

# Why do some international institutions contain strong dispute settlement provisions? New evidence from preferential trade agreements

Todd Allee<sup>1</sup> · Manfred Elsig<sup>2</sup>

Published online: 24 May 2015

© Springer Science+Business Media New York 2015

**Abstract** To understand why some international institutions have stronger dispute settlement mechanisms (DSMs) than others, we investigate the dispute settlement provisions of nearly 600 preferential trade agreements (PTAs), which possess several desirable case-selection features and are evoked more than is realized. We broaden the study of dispute settlement design beyond “legalization” and instead reorient theorizing around a multi-faceted conceptualization of the strength of DSMs. We posit that strong DSMs are first and foremost a rational response to features of agreements that require stronger dispute settlement, such as depth and large memberships. Multivariate empirical tests using a new data set on PTA design confirm these expectations and reveal that depth – the amount of policy change specified in an agreement – is the most powerful and consistent predictor of DSM strength, providing empirical support to a long-positated but controversial conjecture. Yet power also plays a sizeable role, since agreements among asymmetric members are more likely to have strong DSMs due to their mutual appeal, as are those involving the United States. Important regional differences also emerge, as PTAs across the Americas are designed with strong dispute settlement, as are Asian PTAs, which contradicts the conventional wisdom about Asian values and legalization. Our findings demonstrate that rationalism explains much of international institutional design, yet it can be enhanced by also incorporating power-based and regional explanations.

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s11558-015-9223-y) contains supplementary material, which is available to authorized users.

---

✉ Todd Allee  
tallee@umd.edu

Manfred Elsig  
manfred.elsig@wti.org

<sup>1</sup> Department of Government and Politics, University of Maryland, College Park, MD 20742, USA

<sup>2</sup> World Trade Institute, University of Bern, 3012 Bern, Switzerland

**Keywords** Trade · Preferential trade agreements · Institutional design · Dispute settlement · Enforcement

## 1 Introduction

A major development in contemporary international affairs has been the creation of international dispute settlement institutions – international courts, tribunals, and arbitration bodies – to address all types of global disputes. Even prominent international law skeptics have acknowledged that “(i)n the past few years, international dispute resolution has assumed an unprecedented prominence in international politics” (Posner and Yoo 2005, 3). If we look at the economic realm, as one illustration, we can identify many active dispute settlement bodies, both prominent and less prominent. The dispute settlement mechanism (DSM) of the World Trade Organization (WTO) receives the most attention due to its central place in the multilateral trade regime and its unique features (e.g., Bernauer et al. 2012). Yet numerous regional and bilateral dispute settlement institutions also exist, many of which have compiled an impressive record. Most notable is the European Court of Justice (ECJ), whose massive caseload and landmark rulings have been a major part of European integration. Similar courts have arisen in other regions, as 11 “operational copies” of the ECJ now exist worldwide, which collectively have issued more than 2100 rulings (Alter 2012, 136, 140). Likewise, the United Nations-sponsored Integrated Database of Trade Disputes for Latin America and the Caribbean (IDATD) indicates that more than 1,100 disputes have gone before various dispute settlement bodies hemisphere-wide since 1995.<sup>1</sup> The record is impressive elsewhere, such as in eastern and southern Africa, even if it less well known.<sup>2</sup>

The creation of these dispute settlement institutions, and the rules by which they can be used, is spelled out in accompanying international treaties. Most notable international agreements contain dispute settlement provisions, but they vary tremendously. Not all of them create a powerful, heavily-utilized court. Instead, some provide for ad hoc arbitration, specify mediation by a third-party, encourage bilateral negotiations, or various combinations of the above. Moreover, the precise rules governing each settlement option varies much more than is acknowledged. Even among arbitration, the most common option, one observes widespread variation in the specified timelines, rules governing choice of fora, selection of arbitrators, and types of post-award sanctions allowed, among other features.

These design features can be consequential. They may affect, for example, whether states want to pursue formal dispute settlement at all or whether some venues are more likely to be used than others. Furthermore, the mere presence of a powerful judicial body might encourage “out of court” settlements or deter objectionable behavior in the first place. However, the rules also can have unintended consequences if they are used

<sup>1</sup> See [http://idatd.eclac.cl/controversias/index\\_en.jsp](http://idatd.eclac.cl/controversias/index_en.jsp). Last accessed 22 March 2015.

<sup>2</sup> As James Gaathi notes, in a paper aptly titled *The Under-Appreciated Jurisprudence of Africa's Regional Trade Judiciaries*: “...in Africa there has been an exponential use of these judiciaries without much acknowledgment in the academic literature” (Gaathi 2010a, 246).

more frequently than envisioned or entail greater costs than was anticipated.<sup>3</sup> Thus, how future disputes are handled, and whether they are resolved to the satisfaction of aggrieved parties, depends heavily on what is contained in an agreement's dispute settlement provisions.

To understand why states sometimes allow for strong, legal dispute settlement, we investigate variation across the dispute settlements rules of nearly 600 international agreements signed during the past 60 years. Following others, we focus on preferential trade agreements (PTAs), which contain the full landscape of dispute settlement design choices. Additionally, they are numerous, politically relevant, and signed by nearly all countries – which makes them generalizable to other issue areas and types of agreements. Our efforts to explain such wide variation in dispute settlement design are unique in several ways. First, we examine all PTAs since World War II – not just regional trade agreements or those currently in force – and thus we incorporate hundreds more agreements than in other studies. Second, we depart from more simple conceptualizations of DSMs and instead re-orient our theorizing around the notion that DSMs are procedures that help to enforce the terms of agreements. Relatedly, whereas other studies simply examine whether or not selected agreements have a dispute settlement mechanism (DSM) or whether it is “legal” versus “political,” we develop a richer, more multi-faceted empirical measure of DSM strength that draws upon newly compiled data. Finally, we engage and reconcile seemingly disparate explanations for institutional design (rational design, state power, and regionalism), which typically are portrayed as hostile to, or even incompatible with, one another.

Our unifying claim is that stronger DSMs are primarily a rational response by members to the heightened demands of “deeper” agreements with “wider” memberships, but that considerable room exists for signatories to craft agreements to fit the unique preferences of agreement members. Drawing on well-developed, but often untested, arguments about international cooperation, we expect that all states should incorporate stronger dispute settlement rules within agreements that have more members and entail significant obligations. Indeed, the strongest finding from our multivariate empirical tests is that the deepest agreements – those that require the greatest policy change (in our case trade liberalization, market access, and harmonization) – also incorporate the strongest dispute settlement provisions. A similarly robust finding is that DSM strength increases with the number of members. Yet we uncover clear evidence that agreements also reflect power-based considerations. As compared to more balanced agreements, we find that North–South agreements are much more likely to contain strong DSM language as are agreements involving the United States. As for regional patterns, we discover that PTAs signed in the Americas contain stronger dispute settlement rules, as do agreements among Asian partners, contradicting widespread perceptions of a litigation-averse “Asian way.” In sum, treaty features such as depth and membership size play a major role in pushing all states toward stronger dispute settlement rules, but power constellations and regional differences also shape dispute settlement design.

---

<sup>3</sup> One current example is the investor-state dispute settlement provisions that were widely inserted in bilateral investment treaties (BITs). They subsequently were used more frequently than anticipated, in a manner that imposed significant costs on states, thus provoking a recent backlash (see Simmons 2014).

## 2 Theoretical foundations

### 2.1 A wider lens on international dispute settlement

Much scholarship on international dispute settlement focuses quite logically on the record of prominent courts and tribunals. Numerous studies have investigated the workings of bodies like the International Court of Justice (e.g., Paulson 2004), ECJ (e.g., Stone Sweet and Brunell 1999), DSM of the WTO (e.g., Bown and Pauwelyn 2010) as well as the Andean Tribunal of Justice (e.g., Alter and Helfer 2010) and NAFTA dispute settlement panels (e.g., De Mestral 2006). These institutions deserve such attention: they are active, address important disputes, and can impose costs on non-compliant states. But they represent the tip of the iceberg among international dispute settlement institutions – the most visible and active of such bodies in a global ocean of dispute settlement venues. It is important to consider these bodies in a broader context to avoid drawing conclusions that are either biased or perhaps overly optimistic.

One way to broaden our view of international dispute settlement is to consider all of the other, sometimes lesser-known, dispute settlement institutions to which states may turn. Many such options exist, generating what Eric Posner (2009, 150) has called a “...bewildering jungle of judicial, quasi-judicial, and advisory bodies, some global and some regional or bilateral, with overlapping jurisdiction and no hierarchical structure to ensure uniformity in the law.” On the surface, this increasing legalization network may be cause for celebration, since it facilitates rule-based resolution of international disputes based on legal principles. But this potential legal fragmentation could produce competing and perhaps contradictory legal regimes (e.g., Raustiala and Victor 2004; Alter and Meunier 2009) and allow states to forum-shop and to exploit differences in state obligations and also interpretation.<sup>4</sup> Thus the concern is that disputes might not be resolved strictly based on some objective rule of law, but instead by the power politics that determines the rules governing dispute settlement and who can select among competing rules (see also Drezner 2013).

This also suggests the need to understand the origins of these various dispute settlement venues and how the rules for their use are written. Thus a second way of broadening our approach to international dispute settlement is to focus not just on the caseload of dispute settlement bodies, but also earlier on the creation of these bodies, as reflected in the treaty provisions specifying the conditions under which they may be used. For our purposes, we cannot fully understand the efficacy of a dispute settlement mechanism without taking into account its design features (Mitchell 1994). For instance, international agreements designed with a (stronger) dispute settlement mechanism may (Peinhardt and Allee 2012; Berger et al. 2010) or may not (Gray and Slapin 2013; Kono 2007) be more effective, or have different effects, than those without such mechanisms (Haftel 2012).

<sup>4</sup> See the growing work on forum-shopping (Busch 2007; Davis 2009), in which states strategically use the forum that is most likely to rule in their favor (e.g., WTO vs. a regional or bilateral DSM).

## 2.2 International institutional design

There has been a recent proliferation of scholarship on international institutional design; that is, the particular rules and characteristics that member-states include within formal international organizations, international treaties, as well as more informal arrangements. All studies are seemingly motivated by the observation that “major institutions are organized in radically different ways” (Koremenos et al. 2001: 761). Early work lays out theoretical claims and establishes important concepts (e.g., Aggarwal 1998; Haftel and Thompson 2006; Koremenos et al. 2001). A few empirical studies explore design variation across a wide range of international institutions (e.g., Hooghe et al. 2014; Koremenos 2007; Koremenos and Betz 2013), but the dominant trend has been toward empirical work that examines institutional design within particular issue areas or types of institutions. Among the more vibrant domains for empirical institutional design scholarship include: international environmental agreements (e.g., Bernauer et al. 2013; Marcoux 2009), regional organizations (e.g., Acharya and Johnston 2007; Haftel 2012, 2013; Smith 2000) and economic institutions ranging from the WTO (Pelc 2009) to PTAs (Dür et al. 2014) to BITs (Allee and Peinhardt 2010, 2014).

Perhaps the most studied design feature of international institutions is dispute settlement provisions, the literature on which is diverse both methodologically and substantively (e.g., Arnold and Rittberger 2013; Rosendorff 2005; Yarbrough and Yarbrough 1997). The recent trend is toward quantitative work on single issues ranging from maritime boundary agreements (Ásgeirsdóttir and Steinwand 2014) to bilateral investment treaties (Allee and Peinhardt 2010, 2014). Regional and preferential trade institutions are pre-eminent within the dispute settlement design literature, with DSM design being the primary (Jo and Namgung 2012; Smith 2000; also Chase et al. 2013; Porges 2011) or secondary (Mansfield and Milner 2012; Haftel 2012, 2013) dependent variable in many studies. From the few studies that look across all types of agreements, a common finding is that trade and/or regional institutions are more likely than security or environmental organizations to have a dispute settlement provision (Koremenos 2007) or to allow for international arbitration or adjudication (Hooghe et al. 2014; Koremenos and Betz 2013). All of this suggests that much of the variation and innovation in dispute settlement design can be found in economic agreements.

## 2.3 Moving beyond legalization

The literature on international legalization (e.g., Goldstein et al. 2000; Keohane et al. 2000) is a logical starting point for thinking about dispute settlement, particularly the legalization sub-component of “delegation” (Abbott et al. 2000). One might claim that “legal delegation” entails allowing disputes to be resolved by third-parties according to legal principles, as compared to through political or diplomatic interactions between the parties themselves.

Correspondingly, empirical studies have leaned heavily on legalization as the singular defining feature of DSMs (e.g., Jo and Namgung 2012; Kono 2007; Smith 2000; also Chase et al. 2013; Morgan 2008; Porges 2011). These studies ask whether an international treaty or agreement allows for “legal” as compared to “political” (or “diplomatic”) settlement of disputes, and thus the resulting empirical measure is

typically a binary or scaled indicator of whether or not a DSM is “legalistic.” As an example of the most thoughtful treatment in the literature, Smith (2000) measures the DSMs of 62 regional organizations on a scale ranging from “diplomacy” to “legalism.” He incorporates five variables on whether the DSM: allows for legal dispute settlement (third-party review), has a permanent or ad hoc body, whether rulings have direct effect, whether they are binding, and whether private actors have standing. Subsequent work draws heavily on Smith in measuring DSM legalism, whether used as an independent (Kono 2007) or dependent (Jo and Namgung 2012) variable.<sup>5</sup>

The above schemes capture DSM legalism well, but we see some limitations with them. First, there is little variation across some legalism components. For example, explicit mentions that awards are not binding are extremely rare; we find only two PTAs (out of 589) in which the parties explicitly say that awards are not legally binding. Second, the few components that comprise legalization overlap considerably. For instance, a 1) permanent dispute settlement body is by definition one that engages in 2) third-party review and that will issue 3) binding judgments.<sup>6</sup> Third, the legalization concept suggests, intentionally or not, a false dichotomy that does not exist. When legal dispute settlement (arbitration or adjudication) is provided for as an option within an international institutional arrangement, it almost always is specified in tandem with consultations or other “out of court” efforts – as a follow-up option in the event diplomatic efforts fail to produce a solution. Thus the “political” (“diplomatic”) and “legal” options are not mutually exclusive, but work as complements. A fourth and related issue is that contemporary international institutions increasingly provide for at least some type of legal dispute settlement.<sup>7</sup> This “move to law” (Goldstein et al. 2000: 385), in fact, is a major part of what sparked the legalization literature. So the conceptual task is not to think about whether a DSM allows for legal dispute settlement, but instead to consider the various elements of such legalism. Thankfully, many empirical legal studies chronicle important variation across legal dispute settlement mechanisms such as supporting infrastructure, requirements for panelists, decision-making, appeals, and implementation (e.g., Allee and Elsig 2015; Chase et al. 2013; Porges 2011; also Bartels and Ortino 2006; Donaldson and Lester 2009). Therefore, we build on the idea of legalization, but use it primarily as a launch pad for creating a richer conceptualization of dispute settlement mechanisms.

Likewise, we argue that DSMs should be thought of as potential enforcement devices for which the “strength” of the DSM is central. Legalization is a useful typology, but the main purpose of designing a DSM is not to achieve legalization, but to achieve compliance with the obligations enshrined in agreements. Indeed, trade agreements are commonly depicted as prisoners’ dilemmas in which preventing defection is paramount (Yarbrough and Yarbrough 1992, 1997). Many persuasive accounts of WTO dispute settlement portray the DSM as an enforcement body (Maggi 1999; Schropp 2009; also Sattler and Bernauer 2011). Similar dynamics characterize human

<sup>5</sup> Jo and Namgung (2012) collapse Smith’s legalism scale to three components – looking only at the presence of legal dispute settlement (third party review), bindingness of rulings, and the permanence of the body. Kono (2007) uses the same coding as Smith.

<sup>6</sup> From purely a methodological standpoint, one way to deal with this concern is by using factor analysis (Hooghe et al. 2014).

<sup>7</sup> Indeed, our data show that the sizeable majority of PTA signed over the past three decades allow for one or more forms of legal dispute settlement.

rights and environmental agreements, for which the issue of compliance (with the obligations enshrined in a treaty) is crucial (von Stein 2012). Therefore, a major purpose of a DSM is to encourage adherence to an agreement. This, and effective dispute settlement more generally, will be more likely when the DSM is “strong”; that is, provides tools to aggrieved parties to successfully challenge the potential non-compliance of other members.

### 3 Theory and hypotheses

Our explanation for dispute settlement design is informed by existing scholarship, yet we put forward a framework that both draws from and reconciles tensions among existing perspectives. We see a split in the current institutional design literature between rationalist approaches, which treat states as mostly undifferentiated actors that will react similarly to characteristics of the treaty environment, and various other approaches that assume institutional design primarily reflects the individual or collective attributes of the actors – whether they are derived from power, regional identities, or domestic factors. At present the various perspectives ignore, de-emphasize, or criticize the others. We find merit in each as an explanation for DSM design, and believe they can be reinforcing and complementary.

Nevertheless, we posit a hierarchy among institutional design influences. Our overall claim is that DSM design is first and foremost a rational response to treaty characteristics. Thus we expect factors such as agreement depth or membership size to strongly predict dispute settlement design. Yet against this rationalist baseline, powerful states should be able to shape design features to meet their preferences, particularly in asymmetric agreements. Finally, room should exist for regional forces to shape agreement design. In the following sections, then, we articulate the logic of rationalist, power-based, and regional explanations for dispute settlement design and identify the hypotheses that emerge from each approach.

#### 3.1 Rational response to agreement features

One prominent perspective toward international institutional design portrays agreement-design outcomes as a rational response to the nature of the agreement. Some agreements have more members or entail greater obligations, and thus member-states in these settings will design a strong DSM to aid in enforcement. We agree with the basic thrust of this claim, and expect it to explain much of the variation in DSM design. One literature that espouses this type of logic is scholarship on the “rational design of international institutions,” in which institutional design outcomes reflect various “cooperation problems” to which all states will react similarly (Koremenos et al. 2001). Even more directly relevant is game-theoretic work on treaty compliance from the “enforcement” school (e.g., Downs et al. 1996). It maintains that strong DSMs are needed to achieve compliance with agreements that require major policy change, and for which the probability of defection is high.

These arguments are not uncontroversial. First, endogeneity is an ever-present concern in studies which posit that one agreement characteristic should affect another – although



this is less of a concern in our case.<sup>8</sup> Moreover, the “rational design” approach has been criticized on many grounds (e.g., Duffield 2003; Wendt 2001), including its ambiguous and unmeasured concepts, neglect of state power, and lack of quantitative empirical tests (but see Koremenos and Betz 2013; also Copelovitch and Putnam 2014). Also notable are critics from the “managerialist” school of treaty compliance, who reject “enforcement” scholars’ claims that states willfully choose to violate agreements and instead argue that non-compliance results from lack of transparency and limited state capacity (e.g., Chayes and Chayes 1993; Victor et al. 1998). In light of these final two criticisms, we make a notable empirical contribution by carefully testing controversial and unresolved arguments about the relationship between key agreement characteristics (depth, membership size) and dispute settlement design.

The first feature of agreements that should lead to stronger dispute settlement is agreement *depth*. Most closely linked with the work of Downs et al. (1996), the logic is that “deeper” international 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 policy changes, are the ones that require some type of strong dispute settlement mechanism. Such mechanisms can compel states 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. Careful tests of this depth-enforcement connection are elusive, and thus our findings should help to arbitrate the long-standing managerial-versus-enforcement debate. The former would predict no discernible relationship between how “deep” the agreement is and the strength of its DSM, while the latter predicts a strong, positive relationship between the two.

H1a International agreements that are deeper; that is, entail greater obligations, should contain stronger dispute settlement mechanisms.

Beyond depth, a second feature of international institutions that might shape dispute settlement design is the membership-size of the agreement. A frequent claim is that international agreements with more members are more likely to receive strong, legal dispute settlement (Yarbrough and Yarbrough 1992; Koremenos 2007). We flesh out the causal link between number and design outcomes, and see several reasons why agreements with more members should have stronger DSMs. First, greater numbers make monitoring more difficult, and stronger DSMs have devices that help to monitor members’ compliance with an agreement. Second, any disputes that arise in a bilateral arrangement should be easier to deal with directly and informally. By contrast, in agreements with multiple members disputes may entail greater complexity and involve irreconcilable demands, and thus be more difficult to resolve informally. Stronger DSMs can facilitate successful dispute resolution in these scenarios. Likewise, compliance with any dispute settlement outcome (formal or informal) may be more difficult in multi-member agreements, and more advanced DSMs have devices (post-award sanctions, timelines, etc.) that can aid in dispute settlement. In addition, in multi-party

<sup>8</sup> We claim that the number of members and overall depth of the agreement should affect DSM design. These first two elements are established during the early stages of talks, while decisions about how to address future disputes over the agreement’s terms are addressed later in the negotiation process.



agreements the need for uniform interpretation increases, providing an incentive to opt for a more institutionalized dispute settlement process that can provide greater consistency. Finally, setting up more elaborate and formalized arbitration systems can be a financial burden for agreement partners, and with greater numbers the individual contribution decreases, thus making stronger DSMs more likely.

H1b International agreements should contain stronger dispute settlement mechanisms as the number of members increases.

### 3.2 Power and asymmetry

The previous arguments emphasize features of the agreement as the primary determinant of institutional design: *what* does the agreement require members to do, and *how many* actors will be doing it? An important implication is that the states that design the DSMs are undifferentiated. By contrast, all other perspectives emphasize, in one form or another, *who* is part of the agreement and what their specific design preferences might be. Although we believe depth and number of members should be strong, primary determinants, we also expect various characteristics of the members to play a notable role in institutional design. The particular argument in this section differentiates states according to power and emphasizes power asymmetries among members, following previous scholarship on the role of power in shaping international institutions (e.g., Gruber 2000; Drezner 2007; Stone 2011).

We believe that powerful states should prefer stronger dispute settlement mechanisms within international institutions – contrary to much of the prevailing conventional wisdom. Indeed, a long line of scholarship suggests that powerful states will be averse to legal trade dispute settlement because they should fare well in disagreements in the absence of legal institutions (e.g., Gomez-Mera and Molinari 2014; Schneider 1999). Although the above logic is sensible, we find several reasons to believe that powerful states might actually support strong, legal dispute settlement – most of which stem from the general point that powerful states can use international institutions to their advantage (e.g., Brewster 2006; Martin 1992; Gruber 2000; Drezner 2007). First, many DSMs allow for both “diplomatic” and “legal” dispute settlement, so powerful states can benefit from having a menu of dispute settlement options that allows them to forum-shop even within treaties. Second, powerful states can select with whom they sign agreements and what is (and is not) contained in the agreement. So the obligations that are subject to powerful dispute settlement are likely to be desirable elements that powerful states want to have enforced.<sup>9</sup> Third, evidence from other studies of international agreements suggests that powerful states have disproportionate influence on the design of dispute settlement mechanisms.<sup>10</sup> Fourth, power still plays a role, even within legal dispute settlement, since any legal dispute settlement ruling must still be implemented, which may depend on a state’s ability to sanction or impose costs on a non-

<sup>9</sup> For instance, a Congressional Research Service (CRS) analysis of the U.S.-South Korea PTAs notes that “(t)he potential scope of KORUS FTA dispute settlement is limited by the scope of the KORUS FTA obligations that would be taken on by the Parties” (Grimmett 2012, 2).

<sup>10</sup> Gruber (2000) discusses this issue in the context of NAFTA dispute settlement. Also see Allee and Peinhardt (2010, 2014) on the role of powerful states in the design of DSMs in BITs.

compliant party (Bown 2004). For all of the above reasons, then, we expect powerful states to be more likely to include strong dispute settlement in their international agreements.

H2a International agreements that include a powerful state should contain stronger dispute settlement mechanisms.

The preceding discussion of powerful states raises the issue of how broader power dynamics among agreement members might affect institutional design outcomes. Agreements can involve very similar groups of countries, or countries that are quite different. We argue that countries in more asymmetric agreements should prefer stronger DSMs, since it will appeal to states at both ends of the power imbalance. For powerful states, all of the preceding arguments are accentuated in asymmetric agreements: they will have even greater control over the terms of the agreement, the ability to forum-shop, and any post-dispute negotiations over implementation. That said, weaker states enter freely into these arrangements as sovereign states. For the weaker actors in these situations, then, the key imperative is how to ensure that powerful partners will abide by the agreement? We argue that stronger DSMs will be attractive in these situations of greater asymmetry. In the absence of a strong, legally-based DSM, weaker actors have little chance of successfully confronting a more powerful partner. But timely, effective, and legalized dispute settlement can “level the playing field” and give weaker actors the ability to contest policies from the wealthier state that it finds objectionable, or to have a general check on the abuse of power. In sum, both stronger and weaker members of agreements should perceive strong DSMs to be desirable. By contrast, in symmetric agreements members will discuss issues on equal terms and should find it easier to arrive at negotiated settlements to potential disputes.

H2b International agreements among members with significant power asymmetries should contain stronger dispute settlement mechanisms.

### 3.3 Regional patterns

Differences across regions, or what Acharya and Johnston (2007: 245) have labeled “regional exceptionalism” in international institutional design, also should be evident. Once again, this approach posits that different types of members – this time those from particular regions – should have identifiable orientations toward DSM design (Acharya and Johnston 2007). We believe these regional effects also should predict institutional design beyond what is explained by agreement depth or membership size. There might be different regional attitudes toward sovereignty or institutionalization based on varying colonial pasts, collective identity, or other regional characteristics. Drawing upon rich literatures on regionalization, we expect three regions – Asia, Africa, and the Americas – to be somewhat unique in terms of how they approach dispute settlement in their agreements.

Perhaps the strongest candidate for an exceptional region is *Asia*, where *ex ante* we expect international agreements to be less likely to include strong dispute settlement. A

diverse body of scholarship maintains that within Asia there seems to be a long-standing hesitation to adopt formal institutions and legalization (e.g., Acharya 1997; Kahler 2000; Khong and Nesadurai 2007; Luo 2006). Thus international agreements within Asia should have weaker DSMs due to the general preference among states in the region for more informal, and less adversarial, methods of resolving disagreements. The explanation for this expected outcome typically revolves around a discussion of Asian culture or values, or perhaps political heterogeneity (see Kahler 2000).

Another region in which we expect weaker dispute settlement provisions is *Africa*. In general, sovereignty-based concerns should lead states to adopt weaker, more diplomatically-based solutions to dispute settlement (Schneider 1999). These concerns with allowing strong, legal dispute settlement should be most pronounced across the African continent, where the overwhelming majority of countries obtained independence only in the past half-century or so. Indeed, some argue that within Africa greater emphasis is placed on flexibility than on enforcing obligations (Gaathi 2010b). Herbst (2007, 130) likewise claims that there is a “clear style” of African cooperation and questions whether “rational design” predictions can fully explain Africa, noting that many African leaders are more concerned with domestic politics than with international appearances.

The story for Latin America is quite different, and thus we expect agreements within the *Americas* to be more likely to contain strong dispute settlement provisions. One reason is that unlike Africa and to a lesser extent Asia, Latin America is largely free of post-colonial, sovereignty-based fears of transferring power to supranational institutions. In fact, the history of cross-border cooperation in Latin America goes back nearly two centuries (Biukovic 2008; Dominguez 2007). Dominguez notes that regional institutions in the western hemisphere date back to the 1820s, and that across the region “(a) relatively thick array of international institutional rules had emerged by the 1930s” (2007, 83). Moreover, the broadly-defined region contains a general political organization (the Organization of American States) as well as a human rights convention, among other institutions, illustrating the importance of international law in the region. For all of these reasons, we expect states across the Americas to be more comfortable with institutionalized, transnational dispute settlement.

- H3a International agreements involving members from Asia are less likely to have strong DSMs.
- H3b International agreements involving members from Africa are less likely to have strong DSMs.
- H3b International agreements involving members from the Americas are more likely to have strong DSMs.

## 4 Data and measurement

### 4.1 PTAs as a laboratory for understanding dispute settlement design

For a set of cases to test the above hypotheses, we turn to the universe of post-war PTAs, which we define as agreements between states or regional organizations that provide reciprocal preferential market access for members’ goods and services. 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), others 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). These varying features provide useful variation on the dimensions we believe should affect dispute settlement design, such as size of membership, depth, and region. Moreover, by focusing on a consistent type of agreement, albeit one with considerable variation, we hold constant confounding variables associated with issue area, since too much incompatibility across cases can lead to simplistic explanations or garbage-can-type regression models.

We select PTAs not only because they vary on important dimensions, but because they are widespread, policy-relevant, significantly affect trade relations, and have generated thousands of real-world disputes. In the face of continued stagnation at the multilateral level, PTAs are now viewed by many as the pre-eminent institutions for governing global trade for the foreseeable future. All WTO members have officially concluded a PTA, and most governments participate in many, if not dozens of, agreements. Furthermore, they continue to evolve, as mega-PTAs such as the Trans-Pacific Partnership (TPP) and the US-EU Trade and Investment Partnership (TTIP) currently dominate the trade agenda. All told we code the treaty texts of 589 PTAs signed between 1945 and 2009. This list of treaties is drawn from Dür et al. (2014) and combines several lists from international organizations (WTO, Organization of American States) and is complemented by searches of trade and economic ministries' websites.<sup>11</sup>

It is important to note that DSMs are widely used after they are designed. Overall dispute totals are difficult to come by, but recent tallies of PTA-related disputes in the Western Hemisphere number in the thousands.<sup>12</sup> Among individual agreements, NAFTA has three different dispute settlement mechanisms – chapter 11 (investment), chapter 19 (antidumping and countervailing duties), and chapter 20 (general) – and 205 disputes have taken before the agreements' mechanisms through 2014.<sup>13</sup> Likewise, disputes have been initiated against Costa Rica, the Dominican Republic, El Salvador, Guatemala, and the United States as part of DR-CAFTA's investor-state dispute settlement provisions.<sup>14</sup> Various disputes have been taken before lesser-known bodies, such as the court of the Economic and Monetary Community of Central Africa (CEMAC) and the dispute settlement procedures of the Latin American Integration

<sup>11</sup> The project identifies just over 700 post-war treaties, but the number of coded treaties is 589 due to treaty text availability.

<sup>12</sup> Laks-Hutnick (2013), for instance, identifies a total 1,207 trade disputes that were taken before the DSMs of four prominent Western Hemisphere PTAs (North American Free Trade Agreement, NAFTA; the Central American Common Market, CACM; the Common Market of the South, MERCOSUR; Andean Community) between 1995 and 2010.

<sup>13</sup> <https://www.nafta-sec-alena.org/Home/Dispute-Settlement/Status-Report-of-Panel-Proceedings> (last accessed 25 March 2015). For a discussion of the mechanisms and their performance, see De Mestral 2006.

<sup>14</sup> The list of such cases is on the U.S. State Department's website: <http://www.state.gov/s/l/c33165.htm> (last accessed 25 March 2015).

Association (ALADI).<sup>15</sup> Furthermore, the actual number of total PTA disputes goes well beyond the identifiable examples above because many disputes are dealt with through informal procedures *that are specified in DSMs*, but end up being hidden from public view and do not get registered formally.<sup>16</sup>

## 4.2 Conceptualization and measurement of strong dispute settlement mechanisms

We assert that stronger DSMs are those containing legal procedures that provide for speedy litigation free from delay, allow complainants to drive the legal process and make important choices, and facilitate implementation of legal awards.<sup>17</sup> Drawing on established work on international arbitration, adjudication, and courts, we compile precise data on six components that reflect how “strong” or “weak” a particular DSM of a PTA is. The first of these follows directly from existing scholarship on DSM design, but the next five are newly created and represent an important original contribution. These six components are then combined into a primary and two alternate indicators of DSM strength, which are then used in our empirical tests.

The first of six components captures the extent to which dispute settlement authority is delegated to a third-party, legal body – as consistent with existing scholarship on DSM “legalization” in international agreements (e.g., Jo and Namgung 2012; Porges 2011; Smith 2000). The lowest category in the three-category “legal delegation” scale (coded with a 0) includes the scenario in which there are no provisions for dispute settlement or in which dispute settlement provisions exist but specify only consultations and/or mediation. By contrast, a value of 1 is assigned to PTAs that utilize only ad hoc arbitration as a legal dispute settlement option, whereas the largest value of 2 is reserved for PTAs that create a standing body for dispute settlement. The rationale for this last distinction is that by creating a standing body, PTA members delegate substantial authority to a more autonomous, established actor with greater resources, which helps with the carrying out of proceedings and the enforcement of awards. Table 1a contains the complete distribution for this *legal delegation* variable, which shows that ad hoc arbitration is the most commonly specified legal dispute settlement option.

The second component emphasizes the ability of a complainant state to choose the dispute settlement venue. PTAs vary greatly in their language on forum choice (Donaldson and Lester 2009, 378–381). The ability to choose the venue that is viewed as most desirable, or in rare cases to pursue a claim in multiple venues, can be quite advantageous to complainants, particularly powerful ones. PTAs are given a 0 for this *complainant forum choice* variable 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

<sup>15</sup> See Porges 2011, 473, on the former and Chase et al. 2013, 47, on the latter.

<sup>16</sup> For instance, Garcia Berceiro (2006) notes that many trade disputes involving the E.U. are addressed through the various consultations mechanisms that are included in E.U. PTAs. See also Luo (2006, 443) on disputes within the Association of Southeast Asian Nations (ASEAN) and Dominguez (2007) on disputes among CACM members.

<sup>17</sup> We posit that stronger DSMs, which include these design features, should enhance treaty enforcement. Note that our dependent variable measures DSM strength; it is not a measure of actual treaty enforcement or dispute activity.

**Table 1a** Variation in legal delegation across PTAs

Delegation to a legal third party	Number of treaties
0 (none, consultations)	341
1 (ad hoc arbitration)	217
2 (standing body)	31
Total	589

**Table 1b** Variation in complainant forum choice across PTAs

Complainant forum choice	Number of treaties
0 (none)	485
1 (yes, fork-in-the-road)	100
2 (yes, w/o restrictions)	4
Total	589

**Table 1c** Variation in panel chairman selection across PTAs

Selecting panel chairman	Number of treaties
0 (none, bilateral consult)	429
1 (arbitrators select only)	15
2 (by lot or third party)	145
Total	589

**Table 1d** Variation in time limits across PTAs

Time limits	Number of treaties
0 (no)	368
1 (yes)	221
Total	589

**Table 1e** Variation in post-award sanctions across PTAs

Post-award sanctions	Number of treaties
0 (none)	427
1	7
2	60
3	86
4 (greatest # of tools)	9
Total	589

**Table 1f** Variation in areas excluded from dispute settlement across PTAs

Comprehensiveness	Number of treaties
0 (some areas excluded)	420
1 (no areas excluded)	169
Total	589

**Table 1g** Variation across dispute settlement mechanisms in PTAs (Larger numbers indicate stronger dispute settlement)

Overall strength of dispute settlement mechanism	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

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 rare, is when the complainant chooses the venue and there are no restrictions on the use of multiple fora.

A third element is the composition of the judicial body, particularly the issue of *chairman selection*, since judicial panels take decisions by majority and the chair often plays a pivotal role (e.g., Chase et al. 2013; Donaldson and Lester 2009; Porges 2011; Posner 2009; Posner and Yoo 2005). Effective dispute settlement is more likely to occur when there is an unbiased chair or “umpire” (Posner 2009) that is not beholden to the interests of the state parties. Disagreements about panel composition – particularly the selection of the chair – can paralyze proceedings and hinder effective dispute settlement (Chase et al. 2013; Donaldson and Lester 2009; Porges 2011). Defendants in these scenarios have incentives to block or delay, since slowing down dispute settlement proceedings works to their advantage. By contrast, dispute settlement will proceed more efficiently when there are clearly-specified, apolitical rules for selecting a panel chairman. Four options for selecting the chair exist: 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 dispute settlement the most, since the selection process is faster and less subject to pressure by the respondent state. 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. Next in line are the scenarios in which the party-appointed arbitrators select the chairman (coded as 1). All remaining PTAs, coded as 0, specify only bilateral consultations as the method of selection or have no legal dispute settlement.

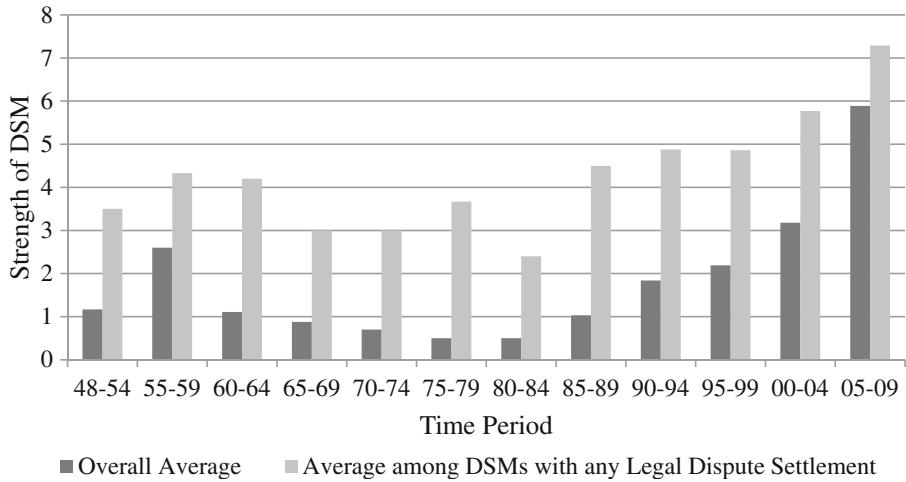
A fourth component captures whether the DSP in a given treaty provides *time limits* for the dispute settlement process, whether overall and/or for particular stages (pre- and post-award). Some of the logic is similar to that for chairman selection, and the omission of both temporal dynamics is a limitation of earlier studies. The specification of time frames encourages a faster dispute settlement process and thus should enhance compliance with obligations. A total of 221 agreements specify time limits (see Table 1d).

The fifth ingredient of a strong DSM is 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 (e.g., Bown and Pauwelyn 2010; Zangl 2008) but are absent from existing empirical studies of PTA DSMs. 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-implementation, 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.<sup>18</sup> The second indicator measures whether the complainant can choose the level of retaliation, which is allowed in 151 PTAs – a large majority 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. 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, do the provisions 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 DSMs, then, are those that lack any such exemptions. This occurs in 169 PTAs, as depicted by “1 s” in Table 1f.

---

<sup>18</sup> We do not count cases where no arbitration or adjudication is specified, but the dispute settlement language allows for trade remedies.



**Fig. 1** Average strength of dispute settlement mechanisms in PTAs over time

We sum these six components to create a simple, additive index of the *strength of dispute settlement* for all 589 PTAs in our data set. We utilize this as our primary measure because it is the most straightforward, transparent, and intuitive of the three slightly different measures we create. This resulting 0–9 index serves as the primary dependent variable in our empirical tests, and Table 1g shows the distribution of this variable across the universe of post-war PTAs.<sup>19</sup>

Figure 1 shows the evolution of DSM strength over the post-war period using this same, primary indicator for DSM strength. Overall there is a clear upward trend, which reflects the general trend toward increased international judicialization over time. As a result, we estimate all of our multiple regression models with controls for 5-year time periods. Nevertheless, we also stress the considerable variation within each decade, each of which contains both weak and strong DSMs. Moreover, we emphasize that many early PTAs also had quite-strong DSMs.

Finally, as the first of two alternate outcome variables used to check for robustness of findings, we also create a more standardized six-category index that forces all subcomponents to have equal weight. In this case we take the six components depicted in Table 1(a to f), but instead of adding the raw values for each we standardize them all on a 0–1 scale, resulting in a more fine-tuned indicator that ranges from 0 to 6.<sup>20</sup> As a final method of measuring the overall strength of dispute settlement, we also combine all of the relevant variables using principal components analysis.<sup>21</sup>

<sup>19</sup> In principle this index can range from 0 to 12, but in reality the largest value maxes out at 9.

<sup>20</sup> 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.

<sup>21</sup> One clear factor emerges, with an Eigenvalue of 3.37. The loadings for the six variables that comprise the factor are as follows: delegation (0.81), forum choice (0.70), chairman selection (0.78), time limits (0.81), sanctions (0.82), and lack of exemptions (0.53).

### 4.3 Conceptualization and measurement of independent variables

In this section we discuss measurement of explanatory variables, including those for primary hypotheses and control variables. We begin by specifying indicators for the easier-to-measure hypotheses from part 3. We then discuss in detail our measures for more complex concepts such as agreement depth and power before introducing some additional control variables. Measuring *number* (*H1b*) is straightforward, as we include a simple count of the number of members of the agreements.<sup>22</sup> Likewise, the measures for region are simple dummy variables that capture whether all PTA members are located in each of the three geographic regions we have singled out: *Asia* (*H3a*), *Africa* (*H3b*), and the *Americas* (*H3c*).<sup>23</sup> This leaves as the comparison group those agreements that span multiple regions and non-EU PTAs in Europe (see below), along with a few agreements in Oceania.

Substantial effort is devoted to collection and measurement of agreement *depth* (*H1a*). Downs et al. (1996, 383) define depth as “the extent to which (an agreement) requires states to depart from what they would have done in its absence.” In the PTA context, “depth” should capture the degree to which commitments are made that can potentially lead to market opening and an increase in the exchange of goods, services and investments. Our operationalization of depth reflects that the effects of depth can be direct (in areas such as tariff liberalization for goods, trade in services, and rules allowing foreign firms to bid for public procurement tenders) as well as indirect (promoting regulatory convergence or harmonizing standards). Given the centrality of this concept and different possible approaches to measurement, we utilize multiple indicators to capture PTA depth, although our findings end up being robust to the decision of which to include. Our primary measure is a simple additive index that potentially ranges from 0 to 7. It tallies (1) whether the PTA reduces substantially all tariffs to zero and (2–7) whether it has a substantive provision 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.<sup>24</sup> A second indicator is a more fine-tuned count of the above, which ranges from 0 to 48 and captures whether the agreement contains certain sub-provisions within each of the above areas. A third indicator uses exploratory factor analysis to combine information across the aforementioned 48 components. Finally, we add a measure of depth that prioritizes the 18 components pertaining to domestic regulation, with the logic being that such commitments to change regulatory rules are particularly difficult to observe, monitor, and enforce – and thus are particularly likely to require stronger dispute settlement.<sup>25</sup>

<sup>22</sup> For purposes of robustness, we also consider various transformations of this variable (log, quadratic) as well as a simple distinction between bilateral (two) and plurilateral (three or more) agreements.

<sup>23</sup> We follow the United Nations classification for matching countries to continents. See <http://unstats.un.org/unsd/methods/m49/m49regin.htm>, last accessed on 31 August 2014.

<sup>24</sup> This, and the next two, indicators of depth are drawn from Dür et al. (2014). For details on the components for the various depth measures, see Appendix 1 of our online appendix, which is available at the journal's website (<http://www.springer.com/social-sciences/journal/11558>).

<sup>25</sup> See Appendix 1 for details.

For the measurement of *H3a*, which singles out powerful states as being more likely to design strong DSMs, we create and include in the model three distinct indicators for the presence of a powerful actor in a given PTA. Most obvious are agreements that include the *United States*. The argument for the U.S. is particularly compelling. It is a highly legalized society with a litigious culture; it was the main actor behind the creation of the WTO DSM in the Uruguay Round (Elsig and Eckhardt 2015), and it has advocated strong dispute settlement within other economic agreements, such as BITs. Similarly, we also include a dummy variable for all agreements involving the *European Union* – the other post-war economic giant – although our expectation for the E.U. is perhaps less definitive. One reason is because of the peculiarities of E.U. trade policy, in particular the multitude of non-economic interests the E.U. pursues with a substantial number of PTAs. Finally, we also include a dummy variable for any agreement that includes a member of the modern-day *Group of 20*, or “G20.” This serves as a much broader measure for the design effects of powerful states, and allows us to compare and contrast agreements involving rising powers such as China and Brazil with the more traditional powers (US, EU). All three indicators are included in our multiple regression models.

To test the related hypothesis about *power asymmetry (H2b)*, we utilize as our primary measure a dummy variable to indicate agreements that involve a country from the global “North” and a country from the global “South.”<sup>26</sup> This North–South indicator has the benefit of being intuitive as well as easy to interpret. It also facilitates comparison with other types of membership combinations such as North–North and South–South agreements, which we believe should have weaker dispute settlement than North–South pairings.<sup>27</sup> Nevertheless, as a robustness check we also create and insert a more continuous measure of power asymmetry based on the difference in GDP from the largest member of the PTA (in terms of economic size) and the smallest member, expressed in billions of dollars.

Finally, we also include a set of control variables for other possible influences on DSM design. A first control variable is for regime type, since studies have found that *democratic states* are more likely to embrace legal dispute settlement to settle disputes (e.g., Allee and Huth 2006; Gomez-Mera and Molinari 2014; Sattler and Bernauer 2011). By extension, one would expect that democratic states will be more receptive to the inclusion of strong, legal dispute settlement language in their agreements. We use the average Polity score among all members of the PTA (Jaggers and Gurr 1995). We also control for whether the signatories are members of the GATT/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. Finally, we

<sup>26</sup> We categorize the United States, Canada, Japan, Australia, New Zealand and all countries in Western Europe as the “North.”

<sup>27</sup> 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 in the absence of any benefit from checking the power of a larger state.

**Table 2** Ordered probit results for the strength of dispute settlement in PTAs

	Coefficient	Robust Std. Error
Agreement Characteristics		
Depth (0–7 index)	0.443	(0.051) ***
Number of Members	0.019	(0.005) ***
Asymmetry and Power		
North–South	0.453	(0.135) ***
United States	0.736	(0.312) **
European Union	–0.177	(0.177)
G20 Member	0.075	(0.130)
Regional Effects		
Asia	0.300	(0.173) *
Americas	0.786	(0.152) ***
Africa	0.123	(0.226)
Control Variables		
WTO members	0.478	(0.120) ***
Democracy	–0.005	(0.011)
PTA Import Volume	–4.31	(8.69)

Time-period effects estimated but not reported

$N=571$  Preferential Trade

Agreements

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < .01S$  (two-tailed)

also control for *PTA trade volume*, since PTAs involving larger amounts of trade might necessitate stronger DSMs (Haftel 2012; Hooghe et al. 2014; Jo and Namgung 2012).<sup>28</sup>

## 5 Primary findings

To evaluate the hypotheses discussed above, we first estimate and interpret a primary model, which is presented in Table 2. In addition, we estimate a series of follow-up models to check for sensitivity of results to alternate independent (Tables 3 and 6) and dependent (Appendices 2 and 3) variable specifications. We also explore the substantive effects of variables found to be statistically significant, and present these in Tables 4, 5, 7, and 8. Ordered probit models are estimated in many cases, but we also

<sup>28</sup> Our empirical measure takes the log of the mean value of imports across all members of the PTA (in millions of dollars), but we also consider exports as well as other methods of aggregation. Trade data are drawn from the International Monetary Fund's Direction of Trade Statistics (DOTS, 2010) as well as Gleditsch (2002).

**Table 3** Ordered probit results for the strength of dispute settlement in PTAs

	Alternate indicators of depth		
Agreement characteristics			
Depth (0–48 index)	0.081 (0.009) ***		
Depth (factor)		0.980 (0.108) ***	
Depth (domestic regulation)			0.191 (0.021) ***
Number of Members	0.017 (0.006) ***	0.013 (0.006) **	0.023 (0.007) ***
Asymmetry and power			
North–South	0.467 (0.131) ***	0.544 (0.129) ***	0.589 (0.137) ***
United States	0.529 (0.333)	0.757 (0.344) **	0.801 (0.393) *
European Union	–0.054 (0.179)	–0.038 (0.174)	0.071 (0.180)
G20 Member	0.101 (0.130)	0.171 (0.130)	0.068 (0.132)
Regional effects			
Asia	0.405 (0.176) **	0.551 (0.183) ***	0.377 (0.173) **
Americas	0.630 (0.146) ***	0.801 (0.152) ***	0.629 (0.150) ***
Africa	0.260 (0.224)	0.409 (0.238) *	0.079 (0.240)
Control variables			
WTO Members	0.506 (0.119) ***	0.432 (0.122) ***	0.467 (0.122) *
Democracy	–0.006 (0.011)	–0.014 (0.011)	–0.002 (0.011)
PTA Import Volume	–3.83 (9.81)	–6.72 (7.80)	–7.63 (9.34)

Time-period effects estimated but not reported. Robust standard errors in parentheses

$N=565\text{--}571$  Preferential Trade Agreements

\* $p<0.10$ , \*\* $p<0.05$ , \*\*\* $p<0.01$  (two-tailed)

utilize ordinary least squares (OLS) regression in appropriate situations.<sup>29</sup> Finally, for all models we include controls for time-period.<sup>30</sup>

The findings from our primary model, shown in Table 2, strongly support our overall expectation of a hierarchy of design influences. The dominant pattern is that features of the agreement are powerful, consistent predictors of strong dispute settlement regardless of the identity or location of the signatories. Yet power dynamics also play an important role: asymmetric, or “North–South,” agreements are far more likely to have strong dispute settlement than all types of more balanced agreements, and U.S. PTAs have stronger DSMs, too. Finally, regional patterns also are evident, with Latin American and Asian agreements exhibiting a greater likelihood of employing strong dispute settlement mechanisms.

<sup>29</sup> Our primary dependent variable (see Table 1g) is ordered but not interval or continuous. Thus when using this primary outcome variable we estimate ordered probit models due to worries that OLS would produce biased and misleading estimates (Peel et al. 1998). We also note that our major findings are robust across estimators.

<sup>30</sup> We include dummy variables for 5 years intervals beginning with years ending in “0” or “5” (except for the period 1948–54). As a sensitivity check, we also substitute narrower (yearly) or broader (decade) time-period dummies, which has no impact on our core conclusions.

**Table 4** Predicted probabilities for strength of dispute settlement in PTAs at varying levels of agreement depth

Depth of agreement	Strength of dispute settlement mechanism									
	0 (weakest)	1	2	3	4	5	6	7	8	9 (strongest)
Very Low Depth (0)	0.870	0.025	0.012	0.043	0.009	0.029	0.006	0.008	0.001	0
Low Depth (2)	0.565	0.049	0.026	0.109	0.027	0.115	0.031	0.060	0.018	0.001
Moderate Depth (4)	0.213	0.038	0.022	0.113	0.034	0.186	0.070	0.192	0.119	0.013
High Depth (7)	0.013	0.005	0.003	0.021	0.009	0.071	0.042	0.212	0.414	0.211

From the core findings in Table 2, and across all of the models we estimate, agreement depth is the strongest and most consistent predictor of DSM design. Overall, across the roughly two dozen models we estimate, the agreement depth variable is positively associated with strong dispute settlement, almost always at the 99 % level of confidence. The finding is notable given the ongoing, aforementioned, debate about whether compliance with agreements should be “managed” or “enforced.” Our empirical findings support the enforcement claim. We find that governments certainly believe that their deeper agreements require stronger dispute settlement, and they design them accordingly.

This strong finding holds regardless of how we conceptualize and measure agreement depth, as Table 3 shows. For instance, similar findings are returned when we substitute the much broader indicator of depth based on the full universe of 48 depth-related indicators (column 1). The results also are nearly identical when we substitute the 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 (column 3), which means that PTAs that require stronger regulatory convergence also receive strong dispute settlement.

To explore the substantive impact of agreement depth, we examine the likelihood of each of the outcomes of DSM strength that comprise our primary dependent variable (and range in increasing order from 0 to 9).<sup>31</sup> Table 4 indicates that for the shallowest PTAs (“very low depth”), agreements have less than a 1 % chance of having a “strong” DSM, defined as a mechanism that lies between 7 and 9 on our 0–9 dependent variable scale. “Low depth” agreements have only an 8 % chance of falling into this range, and are still most likely (61 %) to have the weakest possible DSMs (0 or 1 on the DSM scale). “Moderate depth” agreements become likely to have somewhat stronger dispute settlement procedures, with only a 25 % likelihood of having the weakest DSMs and more than a 60 % having a moderately-strong DSM – defined as lying between 3 and 7 on our 0–9 scale. Finally, Table 4 shows how “high depth” agreements are very likely to have strong DSMs; in fact, they have more than an 80 % predicted likelihood of falling into the highest possible (7–9) range on our scale. By contrast, these deep agreements have only a 2 % predicted chance of falling into one of the weaker DSM categories (0–2).

<sup>31</sup> To do so, we use in Stata 12 the *prvalue* post-estimation command from the *SPost* package (Long and Freese 2014). Unless specified otherwise, we hold all variables at their median values and use 2000 as the year and set depth=3. For tractability, these calculations draw upon a variant of the primary model in which a year variable is inserted in place of fixed time-period effects.



**Table 5** Predicted probabilities for strength of dispute settlement in PTAs at varying levels of membership size

Number of members	Strength of dispute settlement mechanism									
	0 (weakest)	1	2	3	4	5	6	7	8	9 (strongest)
2 Members	0.375	0.048	0.027	0.124	0.034	0.164	0.053	0.120	0.052	0.003
7 Members	0.350	0.047	0.027	0.124	0.035	0.169	0.056	0.130	0.059	0.004
14 Members	0.315	0.046	0.026	0.123	0.035	0.175	0.060	0.144	0.071	0.005
28 Members	0.250	0.041	0.024	0.118	0.035	0.184	0.069	0.174	0.098	0.009
50 Members	0.164	0.033	0.020	0.102	0.032	0.184	0.074	0.217	0.155	0.020

The number of states in a PTA is another feature of international agreements that is a consistent predictor of DSM strength. In Table 2 the *number of members* variable is positively associated with strong dispute settlement at the highest (99 %) level of confidence. Taken together, then, agreement depth and membership size serve as important forces that lead to strong dispute settlement. From a substantive standpoint, the effects of membership size are notable but not particularly substantial. Table 5 shows that compared to bilateral agreements, agreements with 7 and 14 members are about 10 and 20 % more likely to have very strong dispute settlement (8 or 9 on the DSM strength scale) and similarly less likely to have weak dispute settlement (0 or 1 on DSM strength). Agreements with particularly large memberships (50 members) become about twice as likely to have strong DSMs (7–9 on the scale) than bilateral agreements. In some follow-up analyses, we find that the distinction between bilateral (two) and plurilateral (greater than two) agreements is consistently important, but that there are somewhat weak and diminishing effects to adding more members.<sup>32</sup> Thus, although it is clear that agreement features are important and establish a clear baseline for DSM design-strength, depth is a stronger driving force than membership size.

Although membership size and depth, in particular, are strong determinants of dispute settlement design for all states, power dynamics also play an important role. Notable and robust is the finding that agreements with power asymmetries are more likely to contain strong DSMs. The coefficient estimate for asymmetric agreements in Table 2, reflected by the North–South dummy variable, is positive and significant at the 99 % level of confidence. In fact, this relationship is positive across all of the models we estimate, almost always at the same 99 % level. Moreover, this finding also holds when we substitute an alternate measure for asymmetry based on the difference in GDP from the richest to the poorest member of the agreement (see Table 6, column 1). Likewise, when compared to all types of more symmetric agreements, it becomes clear that asymmetric agreements are more disposed to stronger DSMs. In the second column in Table 6 we substitute dummy variables for *North-North* and *South-South* agreements

<sup>32</sup> We substitute a plurilateral (vs. bilateral) variable and find that it is always a positive and statistically significant predictor of DSM strength, with or without the *number of members* variable. We also find that the log of *number of members* is positive and significant when substituted in place of the untransformed *number of members* variable.

**Table 6** Ordered probit results for the strength of dispute settlement in PTAs

	Alternate indicators for asymmetry	
Agreement characteristics		
Depth (0–7 index)	0.463 (0.050) ***	0.450 (0.050) ***
Number of Members	0.023 (0.005) ***	0.019 (0.005) ***
Asymmetry and power		
Difference in GDP	0.053 (0.023) ***	
North–North		–0.648 (0.193) ***
South–South		–0.372 (0.164) **
United States	0.284 (0.199)	0.774 (0.312) **
European Union	–0.072 (0.180)	–0.123 (0.184)
G20 Member	0.018 (0.131)	0.079 (0.130)
Regional effects		
Asia	0.222 (0.169)	0.277 (0.178)
Americas	0.716 (0.145) ***	0.744 (0.161) ***
Africa	0.032 (0.224)	0.091 (0.228)
Control variables		
WTO Members	0.441 (0.122) ***	0.492 (0.121) ***
Democracy	–0.002 (0.011)	–0.002 (0.011)
PTA Import Volume	–9.46 (9.48)	–2.67 (9.22)

Time-period effects estimated but not reported. Robust standard errors in parentheses

$N=571$  Preferential Trade Agreements

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed)

and both return negative coefficient estimates, which can be contrasted with that model's baseline category of *North–South* agreements. In sum, among all power constellations, asymmetric agreements are much more likely to provide for strong dispute settlement. From a substantive standpoint, Table 7 shows that such agreements are about twice as likely as other pairings to have strong dispute settlement (8 or 9 for DSM strength) and about 50 % less likely to have weak DSMs (0–2 for DSM strength).

There is some – albeit less – support for the related claim that powerful states will prefer strong DSMs regardless of the other agreements partners (*H2a*). The most noteworthy finding is that PTAs involving the United States are more likely

**Table 7** Predicted probabilities for strength of dispute settlement in PTAs for asymmetric vs. symmetric agreements

Type of agreement	Strength of dispute settlement mechanism									
	0 (weakest)	1	2	3	4	5	6	7	8	9 (strongest)
Other pairings	0.376	0.048	0.027	0.124	0.034	0.164	0.053	0.120	0.052	0.003
North–South	0.229	0.040	0.023	0.115	0.034	0.185	0.069	0.184	0.110	0.011

**Table 8** Predicted probabilities for strength of dispute settlement in PTAs for various countries and regions

Who is involved in PTA?	Strength of dispute settlement mechanism									
	0 (weakest)	1	2	3	4	5	6	7	8	9 (strongest)
Baseline country/ region	0.376	0.048	0.027	0.124	0.034	0.164	0.053	0.120	0.052	0.003
United States	0.211	0.038	0.022	0.112	0.034	0.186	0.070	0.194	0.122	0.013
Americas	0.154	0.032	0.019	0.100	0.031	0.183	0.074	0.222	0.163	0.021
Asia	0.252	0.042	0.024	0.118	0.035	0.184	0.067	0.173	0.097	0.009

to have strong dispute settlement, *ceteris paribus*. Earlier we argued that the U.S. should be amenable to strong enforcement due to its selectivity in choosing with whom it signs PTAs and the obligations to which it agrees, and well as its overall comfort with litigation. This appears to be the case, as evidenced by the finding in Table 2 in which U.S. agreements are positively associated with strong dispute settlement provisions at the 99 % level of confidence. This relationship is generally strong, although with some qualifiers, as we discuss in the next section. The predicted probabilities displayed in Table 8 show that U.S. PTAs are nearly twice as likely to have strong dispute settlement (DSM strength of 7, 8, or 9) than PTAs involving other countries.

By contrast, we find no evidence that other powerful actors include stronger DSMs in their agreements. The coefficient estimate on the *European Union* variable in Table 2 is not statistically significant, but is negative, as it is across most models we estimate. Upon further exploration, it seems that at least until recently, the E.U. approach to PTA dispute settlement has conformed to the conventional argument about powerful actors' aversion to legally-based dispute settlement.<sup>33</sup> Finally, the *G20* variable exhibits no relationship with DSM design, indicating that powerful states outside of the U.S. exhibit no clear propensity toward strong versus weak dispute settlement.

Several regional patterns also emerge in Table 2, in ways both expected and unexpected. The strongest and most consistent finding is that agreements in the Americas are more likely to have strong dispute settlement procedures. The *Americas* variable is positive and statistically significant at the 99 % level in Table 2, and in nearly all of the sensitivity checks we conduct. Agreements in the Americas have a 65 % likelihood of having at least moderately-strong dispute settlement, defined as 5 or higher on the 0–9 scale (see Table 8). At the low end of the dispute settlement strength spectrum, PTAs in the Americas are less than half as likely (19 % versus 43 %) as

<sup>33</sup> According to some, the E.U. long took a more skeptical position toward legally-based dispute settlement because it felt that encouraging diplomatic dispute settlement with its PTA partners would better serve its interests (e.g., Broude 2004). European Commission official Ignacio Garcia Bercero (2006) argues somewhat differently, claiming that the Community's negative experiences with GATT dispute settlement shaped its preference for diplomatic dispute settlement in its PTAs. Regardless of the reason, close observers now claim that the E.U. has become more receptive to strong, legal dispute settlement in recent years and recent PTAs (e.g., Garcia Bercero 2006; Porges 2011). We find some evidence of such a shift in our design data beginning around the year 2000.

agreements from other regions to have one of the weakest levels of dispute settlement (0 or 1).

Perhaps the most striking finding from our analyses is that agreements in Asia are actually more – not less – likely to have strong dispute settlement mechanisms. The *Asia* variable in our primary model (Table 2) is positive and significant at the 90 % level of confidence. From a substantive standpoint, PTAs in Asia are about 60 % more likely to be in the upper three categories (7–9) of dispute settlement strength than agreements outside Asia.<sup>34</sup> This finding suggests a few reassessments of the prevailing conventional wisdom. One is a rejection of the claim that legalization is weak in Asia due to Asian culture or some inherent regional characteristic. We find merit in Kahler’s (2000) conjecture that any past hesitation by governments in Asia to embrace legalization is not so much due to culture or “Asian values” but instead likely reflects conscious choices made by governments. Moreover, we also draw a distinction between DSM design and DSM use, which can work in symbiotic ways. It may be that governments in Asia have been receptive to the inclusion of legal dispute settlement provisions, even if they have not fully litigated many disputes.<sup>35</sup>

Findings for other variables also deserve a brief mention. Most notably, the evidence suggests that WTO members are more likely to include strong dispute settlement in their PTAs (see Table 2). This finding is quite robust, and likely reflects these states’ familiarity and comfort with legal dispute settlement. In contrast, we find that democratic signatories are neither more nor less likely to include strong DSMs.<sup>36</sup> Moreover, we explore several related variables, such as those for transitioning democracies and strong law-and-order states, and continually find little evidence that these features affect dispute settlement design. Finally, there is no support for the idea that PTAs with greater trade volume have strong dispute settlement once other, more salient, features are taken into account.

We also conduct a series of sensitivity checks using alternate estimators, as well as independent and dependent variables, to assess the stability of findings and to explore further patterns that are suggested by the findings thus far. We first consider different estimators and estimation choices, many of which we touched upon earlier, and present them in online Appendix 2.<sup>37</sup> First we estimate the primary model in Table 2 using OLS and then binary probit, both of which return very similar results to the original ordered probit model.<sup>38</sup> Next we also use OLS to examine the robustness of findings to two of the alternate dependent variables that we discussed earlier: the standardized version of our six-category indicator and the one created using principal components analysis. Appendix 2 also displays the findings from these two models. In general the original

<sup>34</sup> Additionally, we are quite confident in the validity of our findings: there are over 75 Asian PTAs in our data set, and our measure of dispute settlement strength is carefully constructed and multi-faceted.

<sup>35</sup> In the course of our research we have seen references to several disputes in Asian PTAs that were filed before arbitration or adjudication mechanisms, but later were withdrawn, presumably due to an “out of court” settlement. See also Luo (2006: 443).

<sup>36</sup> This is potentially consistent with Jo and Namgung’s (2012) claim that democracies prefer “medium” levels of legalism.

<sup>37</sup> Note that we also explored various selection models, with DSM design as the second (outcome) stage, but in all instances the lack of statistical significance of rho, the correlation between the first and second-stage estimations, led us to reject this approach.

<sup>38</sup> All primary findings and conclusions are unchanged, with the only exception being that the estimate for the U.S. dummy variable remains positive but is not statistically significant.

findings (from Table 2) are overwhelmingly upheld in both models. The findings for most significant variables, including *Depth*, *Number of Members*, *Asymmetry (North–South)*, *Americas*, *Asia*, and *WTO Membership*, are virtually unchanged. The only noteworthy change is that in the two models in Appendix 2 PTAs involving the U.S. are no longer associated with stronger dispute settlement. This lack of support for a previously strong finding is puzzling, and motivates us to delve further into the nuances of our dependent variable.

Although we believe strongly in using a composite measure of the overall strength of dispute settlement, as a final step we unpack our dependent variable and run separate regression models for each of the six dependent variable subcomponents. This allows us to evaluate micro-level relationships between each explanatory variable and each aspect of strong dispute settlement. Appendix 3 presents the results of this endeavor in a six-column table, the contents of which: support our previous findings generally, reveal some important and logical nuances, and clarify the anomalous findings above about U.S. PTAs. A major takeaway from Appendix 3 is that many of the relationships are robust across all components of DSM strength. For instance, the two dominant agreement-feature variables – *depth* and to a lesser extent *number of members* – consistently exhibit a positive and statistically significant relationship to all subcomponents of DSM strength. The same is true for the *Americas*, and to a slightly lesser extent, asymmetric (*North–South*) agreements. These findings increase our already firm belief in the positive relationship between these characteristics and strong dispute settlement in PTAs.

Some other relationships in Appendix 3 largely bolster the pre-existing conclusions, yet also reveal some interesting patterns. Asian agreements are positively associated with all six subcomponents, although not all of the relationships reach conventional levels of statistical significance. E.U. PTAs are negatively associated (at statistically significant levels) with three of the subcomponents (*forum choice*, *panel chair*, *time limits*) and this somewhat mixed pattern is consistent with our earlier interpretation of E.U. PTAs. Finally, the patterns for U.S. PTAs are similarly mixed. U.S. PTAs are positively associated with three of the six subcomponents at the 90 % or greater level of confidence (*forum choice*, *panel chair*, *sanctions*), yet in three other cases there is no discernible relationship. Nevertheless, the particular features of strong dispute settlement the U.S. tends to employ make logical sense. For instance, U.S. agreements typically allow for sanctions in cases of non-implementation, which is consistent with U.S. attitudes toward the WTO DSM and reflects an area of dispute settlement that is particularly amenable to the interests of major powers. Although we strongly advocate thinking about DSMs in a more holistic sense, these fine-tuned relationships are revealing and add to the richness of our study.

## 6 Conclusion

The goal of this research has been to use PTAs as a vehicle to explain why some international institutions have stronger dispute settlement provisions than others. Our study breaks new ground in several ways. One is that we examine a much greater number of agreements than previous studies, capturing considerable variation in nearly 600 PTAs across space and time. Similarly, we collect original data on dozens of

features of dispute settlement provisions in PTAs, which we employ for the first time in these empirical tests. These expanded data are very beneficial in allowing us to investigate a new, richer way of conceptualizing dispute settlement design. We move beyond the conventional notion of legalization and instead theorize DSMs as a comprehensive tool designed to help uphold deep trade obligations. Our core finding is that agreement depth is the strongest of all predictors for dispute settlement design. We offer the most systematic empirical test of the depth-enforcement link that underlies much of the theoretical debate about international institutions. This finding also indicates that governments view strong dispute settlement as necessary for achieving compliance with international obligations, and that compliance is not “managed” but instead is incentivized through deterrence and punishment (Downs et al. 1996). We also reconcile divergent perspectives on institutional design, showing that international agreements can reflect both rational design but also the unique needs and preferences of state and regional actors.

Several of our other findings also have implications beyond the study of PTAs or dispute settlement. We cast doubt on assertions that Asian culture is not amenable to formal dispute settlement or strong legal regimes. We also find that PTAs involving the U.S. are more likely to contain strong DSMs, which challenges the view that the U.S. is skeptical of legalization. Yet it also raises the possibility that state power might creep into realm of legal dispute settlement, which is supposed to neutralize such considerations. Maybe powerful actors design DSMs to suit their expected needs, and then choose from among the menu they have created when future disputes arise? Although this is a concern, we emphasize that asymmetric PTAs have the strongest dispute settlement, suggesting that both powerful and weak states can find mutual benefit from legalistic DSMs. Furthermore, even if some powerful states desire strong dispute settlement, we suspect most other states would be happy to see leading actors bound more closely to the rule of law.

Although our findings clearly demonstrate that post-war DSMs are uniquely designed, we observe some emerging convergence that is worth exploring further in future work. One pattern is that all DSMs are getting stronger over time, with many now including time limits and the ability to retaliate in the event of non-compliance. Perhaps governments are beginning to mimic one another when designing agreements? Some have suggested that institutional design features are likely to diffuse (Jetschke and Lenz 2013). Indeed, recent empirical work shows that treaty language is often replicated from one treaty to the next (Allee and Elsig 2014) and that a handful of models may exist (Baccini et al. 2015a, b).

There now exist hundreds of international agreements worldwide with legal dispute settlement provisions, including various features associated with timely resolution, selection of panelists, forum choice, and sanctioning, among others. These design features are likely to reappear in upcoming decades as more disputes begin to arise and the above features are evoked. Since the obligations in PTAs often take many years to fully “kick in,” some DSMs that are dormant in their early years are likely to become active down the line, as was the case with the ECJ, which was hardly used in the 1950s and 1960s (Alter 2012). Therefore we anticipate that DSMs will be used with increasing frequency in the near future, with disputes being resolved through varied diplomatic and legal channels, revealing further the importance of past design choices and the influences exerted upon them.

**Acknowledgments** We are grateful to the Swiss Science Foundation ([www.nccr-trade.org](http://www.nccr-trade.org)) for financial support and to our research assistants for help in collecting the data used in this paper. For helpful comments we thank Chad Bown, Christina Davis, Jeff Kucik, and the three anonymous reviewers.

**Compliance with ethical standards** This manuscript complies with all specified ethical standards.

## References

- Abbott, K. W., Keohane, R. O., Moravcsik, A., Slaughter, A., & Snidal, D. (2000). The concept of legalization. *International Organization*, 54(3), 401–19.
- Acharya, A. (1997). Ideas, identity, and institution-building: from the ‘ASEAN’ way’ to the “Asia Pacific Way”? *The Pacific Review*, 10(3), 319–46.
- Acharya, A., & Johnston, A. I. (Eds.). (2007). *Crafting cooperation: regional international institutions in comparative perspective*. Cambridge: Cambridge University Press.
- Aggarwal, V. (Ed.). (1998). *Institutional designs for a complex world*. Ithaca: Cornell University Press.
- Allee, T., & Elsig, M. (2014). *Are the contents of international treaties boilerplate? Evidence from preferential trade agreements*. Paper presented at the annual meeting of the American political science association, Washington, D.C., August 28–31, 2014.
- Allee, T., & Elsig, M. (2015). Dispute settlement provisions in PTAs: new data and new concepts. In A. Dür & M. Elsig (Eds.), *Trade cooperation: the purpose, design and effects of preferential trade agreements* (pp. 319–352). Cambridge: Cambridge University Press.
- Allee, T., & Huth, P. K. (2006). International adjudication as domestic political cover. *American Political Science Review*, 100(2), 219–34.
- Allee, T., & Peinhardt, C. (2010). Delegating differences: bilateral investment treaties and patterns of dispute resolution design. *International Studies Quarterly*, 54(1), 1–26.
- Allee, T., & Peinhardt, C. (2014). Evaluating three explanations for the design of bilateral investment treaties. *World Politics*, 66(1), 47–87.
- Alter, K. (2012). The global spread of European style courts. *West European Politics*, 35(1), 135–54.
- Alter, K., & Helfer, L. (2010). Nature or nurture? Judicial lawmaking in the European court of justice and the Andean tribunal of justice. *International Organization*, 64(4), 563–92.
- Alter, K., & Meunier, S. (2009). The politics of international regime complexity. *Perspectives on Politics*, 7(1), 13–24.
- Arnold, C., & Rittberger, B. (2013). The legalization of dispute resolution in mercosur. *Journal of Politics in Latin America*, 5(3), 97–132.
- Ásgeirsdóttir, A., & Steinwand, M. (2014). *Dispute resolution mechanisms and maritime boundary settlements*. Paper Presented at the annual meeting of the international studies association. Toronto, Canada.
- Baccini, L., Dür, A. & Elsig, M. (2015a). The Politics of trade agreement design: revisiting the depth–flexibility Nexus. *International Studies Quarterly*.
- Baccini, L., Dür, A., & Hafel, Y. (2015b). Imitation and innovation in international governance: the diffusion of trade agreement design. In A. Dür & M. Elsig (Eds.), *Trade cooperation: the purpose, design and effects of preferential trade agreements* (pp. 167–194). Cambridge: Cambridge University Press.
- Bartels, L., & Ortino, F. (2006). *Regional trade agreements and the WTO legal system*. Oxford: Oxford University Press.
- Berger, A., Busse, M., Nunnenkamp, P., & Roy, M. (2010). Do trade and investment agreements lead to more FDI? Accounting for key provisions inside the black box. *Kiel working paper, No. 1647*.
- Bernauer, T., Elsig, M., & Pauwelyn, J. (2012). The WTO dispute settlement mechanism: analysis and problems. In A. Narlikar, M. Daunton, & R. M. Stern (Eds.), *Oxford handbook on the world trade organization* (pp. 485–506). Oxford: Oxford University Press.
- Bernauer, T., Kalbhenn, T., & Spilker, G. (2013). Is there a ‘depth versus participation’ dilemma in international cooperation? *Review of International Organizations*, 8(4), 477–97.
- Biukovic, L. (2008). Dispute resolution mechanisms and regional trade agreements: South American and Caribbean modalities. *U.C. Davis Journal of International Law and Policy*, 14(2), 255–96.
- Bown, C. (2004). On the economic success of GATT/WTO dispute settlement. *The Review of Economics and Statistics*, 86(3), 811–23.
- Bown, C., & Pauwelyn, J. (2010). *The law, economics and politics of retaliation in WTO dispute settlement*. Cambridge: Cambridge University Press.



- Brewster, R. (2006). Rule-based dispute resolution in international trade law. *Virginia Law Review*, 92, 251–88.
- Broude, T. (2004). From Pax Mercatoria to Pax Europea: How trade dispute procedures serve the EC's Regional Hegemony. Available at SSRN: <http://ssrn.com/abstract=724641>.
- Busch, M. (2007). Overlapping institutions, forum shopping, and dispute settlement in international trade. *International Organization*, 61(4), 735–61.
- Chase C., Yanovich, A., Crawford, J., & Ugaz, P. (2013). Mapping of dispute settlement mechanisms in regional trade agreements - innovative or variations on a theme? *WTO Staff Working Paper* (ERSD-2013-07).
- Chayes, A., & Chayes, A. H. (1993). On compliance. *International Organization*, 47(2), 175–205.
- Copelovitch, M. S., & Putnam, T. L. (2014). Design in context: existing international agreements and new cooperation. *International Organization*, 68(2), 471–93.
- Davis, C. (2009). Overlapping institutions in trade policy. *Perspectives on Politics*, 7(1), 25–31.
- De Mestral, A. (2006). NAFTA dispute settlement: creative experiment or confusion? In L. Bartels & F. Ortino (Eds.), *Regional trade agreements and the WTO legal system* (pp. 359–81). Oxford: Oxford University Press.
- Dominguez, J. (2007). International cooperation in Latin America: the design of regional institutions by slow accretion. In A. Acharya & A. I. Johnston (Eds.), *Crafting cooperation: regional international institutions in comparative perspective* (pp. 83–128). Cambridge: Cambridge University Press.
- Donaldson, V., & Lester, S. (2009). Dispute settlement. In S. Lester & B. Mercurio (Eds.), *Bilateral and regional trade agreements: commentary and analysis* (pp. 367–414). Cambridge: Cambridge University Press.
- Downs, G. W., Rocke, D. M., & Barsoom, P. N. (1996). Is the good news about compliance good news about cooperation? *International Organization*, 50(3), 379–406.
- Drezner, D. (2007). *All politics is global: explaining international regulatory regimes*. Princeton: Princeton University Press.
- Drezner, D. (2013). The tragedy of the global institutional commons. In M. Finnemore & J. Goldstein (Eds.), *Back to basics: state power in a contemporary world* (pp. 280–312). Oxford: Oxford University Press.
- Duffield, J. (2003). The limits of “rational design”. *International Organization*, 57(2), 411–30.
- Dür, A., Baccini, L., & Elsig, M. (2014). The design of international trade agreements: introducing a new database. *The Review of International Organizations*, 9(3), 353–75.
- Elsig, M., & Eckhardt, J. (2015). The Creation of the Multilateral Trade Court: Design and Experiential Learning. *World Trade Review*.
- Gaathi, J. T. (2010a). The under-appreciated jurisprudence of Africa's regional trade judiciaries. *Oregon Review of International Law*, 12, 245–82.
- Gaathi, J. T. (2010b). African regional trade agreements as flexible legal regimes. *North Carolina Journal of International Law and Commercial Regulation*, 35, 571–667.
- Garcia Bercero, I. (2006). Dispute settlement in European Union free trade agreements: lessons learned? In L. Bartels & F. Ortino (Eds.), *Regional trade agreements and the WTO legal system* (pp. 383–405). Oxford: Oxford University Press.
- Gleditsch, K. (2002). Expanded trade and GDP data. *Journal of Conflict Resolution*, 46(5), 712–24.
- Goldstein, J., Kahler, M., Keohane, R., & Slaughter, A. (2000). Introduction: legalization and world politics. *International Organization*, 54(3), 385–99.
- Gomez-Mera, L., & Molinari, A. (2014). Overlapping institutions, learning, and dispute initiation in regional trade agreements: evidence from South America. *International Studies Quarterly*, 58(2), 269–81.
- Gray, J., & Slapin, J. (2013). Exit options and the effectiveness of regional economic organizations. *Political Science Research and Methods*, 1(2), 281–303.
- Grimmett, J. J. (2012). Dispute Settlement in the U.S.-South Korea free trade agreements (KORUS FTA). *CRS Report for Congress*, March 21, 2012.
- Gruber, L. (2000). *Ruling the world: power politics and the rise of supranational institutions*. Princeton: Princeton University Press.
- Haftel, Y. (2012). *Regional economic institutions and conflict mitigation*. Ann Arbor: University of Michigan Press.
- Haftel, Y. (2013). Commerce and institutions: trade, scope and the design of regional economic organizations. *Review of International Organizations*, 8(3), 389–414.
- Haftel, Y., & Thompson, A. (2006). The independence of international organizations: concept and applications. *Journal of Conflict Resolution*, 50(2), 253–75.

- Herbst, J. (2007). Crafting regional cooperation in Africa. In A. Acharya & A. I. Johnston (Eds.), *Crafting cooperation: regional international institutions in comparative perspective* (pp. 129–144). Cambridge: Cambridge University Press.
- Hooghe, L., Bezuijen, J., Derderyan, S., & Coman, E. (2014). *Designing third party dispute settlement for international organizations*. Paper presented at the 7th Annual conference on the political economy of international organizations, Princeton University, January 16–18, 2014.
- Jagers, K., & Gurr, T. R. (1995). Tracking democracy's third wave with the polity III data. *Journal of Peace Research*, 32(4), 469–82.
- Jetschke, A., & Lenz, T. (2013). Does regionalism diffuse? a New research agenda for the study of regional organizations. *Journal of European Public Policy*, 20(4), 626–37.
- Jo, H., & Namgung, H. (2012). Dispute settlement mechanisms in preferential trade agreements: democracy, boilerplates, and the multilateral trade regime. *Journal of Conflict Resolution*, 56(6), 1041–68.
- Kahler, M. (2000). Legalization as strategy: the Asia-pacific case. *International Organization*, 54(3), 549–71.
- Keohane, R. O., Moravcsik, A., & Slaughter, A. (2000). Legalized dispute resolution: interstate and transnational. *International Organization*, 54(3), 457–88.
- Khong, Y. F., & Nesadurai, H. E. S. (2007). Hanging together, institutional design, and cooperation in southeast Asia: AFTA and the ARF. In A. Acharya & A. I. Johnston (Eds.), *Crafting cooperation: regional international institutions in comparative perspective* (pp. 32–82). Cambridge: Cambridge University Press.
- Kono, D. Y. (2007). Making Anarchy Work: International Legal Institutions and Trade Cooperation. *Journal of Politics*, 69(3), 746–59.
- Koremenos, B. (2007). If Only Half of International Agreements have Dispute Resolution Provisions, Which Half Needs Explaining? *The Journal of Legal Studies*, 36(1), 189–212.
- Koremenos, B., & Betz, T. (2013). The Design of Dispute Settlement Procedures in International Agreements. In J. Dunoff & M. Pollack (Eds.), *Interdisciplinary Perspectives on International Law and International Relations* (pp. 371–93). Cambridge: Cambridge University Press.
- Koremenos, B., Lipson, C., & Snidal, D. (2001). The Rational Design of International Institutions. *International Organization*, 55(4), 761–99.
- Laks-Hutnick, J. A. (2013). *Learning by doing or learning from others? Diffusion of information in international trade dispute settlement mechanisms*. Paper prepared for the 2013 annual meeting of the international studies association, San Francisco, CA, April 4–6.
- Long, J. S., & Freese, J. (2014). *Regression models for categorical dependent variables using Stata* (3rd ed.). College Station, TX: Stata Press.
- Luo, Y. (2006). Dispute settlement in the proposed east Asia free trade agreement: lesson learned from the ASEAN, the NAFTA, and the EU. In Bartels & Ortino (Eds.), *Regional trade agreements and the WTO legal system* (pp. 419–45). Oxford: Oxford University Press.
- Maggi, G. (1999). The role of multilateral institutions in international trade cooperation. *American Economic Review*, 89(1), 190–214.
- Mansfield, E. D., & Milner, H. V. (2012). *Votes, vetoes, and the political economy of international trade agreements*. Princeton: Princeton University Press.
- Marcoux, C. (2009). Institutional flexibility in the design of multilateral environmental agreements. *Conflict Management and Peace Science*, 26(2), 209–28.
- Martin, L. (1992). Interests, power, and multilateralism. *International Organization*, 46(4), 765–92.
- Mitchell, R. (1994). Regime design matters: intentional oil pollution and treaty compliance. *International Organization*, 48(3), 425–58.
- Morgan, D. (2008). *Dispute settlement under PTAs: political or legal? Legal studies research paper 341*. Melbourne: Melbourne University Law School.
- Paulson, C. (2004). Compliance with final judgments of the international court of justice since 1987. *American Journal of International Law*, 98, 434–61.
- Peel, M. J., Goode, M. M. H., & Moutinho, L. A. (1998). Estimating consumer satisfaction: OLS versus ordered probability models. *International Journal of Commerce and Management*, 8(2), 75–93.
- Peinhardt, C., & Allee, T. (2012). Failure to deliver: the investment effects of U.S. preferential trade agreements. *The World Economy*, 35(6), 757–83.
- Pelc, K. (2009). Seeking escape: escape clauses in international trade agreements. *International Studies Quarterly*, 53(2), 349–68.
- Porges, A. (2011). Dispute settlement. In J.-P. Chaffour & J.-C. Maur (Eds.), *Preferential trade agreement policies for development: a handbook* (pp. 467–501). Washington, D.C.: The World Bank.
- Posner, E. A. (2009). *The perils of global legalism*. Chicago: University of Chicago Press.

- Posner, E. A., & Yoo, J. C. (2005). Judicial independence in international tribunals. *California Law Review*, 93(1), 1–74.
- Raustiala, K., & Victor, D. G. (2004). The regime complex for plant genetic resources. *International Organization*, 58(2), 277–309.
- Rosendorff, B. P. (2005). Stability and rigidity: politics and the design of the WTO's dispute resolution procedure. *The American Political Science Review*, 99(3), 389–400.
- Sattler, T., & Bernauer, T. (2011). Gravitation or discrimination? Determinants of litigation in the world trade organization. *European Journal of Political Research*, 50(2), 143–67.
- Schneider, A. K. (1999). Getting along: the evolution of dispute resolution regimes in international trade organizations. *Michigan Journal of International Law*, 20(4), 697–773.
- Schropp, S. (2009). *Trade policy flexibility and enforcement in the world trade organization*. Cambridge: Cambridge University Press.
- Simmons, B. A. (2014). Bargaining over BITs, arbitrating awards: the regime for protection and promotion of international investment. *World Politics*, 66(1), 12–46.
- Smith, J. M. (2000). The politics of dispute settlement design. *International Organization*, 54(1), 137–80.
- Stone, R. W. (2011). *Controlling institutions: international organizations and the global economy*. Cambridge: Cambridge University Press.
- Stone Sweet, A., & Brunell, T. (1999). Constructing a supranational constitution: dispute resolution and governance in the European community. *The American Political Science Review*, 92(1), 63–82.
- Victor, D. G., Raustiala, K., & Skolnikoff, E. B. (Eds.). (1998). *The implementation and effectiveness of international environmental commitments: theory and practice*. Cambridge: MIT Press.
- von Stein, J. (2012). The engines of compliance. In J. Dunoff & M. Pollack (Eds.), *Interdisciplinary perspectives on international law and international relations* (pp. 477–501). Cambridge: Cambridge University Press.
- Wendt, A. (2001). Driving with the rearview mirror: on the rational science of institutional design. *International Organization*, 55(4), 1019–49.
- Yarbrough, B. V., & Yarbrough, R. M. (1992). *Cooperation and governance in international trade*. New York: Columbia University Press.
- Yarbrough, B. V., & Yarbrough, R. M. (1997). Dispute settlement in international trade: regionalism and procedural coordination. In E. D. Mansfield & H. V. Milner (Eds.), *The political economy of regionalism* (pp. 134–63). New York: Columbia University Press.
- Zangl, B. (2008). Judicialization matters! a comparison of dispute settlement under the GATT and WTO. *International Studies Quarterly*, 52(4), 825–54.