a much broader indicator of depth based on the full universe of 48 depth-related indicators (column 1). Similarly, the results are nearly identical when we substitute a measure of depth arrived at by factor analysis (column 2). Finally, our theoretically-driven measure that examines depth in

terms of domestic regulations also is a strong predictor of dispute settlement strength, which means that PTAs that require stronger regulatory convergence also receive strong dispute settlement.

[Table 3 here]

Another distinction within PTAs, namely whether they are bilateral or involve more than two states, also exhibits the predicted effect on the strength of dispute settlement rules. In short, plurilateral PTAs, agreements involving three or more states, are designed with stronger dispute settlement rules. Once again, across all models that we estimate the coefficient estimate for the plurilateral variable is positive at the highest (99%) level of confidence.

Within the world of plurilateral agreements, the number of signatories to the PTA does not seem to matter, however. This variable never reaches statistical significance in any of our empirical tests. The only exception is when we drop the plurilateral (vs. bilateral) variable, in which case the coefficient estimate for number of members is also picking up the distinction between bilateral agreements, with weaker dispute settlement rules, and plurilateral agreement, with stronger rules.

Regional and country patterns are very interesting, producing both expected and unexpected findings. As predicted, PTAs within the Americas are much more likely to contain stronger dispute settlement rules. In all cases, positive coefficient estimates are returned, at the 99% or at times 95% level of confidence. Similarly, PTAs involving the United States also have stronger dispute settlement on average. This is suggestive of arguments that the U.S., with its litigious culture, is comfortable with highly legalistic and enforceable rules. The fact that the supportive findings hold when both an "Americas" and "US" variable are in the model is noteworthy. It suggests that the US finding is distinct: agreements involving the US really do

have uniquely strong rules. But it also means that the "Americas" variable is picking up non-US agreements within the hemisphere. So the Americas finding is not driven by the US agreements (which are controlled for), and thus PTAs across the Americas, and involving various Latin American countries, also exhibit dispute settlement rules that are particularly strong.

The findings for both Asian and European PTAs are interesting and unexpected. Across all models we estimate, there is no evidence whatsoever that PTAs among pairs and groups of Asian countries have weaker dispute settlement rules. In fact, the pattern seems to be suggestive of the opposite, which is that Asian PTAs tend to have stronger dispute settlement. Coefficient estimates for the Asia variable are always positive, with associated p-values or between .10 and .15. These nuances notwithstanding, the bigger takeaway point is that our findings strongly contract the widely-held notion of an "Asian Way" in terms of dispute settlement. We also find European PTAs are less, and not more, likely to have strong dispute settlement rules. The association between European PTAs and "weak" dispute settlement is highly statistically significant as well. The subset of various EU agreements also are not characterized by stronger dispute settlement, since we typically find no consistent pattern for dispute settlement rules in EU agreement. The European and EU finding warrant further investigation, due to their unique nature and potential over-time variation. Overall, though, it seems that non-EU, European agreements have particularly weak dispute settlement rules.

We also uncover some evidence that PTAs among states at differing levels of development are designed with stronger dispute settlement rules. As expected, north-south agreements are more likely to include such rules, and this finding is robust across estimations. We originally did not include indicators for north-north and south-south agreements, so we added indicators for each of those variables into the original specification. Neither of these

agreement-types among developmentally similar countries exhibits any relationship with dispute settlement design, further indicating that the north-south agreements are the unique ones that tend to spur stronger dispute settlement.

None of the findings for remaining variables are particularly noteworthy. There is some suggestive evidence that WTO members are more likely to include strong dispute settlement in their PTAs, but the finding typically hovers short of the 90% level of confidence and the substantive effect of this variable is comparatively weak. Instead, a closer look reveals that PTAs among WTO members tend to include provisions that allow for dispute settlement through the WTO DSM as an external dispute settlement options, unsurprisingly. At this preliminary stage we find little support for arguments that the domestic political characteristics of signatories drives dispute settlement design. We explored many variables that identified strong law and order states, democracies (see Table 2), and transitioning democracies, and found little evidence that these features affect dispute settlement design. Some of this might be due to inherent challenges of determining how to aggregate country-level characteristics into a composite measure, a topic we plan to explore in future research.

Conclusion

Contemporary international relations have been characterized as a "move to law" (Goldstein et al. 2000). As a result, legal dispute settlement through prominent institutions such as the WTO DSM and ECJ, among others, has received considerable scholarly attention. What has received less attention is the logically prior question of why states include strong dispute settlement provisions in some international agreements but not in others. It is the initial design

that creates new courts, specifies important rules that are evoked later, or makes it more (or less) difficult to enforce the commitment enshrined in a treaty.

Our paper emphasizes how international dispute settlement is designed in the first place, and broadens the way we think about dispute settlement institutions. As we have argued in this paper, there is more to treaties' dispute resolution mechanisms than a simple distinction between whether a treaty allows for (quasi-) legal dispute settlement or not. The presence of absence of such "delegation" is important, but is merely a starting point. Dispute settlement authority can be delegated to an ad hoc panel, or a treaty might establish a brand new body to adjudicate claims. Furthermore, we emphasize the various ways in which legal settlement, of whatever type, can transpire and the options available to litigants. Other legal and social science scholarship has investigated important topics such as retaliation (sanctions), forum choice, and timing and delay. We build on various studies of single bodies (like the WTO DSM) and instead conceptualize, code, and explain the inclusion of exclusion of these provisions across hundreds of otherwise similar agreements.

Furthermore, we insert the idea of dispute settlement as "enforcement" of international agreements, to which all of the aforementioned components of dispute settlement speak. Enforcement is a topic that has been emphasized in international relations research for quite some time, but is rarely measured empirically or unpacked in any meaningful way. Our conceptualization and measurement of enforcement therefore is an important contribution.

We also use PTAs as a laboratory to test general arguments about the design of agreements and the conditions under which state will (not) incorporate strong enforcement provisions into their treaties. Studying PTAs has several advantages, notably that one holds common the issue-area within international relations and other conflating factors. PTAs also

provide a very large and heterogeneous laboratory, as we uncover 589 of them – far more than are analyzed in other studies of such agreements. The PTAs in our expanded collection also vary widely in terms of depth, number and type of members, and across space and time (Allee and Elsig forthcoming).

Our analysis uncovers a number of patterns in dispute settlement design related to enforcement, which we think are novel and might apply to other types of agreements. One is that agreement depth is a powerful driver of strong dispute settlement – which is perhaps our most robust and substantively important finding. Other features of the agreement, such as whether it is plurilateral or bilateral, also are quite important. Yet the identities and regions of signatories matter, too – often in unexpected ways – thus casting doubt on a purely functionalist account of dispute settlement design. Agreements within the Americas, and to a lesser extent Asia, surprisingly, exhibit stronger dispute settlement rules, whereas the same is not true of European (and EU) agreements. But US PTAs and those that span the global north and south also include stronger dispute settlement rules, suggesting that power and asymmetries are important factors.

Moving forward, future research will also need to pay closer attention to explaining how design affects actual behavior. In many areas of international politics, we witness that many disputes are settled at pre-legal stages or early on in a dispute settlement process (i.e., consultations). This is also true for PTAs, although it's often not acknowledged and can be difficult to observe. This calls for further inquiry. What is different in cases that are settled early related to cases that are litigated over by using all legal means at disposal? Is it the nature of disputes or the design that affects the decisions of litigating parties? Also, in times where forum shopping becomes more and more a policy option for an increasing number of countries, how

does the WTO dispute settlement system affect choices in PTA disputes? Given the substantive literature on other courts (e.g. WTO, ECJ), there are many interesting questions and possible explanations that lend themselves to testing in the PTA context. Nevertheless, all of these questions emerge from the design of the agreements. Some agreements contain various enforcement-friendly provisions, for reasons we have uncovered, and thus provide states with particular options for dispute settlement. It is the inclusion (or exclusion) of these components that gives rise to all of these questions about the use of international dispute settlement mechanisms.

Table 1: Variation Across Dispute Settlement Provisions in PTAs
(Larger Numbers Indicate ***Stronger Enforcement***)

a) Delegation to a Legal Third Party	Number of Treaties
0	341
1	217
2	31
Total	589

b) Complainant Forum Choice	Number of Treaties
0	481
1	104
2	4
Total	589

c) Selecting Panel Chairman	Number of Treaties
0	429
1	15
2	145
Total	589

d) Time Limits	Number of Treaties
0	368
1	221
Total	589

e) Post-Award Sanctions	Number of Treaties
0	427
1	7
2	60
3	86
4	9
Total	589

f) Comprehensiveness (Lack of Excluded Areas)	Number of Treaties
0	420
1	169
Total	589

g) Overall Strength of Enforcement (raw)	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
Number of Treaties	320	19	11	41	10	49	17	47	52	23	589

h) Overall Strength of Enforcement (standardized)	**0**	**.5**	**1**	**2**	**2.25**	**2.5**	**2.75**	**3**	**3.25**	**3.5**	**3.75**	**4**	**4.25**	**4.5**	**4.75**	**5.25**	**Total**
Number of Treaties	320	1	19	26	3	6	6	29	8	13	7	46	50	36	6	13	589

Table 2: Ordered Probit Results for the Strength of Dispute Settlement in PTAs

	Coefficient	Robust Std. Error	
Agreement Features			
Depth (0-7 index)	.500	(.050)	***
Plurilateral	.724	(.138)	***
Number of Members	.000	(.005)	
Region			
Asia	.473	(.235)	**
Americas	.706	(.228)	***
Europe	-.521	(.245)	**
Intercontinental	.311	(.223)	
Country			
United States	.591	(.303)	*
European Union	-.112	(.060)	*
Asymmetry			
North-South	.330	(.134)	**
Other + Controls			
All WTO members	.258	(.122)	**
Rule of Law (mean)	.007	(.011)	
PTA Import Volume ($x10^6$)	-.271	(.901)	
Time (year)	.020	(.006)	***

N = 571 Preferential Trade Agreements
Wald χ^2 test (14df) = 324.60 (.00)

*p<.10, **p<.05, ***p<.01 (two-tailed)

Table 3: Ordered Probit Results for the Strength of Dispute Settlement in PTAs
(with Alternate Indicators of Depth)

Agreement Features			
Depth (0-48 index)	.098 ***		
	(.009)		
Depth (factor)		1.19 ***	
		(.106)	
Depth (domestic regulation)			.225 ***
			(.021)
Plurilateral	.841 ***	.775 ***	.832 ***
	(.138)	(.139)	(.140)
Number of Members	-.004	-.005	.002
	(.004)	(.005)	(.005)
Region			
Asia	.370 *	.362	.437 *
	(.223)	(.245)	(.231)
Americas	.323	.319	.440 **
	(.212)	(.242)	(.224)
Europe	-.762 ***	-1.08 ***	-.565 **
	(.239)	(.271)	(.243)
Intercontinental	.367	.618 *	.192
	(.326)	(.345)	(.222)
Country			
United States	.367	.618 *	.759 *
	(.326)	(.345)	(.389)
European Union	-.059	-.018	-.024
	(.060)	(.059)	(.058)
Asymmetry			
North-South	.268 **	.360 ***	.403 ***
	(.137)	(.138)	(.138)
Other/Controls			
All WTO Members	.241 **	.159	.210 *
	(.121)	(.131)	(.125)
Rule of Law (mean)	.008	-.001	.013
	(.011)	(.012)	(.011)
PTA Import Volume ($x10^6$)	-.250	-.346	-.818
	(.998)	(.822)	(.964)
Time (year)	.015 **	.007	.018 ***
	(.006)	(.006)	(.006)

N = 565-571 Preferential Trade Agreements *p<.10, **p<.05, ***p<.01 (two-tailed)

Bibliography

Abbott, Kenneth W., Robert O. Keohane, Andrew Moravcsik, Anne-Marie Slaughter and Duncan Snidal. 2000. "The Concept of Legalization." *International Organization* 54 (3): 401-419.

Allee, Todd and Manfred Elsig. 2014. Dispute Settlement Provisions in PTAs: New Data and New Concepts, Manuscript.

Allee, Todd and Paul K. Huth. 2006. "International Adjudication as Domestic Political Cover." *American Political Science Review* 100 (2): 219-234.

Allee Todd and Clint Peinhardt. 2014. "Evaluating Three Explanations for the Design of Bilateral Investment Treaties." *World* Politics 66(1): 47-87.

Allee, Todd and Clint Peinhardt. 2012. "Contingent Credibility: The Reputational Effects of Investment Treaty Disputes on Foreign Direct Investment." *International Organization* 65 (3): 1-26.

Allee, Todd and Clint Peinhardt. 2010. "Delegating Differences: Bilateral Investment Treaties and Bargaining Over Dispute Resolution Provisions." *International Studies Quarterly* 54 (1): 1-26.

Alter, Karen. 1998. "Who Are the 'Masters of the Treaty?: European Governments and the European Court of Justice." *International Organization* 52 (1): 121-47.

Alter, Karen. 2000. "The European Union's Legal System and Domestic Policy: Spillover or Backlash?" *International Organization* 54 (3): 489-518.

Alter, Karen and Laurence Helfer. 2010. "Nature or Nurture? Judicial Lawmaking in the European Court of Justice and the Andean Tribunal of Justice." *International Organization* 64 (4): 563-92.

Baccini, Leonardo. 2010. "Explaining Formation and Design of EU Trade Agreements: The Role of Transparency and Flexibility." *European Union Politics* 11 (2): 195-217.

Baccini, Leonardo, Dür, Andreas, and Manfred Elsig. 2013. "Depth, Flexibility and International Cooperation: The Politics of Trade Agreement Design", Manuscript

Baier, Scott, and Jeffrey Bergstrand. 2007. "Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics* 71 (1): 72-95.

Biukovic, Ljiljana. 2008. "Dispute Resolution Mechanisms and Regional Trade Agreements: South American and Caribbean Modalities." U.C. Davis Journal of International Law & Policy 14 (2): 255-96.

Bown, Chad, and Joost Pauwelyn, eds. 2010. *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*. Cambridge: Cambridge University Press.

Brewster, Rachel. 2006. "Rule-Based Dispute Resolution in International Trade Law." *Virginia Law Review* 92 (2): 251-88.

Busch, Marc. 2007. "Overlapping Institutions, Forum Shopping, and Dispute Settlement in International Trade." *International Organization* 61 (4): 735-61.

Carrubba, Clifford, Matthew Gabel, and Charles Hankla. 2008. "Judicial Behavior under Political Constraints: Evidence from the European Court of Justice," *American Political Science Review* 102 (4): 435-52.

Chase Claude, Yanovich Alan, Crawford Jo-Ann and Ugaz Pamela. 2013. "Mapping of Dispute Settlement Mechanisms in Regional Trade Agreements - Innovative or Variations on a Theme?" WTO Staff Working Paper (ERSD-2013-07).

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

Croome, John. 1995. *Reshaping the World Trading System: A History of the Uruguay Round*. Geneva: World Trade Organization.

Donaldson Victoria and Lester Simon 2009. "Dispute Settlement", in Bilateral and Regional Trade Agreements: Commentary and Analysis, edited by S. Lester and B. Mercurio, Cambridge: Cambridge University Press.

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. 2007. *All Politics is Global: Explaining International Regulatory Regimes*. Princeton: Princeton University Press.

Duffield, John. 2003. "The Limits of "Rational Design." *International Organization* 57 (2): 411-30.

Dunoff, Jeffrey and Mark Pollack, eds. 2013. *Interdisciplinary Perspectives on International Law and International Relations*. Cambridge: Cambridge University Press.

Dür, Andreas, Leonardo Baccini, and Manfred Elsig. 2014. "The Design of International Trade Agreements: Introducing a New Database." Forthcoming in *The Review of International Organizations*.

Elsig, Manfred 2013. "Legalization Leap and International Institutions – The Design of the WTO's Dispute Settlement System" Manuscript.

Elsig, Manfred, and Mark Pollack. 2014. "Agents, Trustees, and International Courts: Nomination and Appointment of Judicial Candidates in the WTO Appellate Body." *Forthcoming in European Journal of International Relations*.

Elsig, Manfred, and Philipp Stucki. 2011. "Low-Income Developing Countries and WTO Litigation: Why Waking Up the Sleeping Dog?" *Review of International Political Economy* 19 (2): 292-316.

Fischer, Dana. 1982. "Decisions to Use the International Court of Justice: Four Recent Cases," *International Studies Quarterly* 26 (2): 251-77.

Goldsmith, Jack L. and Eric A. Posner. 2005. *The Limits of International Law.* New York: Oxford University Press.

Goldstein, Judith, Miles Kahlter, Robert Keohane, and Anne-Marie Slaughter. 2000. "Introduction: Legalization and World Politics." *International Organization* 54 (3): 385-99

Grant, Ruth, and Robert Keohane. 2005. "Accountability and Abuses of Power in World Politics." *American Political Science Review* 99 (1): 29-43.

Hafner-Burton, Emilie. 2008. "Sticks and Stones: Naming and Shaming the Human Rights Problem." *International Organization* 62 (4): 689-716.

Haftel, Yoram. 2007. "Designing for Peace: Regional Integration Arrangements, Institutional Variation, and Militarized Interstate Disputes." *International Organization* 61: 217-237

Helfer, Laurence R and Anne-Marie Slaughter. 1997. "Toward an Effective Theory of Supranational Adjudication." *Yale Law Journal* 107 (2): 273-391.

Horn, Henrik, Giovanni Maggi, and Robert Staiger. 2010. "Trade Agreements as Endogenously Incomplete Contracts." *American Economic Review* 100 (1): 394-419.

Jo, Hyeran and Hyun Namgung. 2012. "Dispute Settlement Mechanisms in Preferential Trade Agreements: Democracy, Boilerplates, and the Multilateral Trade Regime." *Journal of Conflict Resolution* 56 (6): 1041-68.

Kahler, Miles. 2000. "Legalization as Strategy: The Asia-Pacific Case." *International Organization* 54 (3): 549-71.

Karns, Margaret and Karen A. Mingst, *International Organizations*, 2nd edition. Boulder: Lynne Rienner, 2010.

Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55 (4): 761-99.

Koremenos, Barbara. 2007. "If Only Half of International Agreements have Dispute Resolution Provisions, Which Half Needs Explaining?" *Journal of Legal Studies* 36 (1): 189-212.

Kucik, Jeffrey. 2012. "The Domestic Politics of Institutional Design: Producer Preferences over Trade Agreement Rules." *Economics and Politics* 24 (2): 95-118.

Maggi, Giovanni, and Robert Staiger. 2008. "On the Role and Design of Dispute Settlement Procedures in International Trade Agreements." NBER Working Paper No. 14067**.**

Mansfield, Edward D. and Helen V. Milner 1999. "The New Wave of Regionalism." *International Organization* 53 (3): 589-627.

Mansfield, Edward D., and Helen V. Milner. 2012. *Votes, Vetoes, and the Political Economy of International Trade Agreements.* Princeton: Princeton University Press.

Mattli and Burley 1993. "Europe before the Court: A Political Theory of Legal Integration." *International Organization* 47 (1): 41-76.

Milewicz, Karolina and Manfred Elsig. 2014. "The Hidden World of Multilateralism: Treaty Commitments of Newly Democratized States in Europe." *International Studies Quarterly*, forthcoming.

Morgan, David. 2008. "Dispute Settlement under PTAs: Political or Legal?", Legal Studies Research Paper No. 341, Melbourne University: Law School.

Pauwelyn, Joost and Luiz Eduardo Salles. 2009. "Forum Shopping before International Tribunals." *Cornell International Law Journal* 42: 77-118.

Porges, Amelia. 2011. "Dispute Settlement". In *Preferential Trade Agreement Policies for Development: A Handbook*, edited by J.-P. Chauffour and J.-C. Maur Washington DC: The World Bank.

Posner, Eric, and Miguel de Figueiredo. 2005. "Is the International Court of Justice Biased?" *Journal of Legal Studies* 34 (2): 599-630.

Sattler, Thomas, and Thomas Bernauer. 2011. "Gravitation or Discrimination? Determinants of Litigation in the World Trade Organization." *European Journal of Political Research* 50 (2): 143-67.

Smith, James McCall. 2000. "The Politics of Dispute Settlement Design." *International Organization* 54 (1): 137-80.

Spilker, Gabriele, and Tobias Böhmelt. 2013. "The Impact of Preferential Trade Agreements on Governmental Repression Revisited." *Review of International Organizations* 8 (3): 343-61.

Stone Sweet, Alec. 2004. *The Judicial Construction of Europe*. New York: Oxford University Press.

Stone Sweet, Alec and Thomas Brunell. 1999. "Constructing a Supranational Constitution: Dispute Resolution and Governance in the European Community." *American Political Science Review* 92 (1): 63-82.

Voeten, Erik. 2008. "The Impartiality of International Judges: Evidence from the European Court of Human Rights," *American Political Science Review* 102 (4): 417-33.

von Stein, Jana. 2012. "The Engines of Compliance." In. Interdisciplinary Perspectives on International Law and International Relations, edited by J. Dunoff and M. Pollack. Cambridge: Cambridge University Press..

Yarbrough, Beth V., and Robert M. Yarbrough. 1997. Dispute Settlement in International Trade: Regionalism and Procedural Coordination. In *The Political Economy of Regionalism*, edited by E. D. Mansfield and H. V. Milner. New York: Columbia University Press.

Young, Oran. 1994. *International Governance: Protecting the Environment in a Stateless Society*. Ithaca, NY: Cornell University Press.

Zangl, Bernhard. 2008. "Judicialization Matters! A Comparison of Dispute Settlement under the GATT and WTO." *International Studies Quarterly* 52 (4): 825-54.